

XII Congresso Brasileiro de Agroinformática

11 a 14 de Novembro de 2019

ANAIS

INDAIATUBA
Novembro de 2019

ANAIS DO XII CONGRESSO BRASILEIRO DE AGROINFORMATICA

ORGANIZADORES

MARIA FERNANDA MOURA

JAYME GARCIA ARNAL BARBEDO

ALAINE MARGARETE GUIMARÃES

VALTER CASTELHANO DE OLIVEIRA

FATEC/EMBRAPA/UEPG-PPGCA

11 A 14 DE NOVEMBRO DE 2019

Ficha Catalográfica

Congresso Brasileiro de Agroinformática (11.: 2019: Indaiatuaba, SP)

C749 Anais eletrônicos do XII Congresso Brasileiro de Agroinformática. 11 a 14 de novembro de 2019. Indaiatuba, SP; Universidade Estadual de Ponta Grossa/ Maria Fernanda Moura et al. (Org.). Ponta Grossa: SBIAGRO, 2019.

ISBN: 978-65-00-10242-0

1. Agroinformática. 2. Agricultura digital. 3. Internet das coisas. I. Moura, Maria Fernanda (Org.). II. Barbedo, Jayme Garcia Arnal (Org.). III. Guimarães, Alaine Margarete (Org.). IV. Oliveira, Valter Castelhano de (Org.). V. SBIAgro. VI. T.

CDD: 004.22

ISBN: 978-65-00-10242-0
9 786500 102420


Organização:
UEPG
Universidade Estadual de Ponta Grossa
Fatec
Indaiatuba
Dr. Archimedes Lammoglia
Embrapa
Informática Agropecuária
Patrocínio:
VISIONA
Tecnologia Espacial
JOHN DEERE
Apoio:
PREFEITURA DE INDAIATUBA
PENSANDO EM VOCÊ
Museu da Água
Promoção:
SF AGRO

Comissões do Evento

Comissão Organizadora

Presidente: Maria Fernanda Moura (Embrapa Informática Agropecuária)

Vice-Presidente: Valter Castelhano de Oliveira (Fatec)

Comissão de Iniciação Científica

Presidente: Valter Castelhano de Oliveira (Fatec)

Vice-Presidente: Luciana Alvim Romani (Embrapa Informática Agropecuária)

Comissão Científica

Presidente: Jayme Garcia Arnal Barbedo (Embrapa Informática Agropecuária)

Vice-Presidente: Alaine Margarete Guimarães (UEPG)

Adauto Mancini (Embrapa Informática Agropecuária)

Ana Maria Heuminski de Avila (Cepagri/Unicamp)

Ana Paula Lüdtke Ferreira (Unipampa)

Anderson Toledo (IAPAR)

Andre Minitti (Embrapa Informática Agropecuária)

Antonio Lino – (IAC)

Ariovaldo Luchiari Jr. (Embrapa Informática Agropecuária)

Aryeverton Oliveira (Embrapa Informática Agropecuária)

Augusto Guilherme de Araújo (IAPAR)

Bruno Albertini (PCS/USP)

Bruno Nogueira (UFMS)

Carlos Amaral Hölbig (UPF)

Celia Regina Grego (Embrapa Informática Agropecuária)

Celina Takemura (Embrapa Territorial)

Claudio Bazzi (UTFPR)

Daniel Victoria (Embrapa Informática Agropecuária)

Eduardo Augusto Bezerra (UFSC)

Eduardo Speranza (Embrapa Informática Agropecuária)

Emerson Medeiros Del Ponte (UFV)

Francisco Arnold (FT/UNICAMP)

Glauber Vaz (Embrapa Informática Agropecuária)

Hana Karina Salles Rubinsztejn (UFMS)

Helano Póvoas de Lima (Embrapa Informática Agropecuária)

Heloisa Camargo (UFSCAR)

Hevandro Delalibera (IAPAR)

Irineu Sotoma (UFMS)

Jésus Franco Bueno (UFMT)

João Batista Tolentino Júnior (UFSC)

João Camargo Neto (Embrapa Informática Agropecuária)

João Francisco Gonçalves Antunes (Embrapa Informática Agropecuária)

José Carlos Ferreira da Rocha (UEPG)

Juliana Saragiotto Silva (IFMT)

Juliana Yassitepe (Embrapa Informática Agropecuária)

Júlio César Dalla Mora Esquerdo (Embrapa Informática Agropecuária)

Kleber Xavier Sampaio de Souza (Embrapa Informática Agropecuária)

Leonardo Ribeiro Queiros (Embrapa Informática Agropecuária)

Lubia Vinhas (INPE)

Luciana Alvim Santos Romani (Embrapa Informática Agropecuária)

Luciano Gonda (UFMS)

Luciano Koenigkan (Embrapa Informática Agropecuária)

Luiz Camolesi Júnior (FT/UNICAMP)

Marcos Visoli (Embrapa Informática Agropecuária)

Maria do Carmo Fasiaben (Embrapa Informática Agropecuária)

Maria Fernanda Moura (Embrapa Informática Agropecuária)

Maria Salete Marcon Gomes Vaz (UEPG)

Maurício Pereira (UFMT)

Nuno Escudeiro (ISEP)

Rafael Rossi

Raul Teruel dos Santos (UFMT)

Roberto H. Higa (Embrapa Informática Agropecuária)

Sandro Camargo (Unipampa)

Sergio Manuel Serra da Cruz (UFRRJ)

Sonia Ternes (Embrapa Informática Agropecuária)

Tercia Torres (Embrapa Informática Agropecuária)

Thiago Teixeira Santos (Embrapa Informática Agropecuária)

Tiago Carvalho (IFSP)

Valter Castelhano de Oliveira (Fatec Indaiatuba)

Vilmar Rodrigues Moreira (PUC-PR)

Willingthon Pavan (UPF)

Comissão do SBIAgro Conect@

Ariovaldo Luchiari Junior (Embrapa Informática Agropecuária)

Fernando Matsuura (Embrapa - Secretaria de Inovação e Negócios)

Luiz Manoel da Silva Cunha (Embrapa Informática Agropecuária)

Lincon Moreira Peretto (Fatec)

Vinícius Kuromoto (Embrapa Informática Agropecuária)

Comissão de Comunicação

Magda Cruciol (Embrapa Informática Agropecuária)

Flávia Bussaglia Fiorini (Embrapa Informática Agropecuária)

Maria Giulia Croce (Embrapa Informática Agropecuária)

Nadir Rodrigues (Embrapa Informática Agropecuária)

Graziella Galinari (Embrapa Informática Agropecuária)

Vera Scholze Borges (Embrapa Informática Agropecuária)

José Luiz Marques (Fatec)

Felipe Jaconi (Embrapa Informática Agropecuária)

Comissão de Infraestrutura e Logística

Sandro Roberto da Silva Calabrezi (Fatec)

Karina Cavalieri (Fatec)

Comissão de Credenciamento e Suporte

Tercia Torres (Embrapa Informática Agropecuária)

Angelica Felix Lima (Fatec)

Josilma Aparecida da Silva (Fatec)

Kelly Janaine Amaral (Fatec)

Jhenifer Alves Vieira (Fatec)

Zâmbia Taiá Donato dos Santos (Fatec)

Damaris Teixeira de Souza (Fatec)

Andressa Pereira de Oliveira (Fatec)

Gisele Moreno Ferreira Ducatti (Fatec)

Eduarda Ribeiro Luchini Abreu (Fatec)

Paula Valeska dos Santos Mendes (Fatec)

Jonathan Lucas de Matos Diniz (UEPG)

Comissão Financeira e Patrocínio

Maria Fernanda Moura (Embrapa Informática Agropecuária)

Luiz Antonio Daniel (Fatec)

Fernando Matsuura (Embrapa - Secretaria de Inovação e Negócios)

Vinícius Kuromoto (Embrapa Informática Agropecuária)

Luciana Alvim Romani (Embrapa Informática Agropecuária)

Guiomar Alessandra Souza Telles (Embrapa Informática Agropecuária)

Apresentação

O Congresso Brasileiro de Agroinformática (SBIAgro 2019) é o principal evento científico da área de informática aplicada à agricultura no Brasil. Tem como objetivo promover o compartilhamento de resultados de pesquisas, a troca de ideias sobre trabalhos em andamento e a inovação no setor. Na sua 12ª edição, teve como tema principal "IoT na Agricultura", visão atual e futura.

Promovido pela Associação Brasileira de Agroinformática (SBIAgro), o congresso foi realizado entre os dias 11 e 14 de novembro de 2019, na Faculdade de Tecnologia de São Paulo (Fatec), campus de Indaiatuba (SP) - uma região que se destaca pela forte presença de empresas de alta tecnologia e de universidades com reconhecimento internacional.

O evento foi composto por painéis, palestras, apresentação de trabalhos científicos, concurso de teses e dissertações, além de um concurso de iniciação científica. Nesta edição, houve também espaço para resumos, na categoria pôster, a fim de propiciar a apresentação de soluções de ferramentas, processos tecnológicos ou resultados parciais de PD&I.

O SBIAgro 2019 contou com a segunda edição do Conect@, evento paralelo de relacionamento entre instituições, empresas, investidores e desenvolvedores de Tecnologias de Informação e Comunicação (TIC), que buscou proporcionar um ambiente para networking e articulação de negócios e parcerias em pesquisa, desenvolvimento e inovações. A ideia é que a integração de diferentes perfis, resultantes da interdisciplinaridade inerente à Agroinformática, estimule o surgimento de novos projetos e negócios.

A organização do congresso esteve sob responsabilidade da Fatec Indaiatuba, Embrapa Informática Agropecuária e Universidade Estadual de Ponta Grossa (UEPG) representada pelo seu Programa de Pós-Graduação em Computação Aplicada.

Maria Fernanda Moura
Presidente do XII SBIAGRO

SUMÁRIO

ARTIGOS COMPLETOS

Simulação computacional e análise ambiental de processos de valorização do resíduo da indústria de suco de laranja 1
Lorrayne Lins Suzuki, Moisés Teles dos Santos, Cláudia Echevenguá Teixeira

Estação agrometeorológica experimental automatizada 11
Carlos Marcelo Zys, Felipe de Oliveira Gomes, Kelly Lais Wiggers

Rastreando a Cadeia de Carne Brasileira com Blockchain 21
Vasco Varanda Picchi, José Reynaldo Formigoni Filho, Eduardo Ferreira Jucá de Castro, Sérgio Luís Ribeiro

Caracterização morfométrica de uma bacia hidrográfica utilizando técnicas de geoprocessamento 31
Alexandro Gularte Schafer, John Roger Roldan Aleixo, Allan Sampaio Pires

Algoritmos Genéticos para Otimização em Sistemas de Produção de Ruminantes em Pastagens 41
Joel de Oliveira Santos, Laura Emmanuella Alves dos Santos Santana, Henrique Rocha de Medeiros

Predição da germinação de sementes de soja usando aprendizado de máquina 51
Iaslan do Nascimento Paulo da Silva, Wesley Leocádio Barbosa da Silva, Laura Emmanuella Alves dos Santos Santana, Márcio Dias Pereira

Uma API para a classificação de solos do Brasil 63
Glauber José Vaz, Luís de França da Silva Neto, Renato Nunes Lima, Stanley Robson de Medeiros Oliveira

OpenVRT - A smartphone based wireless variable rate solution 73
Luiz Guilherme Fonseca Rosa, Rogério Schneider, Marcio Seiji Oyamada

Rede de Sensores Sem Fio Subterrânea: Análise e Teste de Comunicação 82
Tércio A. S. Filho, Alexandre C. R. Silva, Marcos Napoleão Rabelo, Sérgio F. Silva,Juliano Mendes Ribeiro, José Carlos C. G. Filho

EaD na Agricultura: análise do perfil discente de um curso de compostagem à distância 92
Hugo Rogerio Borges de Freitas, Flávio Lages Monteiro Júnior, Ernani Jardins Reis, Marco Antônio de Almeida Leal, Ana Cristina Siewert Garofolo

Predição da Produção de Leite Caprino utilizando Redes Neurais Artificiais 102
Vitória Mendes da Silva, Lucas Alberto Teixeira de Rezende, João Marcelo Ribeiro, Aline Marques Del Valle

Identificação de pontos de estabilidade e flexibilidade no domínio de formação de preço de venda de produtos da horticultura aplicado à agricultura familiar 112
Giovane Galvão, Simone Nasser Matos

Comunicação modular para uma estação meteorológica 122
Thiago Piovesan, Marcelo Felipe Chan Yu, José Alberto dos Santos, Antonio Marcos Massao Hachisuca

Segmentação de imagens de vants para detecção de ferrugem da soja 131
Gercina Gonçalves da Silva, Ariadne Barbosa Gonçalves, Junior Silva Souza, Diogo Soares da Silva, Nícolas Alessandro de Souza Belete, Denilson de Oliveria Guilherme, Marcelo Theophilo Folhes, Hemerson Pistori

Espacialização de dados de temperatura do ar e soma térmica para o Rio Grande do Sul no contexto do manejo do Capim-Sudão BRS Estribo 142
Allan Sampaio Pires, Gustavo Trentin, Alexandro Gularte Schafer

Modeling technology diffusion of beef cattle in Brazil using a cellular automata 152
Gustavo Gianotto Bratfisch, Luis Gustavo Barioni, Rafael de Oliveira Silva, Célia Regina Grego, Bruno Henrique Pereira, Jair Bertolucci Junior

OpenSoils: Uma Plataforma de Apoio à Ciência do Solo 163
Sérgio Manuel Serra da Cruz, Filipe Klinger, Alexia Duarte, Pedro Vieira Cruz, Elton Carneiro Marinho, Eber Assis Schmitz

Visão Computacional Aplicada ao Monitoramento Bovino: Uma Análise dos Biomas Nacionais 174
Matheus Serrão Marinato, Jayne Ribeiro da Silva Machado, Wanderlany Fialho de Abreu, Márcio José Moutinho da Ponte

Diagnóstico Automático de Ceratoconjuntivite Infecciosa Bovina por meio de Imagens Termográficas e Deep Learning 183
Dhyonatan Santos de Freitas, Sandro da Silva Camargo, Helena Brocardo Comin, Robert Domingues, Emanuelle Gaspar Baldo, Fernando Flores Cardoso

Quantificação da severidade de doenças foliares em plantas e manutenção de um banco de imagens digitais 193
João Batista Tolentino Júnio, Amanda do Prado Matto, Ana Gabriela Mabe Martins, Adriana Terumi Itako

Proposta e avaliação de uma solução baseada em Fog Computing para a Pecuária Leiteira de Precisão 203
Márcio Spenst, Renata W. Suñé, Marcos J. Yokoo, Leonardo B. Pinho

Start Green: Uma proposta de ferramenta pedagógica às crianças para promover interações com a natureza 213
Ana Elisandra Campos da Silva, Leandro da Silva Camargo, Marcel da Silva Camargo

Modelagem e avaliação de desempenho de banco de dados espaço-temporal aplicado a experimento de manejo de pastagem 223
Ricardo Robaina, Alexandro Schaffer, Leonardo Pinho, Gustavo Trentin

POMODORO: Arquitetura para o monitoramento e controle de umidade do solo em um cultivo protegido de tomate 233
Leonardo Rodrigues de Lima Teixeira, Joffrey Peyrac de Albuquerque Galvão Junior, Anderson Patrício Fernandes dos Santos

Abordagem neural para previsão de peso de contusões no processo de transporte bovino 243
Flávia Morales Antunes, Ana Paula Ludtke Ferreira, Tiago Alan Cunha Nardino, Naylor Bastiani Perez, Fabiano Nunes Vaz, Ricardo Zambarda Vaz

Use of artificial intelligence on the agroindustrial domain 252
Roberto Fray da Silva, Fernando Xavier, Gustavo Marques Mostaço, Antonio Mauro Saraiva, Carlos Eduardo Cugnasca

Modelagem Computacional de fluxo em aplicador de agroquímico amparada por simulação com exploração de paralelismo 262
Mateus Specht, Nailor Bastiani Perez, Leonardo Bidesi de Pinho

Labeling images with the LabelIt! App 272
Vitor Trevisan, Thiago Luís Lopes Siqueira

Automação Inteligente para Coletores e Canais de Irrigação com Prototipação Arduíno e Sistema Web 282
André M. Silva, Ezequiel M. Muxito, Miguel M. Awad, Cezar L. Duarte, Matheus V. Silva

HORUS GREEN: Sistema para monitoramento do crescimento da alface (lactuca sativa l.) em cultivo hidropônico utilizando processamento digital de imagens 292
Leonardo Rodrigues de Lima Teixeira, Suéliton dos Santos Silva, Anderson Patrício Fernandes dos Santos

Agriflix plataforma de conhecimento sobre o agronegócio 302
Valter Castelhano de Oliveira, Luiz Antonio Daniel

Uso de redes neurais convolucionais para detecção de laranjas no campo 312
João Camargo Neto, Sônia Ternes, Kleber Xavier Sampaio de Souza, Inácio Henrique Yano , Leonardo Ribeiro Queiros

SOX: UMA PLATAFORMA PARA COLETA MASSIVA DE DADOS E O MONITORAMENTO CONTÍNUO DE VARIÁVEIS AMBIENTAIS 322
Leandro da Silva Camargo, Érico Marcelo Hoff do Amaral, Leandro Bochi da Silva Volk, Marcel da Silva Camargo

Analysis of a ZigBee Wireless Sensor Network Based on IEEE 1451 Standard 332
Tércio A. S. Filho, Juliano Mendes Ribeiro, Alexandre C. R. Silva, Sérgio F. Silva, Antonio Carlos de Oliveira Júnior

Segmentação de Imagem Obtida por RPA Asa Fixa na Detecção de Erosão em Área Agrícola 342
Yasmin Volmer Tullio, Ricardo Kwiatkowski da Silva, Selma Regina Aranha Ribeiro, Maria Salete Marcon Gomes Vaz, Jessyca Kawana Franquito

Modelagem e persistência de dados sensoriais para a pecuária de precisão 351
Leandro Borges da Silva1, Hana Karina Salles Rubinsztejn

Influência das diferentes faces de exposição ao sol nos índices vegetativos e relativo de clorofila em cafés especiais 361
Eduardo Antonio Speranza, Célia Regina Grego, Gustavo Costa Rodrigues, Ariovaldo Luchiari Junior

Comparação do desempenho entre modelos lineares para previsão do preço do café 371
Carolina Deina, Diego Solak Castanho, Matheus Henrique do Amaral Prates, Sergio Luiz Stevan Junior, Fernanda Cristina Corrêa, Hugo Siqueira, Flavio Trojan

Computer Vision and Image Processing for Detecting and Quantifying whiteflies: A Systematic Review 382
João Gabriel Junqueira da Silva, Marcelo Gonçalves Narciso, Cristhiane Gonçalves

Tipologia de municípios produtores de bovinos de corte no bioma amazônia 392
Maria do Carmo Ramos Fasiaben, Stanley Robson de M. Oliveira, Maxwell Merçon Tezolin Barros Almeida, Octávio Costa de Oliveira, Helano Póvoas de Lima

Requirements identification and proposal of a domain model for implementing IoT in the sugar supply chain 402
Roberto Fray da Silva, Gustavo Marques Mostaço, Carlos Eduardo Cugnasca

Difundindo a Agricultura Digital na Agricultura Familiar da Baixada Fluminense 412
Larissa Paiva, Pedro Vieira Cruz, Sérgio Manuel Serra da Cruz, Ana Cláudia de Macedo Vieira

Algoritmo genético aplicado no refinamento de adubação em experimento com alfaces 422
Vinícius Tessele, Augusto Vaghetti Luchese, Arnaldo Candido Junior

Impactos da IoT na Avicultura: um Mapeamento Sistemático 430
Felipe Dias de Oliveira, Alixandre Santana, Luis Filipe Alves Pereira

Agricultura Digital: Levantamento junto ao Produtor Rural na Região Metropolitana de Campinas 440
Maria Angelica de Andrade Leite, Daniel Azevedo Duarte, Silvia Maria Fonseca Silveira Massruhá

Dispositivo para Captura e Processamento de Imagens Térmicas para Detecção de Estresse Hídrico 450
Júlio Anderson de Oliveira Júnior, Marcelo Gonçalves Narciso

Identificação de padrões de produção e intensificação tecnológica de municípios produtores de bovinos de corte no bioma Pampa 459
Vinicius do Nascimento Lampert, Maria do Carmo Ramos Fasiaben, Urbano Gomes Pinto de Abreu, Elisa Kohler Osmari, Stanley Robson de Medeiros Oliveira, Maxwell Merçon Tezolin Barros Almeida, Octávio Costa de Oliveira, Helano Póvoas de Lima

Aplicação de técnicas de classificação em imageamento aéreo com RPA para determinação de classes de produtividade do trigo 469
Jonathan de Matos Diniz, Alessandro Stachak, Alaine Margarete Guimarães, Eduardo Fávero Caires

Proposta de uma estação meteorológica utilizando uma rede de longo alcance: LoRa 477
Guilherme Henrique Alves da Silva, Marco Antonio Franzin Leite, Renan Guides Diniz de Barros, Paulo Roberto Carvalho dos Santos, Juliana Saragiotto Silva

PisciculTec: um sistema WEB para gerenciamento e monitoramento de atividades para piscicultura 489
Mateus Resende Rodrigues, Igor Medeiros Vanderlei, Luis Filipe Alves Pereira, Anderson Fernandes de Alencar

Gerência de Dados de Processos de Genotipagem no GeneMaisLab 499
Eliseu Germano, Luana Rodrigues, Thiago Souza, Tereza Borba, Marcelo Narciso

Sistema Imageador Infravermelho para mensuração do diâmetro e volume de pêssegos 509
Edson d'Avila, Antonio Carlos Loureiro Lino, Inácio Maria Dal Fabbro, Ana Cristina da Silva

RESUMOS

Coletor automático de dados climáticos utilizando redes sem fio e o protocolo MQTT na transmissão das informações 519
Jonas Chalegra Boechat, Otávio Mendonça Dantas, Maurício Fernando Lima Pereira

Um Ambiente para Uso de Outorga Compartilhada de Água 521
Maria Fernanda Moura, Adauto Luiz Mancini , Raysa Masson Benatti, Lineu Neiva Rodrigues, Daniel Victoria, Luciana Alvim Romani, Stanley R Medeiros Oliveira, Leandro Eduardo Annibal Silva, Giovani Anhesini Bezerra, Gabriel Alves Francioli

Análise comparativa entre Sigfox, LoRa e NB-IoT 523
Cleiton Anderson Profilio dos Santos, Eli Aparecido Gabilon Fernandes, Willyan Cândido Silva, Hana Karina Salles Rubinsztejn

Desenvolvimento de metodologia para avaliar a homogeneidade de cobertura de aplicação em folhas de videira 525
Antonio Carlos Loureiro Lino, Antonio Odair Santos

CLIMMAPVIEW: Plataforma de disseminação de dados de precipitação para o estado do Rio de Janeiro 527
Gustavo Garcia dos Reis Nunes, Sidney Sampaio Alves, Gustavo Bastos Lyra, Ednaldo Oliveira dos Santos, Sergio Manuel Serra da Cruz

Modelagem e simulação de um sistema de subirrigação para cana de açúcar 530
Angel Pontin Garcia, Edna Carolina Moriones Polanía, Miguel Ángel Salgado Herrera

Estudos para um controlador de código-aberto em imageador de custo acessível para fenotipagem de plantas 533
Patric Moreto, Luciano Vieira Koenigkan, Thiago Teixeira Santos

Reconhecimento de Bovino Pantaneiro utilizando visão computacional através da Rede Neural Convolucional: Resultados Premilinares 535
Fabricio de Lima Weber, Vanessa Aparecida de Moraes Weber, Geazy Menezes, Adair da Silva Oliveira Junior, Daniela Arestides Alves, Marcus Vinicius Morais de Oliveira, Hemerson Pistori, Urbano Gomes Pinto de Abreu

Classificação de espécies de plantas através de redes neurais convolucionais 537
Luís Guilherme Ribeiro, Renann Rodrigues da Silva, José Carlos Ferreira da Rocha

Módulo sensor autônomo para aplicação de leituras comportamentais de corpos hídricos 539
Camila Cristina Gonçalves Teixeira Leite Moscibrocki, Weliton da Maia, Luiz Alberto Bordignon

Interdisciplinaridades e aplicações didáticas com uso de maquetes voltadas ao Agronegócio 541
Reinaldo Toso Júnior, Wellington Roque, Luis Borges Gouveia

Soluções tecnológicas de baixo custo aplicadas em pequenas propriedades rurais 543
Kelinne de Oliveira Guimarães, Moisés Laurence de Freitas Lima Junior

Arduino da pesquisa ao campo: Um sistema de cultivo de plantas automatizado 545
João Batista Tolentino Junior, Talia Marafon, Willian Macanan

Um sistema para determinação do consumo de água pelas plantas utilizando Arduíno e sensores de baixo custo 547
João Batista Tolentino Júnior, Erick Santo Machado

Identification of sugarcane straw availability through GIS tools 549
Rafaella Pironato Amaro, Bruna Cristina Gama Campagnuci, Lauren Maine Santos Menandro, Sergio Gustavo Quassi de Castro, Thayse Ap. Dourado Hernandes, João Luis Nunes Carvalho, Ana Claudia dos Santos Luciano

Análise das Tendências de Pesquisa no Congresso Brasileiro de Agroinformática por Nuvem de Palavras 551
Bruno Ernesto Techera da Motta, Alfredo Parteli Gomes

Eventos climáticos extremos: Uma análise do clima de Campinas 553
Lucas H. Ueda, Paula D. P. Costa, Ana M. H. de Avila, Eliana C. de Faria

Desenvolvimento de um sistema em rede estrela para administração de estufas inteligentes com IOT 555
Igor Mauro Silva De Almeida,Iraildo Barbosa, Everlaine Leopoldino Dias, Alberto E. Pereira de Araujo

SARIPAN - Sistema de Alerta de Risco de Incêndio para o Pantanal 557
Marcelo Gonçalves Narciso, Balbina Maria Araújo Soriano

Previsão de Indicadores Econômicos Individuais na Pecuária Leiteira 559
Helison Cavalheiro, Thales Maciel, Oldemar Weiller, Gustavo da Silva, Vinicius Lampert

Deep Learning como ferramenta para classificação visual de folhas de soja 561
Eduardo Cintra, Alysson A. Naves Silva, Darlan Einstein do Livramento

Uso do 5G em Área Remota 563
Cledson Akio Sakurai, Cecilia Emi Yamanaka Matsumura

Monitoramento Digital de Doenças em Tomateiros Baseado em Aplicativos Datacêntricos 565
Nicolas T. Maia, Guilherme Andrade; Sérgio Manuel Serra da Cruz

INICIAÇÃO CIENTÍFICA

Utilizando processamento em cascata e agrupamento em imagens para otimizar modelos de classificação de solos 569
Gabriel Teston Vasconcelos, Kleber Xavier Sampaio de Souza, João Camargo Neto, Stanley Robson de Medeiros Oliveira

Estudo de ferramentas para design e modelagem de sistemas de serviço aplicado ao Agronegócio 576
Maria de Jesus de Souza, Amanda Roberta Camargo, Valter Castelhano de Oliveira, Edson Luiz Pereira

Development and analysis of a labeled dataset of Brazilian agricultural market news titles for sentiment analysis 587
Tiago Santa Maria Rodrigues Marto, Roberto Fray da Silva, Angel Felipe Magnossão de Paula, Anna Helena Reali Costa, Carlos Eduardo Cugnasca

SVM com DenseNet para classificação de mudas do guaraná a partir da nervura central do folíolo 597
Brenda Talyne Costa Martins, Marcos Filipe Alves Salame

Detecção de Folíolos do Guaranazeiro Atacados pelo Tripes Utilizando CNNs Pré-Treinadas 605
Richardson Allan Ferreira de Souza, Marcos Filipe Alves Salame

Simulação computacional e análise ambiental de processos de valorização do resíduo da indústria de suco de laranja

Lorrayne Lins Suzuki[1], *Moisés Teles dos Santos*[1], *Cláudia Echevenguá Teixeira*[2]

[1] Departamento de Engenharia Química, Universidade de São Paulo, São Paulo, Brasil, lorraynesuzuki@usp.br, moises.teles@usp.br

[2] Centro de Tecnologias Geoambientais, Instituto de Pesquisas Tecnológicas, São Paulo, Brasil, cteixeira@ipt.br

RESUMO

O Brasil é o maior produtor de laranja e de suco de laranja do mundo. Os resíduos originados na produção do suco de laranja correspondem de 50% a 60% da massa inicial da fruta *in natura*. O grande número de alternativas possíveis para aproveitamento do resíduo da indústria de suco de laranja sugere uma abordagem de síntese de biorrefinarias, capaz de incluir todas as possíveis combinações de tecnologias de processamento em uma única unidade de processamento. O objetivo desse trabalho consiste no emprego de metodologias de simulação de processos e uso de métricas ambientais para um cenário hipotético brasileiro de aproveitamento do resíduo da indústria de suco de laranja, abordando as diversas rotas de aproveitamento. Nesse trabalho é proposto o desenvolvimento de uma biorrefinaria capaz de produzir D-limoneno, pectina, biogás e energia, a partir do aproveitamento das cascas de laranja. A etapa de modelagem e simulação foi desenvolvida no software ASPEN PLUS V. 8.4 e um levantamento dos principais aspectos e impactos dos processos foi realizado para estabelecer os potenciais impactos e significâncias. A partir de 111,1 kg/s de casca de laranja foram produzidos 1,49 kg/s de D-limoneno, 4,59 kg/s de pectina, 16,51 MW no ciclo combinado e 35,38 MW com biometano. Em relação aos impactos ambientais, obteve-se que os processos de reaproveitamento com maiores impactos ambientais foram as produções de pectina, energia, biometano e D-limoneno, respectivamente.

PALAVRAS-CHAVE: Biorrefinaria, Citricultura, Aspen Plus, Aspectos ambientais, Impactos ambientais.

ABSTRACT

Brazil is the largest producer of orange and orange juice in the world. The residues originating from the production of orange juice correspond to 50% to 60% of the initial fruit mass in natura. The large number of possible alternatives for harnessing the residues of orange juice industry suggests a biorefinery synthesis approach, capable of including all possible combinations of processing technologies in a single processing unit. The objective of this work is the use of process simulation methodologies and the use of environmental metrics for a Brazilian hypothetical scenario of waste valorization in the orange juice industry, addressing the different waste recovery routes. This work proposes the development of a biorefinery capable of producing D-limonene, pectin, biogas and energy from the use of orange peels. The modeling and simulation stage was develop in ASPEN PLUS software V. 8.4 and a survey of the main aspects and impacts of the processes was carried out to establish the potential impacts and significance. A mass flow of 111.11 kg/s of orange peel produced 1.49 kg/s of D-limonene, 4.59 kg/s of pectin, 16,51 MW in the combined cycle and 35.38 MW with biomethane. On the environmental impacts, it was obtain that the processes of reutilization with greater environmental impacts were the productions of pectin, energy, biomethane and D-limonene, respectively.

KEYWORDS: Biorefinery, Citrus, Aspen Plus, Environmental aspects, Environmental impacts.

INTRODUÇÃO

O Brasil é um país voltado para o agronegócio e apresenta uma agroindústria robusta, diversificada e lucrativa. Dentre os principais cultivos brasileiros encontra-se o de laranja. De acordo com o Departamento de Agricultura dos Estados Unidos, a produção mundial de laranja referente à safra 2017/2018 foi de 47,75 milhões toneladas, sendo o Brasil responsável por 34% desse valor (USDA, 2018). O Brasil é o maior produtor de laranja e de suco de laranja do mundo e sua citricultura está presente em todos os estados do país. Segundo a Confederação da Agricultura e Pecuária do Brasil, o país responde por 78% da participação no comércio mundial de suco de laranja, consolidando-se como o mais importante fornecedor global do produto (CNA, 2018).

Os resíduos originados no beneficiamento da laranja correspondem de 50% a 60% da massa inicial da fruta in natura. No Brasil, a alternativa encontrada à deposição do resíduo é a produção de ração animal. Entretanto, o subproduto do processamento da laranja contém

óleos essenciais, polifenóis, pectina, açúcares, ácidos, carboidratos insolúveis e pigmentos. Substâncias que podem ser precursoras de produtos mais rentáveis para as indústrias de alimentos, cosméticos e medicamentos (LEITAO et al., 2010; LÓPEZ; LI; THOMPSON, 2010; SHARMA et al., 2017; SATARI; KARIMI, 2018). Além disso, o resíduo pode ser utilizado em processos de conversão termoquímica, nos quais a matéria orgânica presente é convertida em energia (LÓPEZ; LI; THOMPSON, 2010; POURBAFRANI et al., 2010, 2013).

Nesse contexto, este trabalho consiste no emprego de metodologias de simulação de processos e uso de métricas ambientais para o cenário brasileiro de produção de resíduos da indústria de suco de laranja, abordando as diversas rotas de aproveitamento do resíduo. Dessa forma, diferentes alternativas de processos podem ser avaliadas de maneira sistemática levando em consideração a minimização de impactos ambientais associados a cada alternativa possível de valorização do resíduo da citricultura.

MATERIAL E MÉTODOS

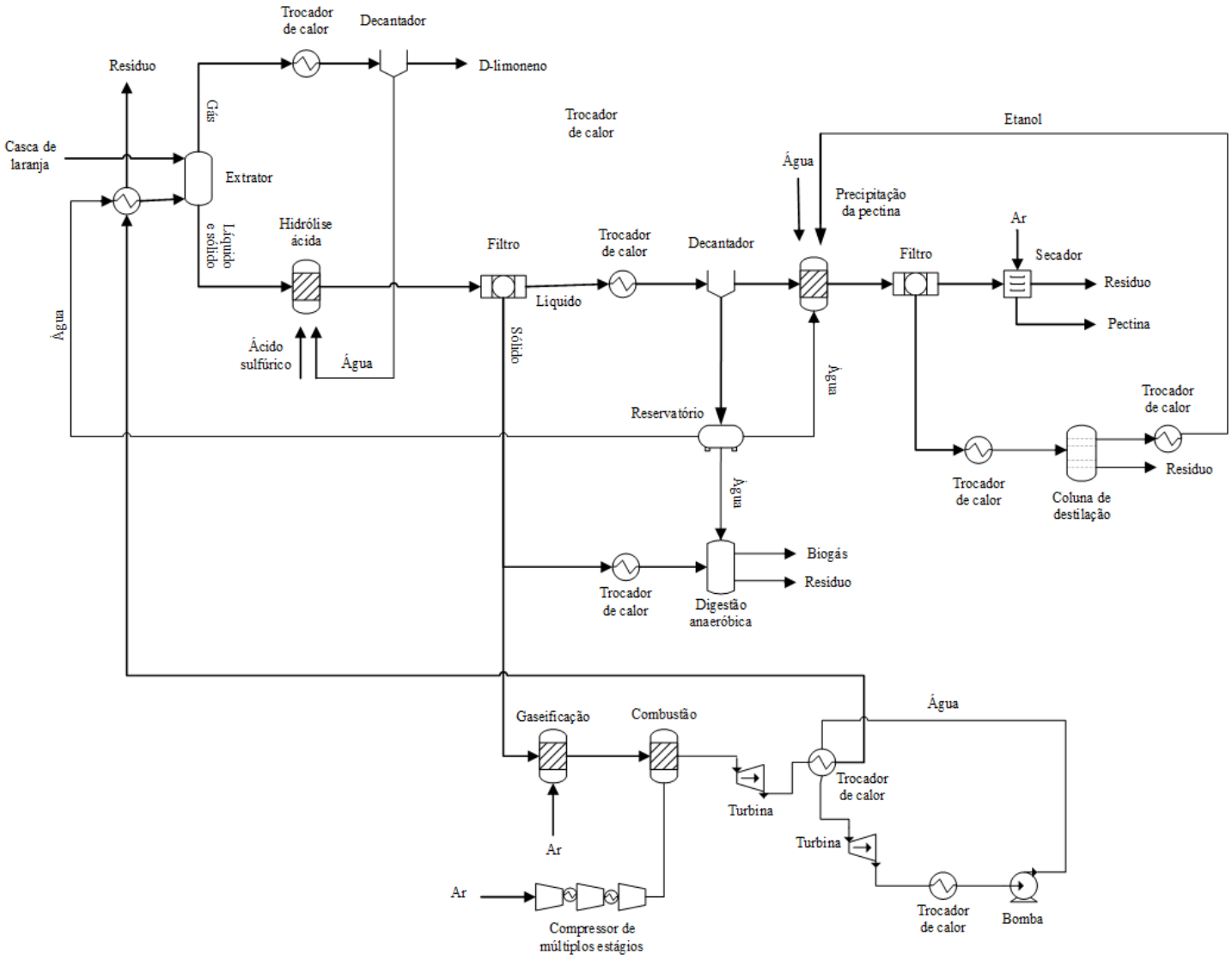
Simulação

Nesse trabalho é proposto o desenvolvimento de uma biorrefinaria capaz de produzir D-limoneno, pectina, biogás e energia. A etapa de modelagem e simulação foi desenvolvida no software ASPEN PLUS V. 8.4 (Aspen Tech, Massachusetts, USA). O dimensionamento da planta foi realizado com base na hipótese de que 5% de toda a casca de laranja gerada no Brasil é reaproveitada. Essa hipótese se traduz em uma vazão de resíduo processado de 111,11 kg/s, valor considerado uma boa estimativa inicial para simulação da produção em escala industrial. Trata-se de um modelo de simulação para avaliar o potencial de aproveitamento do resíduo, estando limitado à etapa conceitual da biorrefinaria. O modelo termodinâmico escolhido foi o *Non-Random Two-Liquid Model* (NRTL). O fluxograma geral do processo integrado pode ser visto na Figura 2.

As cascas da laranja foram alimentadas ao processo e destinadas para a extração do D-limoneno com vapor de água. A corrente de topo do extrator, composta por vapor rico em D-limoneno, foi resfriada até atingir a temperatura de 80ºC. Posteriormente, o líquido foi decantado, onde ocorreu a separação da fase aquosa e do D-limoneno (DÁVILA; ROSENBERG; CARDONA, 2015; SALAZAR; ALZATE, 2011). A corrente de fundo do extrator, contendo os sólidos e líquidos residuais, foi hidrolisada com ácido sulfúrico diluído (0,5% v/v) a 150 ºC e os sólidos foram separados do hidrolisado através de um filtro

(DÁVILA; ROSENBERG; CARDONA, 2015; POURBAFRANI et al., 2010; SALAZAR; ALZATE, 2011).

Figura 2 – Fluxograma da biorrefinaria proposta.



Fonte: Próprios autores.

A corrente líquida foi resfriada até 35 ºC, para preservar as características físico-químicas da pectina. O excesso de água presente no meio foi retirado e a pectina líquida foi precipitada com auxílio de uma solução de etanol 70% (v/v) a 30 ºC. Em seguida, a pectina já em estado sólido foi filtrada e seca com ar ambiente. A corrente líquida proveniente da etapa de filtração foi aquecida a 70 ºC e posteriormente destilada para o reaproveitamento do etanol no processo de precipitação (CASAS-OROZCO et al., 2015; DÁVILA; ROSENBERG; CARDONA, 2015; POURBAFRANI et al., 2010; SALAZAR; ALZATE, 2011).

O resíduo sólido proveniente da etapa de filtração do hidrolisado foi utilizado como insumo para produção de biogás ou energia. No primeiro cenário, o biogás foi produzido a partir do processo de digestão anaeróbia. Nesse caso, a corrente sólida foi resfriada até a temperatura ambiente (25ºC) e diluída até atingir o teor de sólidos equivalente a 25%. Por último, o biogás foi separado do líquido residual (NEGRO; RUGGERI; FINO, 2018; ZEMA et al., 2018).

No segundo cenário, a energia é gerada por meio do ciclo combinado com gaseificação de biomassa integrada (BIGCC). Os sólidos remanescentes da etapa de filtração foram misturados com a corrente contendo oxigênio, sendo a vazão equivalente a 15% do valor da corrente de sólidos. A gaseificação ocorreu a 1100 ºC e 20 bar. Posteriormente, o gás de síntese foi misturado com uma corrente de ar ambiente, com vazão 2,88 vezes maior que a do gás de síntese e junto com o gás de síntese seguiu para a combustão (DARMAWAN et al., 2017; DÁVILA; ROSENBERG; CARDONA, 2015; PRESTIPINO et al., 2017).

A combustão é realizada a 20 bar e 1126,85 ºC. Em seguida, uma turbina a gás foi usada para produzir trabalho de eixo a partir dos gases de exaustão da combustão. O calor proveniente dos gases de exaustão foi utilizado para aquecer uma corrente de água, produzindo vapor superaquecido a alta pressão. O vapor superaquecido é usado para produzir energia adicional na turbina a vapor. Posteriormente, o vapor é totalmente condensado na temperatura de 90 ºC e sua pressão aumentada de 2 para 33 bar na bomba. A água líquida é transformada em vapor superaquecido novamente continuando o ciclo (DARMAWAN et al., 2017; PRESTIPINO et al., 2017).

Métrica ambiental

A planilha de aspectos e impactos ambientais foi a ferramenta utilizada para fundamentar a decisão a respeito das rotas de aproveitamento das cascas de laranja que são conceitualmente mais viáveis ambientalmente. A metodologia consiste na análise técnica qualitativa e quantitativa para caracterizar e avaliar os efeitos das correntes materiais e energéticas que entram e saem dos processos de reaproveitamento das cascas de laranja (SÁNCHEZ, 2006). O objetivo é converter as informações dos balanços de massa e energia em efeitos ambientais potenciais, resultando em uma estimativa qualitativa dos impactos ambientais relacionados aos processos.

A metodologia de construção da planilha de aspectos e impactos se baseou no fluxograma dos processos industriais, a partir desse foram elencados os aspectos ambientais observados nas rotas de aproveitamento. Os impactos ambientais foram avaliados com base

nesses aspectos em termos de severidade, abrangência e probabilidade de ocorrência, conforme as categorias abaixo (CRISTINA; SALAMI, 2013; OLIVEIRA; SILVA, 2016; POTRICH; TEIXEIRA; FINOTTI, 2007):

A severidade corresponde ao grau de intensidade do impacto considerando a capacidade do meio ambiente de suportar ou reverter as consequências do impacto, restabelecendo-se a sua condição original. Os valores para as severidades das matérias-primas, utilidades, efluentes líquidos e emissões atmosféricas se encontram na Tabela 1.

Tabela 1 – Categorias de severidade para as correntes materiais dos processos de reaproveitameto.

	Severidade			
Nível	Matérias-primas e auxiliares (consumo total)	Utilidades (consumo total)	Efluentes líquidos (vazão total)	Emissões atmosféricas
1	0-1%	0-15%	0-7%	Outros
2	2-5%	15-30%	8-14%	Dióxido de carbono
3	5-20%	30-45%	15-20%	Metano

Abrangência diz respeito a extensão do impacto. Abrangências local, regional e global, assumem valores iguais a 1,2 e 3, respectivamente.

A probabilidade de ocorrência corresponde à probabilidade de ocorrência de um acidente ou dano ao meio ambiente. Atribui-se: (1) situações improváveis; (2) situações prováveis de ocorrer; (3) risco iminente de ocorrência.

Usando a categorização feita para a severidade, abrangência e probabilidade de ocorrência é possível obter a importância o impacto, conforme a expressão a seguir:

$$I = (S+A)\times P \quad (1)$$

onde S representa a severidade do impacto, A a sua abrangência e P a probabilidade de ocorrência. De acordo com a pontuação atingida atribui-se: D (desprezível), M (moderado) e C (crítico) para intervalos de 0-4, 5-8 e 9-12, respectivamente.

RESULTADOS E DISCUSSÃO

Na simulação, o D-limoneno produzido alcançou rendimento equivalente a 84,3% e pureza de 94,4%. Já para a pectina o rendimento e a pureza foram 64,8% e 96,5%, respectivamente. O biogás gerado na simulação apresentou rendimento de 0,60 L/g sólidos voláteis totais e a composição obtida na simulação foi de 46,21% CH4 e 45,91% CO2. A partir da vazão mássica do biometano obtido na simulação e do seu poder calorífico, estimou-se que 52,38 MW de energia podem ser gerados.

O calor do gás oriundo da combustão (1126 ºC e HHV igual a 0,15 MJ/kg), foi convertido em trabalho de eixo na turbina a gás, gerando 19,24 MW de energia. Após o processo, o gás apresentou elevada temperatura (631,75 ºC), por isso foi destinado a aquecer a água que participa do ciclo combinado. A água é aquecida até 550 ºC, transformando-se em vapor superaquecido, utilizado como insumo na turbina a vapor e produzindo 4,11 MW de energia adicional.

A planilha de aspectos e impactos foi a ferramenta utilizada para avaliar os impactos ambientais decorrentes dos processos e os resultados encontram-se disponíveis na Tabela 2. A principal finalidade da planilha foi examinar as consequências futuras das ações propostas. Nesta avaliação preliminar e simplificada dos potenciais de impacto dos processos de reaproveitamento das cascas de laranja, obteve-se as rotas mais suscetíveis a geração de danos ambientais, que em ordem decrescente são as produções de pectina, energia, biometano e D-limoneno.

A rota menos viável ambientalmente foi a produção de pectina. Uma coluna de destilação foi incorporada ao processo para reciclar o etanol utilizado na etapa de precipitação, visando minimizar o aspecto e consequentemente o impacto associado ao composto. A corrente a ser purificada é composta por etanol e água, e devido à grande volatilidade do etanol, 99% foram recuperados, valor típico de recuperação por destilação.

A produção de energia elétrica por meio do ciclo combinado foi a segunda rota com maior probabilidade de gerar impactos ao meio ambiente devido a liberação de grandes quantidades de dióxido de carbono (CO_2) e monóxido de carbono (CO), gases que contribuem para o aquecimento global.

No processo de produção de biogás é importante salientar o uso do biometano como potencial fonte de energia para utilização no processo, pois quando o biometano é gerado e não consumido, o gás é categorizado como emissão. O reaproveitamento, além de diminuir o consumo de energia, converte o biometano de emissão para produto, garantindo a não liberação do gás na atmosfera e minimizando o seu impacto ambiental (efeito estufa).

A produção de D-limoneno foi a rota com melhor desempenho ambiental. Os principais motivos foram o processo enxuto, a ausência de consumo de químicos e não geração de efluentes.

Tabela 2 – Tabela de aspectos e impactos para os processos de produção do D-limoneno, pectina, energia e biometano.

Planilha de identificação e avaliação de aspectos e impactos ambientais

Processo/produto	Etapa	Aspecto ambiental	Vazão (kg/s)	Consumo de energia (MW)	Impacto ambiental	Severidade	Abrangência	Frequência	Imporância do impacto[1]
D-limoneno	Extração	Consumo de energia	-	255	Alteração do meio físico e biótico	2	2	2	M
Pectina	Hidrólise ácida	Consumo de metérias-prima (ácido sulfúrico, químico perigoso)	5,15	-	Acidificação do meio ambiente	2	2	2	M
		Consumo de energia	-	390,4	Alteração do meio físico e biótico	3	2	2	C
	Precipitação	Consumo de metérias-prima (etanol)	0,27	-	Contaminação do solo e/ou ambientes aquáticos	1	2	2	M
		Consumo de energia	-	32,3	Alteração do meio físico e biótico	1	2	2	M
	Secagem	Consumo de ar	2,94	-	Alteração do meio físico e biótico	2	2	2	M
		Emissão de gás de efeito estufa	3,26	-	Aquecimento global	2	3	2	C
	Aquecimento da corrente	Consumo de energia	-	5,6	Alteração do meio físico e biótico	1	2	2	M
	Destilação do etanol	Consumo de energia	-	78,6	Alteração do meio físico e biótico	1	2	2	M
		Geração de efluentes líquidos	19,50	-	Acidificação do meio ambiente	3	2	2	C
Biometano	Digestão anaeróbia	Consumo de energia	-	122,9	Alteração do meio físico e biótico	1	2	2	M
		Emissão de gás de efeito estufa	20,97	-	Aquecimento global	3	3	2	C
		Geração de efluentes líquidos	4,51	-	Alteração da qualidade das águas superficiais e subterrâneas	1	2	2	M
Energia	Gaseificação	Consumo de ar	0,99	-	Alteração do meio físico e biótico	1	2	2	M
	Compressão do ar	Consumo de energia	-	7	Alteração do meio físico e biótico	1	2	2	M
		Consumo de ar	21,10	-	Alteração do meio físico e biótico	3	2	2	C
	Combustão	Emissão de gás de efeito estufa	28,46	-	Aquecimento global	2	3	2	C
	Bombeamento da água	Consumo de energia	-	0,02	Alteração do meio físico e biótico	1	2	2	M

[1] M = moderado, C = crítico

CONCLUSÕES

Neste trabalho foi feito uma análise conceitual do potencial dos processos de beneficiamento do resíduo da indústria de suco de laranja. O software Aspen Plus se mostrou uma ferramenta viável para a simulação da biorrefinaria, sendo capaz de redimensionar a capacidade de operação da planta. A partir de 111,1 kg/s de casca de laranja foram produzidos 1,49 kg/s de D-limoneno, 4,59 kg/s de pectina, 16,51 MW no ciclo combinado e 35,38 MW com biometano. Os rendimentos e as purezas dos químicos produzidos estão de acordo com a literatura. No processo de digestão anaeróbia foi produzido biogás com 46,21% de metano e 45,91% de dióxido de carbono. A partir do biometano é possível gerar 0,47 MW por kg de casca de laranja. A quantidade de energia gerada no ciclo combinado com gaseificação de biomassa integrada (BIGCC) mostrou que é possível obter 0,21 MW por kg de casca de laranja. A análise conceitual dos processos de beneficiamento apresentou bons resultados, mostrando o potencial do setor. Outros estudos que contemplem a disponibilidade de utilidades e o custo associado à matéria-prima são encorajadas para a implementação dos processos em escala industrial. Na planilha de aspectos e impactos obteve-se que os processos de reaproveitamento com maiores impactos ambientais foram as produções de pectina, energia, biometano e D-limoneno, respectivamente.

AGRADECIMENTOS

Os autores agradecem o suporte financeiro fornecido pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e o pelo apoio do Instituto de Pesquisas Tecnológicas do Estado de São Paulo (IPT) através do edital 01/2018 do Programa IPT Novos Talentos.

REFERÊNCIAS

CASAS-OROZCO, D. et al. Process development and simulation of pectin extraction from orange peels. **Food and Bioproducts Processing**, v. 96, p. 86–98, 2015.

CNA, C. DA A. E P. DO B. **Exportações de suco de laranja sobem 23% em julho**. Disponível em: <https://www.cnabrasil.org.br/noticias/exportacoes-de-suco-de-laranja-sobem-23-em-julho>. Acesso em: 24 set. 2018.

CRISTINA, S.; SALAMI, S. Universidade Tecnológica Federal Do Paraná Avaliação Dos Aspectos E Impactos Ambientais Dos Laboratórios Da Universidade Tecnológica Federal Do Paraná , Câmpus Campo Mourão , Como Ferramenta Para a Elaboração De Um Modelo De Sistema De Gestão Paraná , Câm. p. 1–90, 2013.

DARMAWAN, A. et al. Enhanced Process Integration of Entrained Flow Gasification and Combined Cycle: Modeling and Simulation Using Aspen Plus. **Energy Procedia**, v. 105, p. 303–308, 2017.

DÁVILA, J. A.; ROSENBERG, M.; CARDONA, C. A. Techno-economic and Environmental Assessment of p-Cymene and Pectin Production froDávila, J. A., Rosenberg, M., & Cardona, C. A. (2015). Techno-economic and Environmental Assessment of p-Cymene and Pectin Production from Orange Peel. Waste and Biomass Val. **Waste and Biomass Valorization**, v. 6, n. 2, p. 253–261, 2015.

LEITAO, V. F. et al. Biomass residues in Brazil: Availability and potential uses. **Waste and Biomass Valorization**, v. 1, n. 1, p. 65–76, 2010.

LÓPEZ, J. Á. S.; LI, Q.; THOMPSON, I. P. Biorefinery of waste orange peel. **Critical Reviews in Biotechnology**, v. 30, n. 1, p. 63–69, 2010.

NEGRO, V.; RUGGERI, B.; FINO, D. Recovery of Energy from Orange Peels Through Anaerobic Digestion and Pyrolysis Processes after d -Limonene Extraction. **Waste and Biomass Valorization**, v. 9, n. 8, p. 1331–1337, 2018.

OLIVEIRA, J. M. DOS S. M.; SILVA, R. G. Aspectos E Impactos Ambientais De Um Laboratório De Biologia. **Holos**, v. 8, p. 368, 2016.

POTRICH, A. L.; TEIXEIRA, C. E.; FINOTTI, A. R. Avaliação de impactos ambientais como ferramenta de gestão ambiental aplicada aos resíduos sólidos do setor de pintura de uma indústria automotiva. **Estudos Tecnológicos em Engenharia**, v. 3, p. 162–175, 2007.

POURBAFRANI, M. et al. Production of biofuels, limonene and pectin from citrus wastes. **Bioresource Technology**, v. 101, n. 11, p. 4246–4250, 2010.

POURBAFRANI, M. et al. Life cycle greenhouse gas impacts of ethanol, biomethane and limonene production from citrus waste. **Environmental Research Letters**, v. 8, n. 1, 2013.

PRESTIPINO, M. et al. Hydrogen rich syngas production by air-steam gasification of citrus peel residues from citrus juice manufacturing: Experimental and simulation activities. **International Journal of Hydrogen Energy**, v. 42, n. 43, p. 26816–26827, 2017.

SALAZAR, I. C.; ALZATE, C. C. Evaluación del proceso integral para la obtención de aceite esencial y pectina a partir de la cáscara de naranja. **Ingeniería y Ciencia - ing.cienc.**, v. 7, n. 13, p. 65–86, 2011.

SÁNCHEZ, L. E. **Avaliação de impacto ambiental: conceitos e métodos**. 3ª ed. São Paulo: 2006, 2006.

SATARI, B.; KARIMI, K. Citrus processing wastes: Environmental impacts, recent advances, and future perspectives in total valorization. **Resources, Conservation and Recycling**, v. 129, n. November 2017, p. 153–167, 2018.

SHARMA, K. et al. Converting citrus wastes into value-added products: Economic and environmently friendly approaches. **Nutrition**, v. 34, p. 29–46, 2017.

USDA, U. S. D. OF A. **Citrus: World Markets and Trade**. Disponível em: <https://apps.fas.usda.gov/psdonline/circulars/citrus.pdf>. Acesso em: 24 set. 2018.

ZEMA, D. A. et al. Anaerobic digestion of orange peel in a semi-continuous pilot plant: An environmentally sound way of citrus waste management in agro-ecosystems. **Science of the Total Environment**, v. 630, p. 401–408, 2018.

Estação agrometeorológica experimental automatizada

Carlos Marcelo Zys[1], *Felipe de Oliveira Gomes*[1], *Kelly Lais Wiggers*[1]

[1]Departamento de Engenharia Elétrica, Faculdade Guarapuava
Guarapuava, Paraná, Brasil
carloszys2@gmail.com, farmaciamonteclaro@gmail.com, kellylaiswiggers@hotmail.com

RESUMO

Uma estação agrometeorológica caracteriza-se como um dispositivo capaz de realizar leituras de grandezas meteorologicas, sendo uma ferramenta fundamental na agricultura. Ela pode auxiliar os agricultores na tomada de decisões pois as variáveis atmosféricas podem influenciar diretamente desde a preparação da terra ao resultado da produtividade. Porém, a disponibilidade de estações em áreas remotas do país ainda é um problema existente e os agricultores de pequeno porte podem não ter acesso em equipamentos desta natureza. Além disso, as estações agrometeorológicas disponíveis no mercados são caras e dispensiosas. Portanto, o objetivo deste trabalho foi desenvolver uma estação agrometeorológica automática e independente de energia elétrica externa. Ela foi instalada em uma área rural para monitoramento da região, utilizou-se a plataforma ESP32 Lora e sensores para coletar informações meteorológicas diariamente, as quais podem ser acessadas remotamente em um servidor de serviços *Web*. Para análise dos resultados, as leituras foram comparadas com os dados disponibilizados de uma estação oficial do INMET e constatou-se que os resultados foram promissores. Além disso, a estação foi desenvolvida com tecnologias *open source* e custo de pelo menos 40% mais barato que estações disponíveis no mercado.

PALAVRAS-CHAVE: Automação, Microcontroladores, Agrometeorologia

ABSTRACT

An agrometeorological station is characterized as a device capable of performing readings of meteorological quantities and, it is a fundamental tool in agriculture. It can assist farmers in decision making as atmospheric variables can influence directly from land preparation to productivity outcome. However, the availability of stations in remote areas of the country is still an existing problem and small farmers may not have access to such equipment. The agrometeorological stations available in the markets are expensive. Therefore, the objective of this paper was to develop an automatic agrometeorological station, independent of external electric energy. The station was installed in a rural area for monitoring of the region, using Lora ESP32

platform and sensors to collect daily meteorological information, which can be accessed remotely on a Web services server. For the experimental results, the meteorological information collected was compared with the data available from an official INMET station and the results are promising. Besides, the station was developed with open source technologies and cost at least 40% cheaper than stations available in the market.
KEYWORDS: Automation, Microcontrollers, Agrometeorology

INTRODUÇÃO

Devido a grande expansão da exportação agrícola no país, o monitoramento das variáveis climáticas são de extrema importância nesse setor com relação direta nos empreendimentos agropecuários. Desta forma, os parâmetros climáticos exercem influência sobre todos os estágios da cadeia de produção agrícola, desde a preparação da terra à comercialização do produto (GALINA; VERONA, 2004).

Nesse sentido, a Agrometeorologia (ou Meteorologia Agrícola) é a área que possibilita este monitoramento das condições atmosféricas e suas conseqüências no ambiente rural (PEREIRA; ANGELOCCI; SENTELHAS, 2007) mediante dados advindos de estações agrometeorológicas, que podem ser pressão atmosférica, temperatura e umidade relativa do ar, precipitação, direção e velocidade do vento. Como consequência, permitem o auxílio na tomada de decisões com relação às diversas práticas agrícolas, sejam as ações a serem tomadas pelo agricultor em caso de imprevistos climáticos nas lavouras, ou mesmo o controle exercido pelo clima sobre um determinado tipo de cultivo e sua viabilidade de plantio.

Existem diversos modelos de estações meteorológicas no mercado, em que o processo de coleta de dados ocorre de forma automatizada. Ou seja, sensores irão emitir sinais elétricos e são capturados por um sistema de aquisição de dados (Datalogger). Como resultado, ocorre o armazenamento e processamento dos dados digitais (GALINA; VERONA, 2004). No entanto, muitas estações utilizam tecnologias normalmente importadas, com alto custo de aquisição e manutenção. Portanto, na agricultura o desenvolvimento de estações com baixo custo tornou-se uma área de pesquisa importante seja para controle de processos de irrigação ou monitoramento da produtividade, como apresentado nos trabalhos de Aita (2017), Andrade et al. (2018), Neto et al. (2018) e Pasternak (2016).

Neste sentido, este trabalho apresenta o desenvolvimento de uma estação automatizada para captação, leitura e armazenamento de dados agrometeorológicos, os quais são disponibilizados para acesso remoto em tempo real por um servidor de serviços *Web*. Além disso, a estação é composta por um sistema de alimentação solar para ser independente de energia elétrica externa.

MATERIAL E MÉTODOS

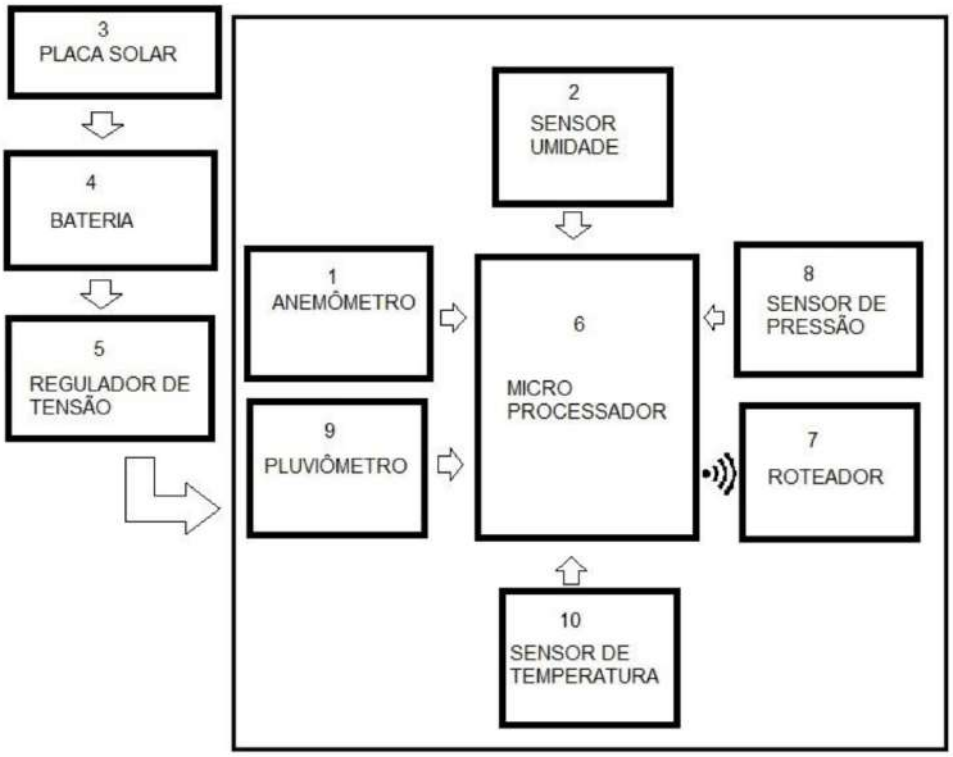
Para atingir o objetivo da meteorologia, é necessário coletar dados que definam a condição física da atmosfera, o que justifica o uso de estações meteorológicas. No Brasil, as redes de estações do Instituto Nacional de Meteorologia (INMET, 2011) são responsáveis pelo monitoramento metereológico de todo o país.

A construção de uma estação agrometeorológica exige a implantação de equipamentos específicos para monitoramento do clima, visto que, a partir da leitura dos sensores, retorna o conhecimento do estado da atmosfera local. O INMET define os elementos básicos para construção de uma estação agrometeorológica automática: (INMET, 2011):

- Sub-sistema de coleta de dados;
- Sub-sistema de controle e armazenamento;
- Sub-sistema de energia (painel solar e bateria);
- Sub-sistema de comunicação.

A figura 1 apresenta a visão geral da estação agrometereológica que foi desenvolvida neste trabalho, contendo os elementos básicos definidos pelo INMET. Na sequência, apresenta-se a descrição destes elementos que compõem a estação desenvolvida.

Figura 1: Visão geral da estação agrometereológica



1. Anemômetro: utilizado o modelo da WRF kits para aferição da velocidade e direção do vento (CARDOSO; SAMPAIO; BIANCO, 2009).
2. Sensor de umidade: modelo DHT11, que apresenta faixa de operação de 20% a 90% e erro de 5% (FILIPEFLOP,).

3. Placa solar: painel de 3w 12v, buscando suprir a necessidade de aproximadamente 0,9W, conforme a equação

$$P = 3,6V * 240 \times 10^{-3}A \quad (1)$$

em que P corresponde a potência máxima que o microprocessador demanda no sistema, com base nos valores de tensão máxima de alimentação V e corrente máxima A fornecida no datasheet do produto

4. Bateria: de 12v, semelhante as baterias utilizadas em nobreaks
5. Regulador de tensão: utilizou-se um regulador CC-CC para adequar a tensão de alimentação das fontes (12V) à tensão de alimentação do microprocessador, que é 3,6V
6. Microprocessador: ESP32 Lora com sistema sem fio embutido. É um componente de baixo custo com vasta aplicação em Internet das Coisas (IoT)(AITA, 2017).
7. Roteador: módulo sem fio do ESP32 que se conecta a um roteador sem fio com acesso a Internet instalado na área de estudo
8. Sensor de Pressão: modelo BMP180 para a aferição de dados relacionados a pressão atmosférica (TENFEN, 2013).
9. Pluviômetro: foi utilizado um pluviômetro de báscula, que transmite o sinal elétrico a cada quantidade específica de chuva precipitada (SANTANA; GUIMARãES; SILVA, 2010).
10. Sensor de Temperatura: modelo DS18B20, que oferece uma grande amplitude de leitura e não há prejuízos relacionados a precisão (em torno de $0,5°C$) (AMORIM; DIAS; SOARES, 2015).

Destaca-se que a definição dos componentes para criação de uma estação agrometeorológica baseou-se em uma pesquisa exploratória nos trabalhos relacionados ao tema, como por exemplo Pasternak (2016), Aita (2017) e Tenfen (2013). O que se diferencia neste trabalho é a busca por desenvolver uma estação de baixo custo, porém com o uso de componentes que retornem boa precisão na leitura dos dados, independência de energia elétrica da rede de distribuição, bem como não necessitar de intervenção humana para coleta e armazenagem dos dados.

A implementação do sistema foi em linguagem C, a qual possibilita o tratamento dos dados obtidos pelos sensores. Após a leitura dos sensores, os dados são enviados a um servidor de serviços *Web*. Nesse servidor são armazenadas, inclusive, as médias de dados coletados pelos sensores, com o objetivo de facilitar análises por parte dos usuários.

Para haver essa comunicação entre estação e servidor, a estação possui um sistema de conexão entre dois microprocessadores ESP32 Lora, que se comunicam via rádio frequência. É um circuito fechado de comunicação, divididos em transmissor (ESP_T) e receptor (ESP_R).

O ESP_T atua como transmissor e é responsável por obter o sinal elétrico de todos os sensores que compõe a estação e também por convertê-los em valores de grandezas meteorológicas. Os dados convertidos são inseridos em uma $string$, que contém os valores das grandezas mensuradas separados por ponto e vírgula. Os cálculos da conversão para temperatura, pressão e umidade estão disponibilizados na biblioteca $OneWire.h$. Este método foi escolhido devido

à facilidade de comunicação, compactando todos os valores em um único envio. A Figura 2 mostra a representação da $string$ de envio dos dados pelo sistema de comunicação.

Figura 2: Exemplo de string para armazenamento de dados coletados pela estação



Cada grandeza medida é identificada por uma sigla, que significa:

- VV - velocidade do vento;
- DV - direção do vento;
- PA - pressão atmosférica;
- TE - temperatura ambiente;
- HU - umidade relativa;
- PC - precipitação de chuva;

O microprocessador ESP_R tem a função de receber a $string$ enviada pelo transmissor, e deve estar alocado onde haja conexão com Internet sem fio, para que os dados coletados pela estação sejam enviados a um servidor de serviços *Web*. Nesse caso, o roteador de Internet já disponível na área de estudo foi utilizado. A comunicação entre o receptor e o servidor Web é realizada via protocolo MQTT, pois é um protocolo de rede leve e flexível, ideal para os desenvolvedores de IoT(YUAN, 2019). O servidor está alocado em um servidor *Google Cloud Plataform*, o qual possui uma máquina virtual com sistema operacional Linux e com o aplicativo *node-red* instalado. Com o uso desta ferramenta é possível obter a visualização dos dados através de um $dashboard$.

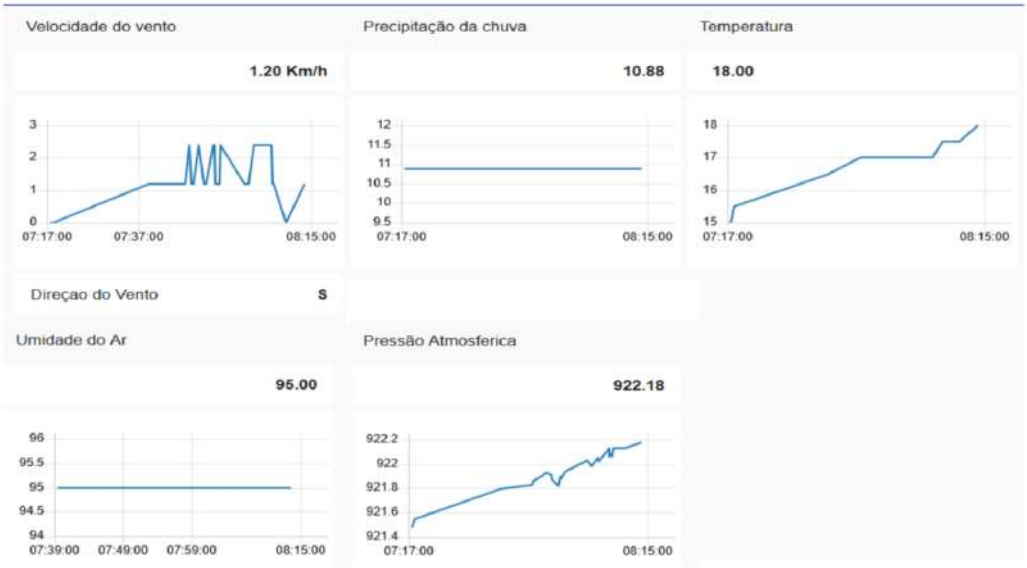
Para acessar esse $dashboard$ do servidor foi definido um endereço Web, e pode ser visualizado de forma remota em qualquer plataforma, seja computador ou smartphones, desde que os mesmos tenham conexão à Internet. A figura 3 apresenta a visualização destas leituras no $dashboard$.

No mesmo ambiente $dashboard$ é possível acessar um menu com opções de amostragem das coletas diárias e mensais. Inclusive, possibilita salvar todos os dados obtidos em um arquivo no formato $.csv$, com envio por e-mail em períodos de 24 horas e 30 dias.

Para atestar a confiabilidade dos dados coletados, primeiramente a estação desenvolvida foi instalada a aproximadamente 250 metros de uma estação meteorológica automática do INMET (veja as figuras 4 e 5), situada na cidade de Laranjeiras do sul, estado do Paraná, nas coordenadas Latitude: $-25.368944°$ e Longitude: $-52.391944°$. Não foi possível instalar a estação mais próxima da INMET devido a limitação de espaço e segurança do local. A estação ficou em funcionamento na área de estudo por 5 dias, com acompanhamento das coletas armazenadas pelo servidor. A frequência da coleta de dados foi realizada diariamente e as leituras armazenadas de forma semelhante aos dados disponibilizados no site do INMET para fins de comparação.

Para a confirmar a confiabilidade dos dados foi efetuada uma média diária das leituras obtidas de cada variável no período de testes, tanto na estação desenvolvida quanto os valores

Figura 3: Exemplo visualização dos dados no dashboard do sistema desenvolvido



disponibilizados pela estação do INMET. Após, foi avaliada a diferença (em %) de todas as médias para cada variável.

Após análise dos resultados, a estação desenvolvida foi instalada para uso definitivo na cidade de Virmond, no estado do Paraná, em uma propriedade rural, com o objetivo de auxiliar o agricultor no monitoramento meteorológico da região.

Figura 4: Distância entre as estações



RESULTADOS E DISCUSSÃO

Montagem da estação

Primeiramente, foi necessária a obtenção de um suporte para fixação do protótipo. Desta forma, todos os componentes da estação desenvolvida estão instalados em um suporte metálico de altura de 1,8 metros conforme a figura 6a. A estação possui uma caixa hermética fixada para abrigar o circuito de controle da mesma, e possui no topo da haste uma placa fotovoltaica.

Para a alimentação da estação foi empregada uma bateria de com tensão de 12V e corrente nominal 7A que fornece uma boa margem para funcionamento da estação com ausência de luz solar. Já para geração de energia foi instalado o painel fotovoltaico de 12V com potência máxima de 3W. Os dois componentes citados acima foram interligados em paralelo, e existe um

Figura 5: Estação INMET de Laranjeiras do Sul



Figura 6: Resultado da montagem da estação

(a) Montagem do suporte e componentes

(b) Instalação interna dos elementos da estação





diodo entre o polo positivo da placa solar e o polo positivo da bateria, o que evita que a corrente retorne para a placa solar quando não existe geração. Também foi empregado um regulador de tensão CC-CC para adequar a tensão de 12V da bateria para a tensão de alimentação do microcontrolador que corresponde a 5V.

Os componentes eletrônicos foram fixados em uma placa e todas as ligações foram soldadas com pequenos fios pela parte interna da mesma (veja a figura 6b), evitando problemas relacionados a perda de contatos entre os componentes.

A figura 7 apresenta o resultado do ambiente $dashboard$ em uma tela de celular, onde é possível acessar um menu com opções de amostragem diárias e amostragem mensais. No mesmo menu está disponível efetuar o $download$ de todos os dados obtidos em um arquivo no formato .csv representado na figura da esquerda. Como pode ser observado, na aba configurações do menu estão as opções de envio dos dados em arquivo por e-mail em períodos de 24 horas e 30 dias.

A figura 8 apresenta um exemplo de como as leituras dos sensores são salvas pelo servidor e disponibilizadas para o usuário. Devido ao grande número de dados coletados, a análise dos experimentos foi realizada pela média das leituras diárias.

Figura 7: Menu de opcões (esquerda) e Menu de configurações (direita)

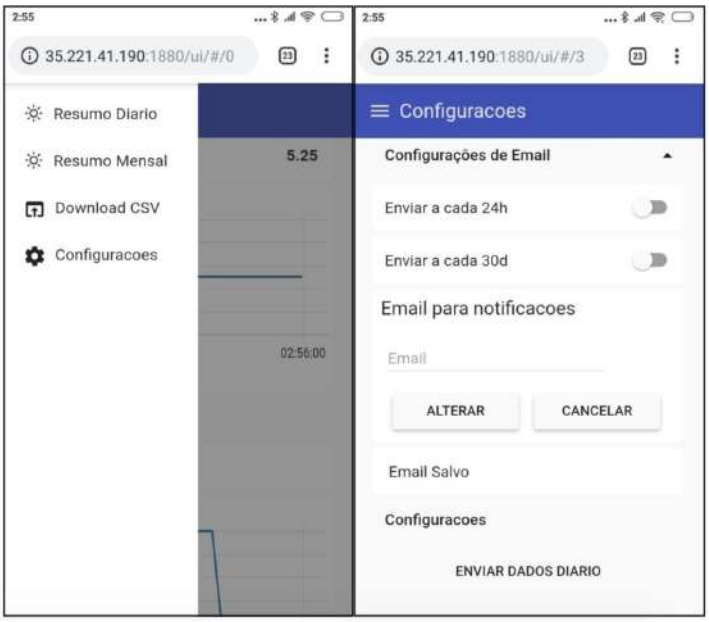


Figura 8: Exemplo de planilha dos leituras salvas pelo servidor

Dados Meteorológicos Sexta-Feira Dia 16/11/2018							
Data	Hora	Vel. vento	Direção do vento	Temperatura	Umidade	Precip. Chuva	Pressão Atm.
16/11/2018	14:47:58	0.00	S	28.50	55.00	0.00	922.65
16/11/2018	14:48:20	0.00	S	28.50	55.00	0.00	922.65
16/11/2018	15:00:53	0.00	S	28.00	62.00	0.00	922.65
16/11/2018	15:00:58	0.00	S	28.00	62.00	0.00	922.65
16/11/2018	15:02:55	0.00	S	28.00	62.00	0.00	922.65

Experimentos

A tabela 1 apresenta as médias dos valores lidos pelas estações para comparativo dos resultados durante os dias de experimentos. Pode-se observar que a diferença de temperatura foi de 7,7% e pressão apenas 0.13%, demonstrando confiabilidade nestes sensores. Já os valores de umidade diferenciaram em 30% e pluviométrico em 31%. Porém, os experimentos foram realizados em época de chuvas de verão, ocorrendo variabilidade nas concentrações de chuva em cada região. Verificou-se que os índices de velocidade e direção do vento não estavam sendo disponibiizadas pelo INMET, portanto não foi possível comparar os resultados.

Tabela 1: Comparativo das leituras da estação desenvolvida com a INMET

Grandeza	Estação desenvolvida	INMET
Temperatura	21,4	19,86
Umidade	53,20	76,5
Pressão	922,67	921,48
Vento	2,41	ND‡
Pluviométrico	10,88	15,8

ND‡ - não disponível

Com relação à autonomia de energia elétrica, os experimentos demontraram resultados satistafórios com os 3W utilizados do painel solar, e a bateria forneceu uma boa margem para funcionamento da estação com ausência de luz solar. Porém, para garantir a completa eficiência energética, visto que ao ser instalada a estação no local definitivo, o tempo de funcionamento é constante, uma placa solar com maior potência pode ser mais adequada para suprir o carregamento da bateria.

Como análise complementar, verificou-se que estação agrometeorológica desenvolvida pode ser considerada de baixo custo. Com base no levantamento dos valores referentes a todos os componentes do protótipo, o custo total desta estação foi de R$1.226,11. No mercado podem ser encontrados diferentes modelos de estações agrometeorológicas. Se forem consideradas estações que não possuem transmissão de dados remotas, e não ofertam possibilidade de exportação de dados, são encontradas por aproximadamente R$2.000,00 (PASTERNAK, 2016). Comparando o custo total do experimento desenvolvido com uma estação meteorológica profissional que efetua medições similares, como por exemplo a Davis Vantage Pro2, disponível para compra em `https://www.davisinstruments.com/solution/vantage-pro2/#compare-buy`, o valor de mercado é de aproximadamente R$7.800,00, porém também há diferenças nos processos de transmissão e exportação dos dados. Diante deste exposto, a estação agrometeorológica desenvolvida apresenta custo pelo menos 40% mais barato, com tecnologias *open source* e independência de rede de distribuição elétrica.

CONCLUSÕES

O presente trabalho apresentou o desenvolvimento de uma estação agrometeorólogica automatizada. Primeiramente, o levantamento bibliográfico realizado possibilitou o entendimento das grandezas que compõem uma estação, bem como a definição dos equipamentos relacionados à IOT, modelos de comunicação e programação dos mesmos.

Com base nos experimentos realizados, a estação desenvolvida demonstrou resultados promissores comparados com uma estação do INMET. Contudo, verifica-se que ainda há bastante espaço para melhorias. A estação também se mostrou autossuficiente no quesito energético, já que utiliza alimentação energética com uma placa fotovoltaica.

Como trabalhos futuros pretende-se avaliar resultados aumentando o período de captura de dados para verificar os desvios obtidos nos testes, e pesquisar outros métodos ou componentes para captura das variáveis que apresentaram valores inconsistentes. Inclusive, pretende-se criar uma rede Lora com $gateway$ Lora para que não seja necessário uma rede de Internet próximo ao local que a estação está instalada, visto que este formato de rede pode se comunicar com um alcance de raio de até 15 quilômetros.

REFERÊNCIAS

AITA, R. H. *Sistema de Irrigação Localizada e Automatizada*. Monografia (Trabalho de

Conclusão de Curso) — Pontifícia Universidade Católica do Rio Grande do Sul, Rio Grande do Sul, 2017.

AMORIM, H. S. do; DIAS, M. A.; SOARES, V. Sensores digitais de temperatura com tecnologia one-wire: Um exemplo de aplicação didática na área de condução térmica. *Revista Brasileira de Ensino de Física*, scielo, v. 37, p. 4310–1 – 4310–9, 12 2015. ISSN 1806-1117.

ANDRADE, J. A. A. et al. Implementing a network of weather and agrometeorological stations: a case study in Timor-Leste. *Finisterra - Revista Portuguesa de Geografia*, scielopt, p. 53 – 70, 08 2018. ISSN 0430-5027.

CARDOSO, C. O.; SAMPAIO, C. A. de P.; BIANCO, D. P. A. O. Construção e calibração de anemômetros de baixo custo. *Revista de Ciências Agroveterinárias*, v. 8, 2009. ISSN 2238-1171.

FILIPEFLOP. *Sensor de Umidade e Temperatura DHT11*. Acesso em: 19-06-2019. Disponível em: <https://www.filipeflop.com/produto/sensor-de-umidade-e-temperatura-dht11/>.

GALINA, M. H.; VERONA, J. A. Fontes de observações meteorológicas no estado de são paulo. *Estudos Geográficos*, p. 107–118, 2004.

INMET. *Rede de Estações Meteorológicas Automáticas do INMET*. [S.l.], 2011.

NETO, E. L. de A. et al. Estação meteorológica wifi de baixo custo baseado em thingspeak. In: *VII Congresso Brasileiro de Energia Solar*. Gramado: VII Congresso Brasileiro de Energia Solar, 2018.

PASTERNAK, W. D. *Estação agrometeorológica experimental automatizada e integrada a softwares para cálculo de estimativas de dormência em pessegueiros e videiras*. Monografia (Dissertação de Mestrado) — Universidade Estadual de Ponta Grossa, Ponta Grossa, 2016.

PEREIRA, A. R.; ANGELOCCI, L. R.; SENTELHAS, P. C. *Metereologia Agrícola*. [S.l.]: Esalq, 2007.

SANTANA, M. A. A.; GUIMARãES, P. L. O.; SILVA, C. M. Pesquisa aplicada e implementação da calibração de sensores e medidores de chuva. In: *Anais do Congresso da Qualidade em Metrologia*. São Paulo: Congresso da Qualidade em Metrologia, 2010.

TENFEN, C. R. *Projeto de um sistema remoto para aquisição de dados meteorológicos*. Monografia (Dissertação de Mestrado) — Universidade do Vale do Itajaí, São José, 2013.

YUAN, M. *Conhecendo o MQTT*. 2019. Acesso em: 19-06-2019. Disponível em: <https://www.ibm.com/developerworks/br/library/iot-mqtt-why-good-for-iot/index.html>.



Rastreando a Cadeia de Carne Brasileira com Blockchain

Vasco Varanda Picchi[1], *José Reynaldo Formigoni Filho*[2],
Eduardo Ferreira Jucá de Castro[2], *Sérgio Luís Ribeiro*[2]

[1] Safe Trace, Itajubá, Minas Gerais, Brasil, vasco@safetrace.com.br
[2] Gerência de Soluções Blockchain, CPQD, Campinas, São Paulo, Brasil,
reynaldo@cpqd.com.br, jucah@cpqd.com.br, sribeiro@cpqd.com.br

RESUMO

Esse trabalho apresenta a aplicação da tecnologia blockchain no atendimento dos desafios da cadeia produtiva da carne, destacando-se a rastreabilidade, do consumidor ao produtor. Na introdução o problema é identificado e contextualizado, listando os desafios, tais como práticas ilegais de desmatamento, trabalho escravo, embargos ambientais, entre outros, nessa cadeia. Em seguida, a tecnologia blockchain é mostrada como solução das necessidades de rastreabilidade, com confiabilidade e transparência, para todos os atores envolvidos nessa cadeia. A arquitetura utilizada na solução é descrita, com suas principais funcionalidades e tecnologias. E, finalmente, os resultados e conclusões são apresentadas.

PALAVRAS-CHAVE: Rastreabilidade, Blockchain, Contratos Inteligentes, Aplicações Descentralizadas.

ABSTRACT

This work presents the application of blockchain technology to address the challenges of the meat production chain, highlighting the traceability from the consumer to the producer. In the introduction, the problem is identified and contextualized listing the challenges, such as illegal deforestation, slave labor, among others, in this chain. Following, the blockchain technology is presented as a solution to traceability needs, with reliability and transparency for all actors involved in this chain. The architecture used in the solution is described with its main features and technologies. Finally, the results and conclusions are presented.

KEYWORDS: *Traceability, Blockchain, Smart Contracts, Decentralized Applications*.

INTRODUÇÃO

Impedir a disseminação da Encefalopatia Espongiforme Bovina (BSE) na Europa "O mal da vaca louca" foi o que motivou, a partir de 1998, a aplicação de maneira efetiva do conceito de rastreabilidade na cadeia produtiva da carne.

Diante dos riscos identificados à saúde e o comprometimento da capacidade de abastecimento interno de carne bovina, os países membros da União Europeia passaram a exigir de seus fornecedores internacionais a adoção, até o ano 2000, de sistemas comprovadamente confiáveis de identificação e registro de animais para garantir a procedência da carne.

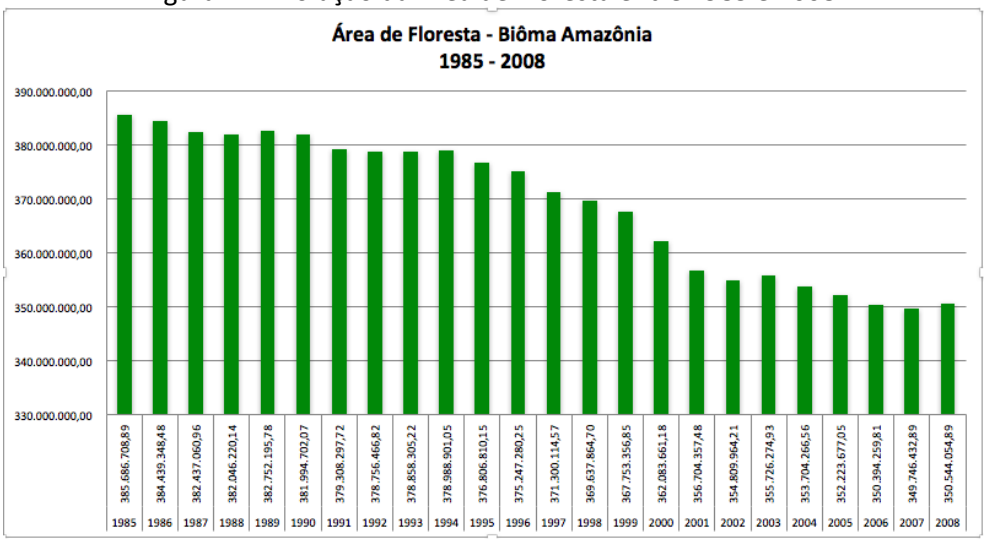
No Brasil, para atender às exportações à União Europeia, o Ministério da Agricultura, Pecuária e Abastecimento (MAPA), criou em 2000 o Sistema Brasileiro de Identificação e Certificação de Origem, atualmente denominado Sistema Brasileiro de Identificação Individual de Bovinos e Búfalos (SISBOV).

Devido ao alcance restrito desta norma, a lei 12.097/09 estabelece as normas para a aplicação de rastreabilidade na cadeia produtiva das carnes de bovinos e de búfalos, bem como as ferramentas que podem ser utilizadas [1]. Até então, a rastreabilidade era principalmente aplicada no centro oeste, sudeste e sul do Brasil, por serem áreas livres de febre aftosa com vacinação e contarem com as indústrias mais focadas na exportação, em especial para o continente Europeu, entretanto em 2008 a organização não governamental Greenpeace publicou o documento "A Farra do Boi na Amazônia", retratando o avanço do desmatamento na região da Amazônia, e atribuindo este à atividade pecuária.

Por meio de séries temporais de imagens via satélite, é possível confirmar que houve perda de 10% de toda a cobertura florestal entre 1985 (início da série histórica) e 2008 (Figura 1) [2]. O relatório representou um duro golpe às indústrias exportadoras de carne que atuavam na região, pois pela primeira vez a investigação conseguiu rastrear o destino dos produtos de abates em fazendas com práticas ilegais de desmatamento, trabalho escravo e embargos ambientais levando a responsabilidade a marcas locais e globais de lanchonetes, supermercados, artigos esportivos e até automóveis a serem corresponsáveis pelo desmatamento na região.

Com o objetivo de frear o avanço do desmatamento, o Ministério Público Federal passou a tomar medidas de comando e controle, propondo aos frigoríficos a assinatura de um Termo de Ajuste de Conduta (TAC) [3]. Por esse termo os frigoríficos se comprometeram a não comprar carne de produtores que tivessem realizado desmatamento de suas propriedades após 22 de julho de 2008, constassem na lista de propriedades embargadas pelo Ibama ou na lista de empregadores flagrados com trabalhadores em condição análoga à escravidão do Ministério do Trabalho.

Figura 1 – Evolução da Área de Floresta entre 1985 e 2008.



Fonte: Mapbiomas 3.0 - Área de Floresta (Ha)

A tecnologia trouxe apoio para o cumprimento do acordo por meio de duas importantes ferramentas: A guia de transporte animal eletrônica (e-GTA) e a aplicação do projeto PRODES do INPE (Instituto Nacional de Pesquisas Espaciais) [4], que realiza o monitoramento por satélite do desmatamento por corte raso (*Figura 2*) na Amazônia Legal e produz, desde 1988, as taxas anuais de desmatamento na região, entretanto, havia um grande gargalo prático em sua aplicação: A falta de ordenamento fundiário, que não permitia correlacionar o desmatamento a uma propriedade específica por problemas de titularidade da posse.

Figura 2 – Exemplo de identificação de área desmatada por corte raso.



O desmatamento é ainda acompanhado de outras práticas ilícitas como a invasão de Unidades de Conservação e Terras Indígenas, a grilagem de terra e o garimpo ilegal. A falta de titularidade das terras intensifica ainda mais a dificuldade de identificar os responsáveis leva à impunidade e estimula a ação dos infratores.

A Lei nº 12.651, de 25 de maio de 2012, conhecida como novo código ambiental, trouxe uma nova ferramenta com potencial de solucionar esse problema: o Cadastro Ambiental Rural (CAR). Regulamentado pela Instrução Normativa MMA nº 2, de 5 de maio de 2014, o CAR é um registro público eletrônico de âmbito nacional, obrigatório para todos os imóveis rurais,

com a finalidade de integrar as informações ambientais das propriedades e posses rurais referentes às Áreas de Preservação Permanente - APP, de uso restrito, de Reserva Legal, de remanescentes de florestas e demais formas de vegetação nativa, e das áreas consolidadas (Figura 3).

Figura 3 – Exemplo de áreas com limites cadastrados no CAR.



A aplicação gradativa dessas ferramentas passou a ordenar os processos de compras das indústrias que atuam na região da Amazônia, que após um período de acomodação, passou a ser realizada de forma estruturada, realizando o cadastro e análise prévia da conformidade ambiental e trabalhista por meio de prestadores de serviços de geomonitoramento, puxada principalmente pelas indústrias exportadoras da região, que além da TAC, assinaram acordos individuais de desmatamento zero com o Greenpeace e se comprometeram a realizar auditorias externas do cumprimento das regras.

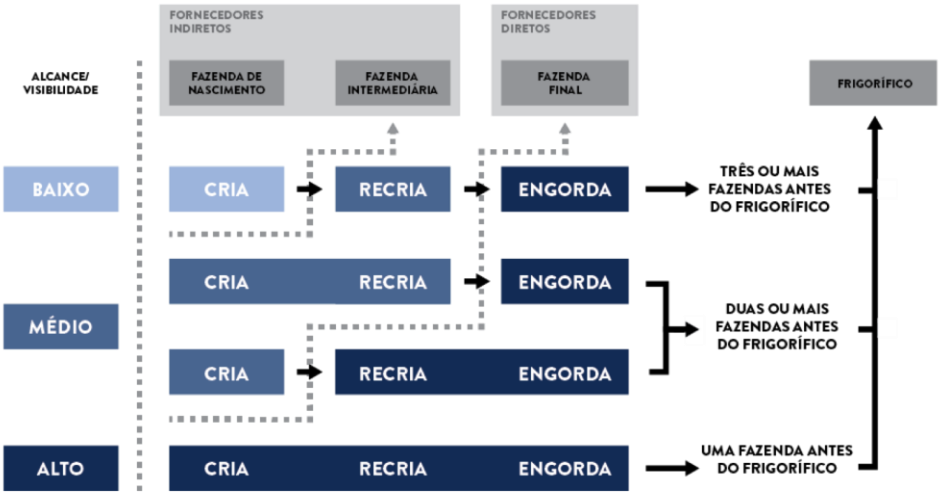
Ainda em 2010 o varejo deu início à utilização da rastreabilidade como forma de diferenciar os produtos com o lançamento da linha Garantia de Origem Carrefour [5], dando visibilidade ao programa e os diferenciais de seus produtos, todos com auditoria de terceira parte. A falta de fiscalização do Ministério Público Federal sobre o cumprimento dos TAC's, o efeito do acordo sobre o desmatamento se mostrou limitado e a tendência de queda dos índices de desmatamento cessou e passou a apresentar nova tendência de alta desde 2013.

Com base em pressões do terceiro setor e consumidores, o Walmart foi o primeiro varejista no Brasil a se comprometer com uma política de compras, seguido por GPA e Carrefour. Walmart e Carrefour foram inicialmente pelo mesmo caminho dos frigoríficos, contratando ferramentas de geomonitoramento, enquanto o GPA procurou uma nova abordagem sobre o problema por entender que o geomonitoramento deve ser uma ferramenta dos frigoríficos e que por si só não resolve o problema.

Mesmo com a entrada dos varejistas pressionando seus fornecedores a adotar o geomonitoramento ou adotando por sua própria conta, a evolução dos dados de desmatamento mostra a elevação do corte raso da floresta nesse período. O monitoramento dos frigoríficos só é realizado em fazendas fornecedoras diretas, ignorando elos anteriores da cadeia de produção. Isso faz com que fazendas que estão indiretamente ligadas à produção dos animais e que estejam

realizando desmatamento ilegal não sofram restrições à comercialização dos seus animais junto aos frigoríficos signatários da TAC.

Figura 4 – Visibilidade da cadeia de produção pecuária.



A complexidade da cadeia de produção faz com que a rastreabilidade apareça novamente como parte fundamental na capacidade de visibilidade da cadeia (Figura 4) e de cumprimento de requisitos socioambientais.

Em 2017, quase simultaneamente à operação Carne Fraca que foi exaustivamente noticiada, ocorreu a Operação Carne Fria do Ibama. Para essa operação foram verificadas nas bases governamentais as fazendas com maior volume de fornecimento dos frigoríficos e a relação entre sua área e o volume de animais vendidos. Como a produtividade média das propriedades na região amazônica é de 1,25 Unidades Animais por hectare, foi dada uma tolerância de quase 10 vezes esse valor para identificar quais propriedades poderiam estar burlando os sistemas de geomonitoramento dos frigoríficos, triangulando animais de áreas de desmatamento e embargadas pelo Ibama.

Como resultado o Ibama autuou 14 frigoríficos no Pará, Bahia e Tocantins que compraram 58 mil cabeças de gado, produzidas em 26 fazendas com áreas embargadas por desmatamento ilegal na Amazônia. Na prática, a operação replicou em escala a mesma estratégia de análise de fazendas indiretas aplicada no programa do campo à mesa em SFX [6]. Essa estratégia de monitoramento para a cadeia completa de produção não pode ser reproduzida em escala pelo setor produtivo devido ao acesso das informações da GTA não ser público. Entretanto, em reação à operação Carne Fria foi instituído no Ministério do Meio Ambiente um GT da Pecuária Sustentável, com a participação de frigoríficos, varejistas e Governo Federal, onde a liberação mediante salvaguardas aos dados de sigilo dos produtores [7].

Mas diante de tanta complexidade, há uma tecnologia capaz de registrar todo o histórico de eventos de cadeias produtivas tão complexas a ponto de permitir atestar a origem de um alimento?

Para responder a essa pergunta com total segurança, a tecnologia blockchain se mostra como uma ferramenta poderosa na capacidade de registrar o percurso de um alimento e todos os eventos relacionados à sanidade e qualidade que podem influenciar na escolha do consumidor.

CONCEITOS E APLICAÇÕES BLOCKCHAIN PARA O SETOR AGROPECUÁRIO

Existem várias definições para a tecnologia blockchain. Pode-se dizer que se trata de uma base de dados, mantida e gerida de forma compartilhada e descentralizada (através de uma rede *peer to peer* - P2P), onde todos são responsáveis por armazenar e manter esta base de dados e que tem como principais pilares: segurança, arquitetura descentralizada, integridade de dados e imutabilidade [8].

Do ponto de vista de aplicação, a tecnologia blockchain passou por uma grande evolução com a possibilidade de uso dos contratos inteligentes. Tais contratos são programas de computador que espelham um contrato tradicional [9]. Eles são replicados e executados por todos os nós da rede ou por um conjunto predeterminado de nós denominados validadores. Algumas plataformas de desenvolvimento comportam programas mais sofisticados. Na plataforma Ethereum um programa pode conter uma mini base de dados. Aplicações baseadas em contratos inteligentes são chamadas *Decentralized App* ou DApp [10].

Nestes pouco mais de 10 anos de história as tecnologias associadas ao blockchain evoluíram, permitindo o desenvolvimento de um grande número de aplicações para vários setores da economia tais como financeiro, governo, energia, saúde, telecomunicações e também o setor agropecuário [11].

Especificamente para o setor agropecuário, vislumbra-se a utilização da tecnologia blockchain nas seguintes áreas [12]:

- Proveniência: possibilidade de rastreamento de produtos, desde a sua origem, pontos e formas de armazenamento, e datas e meios de transporte até o seu destino final. Tudo isso, com total controle de acesso às informações para os participantes da rede, desde a total transparência dos registros até o acesso com diferentes níveis de acesso, baseado nas políticas de segurança estabelecida pelos requisitos de segurança implementados;
- Acessos mais eficientes aos meios de pagamento e financiamento: uso de plataformas de pagamentos móvel seguro, sem intermediários, assim como acesso ao crédito por

parte dos produtores de forma rápida com menores taxas começam a serem viabilizadas com a utilização da tecnologia blockchain. Principalmente nos países em desenvolvimento, o acesso rápido aos investimentos é um grande desafio;

- Gestão em tempo real das transações em toda a cadeia de suprimentos e financiamento: blockchain pode melhorar o processo de liquidação de contas para agricultores, compradores e financiadores, como bancos. Isso é possível através da criação de registros de diferentes tipos de transação (financeiras e não financeiras) de forma confiável e auditável em todas as etapas da cadeia produtiva.

Recentemente, a *Food and Agriculture Organization of the United Nations* (FAO) e a *International Telecommunication Union* (ITU) disponibilizaram um extenso e abrangente relatório denominado "*e-Agriculture in Action: Blockchain for Agriculture – Opportunities and Challenges*" [13]. O objetivo da publicação é desmistificar a tecnologia, fornecer algumas reflexões sobre as oportunidades e desafios na implementação de sistemas baseados na tecnologia blockchain, bem como documentar alguns estudos de caso sobre o seu uso para a agricultura. Na visão dos autores, "os contratos inteligentes autoexecutáveis, juntamente com os pagamentos automatizados, seriam o fator decisivo do jogo. O papel dos contratos inteligentes, especialmente em seguros agrícolas, títulos verdes e rastreabilidade, pode ser muito eficaz. Seguro agrícola construído sobre a tecnologia blockchain com incidentes meteorológicos chave e pagamentos elaborados em um contrato inteligente, vinculados a carteiras móveis, com dados meteorológicos sendo fornecidos regularmente por sensores no campo e correlacionados por dados de estações meteorológicas de proximidade facilitariam o pagamento imediato no caso de uma seca ou inundação no campo".

ARQUITETURA E FUNCIONALIDADES

Embora o foco seja a cadeia de carne brasileira, a arquitetura foi concebida com uma abordagem mais genérica, visando atender também aos diversos cenários do agronegócio e ao segmento de *Supply Chain* como um todo. Assim, o objeto principal é um *asset*, podendo representar um boi, uma carcaça bovina, um corte de carne, uma saca de café, etc.

Outra diretriz que orientou a concepção foi o maior desacoplamento possível entre os componentes da tecnologia blockchain e a aplicação. Isso permite que a tecnologia seja utilizada por mais participantes sem que haja necessidade de treinamento ou aprofundamento nessa tecnologia e também, viabiliza a integração com sistemas legados.

Para cada operação sobre um *asset*, há uma funcionalidade correspondente. Tomando a cadeia produtiva da carne para exemplificar, um *asset* é criado (***create***) para representar um

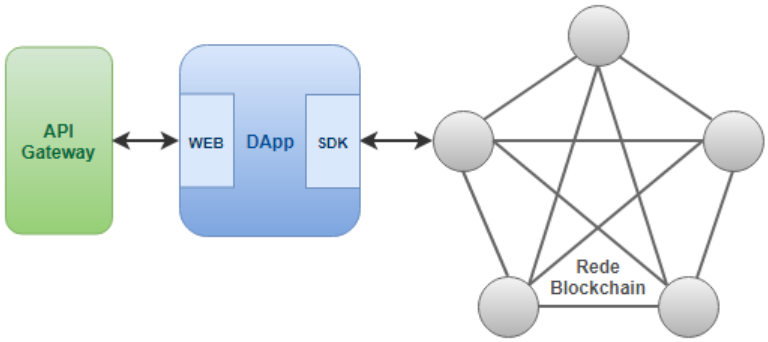
bovino no produtor; no frigorífico esse *asset* é desmembrado (***split***) em vários outros, para representar cada parte da carcaça; em seguida esses *assets* podem ser transferidos (***transfer***) para um mercado; o mercado pode desmembrar (***split***) cada um desses *assets* em novos, representando cada corte de carne; e, como cada novo *asset* está associado ao que o originou, toda a cadeia pode ser rastreada até a origem no produtor.

Atualmente é possível identificar algumas dezenas de plataformas blockchain que estão desenvolvidas e em operação, e outras dezenas ainda em fase de experimentação. A identificação da plataforma blockchain que mais se adequa a determinado cenário da aplicação a ser desenvolvida ainda é um fator chave e não é uma tarefa trivial.

Para isso, o CPQD desenvolveu o "Método e Sistema de Análise e Definição de Plataforma Blockchain Baseado em Contexto – MetaBlockX [14]", que com base na análise qualitativa dos principais aspectos que exercem influência direta sobre a plataforma e principalmente levando em consideração o contexto da aplicação ou serviço que será disponibilizado pela plataforma. Fornecendo de forma clara e objetiva a plataforma blockchain que mais se adequa a tal contexto, eliminando assim processos empíricos na escolha da plataforma para o desenvolvimento.

Com a aplicação e resultado da MetaBlockX temos que para a Rede Blockchain, núcleo da solução, foi escolhido o *framework* Fabric, do projeto Hyperledger, iniciativa de código aberto mundial com mais de 270 filiados, hospedado pela Fundação Linux [15]. A arquitetura da solução está representada, de forma simplificada, na *Figura 5*.

Figura 5 – Arquitetura.



No bloco DApp, uma aplicação descentralizada, é onde estão implementadas as funcionalidades da solução. Há uma interface Web e a integração com a Rede Blockchain é realizada através de um SDK em Java disponibilizado para o Hyperledger Fabric [16]. A Dapp foi implementada na linguagem Java. Um *API Gateway* está à frente da solução, tanto para a segurança quanto para escalar a solução sem impacto aos clientes. Como *API Gateway* foi usado o Kong, de código aberto.

RESULTADOS E DISCUSSÃO

A solução apresentada foi implantada em *multicloud*, com duas máquinas na AWS e uma na SoftLayer. A distribuição dos elementos da arquitetura levou em consideração a separação lógica de funções, sendo uma máquina com o API Gateway e a DApp e as outras duas com os elementos da Rede Blockchain (*orderer, peers, CA servers, databases*). Todos os elementos foram implantados com o uso de *containers* Docker.

No processo de validação, parceiros foram selecionados para exercitar a solução. Dados de operações reais foram inseridos na aplicação, usando as funcionalidades disponibilizadas pela solução. A comparação com o uso de sistemas tradicionais tem sido usada para demonstrar a esses parceiros a viabilidade e as vantagens da adoção da solução com a tecnologia blockchain. Em uso para testes e demonstrações desde de janeiro de 2019, pode-se verificar que a tecnologia blockchain satisfaz plenamente as necessidades de rastreabilidade com confiabilidade e transparência para todos os atores envolvidos na cadeia, produtores, frigoríficos, comerciantes, consumidores e órgãos fiscalizadores.

CONCLUSÕES

Dada a complexidade da cadeia produtiva da carne que envolve desafios como práticas ilegais de desmatamento, trabalho escravo, embargos ambientais, entre outros, a rastreabilidade aparece como parte fundamental na capacidade de visibilidade da cadeia e de cumprimento de requisitos socioambientais. Também, a segurança, a confiabilidade e a transparência nas operações são fatores decisivos na adoção de uma solução para a rastreabilidade.

A solução desenvolvida baseada na tecnologia blockchain se mostra uma poderosa ferramenta na capacidade de registrar o percurso de um alimento e todos os eventos relacionados à sanidade e qualidade que podem influenciar na escolha do consumidor.

Como trabalho futuro, deve-se abordar a integração com dispositivos IoT (*Internet of Things*) e o uso de Inteligência Artificial, automatizando a coleta e o tratamento de dados, como *tags* em bovinos, imagens, temperatura de transporte e armazenamento, entre outros.

AGRADECIMENTOS

Os autores reconhecem o apoio financeiro dado a este trabalho, com o apoio do Funttel, o Fundo de Desenvolvimento de Tecnologia de Telecomunicações. Este documento reflete apenas as opiniões dos autores e as Agências não são responsáveis por qualquer uso que possa ser feito das informações nele contidas.

REFERÊNCIAS

[1] MAPA. Instrução normativa. Disponível em: http://www.agricultura.gov.br/acesso-a-informacao/participacao-social/consultas-publicas/documentos/IN_SISBOV_jan18 ajustadaCONJUR1.pdf. Acesso em: 18/06/2019.

[2] TERRABRASILIS – PRODES (Deforestation). Disponível em: http://terrabrasilis.dpi.inpe.br/app/map/deforestation?hl=pt-br. Acesso em: 18/06/2019.

[3] MP diz que produtores e frigoríficos devem cumprir exigências. Disponível em: http://www.senado.gov.br/noticias/agencia/quadros/qd_062.htm. Acesso em: 18/06/2019.

[4] INPE. PRODES - Amazônia. Disponível em: http://www.obt.inpe.br/OBT/assuntos/programas/amazonia/prodes. Acesso em: 18/06/2019.

[5] CARREFOUR. Produtos Garantia de Origem Carrefour. Disponível em: https://www.carrefour.com.br/institucional/produtos-carrefour/garantia-de-origem/produtos-garantia-de-origem. Acesso em: 24/06/2019.

[6] OPERAÇÃO carne fraca, carne fria. Disponível em: https://amazonia.org.br/2017/05/carne-fraca-carne-fria-e-carne-na-mesa/. Acesso em: 24/06/2019.

[7] GRUPO DE TRABALHO DE PECUÁRIA SUSTENTÁVEL. Disponível em: http://gtps.org.br. Acesso em: 24/06/2019.

[8] FORMIGONI FILHO, J. R. et al. Tecnologia Blockchain: Uma visão Geral. Disponível em: https://www.cpqd.com.br/wp-content/uploads/2017/03/cpqd-whitepaper-blockchain-impresso.pdf. Acesso em 24/06/2019.

[9] WERBACH, Kevin, CORNELL, Nicolas. CONTRACTS EX MACHINA. Disponível em: https://scholarship.law.duke.edu/cgi/viewcontent.cgi?article=3913&context=dlj. Acesso em: 24/06/2019.

[10] Ethereum Foundation. A Next-Generation Smart Contract and Decentralized Application Platform. Disponível em: https://github.com/ethereum/wiki/wiki/White-Paper. Acesso em: 24/06/2019.

[11] RIBEIRO, Sérgio Luís. Tecnologia Blockhain: Aplicações e Iniciativas. Disponível em: https://www.cpqd.com.br/wp-content/uploads/2017/09/whitepaper_aplicacoes_e _iniciativas_final.pdf. Acesso em: 24/06/2019.

[12] WESTON, Emma, NOLETE, Sarah. Bitcoin to Agriculture: How Can Farmers Benefit from Blockchain? Disponível em: https://agfundernews.com/from-bitcoin-to-agriculture-how-can-farmers-benefit-from-blockchain6380.html. Acesso em: 24/06/2019.

[13] FAO and ITU. e-Agriculture in Action: Blockchain for Agriculture – Opportunities and Challenges. Disponível em: http://www.fao.org/3/ca2906en/CA2906EN.pdf. Acesso em: 25/06/2019.

[14] RIBEIRO, S.L., NAKAMURA, E.T. Context-Based Blockchain Platform Definition and Analysis Methodology. The 18th International Conference on Security and Management (SAM19), Las Vegas, United States, 2019.

[15] HYPERLEDGER. Hyperledger Fabric. Disponível em: https://www.hyperledger.org/projects/fabric. Acesso em: 18/06/2019.

[16] HYPERLEDGER. Java SDK for Hyperledger Fabric 1.2. Disponível em: https://github.com/hyperledger/fabric-sdk-java/tree/v1.2.2. Acesso em: 18/06/2019.

Caracterização morfométrica de uma bacia hidrográfica utilizando técnicas de geoprocessamento

Alexandro Gularte Schafer[1], *John Roger Roldan Aleixo*[2], *Allan Sampaio Pires*[1]

[1] Programa de Pós-Graduação em Computação Aplicada, Universidade Federal do Pampa, Bagé, Rio Grande do Sul, Brasil, Alexandro.schafer@unipampa.edu.br, allan.spires@gmail.com

[2] Campus Bagé, Universidade Federal do Pampa, Bagé, Rio Grande do Sul, Brasil, johnrroldan@gmail.com

RESUMO

O objetivo principal desta pesquisa é caracterizar a morfometria da bacia hidrográfica do Arroio Bagé, localizada no município de Bagé, Rio Grande do Sul, com base em parâmetros físicos. Para atingir o objetivo proposto, fez-se uso de dados de altimetria oriundos de imagens TanDEM-X, com resolução espacial de 90m, utilizando rotinas de geoprocessamento em um software de Sistemas de Informações Geográficas (SIG). O estudo resultou na delimitação da bacia hidrográfica e na determinação de índices referentes à rede de drenagem e a parâmetros morfométricos. Os resultados evidenciam que a bacia hidrográfica possui área de drenagem de 137km², com perímetro de 109km e altitude média de 252m. O fator de forma de 0,14 em conjunto com um coeficiente de compacidade de 2,63 e um índice de circularidade de 0,16 indicam que a bacia hidrográfica do Arroio Bagé tem forma alongada, o que contribui com a vazão de escoamento e para uma baixa suscetibilidade de ocorrência de inundações. A densidade de drenagem calculada foi de 2,19km/km², evidenciando que a bacia tem capacidade média de drenagem e com predominância de canais intermitentes e de primeira ordem.

PALAVRAS-CHAVE: Geoprocessamento, Geomorfologia, Arroio Bagé-RS.

RESUMO

The main objective of this research is to characterize the morphometry of the Bagé River watershed, located in Bagé county, Rio Grande do Sul, based on physical parameters. In order to reach the proposed objective, we used altimetric data from TanDEM-X images, with a spatial resolution of 90m, using geoprocessing routines in Geographic Information Systems (GIS) software. The study resulted in the delimitation of the watershed and in indices related to the drainage network and the morphometric parameters. The results show that the watershed has a drainage area of 137km², with a 109km perimeter and an average elevation of 252m. The form factor of 0.14, a compacity coefficient of 2.63 and a roundness index of 0.16 shows that the Bagé river watershed has an elongated shape, contributing to the outflow and to the low susceptibility to flood occurrence. The calculated drainage density was 2.19km/km², evidencing that the watershed has medium capacity of drainage and with predominance of intermittent and first order channels.

KEYWORDS: Geoprocessing, Geomorphology, Bagé River.

INTRODUÇÃO

O uso inadequado dos recursos naturais resulta na degradação do meio ambiente, pois pode provocar, dentre outros problemas, o assoreamento dos cursos d'água, menor disponibilidade de água para irrigação, a redução da produtividade agrícola (CANDIDO et al., 2008). De acordo com Guerra e Cunha (1996), as bacias hidrográficas constituem excelentes unidades de gestão de elementos naturais e sociais, possibilitando o acompanhamento das mudanças introduzidas pelo homem e as respectivas respostas da natureza. O gerenciamento de bacias hidrográficas é considerado por Lanna (2000) como um instrumento orientador das ações do poder público e da sociedade em longo prazo, e tem como objetivo garantir a adequação dos meios de exploração dos recursos naturais pelo homem, visando ao desenvolvimento sustentável.

A compreensão dos processos hidrológicos de uma bacia hidrográfica tem grande relevância em estudos ambientais e na gestão dos recursos hídricos. As características físicas e bióticas da bacia hidrográfica têm papel relevante nos processos do ciclo hidrológico, influenciando, dentre outros, na infiltração, na quantidade de água produzida como deflúvio, na evapotranspiração e nos escoamentos superficial e sub-superficial (TONELLO, 2005). As características físicas de uma bacia hidrográfica estão diretamente relacionadas aos seus parâmetros morfométricos. A análise morfométrica de bacias hidrográficas pode ser definida como a "análise quantitativa das interações entre a fisiografia e a sua dinâmica hidrológica"

que permite um conhecimento da dinâmica fluvial, bem como das relações existentes entre ela e os diversos componentes do meio físico e biótico (FARIA et al., 2009). Para Lindner et al. (2007) os índices morfométricos são importantes pressupostos para a prevenção de eventos hidrometeorológicos, como enchentes e estiagens, podendo ainda ser utilizados para apontar áreas de maior suscetibilidade a processos erosivos. Como instrumento, os indicadores morfométricos justificam a sua importância na gestão dos espaços urbanos e rurais, e podem contribuir para melhor aproveitamento e prevenção da degradação dos recursos naturais (TRAJANO et al., 2002).

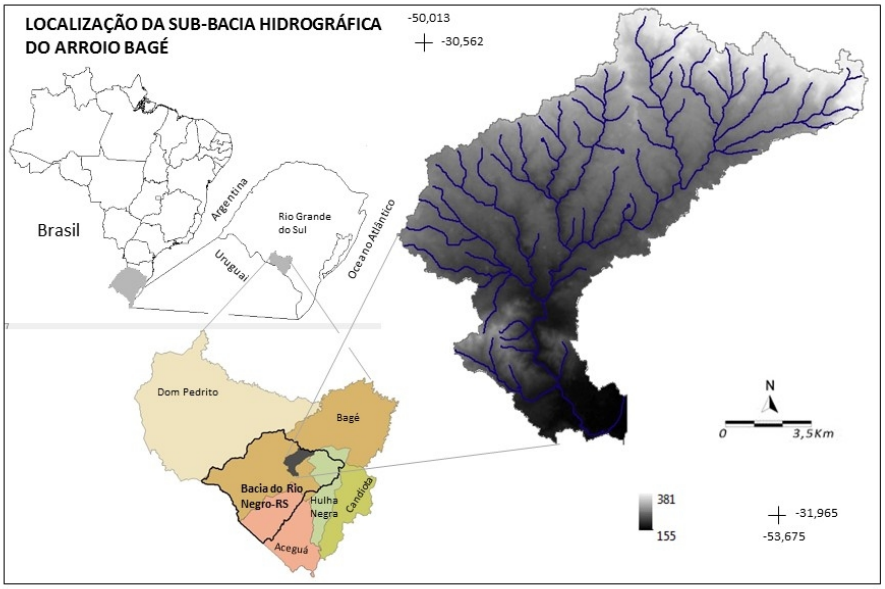
A bacia hidrográfica do Arroio Bagé, localizada no município de Bagé-RS, apresenta considerável taxa de urbanização e de atividade agropecuária, com suas águas sendo utilizadas para irrigação em áreas agrícolas e de pecuária, bem como para o abastecimento humano e de pequenas indústrias. Registram-se na bacia eventos esporádicos e localizados de inundação e problemas relacionados à qualidade de suas águas. Além disso, registram-se frequentes e prolongados períodos de estiagem na região em que a bacia hidrográfica está inserida. Desta maneira, entende-se que a realização de pesquisas que busquem avançar no conhecimento de suas características físicas com o intuito de relacioná-las à quantidade e qualidade de suas águas superficiais se constitua de grande relevância. Nesse sentido, o objetivo geral do presente trabalho é caracterizar a morfometria da bacia hidrográfica do Arroio Bagé, localizada no município de Bagé-RS, com base em parâmetros físicos, utilizando técnicas e ferramentas de geoprocessamento.

MATERIAL E MÉTODOS

Localização da área de estudo

A área de estudo compreende a bacia hidrográfica do arroio Bagé (Figura 1), localizada no município de Bagé-RS e que está inserida na bacia hidrográfica do Rio Negro-RS.

Figura 1 - Localização bacia hidrográfica do Arroio Bagé.



Método

A bacia hidrográfica foi delimitada utilizando técnicas e ferramentas de geoprocessamento da extensão ArcHydro (MAIDMENT, 2002) do software ArcGIS (ESRI, 2015), versão 10.3, com base em imagens obtidas por interferometria de radar da missão TanDEM-X (WESSEL et al., 2018; RIZZOLI et al., 2017), com resolução espacial de 90 metros. A delimitação se deu em nove etapas: a) preenchimento dos sumidouros, b) direção de fluxo, c) acumulação de fluxo, d) definição da rede de drenagem, e) segmentação da rede de drenagem, f) geração da grade de micro-bacias, g) geração das linhas de drenagem, h) junção das linhas de drenagem e micro-bacias, i) delimitação da sub-bacia. A partir da realização desse conjunto de procedimentos, definiu-se também a rede de drenagem.

A determinação das características físicas da bacia do Arroio Bagé foi realizada com base na quantificação: da área da bacia, do perímetro, do comprimento do rio principal, da ordem da bacia, da declividade, da hipsometria, do fator de forma (Kf), da razão de bifurcação (Rb), da razão entre os comprimentos médios dos canais de cada ordem (Lm), da densidade hidrográfica (Dh), da densidade de drenagem (Dd), do coeficiente de manutenção (Cm) e do Coeficiente de Compacidade (Kc). No quadro 1 apresentam-se as equações para a determinação de alguns dos parâmetros morfométricos da bacia hidrográfica, bem como a sua definição.

Quadro 1 – Equações e significado de parâmetros morfométricos.

Parâmetro	Equação	Significado dos índices

Fator de Forma (Kf)	$F = \frac{A}{L^2}$	A: área de drenagem (km^2); L: comprimento do eixo da bacia (km)
Razão de bifurcação (Rb)	$Rb = \frac{Nu}{Nu+1}$	Nu: número de segmentos de uma determinada ordem; Nu+1: número de segmentos da ordem imediata superior
Densidade hidrográfica (Dh)	$Dr = \frac{N}{A}$	N: número total de rios; A: área de bacia (km^2);
Densidade de drenagem (Dd)	$Dd = \frac{Lt}{A}$	Lt: comprimento total dos canais; A: área de bacia (km^2)
Coeficiente de manutenção (Cm)	$c_m = \frac{1}{D_d} \times 1000$	Dd: densidade de drenagem.
Coeficiente de compacidade (Kc)	$Kc = 0{,}28 \times \frac{P}{\sqrt{A}}$	Kc: coeficiente de compacidade; P: perímetro da bacia (km); A: área de bacia (km^2);

A classificação dos rios quanto à ordem reflete-se no grau de ramificação ou de bifurcação dentro de uma bacia. A ordem da rede de drenagem foi determinada de forma automática, segundo os critérios de Strahler (1957). A densidade de drenagem, expressa pela razão entre o comprimento total dos canais (Lt) e a área da sub-bacia (A), indica o grau de desenvolvimento do sistema de drenagem, fornecendo uma indicação da eficiência de drenagem da bacia. A partir da densidade de drenagem foi possível calcular o coeficiente de manutenção (Cm), que representa a área necessária para manter perene cada metro de canal de drenagem na bacia hidrográfica. A forma da bacia, que tem efeito sobre o seu comportamento hidrológico, foi definida a partir do Coeficiente de compacidade (Kc) e do Fator de Forma (Kf). Kc é a relação entre o perímetro da sub-bacia (P) e a circunferência de um círculo com área igual à da bacia (A). Kf é a razão entre a área da sub-bacia (A) e o comprimento axial da mesma (L). A hipsometria, que segundo Guerra e Guerra (2006) "é a representação altimétrica do relevo de uma área, através do uso de cores convencionais", foi determinada de acordo com as variações de altitude do terreno e representada pela carta hipsométrica e pela curva hipsométrica. A declividade de uma bacia hidrográfica consiste na inclinação maior ou menor do relevo, sendo um parâmetro de grande relevância por estar associado ao tempo de duração do escoamento superficial e de concentração da precipitação nos leitos do curso d'agua. A declividade foi classificada de acordo com os intervalos apresentados na tabela 1.

Tabela 1 - Classificação do relevo de acordo com a declividade. Adaptado de Embrapa (1979).

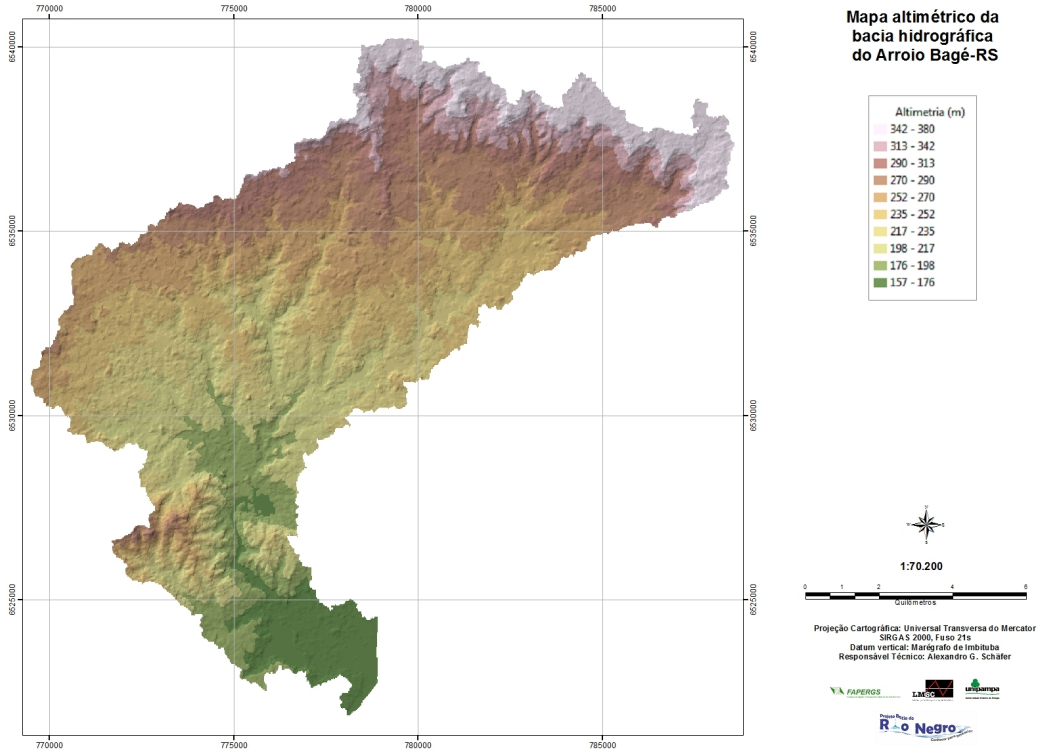
Declividade (%)	**Classificação do relevo**
0 – 3	Plano

3 -8	Suave ondulado
8 – 20	Ondulado
20 – 45	Fortemente ondulado
45 – 75	Montanhoso
>75	Fortemente montanhoso

RESULTADOS E DISCUSSÃO

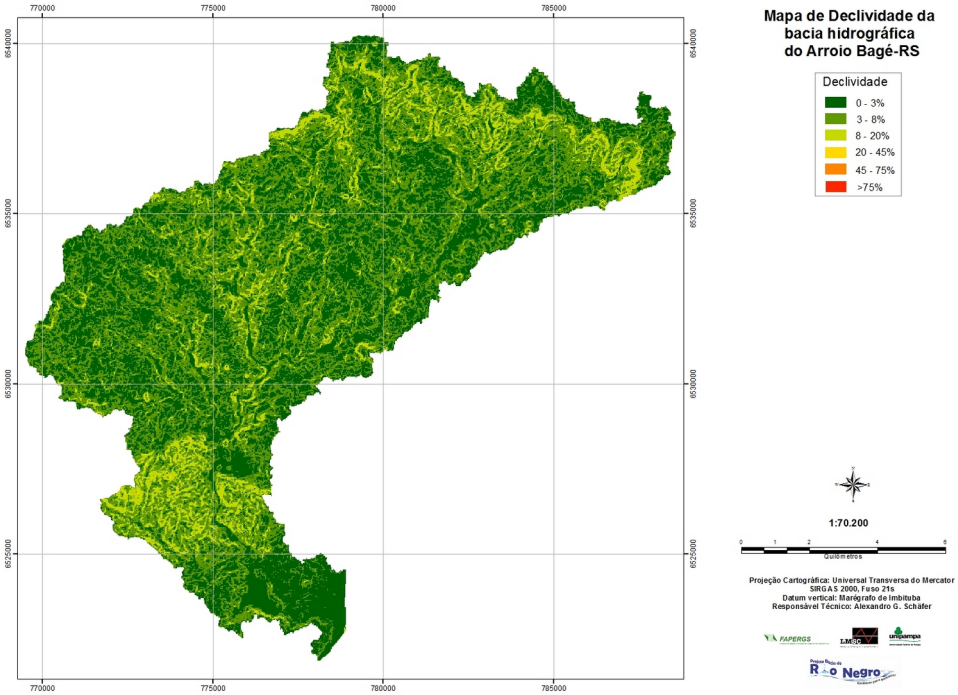
A área de drenagem da bacia hidrográfica do Arroio Bagé obtida por geoprocessamento possui 137 km^2 e 109 km de perímetro. Na figura 2 apresenta-se o mapa altimétrico da bacia.

Figura 2 - Mapa altimétrico da bacia hidrográfica do Arroio Bagé elaborado a partir das imagens TanDEM-X.



É possível verificar, na figura 2, que as altitudes variam de 157m a 380m. A altitude média é de 252m. Na figura 3 apresenta-se o mapa de declividade da bacia hidrográfica do Arroio Bagé.

Figura 3 - Mapa de declividade da bacia hidrográfica do Arroio Bagé



É possível verificar, na figura 3, que a bacia hidrográfica apresenta relevo plano a suave ondulado na maior parte de sua área de abrangência e relevo ondulado e fortemente ondulado em menor proporção. Na tabela 2 apresentam-se os índices físicos calculados a partir das equações apresentadas no quadro 1.

Tabela 2 - Índices físicos da bacia hidrográfica do Arroio Bagé

Área	137	km^2
Perímetro	109	km
Altitude média	252	m
Coeficiente de compacidade	2,63	
Fator de forma	0,14	
Densidade de drenagem	2,185	km^{-1}
Ordem da rede hídrica	5	
Comprimento da rede hídrica	334,15	km
Comprimento do Rio principal	25,41	km
Extensão do percurso superficial	20,54	m
Índice de Circularidade	0,16	
Razão de Bifurcação	8,32	
Densidade hidrográfica	3,04	km^{-2}
Coeficiente de manutenção	456,62	

De acordo com os resultados apresentados na tabela 2, a bacia hidrográfica do Arroio Bagé é de 5º ordem, possuindo 417 segmentos, o que indica grau médio de ramificação. O

elevado valor do coeficiente de compacidade (Kc = 2,63), associado ao índice de circularidade de 0,16 indica que a bacia possui formato alongado, o que está associado a um bom processo de escoamento e, consequentemente, baixa propensão à ocorrência de inundações. Além disso, nesse tipo de configuração de bacia, os tributários encontram o curso d'água principal em diversos pontos, distribuindo as vazões ao londo de sua extensão. O formato alongado denota ainda menor possibilidade de precipitações intensas cobrirem simultaneamente toda sua extensão, reduzindo a probabilidade de grande quantidade de água vinda de diferentes tributários contribuírem, simultaneamente, à calha principal, reduzindo o risco de extravasamento da água no canal (FRAGA et al., 2014). A densidade de drenagem foi de 2,185 km/km^2, o que indica uma bacia medianamente drenada, segundo valores apresentados por Villela e Mattos (1975).

A bacia apresenta densidade de drenagem (Dd) igual a 2,19 km/km². Segundo Villela e Mattos (1975), esse índice pode variar de 0,5km/ Km2 em sub-bacias com drenagem pobre a 3,5km/ Km2, ou mais para sub-bacias bem drenadas, portanto a sub-bacia do arroio Bagé pode ser considerada de média capacidade de drenagem.

A densidade hidrográfica (Dh) ficou em 3,04 canais/km², considerado um valor médio a alto. Esse parâmetro indica a média da quantidade de canais por quilômetro quadrado de uma bacia, demonstrando sua capacidade hídrica e competência na formação de novos canais fluviais (SANTOS; MORAIS, 2012). Com relação ao coeficiente de manutenção (Cm), a bacia hidrográfica necessita de área equivalente a 456,62 m²/m para a manutenção de um metro do canal de escoamento, sugerindo que os corpos hídricos de primeira ordem apresentam propensão à características de intermitência.

CONCLUSÕES

A partir da análise morfométrica da bacia hidrográfica do Arroio Bagé verificou-se, entre outras questões, que a drenagem apresenta baixa hierarquia fluvial e que os canais fluviais de primeira ordem, em sua maioria, são intermitentes. Isto indica que a rede de drenagem pode ser mais sensível à alterações hidrológicas. Desta maneira, é importante que seja evitada, na medida do possível, a ocorrência de alterações nos canais de drenagem de origem antrópica.

A partir da realização da pesquisa, espera-se avançar no conhecimento dos aspectos físicos da bacia hidrográfica do Arroio Bagé, obtendo um conjunto de dados e de informações

que subsidiem futuras pesquisas hidrológicas e de drenagem. Além disso, espera-se que esses dados e informações venham a auxiliar na identificação e na manutenção dos arroios e dos sistemas de drenagem existentes, além de subsidiarem questões referentes ao processo de planejamento e de gestão territorial do município de Bagé-RS.

AGRADECIMENTOS

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPQ, através do Programa Institucional de Bolsas de Iniciação Científica – PIBIC.

REFERÊNCIAS

CANDIDO, Humberto Gois et al. Degradação ambiental da bacia hidrográfica do rio Uberaba-MG. 2008. 96 f. Tese (Doutorado) - Unesp, Jaboticabal, 2008.

CHRISTOFOLETTI, A. Geomorfologia. São Paulo / SP, Ed. Editora Edgard Blücher, 2ª. ed., 1980.

EMBRAPA. Serviço Nacional de Levantamento e Conservação de Solos (Rio de Janeiro, RJ). Súmula da 10. Reunião Técnica de Levantamento de Solos. Rio de Janeiro, 1979, 83p.

ESRI. Environmental Systems Research Institute. ArcGIS Professional GIS for the desktop, version 10.3. Software. 2015.

FARIA, M. M.; ZACCHI, R. C.; FERREIRA, E. S. Caracterização morfométrica e biológica da Bacia Hidrográfica do Córrego Serafim, Sub-Bacia do Rio Paraibuna, Juiz de Fora MG. In: Anais do SIMPÓSIO BRASILEIRO DE GEOGRAFIA FÍSICA APLICADA, 13, 2009.

FRAGA, M. S.; FERREIRA, R. G.; SILVA, F. B.; VIEIRA, N. P. A.; SILVA, D. P.; BARROS, F. M.; MARTINS, I, S, B. Caracterização morfométrica da bacia hidrográfica do rio Catolé Grande, Bahia, Brasil. Nativa, v. 2, n. 4, 2014, p. 214-218.

GUERRA, A.J.T.; CUNHA, S. B. Degradação ambiental. In: CUNHA, S. B. Geomorfologia e meio ambiente. Rio de Janeiro: Bertrand Brasil, 1996. p. 337-339.

GUERRA, A. T.; GUERRA, A. J. T. Novo dicionário geológico-geomorfológico. 5.ed. Rio de Janeiro: Bertrand Brasil, 2006.

LANNA, A. E. A inserção da gestão das águas na gestão ambiental. In: MUÑOZ, H. R. (Org.). Interfaces da gestão de recursos hídricos: desafios da lei de águas de 1997. 2.ed. Brasília, DF: Secretaria de Recursos Hídricos, 2000. p. 75-108.

LINDNER, E. A.; GOMIG, K.; KOBIYAMA, M. Sensoriamento remoto aplicado à caracterização morfométrica e classificação do uso do solo na bacia rio do Peixe/SC. In:

Anais do SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 13, 2007, Florianópolis. p. 3405-3412.

MAIDMENT, D. (2002) ArcHydro GIS for Water Resources. ESRI Press, Redlands, CA.

RIZZOLI, P., MARTONE, M., GONZALEZ, C., WECKLICH, C., BORLA TRIDON, D., BRÄUTIGAM, B., BACHMANN, M., SCHULZE, D., FRITZ, T., HUBER, M., WESSEL, B., KRIEGER, G., ZINK, M., MOREIRA, A. Generation and performance assessment of the global TanDEM-X digital elevation model. ISPRS Journal of Photogrammetry and Remote Sensing, 2007, Vol 132.

SANTOS, D. A. R.; MORAIS, F. Análise morfométrica da bacia hidrográfica do rio Lago Verde como subsídio à compartimentação do relevo da região de Lagoa da Confusão – TO. Revista Geonorte, v. 3, n.4, p. 617-629, 2012.

STRAHLER, A.N. Quantitative analysis of watershed geomorphology. New Halen: Transactions: American Geophysical Union, v. 38, 1957, p. 913-920.

TONELLO, K.C. Análise hidroambiental da bacia hidrográfica da cachoeira das Pombas, Guanhães, MG. Tese (Doutorado em Ciências Florestal) – Universidade Federal de Viçosa, Viçosa, 2005. 69p.

TRAJANO, S. R. R.; SPADOTTO, C. A.; HOLLER, W. A.; DALTIO, J.; MARTINHO, P. R.; FOIS, N. S.; SANTOS, B. B.; TOSCHI, H. H.; LISBOA, F. S. Análise Morfométrica de Bacia Hidrográfica - Subsídio à Gestão Territorial: Estudo de Caso no Alto e Médio Mamanguape. Boletim de Pesquisa e Desenvolvimento 2. Embrapa Gestão Territorial, Campinas, 2012.

VILLELA, S. M.; MATTOS, A. Hidrologia aplicada. São Paulo: McGraw-Hill do Brasil, 1975. 245 p.

WESSEL, B., HUBER, M., WOHLFART, C., MARSCHALK, U., KOSMANN, D., ROTH, A. Accuracy Assessment of the Global TanDEM-X Digital Elevation Model with GPS Data. ISPRS Journal of Photogrammetry and Remote Sensing. Vol. 139, 2018, pp. 171-182.

Algoritmos Genéticos para Otimização em Sistemas de Produção de Ruminantes em Pastagens

Joel de Oliveira Santos, Laura Emmanuella Alves dos Santos Santana, Henrique Rocha de Medeiros

Escola Agrícola de Jundiaí, Universidade Federal do Rio Grande do Norte, Macaíba, RN, Brasil, joeloliveira56@gmail.com, lauraemmanuella@eaj.ufrn.br, hrdemedeiros@yahoo.com.br

RESUMO

Os animais, o clima, o solo e as plantas são elementos de um sistema de produção de ruminantes e todos eles devem ser observados nos processos de tomada de decisão do sistema, especialmente porque o processo de produção é longo e, em sua maioria, ininterruptível. Tais decisões são impregnadas de riscos e incertezas, uma decisão errada pode acarretar uma grande perda de produtividade. Para auxiliar no planejamento dessas decisões adota-se uma ferramenta conhecida como orçamento forrageiro que é uma ferramenta gerencial de planejamento alimentar que possibilita ao produtor fazer a quantificação das fontes de alimento e das necessidades do rebanho garantindo um equilíbrio entre a produção e a demanda de forragem. Este trabalho apresenta a utilização de algoritmos genéticos para gerar cenários produtivos ótimos no contexto de um sistema de produção de ruminantes em pastagens possibilitando a tomada de decisão em relação à quantidade de ruminantes que poderão pastejar em uma determinada área em função da oferta de forragem. Os cenários ótimos apresentam a quantidade máxima de dias e de animais em uma determinada área pastejável até o ponto em que não há a necessidade de acrescentar suplementação alimentar para os animais. Os resultados desse trabalho foram comparados com trabalhos envolvendo orçamento forrageiro, mostrando que o algoritmo genético modela o sistema de forma correta com a vantagem de gerar vários cenários possíveis que minimizam o saldo de forragem.

PALAVRAS-CHAVE: Orçamento forrageiro, Computação evolucionária.

ABSTRACT

Animals, climate, soil and plants are all elements of a ruminant production system and all of them must be observed in the system's decision-making processes, especially since the production process is long and mostly uninterrupted. Such decisions are fraught with risks and incidents, one wrong decision can cause a great loss of execution. To assist in the planning of these decisions, a tool known as forage budgeting is adopted, which is a managerial food planning tool that allows for the quantification of food sources and herd requirements, using the balance between forage production and demand. This paper presents the use of genetic algorithms to generate the best resources in the context of a pasture ruminant production system, allowing a decision to be made regarding the number of ruminants that can graze in a specific area as a function of supply forage. Ideal scenarios present the maximum number of days and animals in a specific area pasted to the point where there is no need to increase feed supplementation for the animals. The results of this work were compared with the forage budget works, showing which genetic algorithm of model or system correctly with the advantage of generating several possible results that minimize the forage balance.

KEYWORDS: Forage Budget, Evolutionary computing.

INTRODUÇÃO

A agropecuária possui grande importância na economia do Brasil, tendo forte relevância nas exportações e no abastecimento do mercado interno. Mundialmente, a pecuária brasileira destaca-se, tornando-se cada vez mais produtiva, eficiente e sustentável devido a grandes avanços nos campos de estudo da alimentação, genética, sanidade e da reprodução (EMBRAPA, 2016).

Nos sistemas agropecuários (SA) o clima, o solo, os animais e as plantas são componentes inerentes ao processo de produção (MEDEIROS; PEDREIRA; MELLO 2006). Como o número de variáveis que influenciam na produção é grande, faz-se necessário um planejamento adequado para auxiliar o produtor pecuário a tomar as devidas decisões. O orçamento forrageiro é um método de planejamento estratégico dos sistemas de produção animal em pastagens que permite ao pecuarista fazer uma demanda adequada da suplementação dos animais ao longo do ano (MEDEIROS et al., 2008), permitindo que aconteça um planejamento e auxílio nas tomadas de decisões quanto à taxa de lotação e manejo dos recursos forrageiros.

O orçamento forrageiro é comumente acompanhado a partir do uso de planilhas, sendo feitas simulações a partir de entradas testadas pelo produtor. Neste sentido, as técnicas computacionais de otimização podem auxiliar o processo decisório dos SA e na otimização de processos, uma vez que essas técnicas resultam em valores ótimos para as variáveis de entrada do sistema que minimizam custos ou maximizam lucros.

Na resolução de problemas de otimização, existe uma grande relevância no uso de técnicas computacionais que visam encontrar soluções aproximadas, como os Algoritmos Genéticos (AG). Isso ocorre pelo fato de lidarem com qualquer problema de otimização, não estando atrelados a problemas específicos de derivadas (MIRANDA, 2009).

O algoritmo genético tem como finalidade realizar uma busca estocástica para encontrar as melhores soluções para determinado problema, para isso o algoritmo mantém a cada iteração uma população de possíveis soluções. Essas soluções são avaliadas e submetidas a um processo de seleção, com maior probabilidade para a escolha das melhores soluções. As soluções selecionadas são combinadas de forma a gerar soluções ainda melhores, repetindo esse processo até que se tenha passado tempo suficiente ou se tenha obtido uma solução satisfatória para o problema (LINDEN, 2008).

Este trabalho propõe o uso de AG para gerar cenários ótimos em sistemas de produção de ruminantes em pastagens, de forma que possibilite a determinação da quantidade máxima de dias e ruminantes que poderão alimentar-se em uma área pastejável para fins de planejamento de produções. São relevantes para este objetivo, a oferta de forragem (plantas utilizadas na alimentação animal) disponível para consumo e algumas características em relação ao consumo e conversão alimentar dos ruminantes.

Na literatura encontram-se alguns trabalhos que aplicaram algoritmos genéticos no agronegócio e se assemelham ao objetivo do trabalho proposto, mas com abordagens diferentes. Ferreira (2015) utilizou em sua pesquisa o uso de AG para encontrar soluções que diminuíssem os custos de implantação e irrigação das lavouras de arroz. Vieira et al. (2010) aplicou AG como tomada de decisão no planejamento de um agronegócio, visando o maior retorno econômico da atividade exercida. Já Câmara Júnior (2005), utilizou AG otimizando a determinação de áreas, definindo as culturas a serem cultivadas numa propriedade agrícola, visando maximizar a receita líquida de planos agrícolas.

MATERIAL E MÉTODOS

O desenvolvimento do trabalho constituiu-se em quatro etapas: (i) levantamento de informações necessárias para o desenvolvimento do AG, (ii) formulação do problema, (iii) implementação do AG e (iv) validação dos resultados obtidos.

Levantamento de informações

Essa etapa do desenvolvimento teve o intuito de reunir todas as informações necessárias em relação ao sistema de produção de ruminantes em pastagens. O AG pode ser modelado para qualquer sistema de produção extensiva em pastagens, neste caso específico foram reunidas informações ligadas à produção de ovinos. Algumas informações que são relevantes para a elaboração do algoritmo seguem listadas abaixo:

- TMF = Taxa média de acúmulo da forragem medida em kg;
- GMD = Ganho de peso médio diário por animal, medido em kg/animal/dia;
- PV = Peso vivo, medido em kg/animal;
- CV = Consumo voluntário, medido em kg de alimento/animal/dia;
- MF = Massa de forragem, alimento disponível para os animais, medida em kg;
- SF = Saldo de forragem, medido em kg.

A TMF é obtida através do acúmulo de forragem durante o dia, numa relação entre a massa de forragem inicial e a massa de forragem final de um determinado período, sua medida é estimada em kg e varia em relação ao clima, solo e espécie de forrageira.

O GMD é um valor constante medido em kg, que expressa a relação entre o consumo e a conversão do alimento em produto animal.

O PV representa o peso vivo de um animal adulto. Sua medição é expressa em kg numa relação entre seu peso inicial, GMD e quantidade de dias do pastejo.

O CV é um cálculo que estima a quantidade de matéria seca que os animais irão consumir diariamente. Sua medição é obtida em kg numa relação entre a variação de PV e do GMD. Para encontrar o CV foi utilizado o modelo proposto por Cabral et al. (2008) onde é estimado o consumo dos animais em função do seu PV e do GMD. A seguir, na Equação 1, o cálculo que expressa esta relação.

$$CV = 0{,}311 + (0{,}0197 * PV) + (0{,}682 * GMD) \text{¿} \quad (1)$$

MF representa a quantidade de alimento (forragem) disponível para o consumo dos animais, é medida em kg.

SF é a quantidade de alimento que restou após o consumo dos animais. É medido em kg e obtido numa relação entre a TMF, MF e o CV.

Formulação do problema

A fase de formulação do problema é muito importante para o desenvolvimento do trabalho. Está ligada diretamente com o AG, pois nessa etapa são esclarecidos alguns questionamentos que são as peças chave para a construção do AG. Sendo elas: (i) o que será otimizado, (ii) qual a função objetivo e (iii) quais as restrições do problema.

O objetivo do problema é encontrar uma solução ótima para a quantidade de dias e animais no sistema de produção de modo que minimize o saldo de forragem em função do consumo, de acordo com a função abaixo:

$$Minimizar\, f(NA, ND) = (SF + TMF * ND) - (CV * NA * ND) \quad (2)$$

Sujeito a:

$SF \geq 0, TMF \geq 0, CV \geq 0$

$Mínimo \leq ND \leq Máximo$

$Mínimo \leq NA \leq Máximo$

Onde o SF é o saldo de forragem após o consumo dos animais, ND é a quantidade de dias e NA é quantidade de animais. A quantidade de dias e a quantidade de animais estará dentro de um intervalo definido pelo produtor.

A função objetivo do problema é expressa numa relação entre o consumo e a oferta de forragem.

Faz-se o uso da Equação 1 para calcular o consumo dos animais em função dos dias, calcula-se a acumulação de forragem ao longo dos dias e subtrai-se o consumo total dos animais pelo saldo de forragem final.

Será otimizado a quantidade de dias e animais, sendo que os valores de TMF, GMD, PV e MF deverão ser informados pelo produtor e CV deverá ser calculado pelo sistema.

Essa configuração deverá retornar soluções que otimizem a quantidade de dias e animais, onde o saldo de forragem seja o mínimo possível acima ou igual a zero. A partir do momento que o AG encontrar esses valores deverá finalizar sua execução e retornar as soluções encontradas.

Construção do Algoritmo

A construção do AG deu-se após o levantamento dos requisitos necessários em relação aos animais e a forragem. O cromossomo que contém a solução é constituído por duas informações: (i) número de animais e (ii) quantidade de dias do pastejo.

O algoritmo implementado está apresentando na Figura 1.

Figura 1 – Fluxo do algoritmo genético implementado

Algoritmo Fluxo básico de um algoritmo genético simples

Seja $S(t)$ uma população de cromossomos na geração t;
$t \leftarrow 0$;
Inicializar $S(t)$;
Avaliar $S(t)$;
while o critério de parada não for satisfeito **do**
  $t \leftarrow t + 1$;
  Selecionar $S(t)$ a partir de $S(t-1)$;
  Aplicar operadores genéticos sobre $S(t)$;
  Avaliar $S(t)$;
end while
Retornar melhor solução encontrada.

Fonte: O próprio autor.

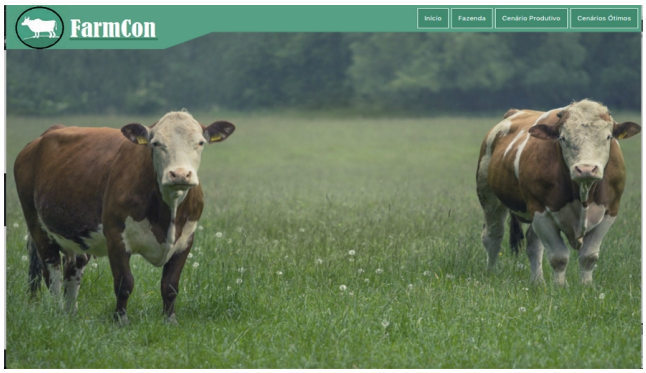
Foi utilizada a seguinte configuração para obtenção da solução.

- População inicial = 100 indivíduos;
- Taxa de cruzamento = 90%;
- Elitismo = Presente;
- Taxa de mutação = 5%;
- Número máximo de gerações = 1000;
- Técnica de seleção = Torneio binário;
- Técnica de cruzamento = 1 ponto de corte;
- Critério de parada = Estagnação ou número máximo de gerações.

RESULTADOS E DISCUSSÃO

O AG desenvolvido foi executado em um aplicativo denominado FarmCon (ARGOLO; MATIAS, 2018) desenvolvido por alunos do curso de informática da Escola Agrícola de Jundiaí/UFRN.

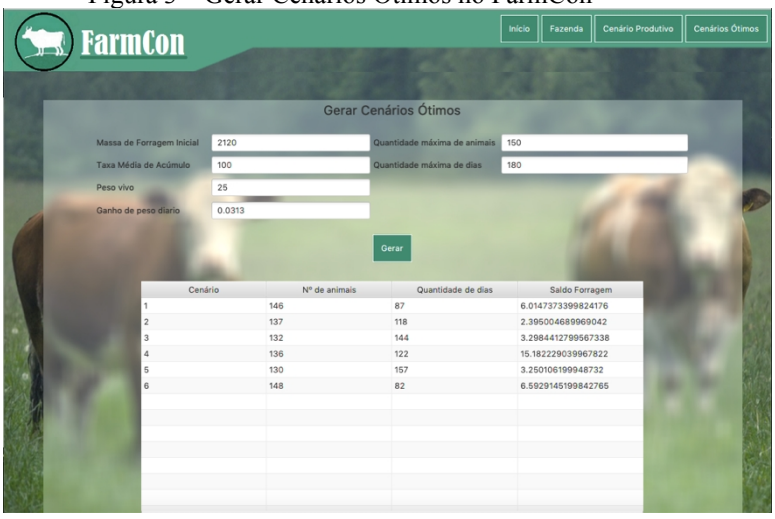
Figura 2 – Aplicativo FarmCon



Fonte: (ARGOLO; MATIAS, 2018).

Nesse aplicativo, a opção Cenários Ótimos apresentada na Figura 2 permite ao produtor gerar vários cenários de orçamento forrageiro para uma determinada configuração dada, como pode ser observado na Figura 3.

Figura 3 – Gerar Cenários Ótimos no FarmCon



Fonte: (ARGOLO; MATIAS, 2018).

Para validar os resultados obtidos pelo AG foram utilizados os dados contidos no trabalho realizado por Carnevalli et al. (2001), onde o objetivo do estudo foi detalhar o conhecimento de como o desempenho animal e características do pasto respondem a mudanças na altura do pastejo.

Do trabalho acima citado, foram utilizados os seguintes dados para configuração do AG:

- MF;
- TMF;

- PV;
- GMD.

Para extrair o valor da MF foi utilizada a seguinte fórmula:

$$MF = A * D \quad (3)$$

Onde A é a altura da forragem expressa em milímetros e D é a densidade da forragem expressa em quilogramas por centímetro.

Foram desenvolvidas planilhas de orçamento forrageiro em relação à quantidade de dias e animais que foram utilizados no experimento de Carnevalli et al. (2001), e outra em relação ao resultado obtido pelo AG com intuito de comparar os resultados a fim de verificar se os resultados que o AG encontrou possuem o mesmo comportamento que o do experimento em comparação.

A planilha de orçamento forrageiro do trabalho adotado para comparação foi desenvolvida em relação aos períodos estacionais que a produção foi realizada e altura do pasto, considerando a quantidade de dias para cada estação onde a produção foi implementada, foi definido um contingente de 180 dias englobando a estação primavera e verão, e uma quantidade de animais igual a 96 para uma pastagem de altura igual a 100 mm.

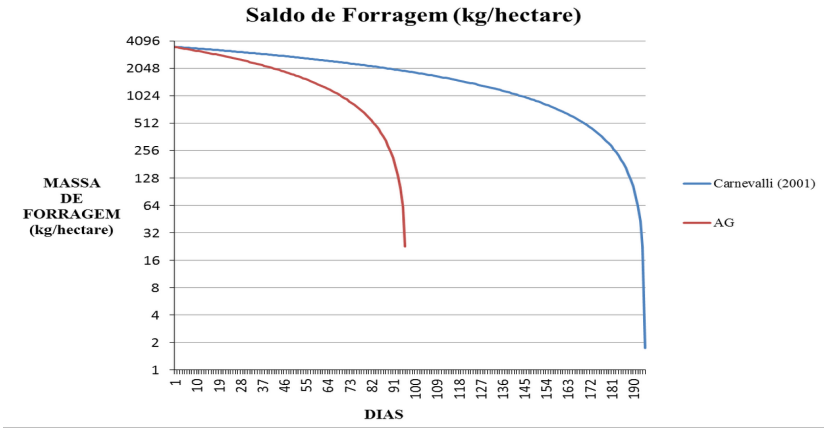
A quantidade de animais foi obtida a partir da informação referente ao ganho por área, que está expressa na formula a seguir:

$$Ganho\ por\ área = NA * GMD \quad (4)$$

Onde o NA é o número de animais.

O AG gerou uma configuração ótima dada a configuração vista anteriormente retornando os seguintes valores: 120 animais e 96 dias minimizando o saldo de forragem final. Após isso também foi gerada uma planilha com o orçamento forrageiro a fim de verificar se a solução ótima obtida pelo AG atendeu a expectativa de minimizar o saldo de forragem. O resultado obtido pelo AG e o comportamento em relação ao experimento escolhido pode ser observado no gráfico da Figura 4.

Figura 4: Gráfico de comparação



Fonte: O próprio autor.

Pode-se observar que as funções possuem o mesmo comportamento, o AG retornou um saldo de 22,74 kg de forragem para o dia 96 minimizando o saldo de forragem atendendo a expectativa, pois no dia posterior (97) o saldo de forragem estaria em -15,55 kg, ou seja, faltaria alimento para os animais após o dia 96.

Quando comparado o resultado obtido pelo AG com o do experimento de Carnevalli et al. (2001), pode-se notar que existe uma diferença em relação a quantidade de dias. O experimento tem uma quantidade de dias maior, pois existe um contingente menor de animais, onde o mesmo retornou uma sobra de alimento, já o AG retornou uma solução ótima que minimizou a sobra de alimento mantendo uma relação inversamente proporcional para a quantidade de dias e animais da produção.

CONCLUSÕES

O AG apresentado neste trabalho tem o objetivo de realizar otimizações de áreas pastejáveis de forma que possibilite a criação de cenários produtivos para fins de planejamento de produções de ruminantes em pastagens, determinando a quantidade de animais e dias que se pode ter na produção sem a necessidade de acrescentar ração para os animais, proporcionando um auxílio na tomada de decisões para a criação de sistemas de produções de ruminantes em pastagens.

A possibilidade de analisar diversas configurações de cenários de produção proporciona que haja um melhor planejamento de produções, auxiliando na definição de qual produção melhor se adapta a problemática que está inserida, suprindo as dificuldades encontradas no planejamento de produções por se tratarem de períodos longos e que se eventualmente forem alterados ao longo da produção ocasionarão em uma perda de produtividade.

A utilização de AG se mostrou satisfatória na pesquisa tendo em vista que lidam com qualquer problema de otimização. O AG desenvolvido atendeu o objetivo proposto, onde as soluções encontradas proporcionaram a otimização da quantidade de dias e animais, minimizando o saldo de forragem em função do consumo animal, gerando vários cenários possíveis que minimizam o saldo de forragem.

REFERÊNCIAS

ARGOLO, W.; MATIAS, Y. **Farmcom: Planejamento e Simulação de Sistemas de Produção de Ruminantes em Pastagens**. 2018. 68f. Trabalho de Conclusão de Curso (Técnico em Informática) – Escola Agrícola de Jundiaí – UFRN, Macaíba, 2018.

CABRAL, L. S. et al. Estimativas dos requisitos nutricionais de ovinos em condições brasileiras Nutrients. **Revista Brasileira de saúde e produção animal**, v. 9, n. 3, p. 529–542, 2008.

CÂMARA JÚNIOR; A. C. **Suporte à decisão para o planejamento agrícola utilizando algoritmos genéticos**. 2005. 155f. Dissertação (mestrado) – Universidade Federal de Minas Gerais, Escola de Engenharia.

CARNEVALLI, R. A. et al. Desempenho de ovinos e respostas de pastagens de tifton 85 (Cynodon spp.) sob lotação contínua. **Scientia Agricola**, v. 58, n. 1, p. 7–15, 2001.

EMBRAPA. **Grupo de Produção Animal (GPA)**. Disponível em <https://www.embrapa.br/gado-de-corte/pesquisa-e-desenvolvimento/grupo-de-producao-animal>. Acesso em outubro de 2016.

FERREIRA, A. et al. Algoritmos evolutivos no projeto de irrigação de lavouras de arroz. In: Congresso Brasileiro de Agroinformática, X, 2015, Ponta Grossa. **Anais...**Ponta Grossa: UEPG, 2015. Disponível em: < http://eventos.uepg.br/sbiagro/2015/anais/SBIAgro2015/index.htm>

LINDEN, R. **Algoritmos genéticos**. Rio de Janeiro: Brasport, 2008.

MEDEIROS, H. R.; PEDREIRA, C. G. S.; MELLO, A. C. L. Avaliação do modelo Stockpol em sistemas de produção. **Pasturas Tropicales**, v. 28, n. 2, 2006.

MEDEIROS, H. R. et al. Planejamento de Sistemas de Produção de Caprinos e Ovinos Utilizando Orçamento Forrageiro. **Embrapa**, p. 1–7, 2008.

MIRANDA, M. N. de. **Algoritmos Genéticos: Fundamentos e Aplicações**. Disponível em <http://www.gta.ufrj.br/~marcio/genetic.html>.

VIEIRA, D. A. H. et al. Algoritmos Genéticos como Ferramenta Auxiliar na Tomada de Decisão em Atividades de Gestão Agroindustrial. **Informe Gepec**, Toledo, v. 14, n. 1, p. 113-126, jan./jun. 2010

Predição da germinação de sementes de soja usando aprendizado de máquina

Iaslan do Nascimento Paulo da Silva, Wesley Leocádio Barbosa da Silva, Laura Emmanuella Alves dos Santos Santana, Márcio Dias Pereira

Escola Agrícola de Jundiaí, Universidade Federal do Rio Grande do Norte, Macaíba, RN, Brasil, iaslannascimento@gmail.com, silvawesley374@gmail.com, lauraemmanuella@eaj.ufrn.br, marcioagron@gmail.com

RESUMO

A soja (*Glycine max*) é uma das mais importantes culturas na economia mundial, porém o sucesso da lavoura de soja depende de diversos fatores, sendo o mais importante deles a utilização de sementes de elevada qualidade, capazes de gerar plantas de alto vigor. A avaliação da qualidade de sementes é feita por testes de germinação e de vigor que ajudam na identificação de lotes com melhor desempenho. Um dos desafios encontrados pelos produtores de soja é que essas análises para avaliação da qualidade da semente são destrutivas e demandam tempo, que é considerado longo e caro na cadeia que envolve a produção e comercialização das sementes. Uma das formas de se aprimorar e tornar rápido o processo de avaliação é através da análise de imagens por computador. Alguns trabalhos têm sido desenvolvidos nesse sentido, utilizando a análise de imagens digitais para classificação da morfologia de sementes e avaliação fisiológica. O presente trabalho propõe a utilização de técnicas de aprendizado de máquina para predizer a germinação das sementes de soja a partir de características físicas da semente obtidas por meio da análise digital de imagens. Busca-se com isso maior agilidade no processo de avaliação da qualidade de sementes de soja, além de caracterizar um processo de avaliação da semente não destrutivo.

PALAVRAS-CHAVE: Avaliação da qualidade de sementes, Análise de imagens digitais, Seleção de atributos, Algoritmos genéticos.

ABSTRACT

Soybean (Glycine max) is one of the most important crops in the world economy, but the success of soybean crops depends on several factors, the most important being the use of high quality seeds capable of generating high vigor plants. The quality of seeds is evaluated by germination and vigor tests that help in the identification of better performance lots. One of the challenges encountered by soybean farmers is that these analyzes for assessing seed quality are destructive and time consuming, which is considered long and expensive in the chain that involves the production and commercialization of the seeds. One of the ways to improve and speed up the evaluation process is through computer image analysis. Some works have been developed in this sense, using digital image analysis to classify seed morphology and physiological evaluation. The present work proposes the use of machine learning techniques to predict the germination of soybean seeds from the physical characteristics of the seed obtained through digital image analysis. This research seeks to increase agility and reliability in the process of evaluating the quality of soybeans, in addition to characterizing a non-destructive seed evaluation process.

KEYWORDS: Seed quality assessment, Digital imaging analysis, Feature selection, Genetic Algorithms.

INTRODUÇÃO

A soja (*Glycine max*) é uma das mais importantes culturas na economia mundial. Seus grãos são muito usados pela agroindústria (produção de óleo vegetal e rações para alimentação animal), indústria química e de alimentos, além disso, seu uso como fonte alternativa de biocombustível é crescente (COSTA NETO; ROSSI, 2000). O Brasil é o segundo maior produtor do grão com a produção de 95 milhões de toneladas (EMBRAPA, 2016).

O sucesso da lavoura de soja depende de diversos fatores, porém o mais importante deles é a utilização de sementes de elevada qualidade, capazes de gerar plantas de alto vigor. Essas sementes de melhor qualidade propiciam a germinação e a emergência de plântulas em campo de maneira rápida e uniforme, resultando na produção de plantas de alto desempenho, que têm um potencial produtivo mais elevado (FRANÇA NETO et al., 2015).

A avaliação da qualidade de sementes é feita por testes de germinação e de vigor que ajudam na identificação de lotes com melhor desempenho. Os principais testes aplicados nesse contexto são definidos pelas Regras de Análise de Sementes (BRASIL, 2009) e pelo manual de vigor da Associação Brasileira de Tecnologia de Sementes, a ABRATES (KRZYZANOWSKI et al., 1999). Um dos desafios encontrados pelos produtores de soja é que essas análises para avaliação da qualidade da semente são destrutivas e demandam tempo, que é considerado longo e caro na cadeia que envolve a produção e comercialização das sementes (MASSETO et al., 2008). Uma das formas de se aprimorar e tornar rápido o processo de avaliação é através da análise de imagens por computador; trata-se de uma técnica de inspeção rápida, econômica e objetiva que tem-se expandido em diversos meios do setor industrial e onde a precisão e velocidade devem satisfazer uma exigência crescente de produção e qualidade (BROSNAM, 2002).

A análise digital de imagens de sementes tem sido pesquisada a algum tempo, como por exemplo em Guedes et al. (2011), onde os autores validaram o uso da análise de imagem da semente de soja para avaliar as características físicas das sementes, verificando que esse método é tão eficaz quanto os métodos de avaliação tradicional, a saber, o método com paquímetro e o método com projeção da imagem em papel milimetrado, com a vantagem de ser não destrutivo e mais rápido.

Mais recentemente, a análise de imagens de sementes também tem sido pesquisada em conjunto com algoritmos de aprendizado de máquina (AM) que permitem ao sistema computacional adquirir conhecimento sobre um domínio desejado através de amostras desse domínio (FACELI et al., 2017). Esse aprendizado é dito supervisionado quando a base de dados com as amostras possui um atributo alvo, que determina o que o algoritmo deve aprender. Na área de AM os dados são muito importantes, pois sua qualidade impacta diretamente na qualidade do aprendizado e consequente desempenho dos algoritmos, dessa forma, os atributos que constituem a base de dados devem ser relevantes, não redundantes e não ruidosos. Existem várias técnicas que podem ser adotadas para selecionar subconjuntos mais adequados de atributos que promovam nos algoritmos de AM um melhor aprendizado sobre os dados, dentre elas uma abordagem que tem demonstrado bom desempenho em comparação com as demais é a metaheurística Algoritmos Genéticos (AG) que realiza uma busca pelo subconjunto de atributos ótimo que maximize a acurácia do algoritmo de AM

adotado (JOVIC; BRKIC; BOGUNOVIC, 2015). Exemplos de trabalho que usaram análise de imagens em conjunto com AM para avaliação da qualidade de sementes são citados a seguir.

Em Mahajan, Mittal e Das (2018), os autores validaram o uso da análise de imagem da semente de soja para avaliar a pureza física, testes de viabilidade e vigor das sementes a partir de descritores de forma, tamanho, cor e textura. Os atributos gerados foram reduzidos usando a técnica de análise dos componentes principais (PCA) e algoritmos de aprendizado supervisionado foram usados para fazer a classificação das sementes em saudáveis e não saudáveis. Já em Pereira, Bugatti e Saito (2016) os autores apresentam uma metodologia para classificar sementes de soja de acordo com o nível, tipo e local do dano usando descritores de cor em diversos algoritmos de classificação a partir de um aprendizado ativo para a seleção das melhores amostras para o processo de aprendizado dos classificadores.

Diante deste contexto, o presente trabalho propõe a utilização de técnicas de aprendizado de máquina para predizer a germinação das sementes de soja a partir de características físicas da semente obtidas por meio da análise digital de imagens. Busca-se com isso maior agilidade no processo de avaliação da qualidade de sementes de soja, além de caracterizar um processo de avaliação da semente não destrutivo.

MATERIAL E MÉTODOS

Para o desenvolvimento deste projeto foram utilizados 3 lotes de sementes de soja adquiridos na região de Macaíba/RN. A partir desses lotes foi feita a aquisição de 300 imagens de sementes de soja no laboratório de sementes da Escola Agrícola de Jundiaí (EAJ). A aquisição foi feita com uma câmera digital com resolução 4320 x 2432 pixels em formato JPEG. Após a aquisição das imagens, foi feito o pré-processamento, segmentação e descrição das imagens, gerando os 50 atributos apresentados na Tabela1, que dizem respeito a características físicas da semente.

É importante destacar que para o processo de aquisição das imagens, as sementes foram identificadas recebendo uma numeração, pois posteriormente foram colocadas para germinar em um papel próprio para a germinação chamado de germitex e levados para o germinador do tipo BOD (*Biochemical Oxygen Demand*) a uma temperatura de 28°C e após 7 dias verificado

quais germinaram. Essa informação foi colocada na base de dados construída, sendo o atributo alvo para os algoritmos de aprendizado de máquina utilizados.

Após a construção da base de dados, algoritmos de aprendizado de máquina foram treinados para realizar a predição do teste de germinação. Algoritmos genéticos foram aplicados à base de dados a fim de selecionar um subconjunto de atributos mais eficiente para os algoritmos de aprendizado de máquina adotados.

Tabela 1 – Atributos da base de dados de soja.

DESCRITORES DE FORMA	DESCRITORES DE COR
Alongamento	Cor_Vermelha
Area	Cor_Verde
Area_Convexa	Cor_Azul
Area_Preenchida	Cor_Media
Bordas	Cor_Dominante
Centroide_Eixo_X	DESCRITORES DE TEXTURA
Centroide_Eixo_Y	Brilho
Centroide_Ponderado_Eixo_X	Energia
Centroide_Ponderado_Eixo_Y	Entropia
Circunferencia	Homogeniedade
Comprimento	Correlacao
Comprimento_De_Eixo_Maior	Contraste
Comprimento_De_Eixo_Menor	Intensidade_Maxima
Diametro	Intensidade_Minima
Diametro_Equivalente	Intensidade_Media
Esfericidade	DESCRITORES MATEMÁTICOS
Excentricidade	Convexidade
Extensao	Desvio_Padrao
Forma	Media_De_Pixel_Do_Elemento
Largura	Numero_De_Euller
Manchas	Orientacao
Perimetro	Diametro_Maximo_De_Feret
Quantidade_De_Manchas	Diametro_Mínimo_De_Feret
Raio	Diametro_Médio_De_Feret
Redondeza	Relacao_De_Feret
Solidez	
Volume	

Construção da base de dados

Em processamento digital de imagens é possível extrair características de uma imagem, onde cada característica é chamada de descritor. Os descritores utilizados neste trabalho foram separados em quatro classes, sendo elas: forma, textura, cor e matemáticos, conforme pode ser visto na Tabela 1. Para esta etapa do trabalho foi utilizada a ferramenta MATLAB[1]. Os 50 descritores aplicados às imagens de soja e mais a informação sobre a germinação (sim ou não) das sementes constituem a base de dados de soja construída para este trabalho.

Aprendizado de máquina

Para realizar a predição da germinação da semente de soja foi utilizada a ferramenta WEKA (*Waikato Environment for Knowledge Analysis*) (HALL et al., 2009) que implementa uma coleção de algoritmos de aprendizado de máquina e é bastante difundida e aceita entre os pesquisadores da área. Os algoritmos do pacote utilizados foram os seguintes:

- NaiveBayes;
- IBk, implementa o algoritmo k-vizinhos mais próximos (KNN);
- SMO, implementa o algoritmo de máquina de vetores de suporte (SVM);
- MultilayerPerceptron, implementa uma rede neural Perceptron de múltiplas camadas (MLP).

Os algoritmos foram utilizados com os seguintes parâmetros:

- NaiveBayes com estimador de kernel para atributos numéricos em vez de uma distribuição normal;
- KNN com k=1;
- SVM com C = 10 e Kernel = puk;
- MLP com taxa de aprendizado = 0.3, momento = 0.2 e 500 iterações. Foi utilizada 1 camada escondida, usando o parâmetro a (quantidade de atributos + quantidade de classes) para definição da quantidade de neurônios.

A amostragem dos dados foi feita separando-se 33% dos dados para teste e usando o restante para treinar os algoritmos. Além disso, a base de dados foi normalizada para evitar impactar negativamente o desempenho dos algoritmos.

1 MATLAB and Statistics Toolbox Release 2012b, The MathWorks, Inc., Natick, Massachusetts, United States

Seleção de atributos

Para melhorar o desempenho dos algoritmos de aprendizado de máquina para a base de dados de soja foi aplicada a técnica de seleção de atributos embutida usando algoritmos genéticos. O algoritmo implementado está apresentando na Figura 1.

Figura 1 – Fluxo do algoritmo genético implementado

Algoritmo Fluxo básico de um algoritmo genético simples

Seja $S(t)$ uma população de cromossomos na geração t;
$t \leftarrow 0$;
Inicializar $S(t)$;
Avaliar $S(t)$;
while o critério de parada não for satisfeito **do**
  $t \leftarrow t + 1$;
  Selecionar $S(t)$ a partir de $S(t-1)$;
  Aplicar operadores genéticos sobre $S(t)$;
  Avaliar $S(t)$;
end while
Retornar melhor solução encontrada.

Fonte: O próprio autor.

Foram adotados os seguintes parâmetros neste experimento:

- População inicial com 30 indivíduos;
- Seleção feita com torneio binário e elitismo;
- Os operadores genéticos de mutação e cruzamento com taxas de 0,5 e 0,9, respectivamente;
- Cruzamento foi realizado com apenas um ponto de porte.

RESULTADOS E DISCUSSÃO

Os resultados dos experimentos feitos para predição do teste de germinação com a base de dados de sementes construída estão apresentados na tabela abaixo, usando como métrica a acurácia dos classificadores (percentual de instâncias classificadas corretamente):

Tabela 2 – Acurácia dos classificadores construídos

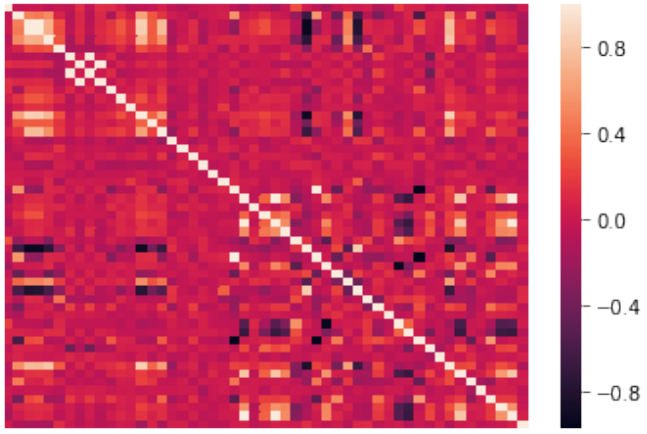
Classificadores	Naive	KNN	SVM	MLP
Base completa com 50 atributos				
Acurácia (%)	65,68	61,76	59,00	56,86
Base reduzida usando algoritmos genéticos				
Acurácia (%)	73,52	79,41	75,49	75,49
Qtd de atributos	11	22	13	8

Como pode ser observado na Tabela 2, a acurácia dos classificadores usando a base de dados de soja com os 50 atributos construídos pela etapa de processamento digital de imagens

está entre 56,86% e 65,68%, sendo o classificador Naive o que classificou o maior percentual de instâncias corretamente. Quando a base de dados é reduzida pelo algoritmo genético observa-se um aumento significativo na acurácia dos classificadores, sendo o KNN e a MLP os classificadores com maior aumento na acurácia, chegando a uma diferença de 17,6 pontos percentuais. O Naive teve um aumento menor, de 7,8 pontos percentuais, e ficou com o menor desempenho nessa comparação.

Observa-se que os atributos inicialmente definidos para a descrição da base de dados de sementes de soja devem apresentar atributos redundantes e irrelevantes que impactam no desempenho dos classificadores, uma vez que a seleção trouxe uma melhoria significativa. Isso pode ser bem observado pelo desempenho do KNN que é um algoritmo muito influenciado pela presença de atributos redundantes e irrelevantes e foi o que obteve maior acréscimo na acurácia após a seleção dos atributos. Para investigar a presença de atributos redundantes, a correlação de Pearson foi calculada entre os 50 atributos da base completa. A matriz com as correlações está demonstrada na Figura 2 a partir do mapa de calor, onde os valores mais claros e mais escuros indicam a presença de correlações positivas e negativas entre os atributos, respectivamente. O mapa indica a presença de atributos correlacionados que geram informação redundante e também influencia negativamente o desempenho do Naive, uma vez que este classificador assume que os atributos não são correlacionados entre si.

Figura 2: Mapa de calor das correlações entre os atributos da base de dados de sementes de soja



Fonte: Próprio autor.

As Tabelas 3 e 4 apresentam a matriz de confusão para cada classificador gerado, onde as colunas apresentam como o estimador classificou as instâncias e as linhas mostram qual a real classe da instância.

Tabela 3 – Matriz de confusão dos classificadores com a base completa

	Não Germinou	Germinou
	Naive	
Não Germinou	9	31
Germinou	4	58
	KNN	
Não Germinou	24	16
Germinou	23	39
	SVM	
Não Germinou	31	9
Germinou	34	28
	MLP	
Não Germinou	25	15
Germinou	29	33

Tabela 4 – Matriz de confusão dos classificadores com a base reduzida

	Não Germinou	Germinou
	Naive	
Não Germinou	27	13
Germinou	14	48
	KNN	
Não Germinou	28	12
Germinou	9	53
	SVM	
Não Germinou	26	14
Germinou	11	51
	MLP	
Não Germinou	32	8
Germinou	17	45

A matriz de confusão serve para verificar o desempenho do classificador em relação a cada classe, distinguindo os erros do classificador, ou seja, verificando em que classe o classificador mais erra. Isso pode ser inferido pelo percentual de falsos positivos e negativos, ou seja, o percentual de exemplos que eram de uma classe, mas foram incorretamente classificados como sendo de outra. Dessa análise pode-se extrair, por exemplo, se há maior necessidade de exemplos de uma classe específica.

Os resultados apresentados na Tabela 3 mostram uma grande quantidade de falsos positivos e negativos, especialmente no Naive que apresenta baixo desempenho para classificar as instâncias da classe 'Não Germinou'. Já os demais classificadores tiveram baixo desempenho para classificar as instâncias da classe 'Germinou'. O que indica uma necessidade de aumentar exemplos de sementes que germinam.

A Tabela 4 mostra que os classificadores conseguiram separar melhor as instâncias com a base reduzida, uma vez que a quantidade de falsos diminuiu notadamente, indicando que com a seleção dos atributos, os classificadores aprenderam melhor como separar as classes. A matriz de confusão dos classificadores com a base reduzida mostra que a precisão foi maior para a classe 'Germinou'.

Estes resultados são indício de que existe a viabilidade da automatização no processo de teste de germinação por aprendizado de máquina, uma vez que a base possua quantidade suficiente de exemplos para as duas classes e atributos representativos para a tarefa.

CONCLUSÕES

O presente trabalho teve por objetivo analisar a viabilidade de predição da germinação de sementes de soja a partir de técnicas de aprendizado de máquina. Para isto, foi construída uma base de dados com informações sobre características físicas da semente, extraídas via processamento digital de imagens, e a supervisão de um atributo alvo indicando as sementes que germinaram ou não.

A base de dados foi utilizada para realizar o treinamento de 4 algoritmos de aprendizado de máquina, Naive, KNN, SVM e MLP. Após estes experimentos iniciais, a base de dados foi reduzida utilizando um algoritmo genético para seleção de atributos. Os resultados mostraram que a seleção de um subconjunto de atributos relevante melhora significantemente o desempenho dos classificadores indicando que o melhoramento da base pode viabilizar a predição sobre a germinação das sementes de soja. Isto pode ser feito pela adição de novos exemplos, bem como, pela definição de novos descritores mais relevantes ao processo preditivo.

REFERÊNCIAS

BRASIL. Ministério da Agricultura, Pecuária e Abastecimento. **Regras para análise de sementes**. Secretaria de Defesa Agropecuária. Brasília: Mapa/ACS, 2009.

BROSNAM, T.; SUN, D**.** Inspection and grading of agricultural and food products by computer vision systems a review. **Computers and Electronics in Agriculture**, Dublin, v.36, n.2, p.193-213, 2002.

COSTA NETO, P. R.; ROSSI, L. F. S. Produção de biocombustível alternativo ao óleo diesel através da transesterificação de óleo de soja usado em fritura. **Química Nova**, v.23, p. 4, 2000.

EMBRAPA. **A Soja no Sistema de Cultivo.** Disponível em < https://www.embrapa.br/busca-de-noticias/-/noticia/8901995/artigo---a-soja-no-sistema-de-cultivo>. Acesso em: 1 jan 2016.

FACELI, K. et al. Inteligência Artificial: Uma abordagem de Aprendizado de Máquina. Rio de Janeiro: LTC, 2017.

FRANÇA-NETO, J.B. et al. Padrões de qualidade de sementes no Brasil: A importância do uso de sementes de soja de alta qualidade e os seus efeitos sobre a produtividade. **Informativo ABRATES**, v. 25, n. 2, p. 27, 2015.

GUEDES, M. A. et al. Caracterização Física de Grãos de Soja Utilizando-se Processamento Digital de Imagens. **Revista Brasileira de Produtos Agroindustriais**. Campina Grande, v.13, n.3, p.279-294, 2011.

HALL, M. et al. The weka data mining software: An update. **SIGKDD Explorations**, v. 11, n. 1, p. 10–18, 2009.

KRZYZANOWSKI, F.C.; VIEIRA, R.D.; FRANÇA NETO, J.B. **Vigor de sementes: conceitos e testes.** Londrina: ABRATES. p.3.2-3.4. 1999.

JOVIC, A.; BRKIC, K.; BOGUNOVIC, N. A review of feature selection methods with applications. In: International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), 38., 2015, Opatija. **IEEE**. Disponível em: < https://ieeexplore.ieee.org/document/7160458>.

MAHAJAN, S.; MITTAL, S. K.; DAS, A. Machine vision based alternative testing approach for physical purity, viability and vigour testing of soybean seeds (Glycine max). **Journal of food science and technology**, v. 55, n. 10, p. 3949-3959, 2018.

MASSETO, T. E.; FARIA, J. M. R.; QUEIROZ, S. E. E. Evaluation of the seed quality in Cedrela fissilis - (Meliaceae) by X-ray test. **Ciência e Agrotecnologia**, Lavras, v. 32, n. 6, p. 1708-1712, 2008.

PEREIRA, D. F.; BUGATTI, P. H.; SAITO, P. T. M. Aprendizado Ativo para Classificação do Vigor de Sementes de Soja. In: Conference on Graphics, Patterns and Images, 2016, São José dos Campos. **Electronic Proceedings of the 29th Conference on Graphics, Patterns and Images (SIBGRAPI'16)**.

Uma API para a classificação de solos do Brasil

Glauber José Vaz[1], Luís de França da Silva Neto[2], Renato Nunes Lima[3], Stanley Robson de Medeiros Oliveira[1]

[1] Embrapa Informática Agropecuária, Campinas, SP, Brasil,

glauber.vaz@embrapa.br, stanley.oliveira@embrapa.br

[2] Embrapa Solos, Recife, PE, Brasil, luis.franca@embrapa.br

[3] Unisal, Campinas, SP, Brasil, renato.nlma@gmail.com

RESUMO

O objetivo deste trabalho é disponibilizar por meio de uma API um sistema para a classificação de solos do Brasil, a fim de auxiliar no avanço da qualidade dos dados de solos e apoiar iniciativas de melhoria na gestão deste recurso. A API de classificação de solos facilita a curadoria de dados de solos e pode orientar um registro mais adequado destes dados. Com auxílio desta ferramenta, uma análise detalhada sobre a classificação de solos foi realizada em uma das mais importantes bases de dados de solos no país. Os resultados mostram que há uma necessidade muito grande de curadoria desses dados por parte dos especialistas do domínio e que a ferramenta considerada é poderosa aliada na melhoria da qualidade de dados de solos do Brasil.

PALAVRAS-CHAVE: SiBCS, Sistema especialista, Dados de solos, Curadoria de dados.

ABSTRACT

The objective of this work is to provide an API for the classification of soils of Brazil, in order to help improve the quality of soil data and support initiatives to improve the management of this resource. The soil classification API supports soil data curation and can guide a more appropriate record of these data. With this software, a detailed analysis on soil classification was performed in one of the most important soil databases in Brazil. The results show that there is a great need for the curation of this data by the domain specialists and that the tool considered is powerful in helping to improve the quality of soil data in Brazil.

KEYWORDS: SiBCS, Expert system, Soil data, Data curation.

INTRODUÇÃO

O Sistema Brasileiro de Classificação de Solos (SiBCS) é o sistema taxonômico oficial de classificação de solos do Brasil. Hierárquico, multicategórico e aberto, o sistema possibilita a classificação de todos os solos existentes no território nacional (DOS SANTOS et al., 2018). Em sua quinta edição, o SiBCS conta com 13 classes no primeiro nível categórico. Embora o sistema envolva seis níveis categóricos, o quinto e o sexto níveis não são obrigatórios.

Vaz et al. (2018) construíram um sistema especialista baseado nas regras do SiBCS para auxiliar profissionais que necessitem fazer a classificação de solos brasileiros. O sistema simula o raciocínio de um profissional especialista do domínio ao realizar a classificação de perfis de solo. Assim, pode ser utilizado para classificar perfis segundo o SiBCS em diferentes níveis categóricos, de acordo com os dados fornecidos, e também para validar perfis previamente classificados. Embora o SiBCS tenha sido publicado há muitos anos, não há outro programa de computador que faça essa classificação automática de solos brasileiros.

Para se obter maior alcance deste sistema especialista, ele foi disponibilizado via API (*Application Programming Interface*), uma forma de comunicação entre aplicações computacionais através de uma rede, predominantemente a Internet, e com o uso de uma linguagem comum (JACOBSON et al., 2011).

Por meio da AgroAPI, a Embrapa (2019) já vem disponibilizando algumas APIs para uso interno, negócios com organizações parceiras ou acesso aberto. Assim, promove a criação de valor na Agricultura Digital por meio de APIs, que possibilitam o compartilhamento de dados e serviços, facilitam a integração de sistemas de informação e o estabelecimento de acordos entre organizações e, ainda, viabilizam maior alcance dos resultados obtidos pela empresa e seus parceiros (VAZ et al., 2017). Embora com acesso restrito em um primeiro momento, a API criada neste trabalho também está disponível nesta plataforma de APIs da Embrapa.

Esta API pode auxiliar inciativas que visem à melhoria da gestão do solo tanto nacionais quanto internacionais. O Programa Nacional de Solos do Brasil (Pronassolos), por exemplo, tem por principais objetivos a retomada da realização dos levantamentos pedológicos e o estabelecimento de uma base de dados integrada na qual as informações de solos estejam organizadas e sistematizadas para a consulta do público em geral (POLIDORO et. al, 2016). Esta API pode auxiliar no Pronasolos facilitando a curadoria dos dados e indicando como os dados podem ser coletados a fim de facilitar seu processamento computacional. No contexto internacional, a Global Soil Partnership (GSP) foi estabelecida

para melhorar a governança e promover a gestão sustentável de solos (FOOD AND AGRICULTURAL ORGANIZATION OF THE UNITED NATIONS, 2019). Um de seus pilares de ação é avançar na qualidade e na quantidade de informações e dados de solos, mas, para isso, conta com a colaboração e o compartilhamento de informações locais. Assim, a API de classificação de solos também pode contribuir com esta e outras iniciativas globais.

A fim de se conduzir um teste preliminar no sistema, seus resultados de classificação foram comparados às classificações fornecidas pelo Banco de Dados de Informações Ambientais (BDiA). Este banco de dados é disponibilizado pelo Instituto Brasileiro de Geografia e Estatística (IBGE) e reúne bases temáticas de recursos naturais do território brasileiro, dentre as quais a de Pedologia (IBGE, 2019).

O objetivo deste trabalho, portanto, é disponibilizar por meio de API um sistema especialista para a classificação de solos brasileiros desenvolvido com base no SiBCS, a fim de apoiar iniciativas de melhoria na gestão de solo, auxiliando, principalmente, no avanço da qualidade dos dados. Além disso, uma análise de consistência foi realizada com o auxílio desta ferramenta sobre os dados de uma das principais bases de solos do Brasil.

MATERIAL E MÉTODOS

O sistema especialista, desenvolvido na linguagem SWI Prolog, pode ser acessado via uma API disponível na AgroAPI, plataforma implementada com tecnologia para gerenciamento de APIs. Segundo Heffner (2018), este tipo de solução atende a, principalmente, três necessidades: (i) a otimização de valor por parte dos provedores de APIs, (ii) a gestão do relacionamento entre os provedores e os consumidores de APIs, e (iii) o estabelecimento de acordos sobre o uso e a segurança das APIs.

A comunicação com a API de classificação de solos é realizada com dados em JSON (*JavaScript Object Notation*), um formato leve para troca de dados que é fácil de ler e escrever para humanos e de interpretar e gerar por computadores. Neste formato, os dados são representados em texto e organizados em listas de valores e em coleções de pares formados por nome e valor (JSON.ORG, 2019).

Embora um arquivo JSON possa ser lido por um ser humano, sua edição direta pode não ofecerer uma experiência boa para o usuário e ainda aumenta a probabilidade de erros nos dados. Assim, foi construída uma interface de usuário simples para a geração padronizada dos dados de entrada para a API. Por meio desta interface, é possível preencher os campos correspondentes aos dados de um perfil de solo e salvá-los em um arquivo JSON. Também é

possível carregar dados de um arquivo neste formato nos campos da interface para edição dos valores. Além de possibilitar a chamada ao sistema especialista de classificação de solos com os dados do formulário como entrada, esta interface faz alguns testes de consistência de valores para os campos. A ideia, no entanto, é que os clientes da API desenvolvam suas próprias interfaces para atendimento a necessidades específicas.

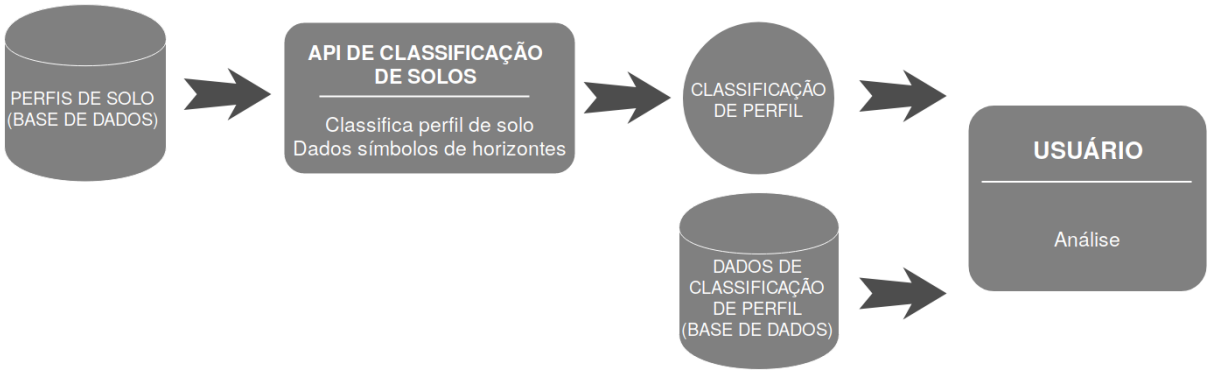
Os dados do BDiA foram utilizados para validar o sistema especialista. Eles são públicos, facilmente obtidos na Internet e contam com classificações para os perfis do primeiro ao quarto nível categórico do SiBCS. No entanto, devido à grande quantidade de perfis, uma análise mais detalhada foi realizada considerando-se apenas o primeiro nível. A fim de servir de entrada para o sistema especialista, os dados do BDiA foram convertidos para o formato JSON com várias alterações em seus campos e em seus conteúdos, de maneira a simplificar o processamento computacional e a padronizar os valores desses campos.

RESULTADOS E DISCUSSÃO

O sistema especialista disponibilizado via API assume que os valores dos atributos de solos são corretamente fornecidos, assim como as representações dos símbolos dos horizontes dos perfis de solo. Mesmo nos casos em que seria possível classificar o perfil independentemente dos símbolos de horizonte, estes são verificados e precisam estar corretamente especificados. Além disso, o sistema realiza a classificação de acordo com a quinta e última versão do SiBCS, podendo haver incompatibilidade com perfis que foram classificados com regras de edições anteriores.

A Figura 1 representa o processo de análise conduzido neste trabalho. Os resultados gerados pelo sistema disponibilizado via a API são comparados aos dados registrados. A Tabela 1 exibe, para cada classe, a proporção de perfis cujas classificações obtidas via a API coincidem com as previamente anotadas no banco de dados utilizado.

Figura 1 – Análise de classificação de perfis de solo



Fonte: Adaptado de Vaz et al. (2018)

Tabela 1 – Resultados do sistema especialista com os dados do BDiA

CLASSE	TOTAL DE PERFIS	COINCIDE	NÃO COINCIDE	% COINCIDE
ORGANOSSOLOS	13	8	5	61.54 %
ESPODOSSOLO	65	45	20	69.23 %
PLANOSSOLOS	200	114	86	57 %
PLINTOSSOLO	247	226	21	91.5 %
VERTISSOLO	33	0	33	0 %
CHERNOSSOLO	86	23	63	26.74 %
GLEISSOLO	204	167	37	81.86 %
LUVISSOLO	76	51	25	67.11 %
NITOSSOLO	128	29	99	22.66 %
ARGISSOLO	1428	785	643	54.97 %
LATOSSOLO	1157	704	453	60.85 %
CAMBISSOLO	365	243	122	66.58 %
NEOSSOLO	465	437	28	93.98 %
TODOS	4467	2832	1635	63.4 %

Nota-se que em grande parte dos casos, há diferença na classificação do solo. Esse índice ainda é superior a uma análise prévia realizada, devido ao relaxamento de algumas restrições. Na verificação de planossolos, nitossolos e chernossolos, por exemplo, atributos de estrutura e de cor não foram verificados devido à ausência desses atributos em grande quantidade dos perfis. No teste de luvissolos, as regras referentes a planossolos foram desativadas, pois estrutura e cor são elementos importantes para diferenciar estas duas classes e, por isso, muitos dos perfis de luvissolos eram classificados como planossolos. Essa

supressão de algumas regras foi feita para verificar se as outras condições seriam atendidas nos casos em que os atributos mencionados fossem adequadamente anotados e satisfizessem às restrições.

Em outros casos, as regras foram mantidas mesmo quando tiveram grande impacto nos resultados. Por exemplo, o carbonato de cálcio equivalente é muito importante na verificação de chernossolos, mas a grande maioria dos perfis não trazem valores para este atributo. Ainda assim, as condições envolvendo-o foram mantidas. Se todos os perfis que dependem deste valor para ser classificado como chernossolo atendessem às restrições de carbonato de cálcio, as discrepâncias seriam reduzidas de 63 para 39. E se a verificação de planossolo também fosse desativada nestes mesmos casos, esse valor seria ainda menor, 29. Portanto, a anotação incorreta ou ausência de valores para um único atributo podem gerar grande impacto na classificação de solos.

A Tabela 1 também mostra a ordem com que as classes são testadas pelo sistema e, portanto, sua precedência. Assim, se determinado perfil atende às restrições de diferentes classes, o resultado corresponde à primeira delas. Um perfil sempre é classificado entre as treze classes consideradas, de maneira que, caso não atenda às condições de nenhuma das doze primeiras, automaticamente, é tratado como neossolo. Por exemplo, no conjunto de dados utilizado, há inúmeros perfis que não apresentam os dados de horizontes. Todos eles são classificados como neossolos.

As diferenças entre as classificações do banco de dados e do sistema especialista poderiam ser ocasionadas tanto pela implementação equivocada das regras do SiBCS no programa de computador quanto por inconsistências nos dados, como, por exemplo, valores incompatíveis para atributos específicos, simbologia de horizontes equivocados, ausência de valores para determinados atributos, não atendimento a restrições de determinada classe de solo e incompatibilidade de versões do SiBCS, entre outras.

Então, análises mais detalhadas foram realizadas para checar as causas dessas discrepâncias nas classificações. Nos casos em que foram identificados erros nas regras do software, estes foram corrigidos e os resultados esperados passaram a ser gerados. No entanto, em grande quantidade de perfis, o programa continuou a exibir classificações distintas das previamente fornecidas. A Tabela 2 apresenta os motivos porque cada perfil previamente anotado em determinada classe não foi classificado conforme a anotação.

Tabela 2 – Motivos pelos quais há diferença entre a classe anotada e a classe obtida pelo sistema especialista

CLASSE	MOTIVOS DE DIFERENÇA NA CLASSIFICAÇÃO
ORGANOSSOLOS	1 sem horizontes 1 sem horizonte H ou O 3 com quantidade de carbono orgânico insuficiente
ESPODOSSOLO	1 sem horizontes 19 sem horizonte Bh ou Bs
PLANOSSOLOS	14 sem horizontes 18 sem transição abrupta ou mudança textural abrupta 6 sem relação textural B/A de horizonte B plânico 22 sem horizonte textural Bt 14 sem horizonte A ou E antes de Bt 2 horizonte A ou E não está imediatamente antes de Bt 10 em classe precedente (plintossolos em casos com caráter plânico sem carater sódico)
PLINTOSSOLO	5 sem horizontes 10 sem horizonte plíntico, litoplíntico ou concrecionário 6 horizonte anterior ao horizonte plíntico, litoplíntico ou concrecionário não atende
VERTISSOLO	33 sem horizonte com sufixo ‘v’
CHERNOSSOLO	5 sem horizontes 13 em classes precedentes (12 planossolo, 1 vertissolo) 13 não tem profundidade de A suficiente 3 não tem horizontes depois de A chernozêmico 6 não satisfazem restrições de saturação por bases 23 não tem Bi ou Bt, ou não atende a condições de rêndzico (sem dados de CaCO3)
GLEISSOLO	5 sem horizontes 10 em classe precedente (9 plintossolo, 1 chernossolo) 14 sem horizonte com sufixo ‘g’ 5 horizonte anterior ao glei não atende 3 sem profundidade suficiente de horizonte glei
LUVISSOLO	9 em classe precedente (plintossolos) 7 não atende a condições de saturação por bases alta ou argila de atividade alta 8 sem horizontes Bt 1 sem profundidade suficiente de Bt
NITOSSOLO	2 sem horizontes 10 em classes precedentes (6 planossolo, 4 luvissolo) 50 sem horizontes Bt 9 sem profundidade suficiente de Bt 1 sem textura argilosa (sem dados de argila) 23 sem registro de cerosidade 4 com cerosidade que não representa nitossolo
ARGISSOLO	70 sem horizontes 436 em classes precedentes (255 plano, 127 plinto, 44 nito, 6 cherno, 2 luvi, 2 gleissolo) 135 não tem Bt imediatamente abaixo de A ou E 2 não apresenta argila baixa ou saturação por bases baixa ou caráter alumínico
LATOSSOLO	87 sem horizontes 44 em classes precedentes (1 plano, 25 plinto, 16 argi, 2 nito) 322 não tem Bw abaixo de A
CAMBISSOLO	37 sem horizontes 28 em classe precedente (1 gleissolo, 21 plintossolo, 6 argissolo) 57 sem horizonte Bi imediatamente abaixo de A, E, O ou H
NEOSSOLO	28 em classes precedentes (2 plano, 11 plinto, 5 glei, 2 nito, 2 argi, 5 cambi, 1 latossolo)

Para as classes listadas de organossolos a nitossolos, a análise foi bastante minuciosa, de maneira que os motivos das discordâncias nas classificações foram até mais detalhadas. A partir de argissolos até neossolos, devido ao maior número de perfis, os motivos levantados são um pouco mais genéricos e contaram com um maior grau de automação. Porém, em todos os casos, houve um trabalho humano na análise desses dados. Portanto, embora sujeita a erros, essa análise fornece um retrato geral da consistência dos dados utilizados e de suas classificações.

Observa-se uma grande quantidade de perfis que não possuem dados relacionados a seus horizontes, o que caracteriza falha no registro dessas informações. Nos perfis classificados como neossolos, ocorrências desse tipo também foram verificadas, mas são considerados neossolos porque não atendem às restrições de nenhuma das outras classes. Diferenças nas classificações devido à simbologia de horizontes também foram bastante comuns neste conjunto de dados. Em trabalhos futuros, isso pode ser amenizado porque os símbolos também podem ser inferidos a partir dos dados coletados no campo. Também foram comuns situações em que os perfis assumiram classes testadas previamente pelo programa, quando também atenderiam às restrições da classe anotada. A desativação de algumas regras importantes na distinção das classes explica a maior parte desse efeito, porém, esta medida somente foi tomada porque valores de atributos essenciais não foram registrados para uma grande quantidade de perfis, com destaque para os de cor e de estrutura. A Tabela 2 ainda exibe uma série de informações com riqueza de detalhes e que possibilitam a compreensão das inconsistências dos dados de solos registrados.

Este trabalho de análise dos dados contou com auxílio de pedólogos, mas foi realizado predominantemente por especialistas de computação e análise de dados, leigos no domínio de solos. Portanto, tanto o software quanto esta análise demandam uma validação mais abrangente por parte dos especialistas em classificação de solos.

CONCLUSÕES

Os resultados mostrados neste trabalho fornecem um quadro geral da qualidade dos dados considerados. A base utilizada é uma das mais importantes, mais utilizadas e de maior qualidade para se obter dados de perfis de solos do Brasil e suas classificações. No entanto, ainda há uma necessidade muito grande de curadoria desses dados por parte dos especialistas desse domínio. Também, é importante que novos levantamentos de solos sejam realizados de maneira que os dados obtidos sejam registrados de uma maneira mais padronizada e mais

adequada para o processamento computacional, reduzindo-se assim esses erros tão presentes nas bases de dados disponíveis atualmente.

A API de classificação de solos considerada neste trabalho pode auxiliar na curadoria dos dados existentes e na verificação da consistência de novos dados coletados. O padrão de entrada necessário para o uso correto desta API também pode nortear o registro dos dados de forma padronizada e de maneira a abranger todos os dados necessários para a correta classificação do solo. Portanto, embora ainda demande uma maior validação por parte dos especialistas em classificação de solos, trata-se de uma poderosa ferramenta para a melhoria da qualidade de dados de solos no Brasil.

AGRADECIMENTOS

Esta pesquisa foi apoiada pela EMBRAPA. Projeto SmartSolos [SEG 23.14.01.017.00.00].

REFERÊNCIAS

DOS SANTOS, H. G. et al. Sistema brasileiro de classificação de solos. Brasília, DF: Embrapa, 2018.

EMBRAPA. AgroAPI: Plataforma de APIs da Embrapa. Disponível em: <https://www.agroapi.cnptia.embrapa.br>. Acesso em: 27 mai. 2019.

FOOD AND AGRICULTURAL ORGANIZATION OF THE UNITED NATIONS. Global Soil Partnership. Disponível em: <www.fao.org/global-soil-partnership>. Acesso em: 18 jun. 2019.

HEFFNER, R. (2018). The Forrester Wave TM: API Management Solutions, Q4 2018. Forrester Research, Inc, 2018.

IBGE. BDiA: Banco de dados de informações ambientais. Disponível em: <https://bdiaweb.ibge.gov.br>. Acesso em 31 mai. 2019.

JACOBSON, D.; WOODS, D.; BRAIL, G. APIs: A strategy guide. O'Reilly Media, Inc., 2011.

JSON.ORG. Introducing JSON. Disponível em: <http://json.org/>. Acesso em 28 jun. 2019.

POLIDORO, J. C. et al. Programa Nacional de Solos do Brasil (PronaSolos). Rio de Janeiro: Embrapa Solos, 2016.

VAZ, G. J.; APOLINÁRIO, D. R. de F.; CORREA, J. L.; VACARI, I.; GONZALES, L. E.; DRUCKER, D. P.; BARIANI, J. M.; EVANGELISTA, S. R. M.; ROMANI, L. A. S. AgroAPI: criação de valor para a Agricultura Digital por meio de APIs. In: CONGRESSO BRASILEIRO DE AGROINFORMÁTICA, 11., 2017, Campinas. Ciência de dados na era da agricultura digital: anais. Campinas: Editora da Unicamp: Embrapa Informática Agropecuária, 2017. p. 59-68. SBIAgro 2017.

VAZ, G. J.; SILVA NETO, L. de F. da; OLIVEIRA, S. R. de M.; BOTELHO, F. P.; ARAUJO FILHO, J. C. de. Development of an expert system for classification of Brazilian soil profiles. In: WORLD CONGRESS OF SOIL SCIENCE, 21., 2018, Rio de Janeiro. Soil science: beyond food and fuel: abstracts. Viçosa, MG: SBCS, 2018. Não paginado. WCSS 2018.



OpenVRT - A smartphone based wireless variable rate solution

Luiz Guilherme Fonseca Rosa[1], *Rogério Schneider*[2], *Marcio Seiji Oyamada*[1]

[1]Curso de Ciência da Computação, Universidade Estadual do Oeste do Paraná
Cascavel, Paraná, Brasil
luizguilhermefr@gmail.com, marcio.oyamada@unioeste.br

[2]Agromatic
Cascavel, Paraná, Brasil
agromatic@agromatic.com.br

ABSTRACT

The Precision Agriculture (PA) concept is ceasing to be a competitive differential to become an industrial pattern. Focused on the exploit of temporal and spatial variability, it remains a difficulty to medium and small farmers, mostly due to the high costs of the equipment. This work presents an alternative of low-cost Variable Rate Technology (VRT) based on the use of an Android smartphone and an Arduino microcontroller. The system was developed using a Proportional Integral and Derivative (PID) controller and uses a treadmill miniature to simulate a fertilizer or lime spreader. The results showed that nowadays smartphones are capable in terms of performance and connectivity to ensure a precise and reliable application.
KEYWORDS: Variable Rate Technology, Mobile Applications, Embedded Systems

INTRODUCTION

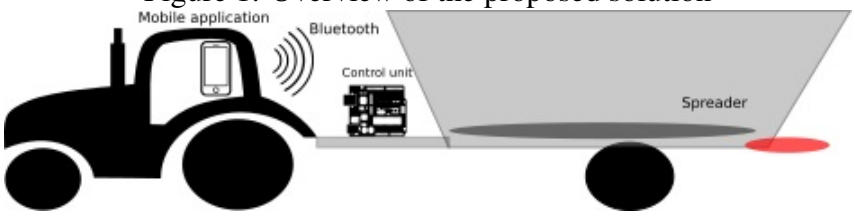
The Precision Agriculture (PA) is one of solutions to deal with the environmental legislation, each day more rigorous, the global worries about the agrochemicals excessive usage and the return of the investment applied (MAPA, 2017). However, it is still hard for medium to small farmers to apply this concept, due to the high costs involved in the equipment and embedded electronics that acquire, process, store and exchange data (ZHANG; WANG; WANG, 2002). Today's machinery comes with Variable Rate Technologies (VRT) directly from the factory. These technologies focus on the controlled variation of agrochemicals, seeds, and fertilizers across the management zone (PEREIRA, 2008), applying strictly the needed and avoiding lack and waste of outputs. To achieve this, it usually counts on features like automatic speed control and Global Navigation Satellite System (GNSS) receivers, interconnected with a computational system (SOUSA; LOPES; INAMASU, 2014). Also, these technologies can be based on maps, previ-

ously acquired by the farmer (e.g. via a crop map or soil analysis), or *on-the-go*, using sensors that determines the rate in real-time, during the application (PEREIRA, 2008; ZHANG; WANG; WANG, 2002). This work presents an alternative for machinery that does not come with VRT capabilities, either because it is a legacy machine or because it was an optional item. The solution is composed of smartphone that stores the application map and is responsible for the geolocation and a control unit connected wirelessly. This system is intended to be generic and versatile, where the flow rate can be of any kind of input, requiring only small adaptations. The main contributions of this work are: the development of a desktop application for visualization and conversion between the map formats Shapefile and GeoJSON; the evaluation of the use of smartphone as an in-field operator console; and a prototype system that can be used for tests and educational purposes.

Materials and methods

Figure 1 presents an overview of the proposed solution. The main components are the mobile application with the prescription map and the control unit that actuates in the equipment to allow the VRT solution. In this section we describe firstly the prescription map format used in the proposed solution. Secondly, the mobile application and the hardware of the control unit are presented. Finally, is presented a treadmill prototype, developed to simulate a fertilizer or lime spreader, and used in the validation tests.

Figure 1: Overview of the proposed solution



Prescription maps

In this work, the mobile application accepts the prescription map only in GeoJSON format. This format is defined publicly in the Request for Comments (RFC) 7946 (INTERNET ENGINEERING TASK FORCE, 2016) and supports geometries and attributes associated with these geometries, in such a way that if we split a productive area into multiple polygons and associate a rate value to each one, we have a prescription map. The more polygons we have, the more precise the prescription will be.

However, one of the most common prescription map format is the Shapefile. It stores non-topological geometric data and the attributes related of each area (ESRI, 1998). The geometric data delimits the area in the field, and the attributes stores the application rate in this area. To support prescription maps described in Shapefile format, we have developed a desktop

application (Figure 2), that reads the prescription map in Shapefile format and shows the areas and application rate found in the file. The desktop application is also responsible to generate the GeoJSON file from the Shapefile, maintaining the support to Shapefiles already available. The GeoJSON format uses a textual format, and it is expected the file size will be bigger than an equivalent in Shapefile format.

Figure 2: Desktop application for visualization and conversion



Mobile Application

For the mobile application, the Android platform was chosen, due to its high-quality documentation and large adoption, especially in Brazil, where it had almost 94 % of market share on December 2018 (KANTAR, 2018). For the actuator device, the Arduino micro-controller was chosen, due to its low cost and flexibility, suitable for fast prototyping. The communication between the two platforms is established via Bluetooth, a low-energy wireless protocol. The purpose of the mobile application is control and monitor the variable rate flow without the need of a new console on the tractor cab, using then a device that is already present in many people's pockets. The mobile application usage flow is as follows:

- Start the application
- Select a map (all files with ".geojson" and ".json" extensions are discovered and listed automatically)
- Pair with the control unit using Bluetooth
- Define the rate measurement (mass unity per space unity) and the machine width
- Start applying

The application communicates with the control unit by sending the current desired rate, based on the current geolocation and prescription map. The Figure 3 shows the application screenshot. Mobile devices tend to have an average geolocation precision of three meters at open environments (SHANER, 2013), which is not as precise as professional PA equipment, but enough for many use cases. The user can also track the current location, the precision of this location and the expected rate real-time on the phone. When the precision is degraded, reaching values higher than 10 meters, an alert is raised, informing the user that the conditions are not ideal.

Figure 3: Mobile application screenshot



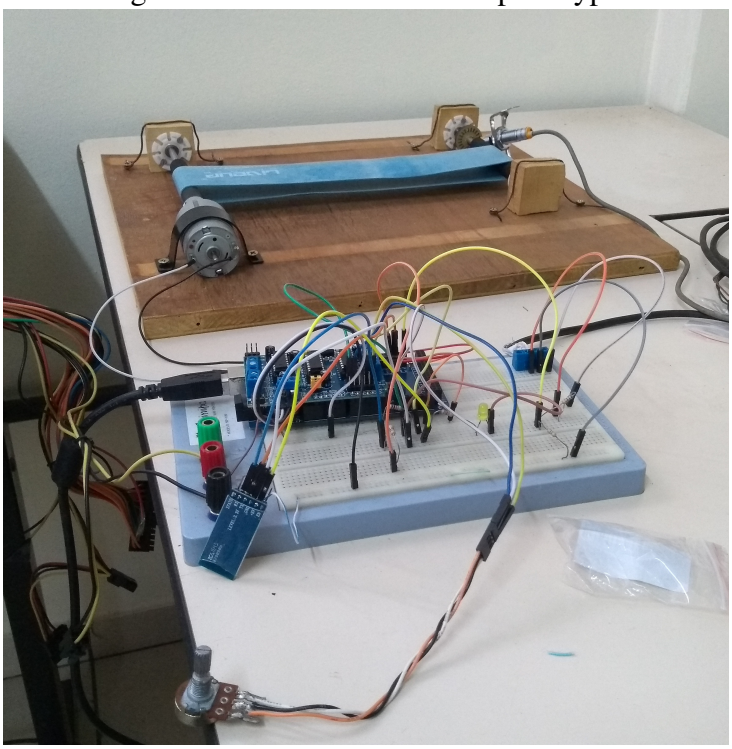
Hardware Platform

The control unit is an embedded device responsible for being the bridge between mobile device and actuators, measuring the actual flow and adjusting the output signals to reach a rate flow closer to the desired one. It was built with an Arduino Mega 2560 and an HC-05 Bluetooth module to communicate with the mobile application. An inductive proximity sensor is used to obtain the actual flow rate. In a practical scenario, it can be positioned for instance, in the spreader mechanism and obtain the speed and consequently the flow rate. The actuation is concepted to be a PWM (Pulse Width Modulation) signal, that varies the speed to achieve the desired flow rate. In a real equipment, the PWM signal will actuate for instance in the hydraulic valve to control the application speed (QUEIROZ, 2010).

For the system validation, a treadmill miniature prototype was used with a DC motor, simulating a fertilizer or lime spreader (Figure 4). The Android application sends a message to the hardware platform every time the desired rate changes, according to the geolocation. The hardware platform then adjusts the treadmill speed to meet the desired rate.

In a real situation a calibration phase must be accomplished. In the simulation, fixed values varying from 0 to 1 kilogram per second (*k/s*), corresponding to 0 to 1,344 rotations per minute, linearly, was considered. To manage the DC motor's speed, Arduino's Proportional Integral Derivative (PID) standard library was used. The actual axis rotation is measured using

Figure 4: Simulator: treadmill prototype



a LJ12A3-4-Z/BY inductive proximity sensor, coupled next to the axis. At each rotation, the sensor is triggered and at each 5 seconds period, the moment's RPM is calculated. This value is converted into *kg/s*, according to the linear correlation explained previously, and the error can be calculated subtracting the actual value by the desired, so the PID controller can adjust the output signal.

Communication protocol

The communication protocol between the mobile application and the control unit is similar to the TCP, used widely in internet services, like HTTP for web and SMTP for e-mail. The protocol uses the strategy "stop-and-wait", which the transmitter sends a message and waits the confirmation message, called "ACK". There are three different scenarios in the protocol:

- No error: the transmitter sends a message and receives the ACK.
- Transmission error: the transmitter sends a message and it is lost. The transmitter will not receive the ACK message, and after the timeout the message will be sent again.
- ACK error: The transmitter sends a message and it is received. The receiver sends the ACK message, but it is lost. The transmitter will not receive the ACK message, and after the timeout the message will be sent again.

The protocol is configured to a limit of retries and timeout. If the retries limit is reached an error message is presented in the mobile application screen and the control unit enters in emergency mode.

RESULTS AND DISCUSSIONS

In order to find the PID tuning, a trial-and-error method was used. The PID controller uses three constants: K_p (proportional), K_i (integral) and K_d (derivative) factors. The constants values leverages the response time and oscillations of the system. Ten different configurations were tested, and the response was compared using the Root Mean Square Error (RMSE). The tests were carried out using a displacement speed of 5 *m/s* and rate from 0 to 1,856.54 kg/ha. The tests with the worst performance (highest RMSE), are those with does not match the expected rate in a fast way and without oscillations. For instance, the Figure 6 presents high RMSE configuration, about 0.0773, with the constants $K_p = 200$, $K_i = 35$ e $K_d = 20$.

The selected configuration was obtained with the values of K_p = 500, K_i = 200 and K_d = 50, obtaining a Root Mean Square Error of 0.0439 and the PID controller response show in the Figure 5.

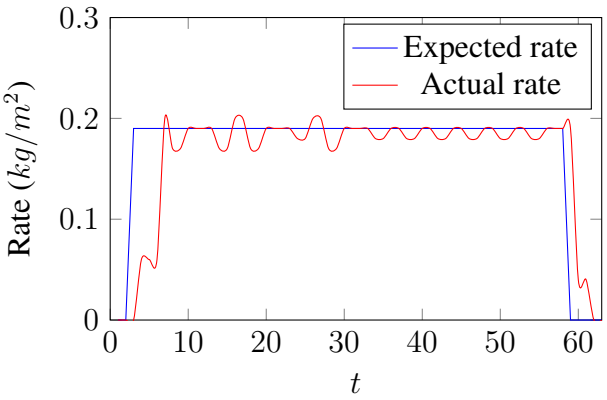


Figure 5: Expected and Actual rate test using the constants $K_p = 500$, $K_i = 200$ e $K_d = 50$ RMSE = 0.0439

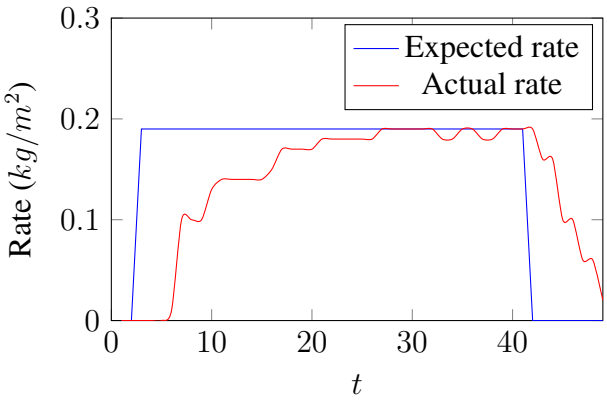


Figure 6: Expected and Actual rate test using the constants $K_p = 200$, $K_i = 35$ e $K_d = 20$ RMSE= 0.0773

The system was validated in the laboratory using the "Mock Locations" application[1],

[1] https://play.google.com/store/apps/details?id=ru.gavrikov.mocklocations

that allowed us to simulate fake routes at the smartphone, without having to actually travel them. This allowed to perform different tests at different rates, enabling the test of different scenarios. To verify the response time of the mobile application and the capacity to correlate the geolocalization in the prescription map, the simulated regions where covered at speed of $\approx$ 15 *km/h* (4.16 *m/s*). The Figure 7 presents the actual and desired rate obtained in the simulation using the miniaturized prototype. As one can see, in the period between $t = 60$ and $t = 70$, where the rate is 3,000 *kg/ha*, with a simulated tractor speed of 4.16 *m/s*, the maximum rate is $\approx$ 2,403 *kg/ha*. In this case, the only possible solution is to decrease the tractor speed to achieve the desired application rate.

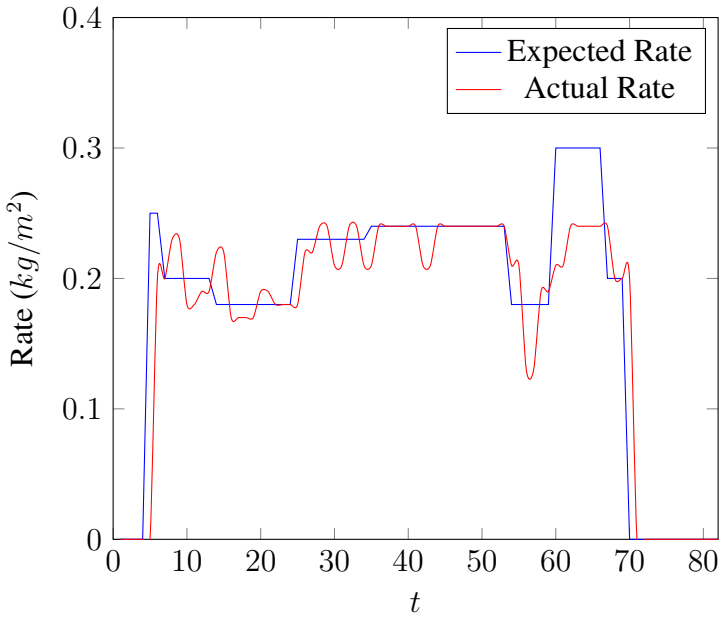


Figure 7: Actual and desired rate, between 0 a 3,000 *kg/ha*

Performance of the mobile application

The tests are carried out in a Motorola Moto G5S XT1792 device, with an Android Android 8.1 (Oreo), 2 GB memory and 8 cores running at 1.40 GHz. The mobile application was executed without freezes, even using high cost resources in terms of energy like GPS and Bluetooth communication. The mobile data network is not mandatory, but when available permits the download of the background map of the area. We have collected the performance numbers of CPU usage and memory (Figure 8), when the displacement over the area was simulated. The CPU usage is around 50.19% and the memory around 128 MBytes. These performance numbers shows that the mobile application can run in mid-range devices without performance problems.

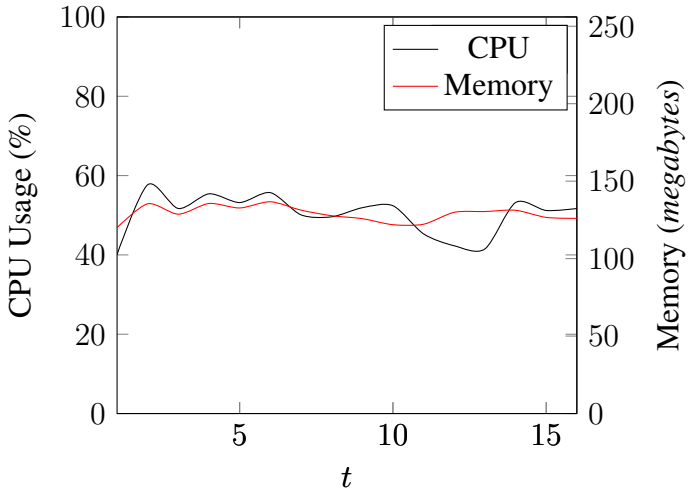


Figure 8: CPU and Memory usage

CONCLUSIONS

This work presented a generic alternative to implement the variable rate technology in legacy machines. There are some available solutions in the market that proposes the same solution[2], but using the traditional operator console. Others solutions are available only in foreign countries, as the GreenStar[3]. This work proposes a desktop application to visualize the prescription maps, compatible with the Shapefile and GeoJSON standards, and a mobile application that works in jointly with a control unit implemented in Arduino. The control unit is developed to be integrated in the equipment using a PWM signal to allow the variable rate. The software components of the proposed work are available at GitHub[4].

The tests were carried out using a treadmill to simulate a fertilizer or lime spreader. This prototype can be used for educational purposes to show the variable rate technology, and also to test the prescription maps, enabling the verification of the rates and the transitions over the field. The embedded firmware was implemented using a PID controller, and was calibrated to allow smooth transitions between the rates and reducing the mean error compared to the expected rate.

As the future work, we are planning the validation of the system in a real machinery, including the calibration process. Another future work is the inclusion of an interpolation method in the mobile application, in order to obtain a smooth rate transition. This can be implemented using some methods such as normal distance or ordinary kriging(COUTO; SCARAMUZZA; MARASCHINI, 2002).

[2]http://www.mfrural.com.br/detalhe/kit-taxa-variavel-agropioneiro-243454.aspx

[3]https://www.deere.com.br/pt/tecnologia-de-produtos/agricultura-de-precisão/soluções-para-operações-no-campo/taxa-variada-para-líquidos-greenstar/

[4]https://github.com/luizguilhermefr

REFERENCES

COUTO, E. G.; SCARAMUZZA, J. F.; MARASCHINI, L. Influência dos métodos de interpolação dos dados nos mapas usados na agricultura de precisão. *Simpósio Internacional de Agricultura de Precisão*, v. 2, p. 1–5, 2002.

ESRI. *Shapefile Technical Description*. [S.l.], 1998. Acesso em 08 jun. 2018. Disponível em: <https://www.esri.com/library/whitepapers/pdfs/shapefile.pdf>.

INTERNET ENGINEERING TASK FORCE. *The GeoJSON Format*. [S.l.], 2016. Acesso em 08 jun. 2018. Disponível em: <https://tools.ietf.org/html/rfc7946>.

KANTAR. *Kantar Worldpanel - Android vs iOS - Smartphone OS sales market share evolution*. 2018. Acesso em 01 jun. 2018. Disponível em: <https://www.kantarworldpanel.com/global/smartphone-os-market-share/>.

MAPA. *Principais Conceitos e Expressões Utilizados na Agricultura de Precisão*. 2017. Acesso em 29 mar. 2018. Disponível em: <http://www.agricultura.gov.br/assuntos/sustentabilidade/tecnologia-agropecuaria/agricultura-de-precisao-1/arquivos-de-agricultura-de-precisao/conceitos-relacionados-a-agricultura-de-precisao-1.pdf/view>.

PEREIRA, R. R. D. *Protocolo ISO 11783: procedimentos para comunicação serial de dados do controlador de tarefa*. Tese (Doutorado) — Universidade de São Paulo, 2008.

QUEIROZ, C. d. A. *Implementação de sistema de aplicação de calcário com taxa variável*. 2010. Trabalho de Conclusão de Curso- Universidade de São Paulo.

SHANER, J. *Smartphones, Tablets and GPS Accuracy*. 2013. Acesso em 29 ago. 2018. Disponível em: <https://www.esri.com/arcgis-blog/products/arcgis-online/field-mobility/smartphones-tablets-and-gps-accuracy/>.

SOUSA, R. V. de; LOPES, W. C.; INAMASU, R. Y. Automação de máquinas e implementos agrícolas: eletrônica embarcada, robótica e sistema de gestão de informação. *Embrapa Informática Agropecuária-Capítulo em livro científico (ALICE)*, In: MASSRUHÁ , SMFS; LEITE, MA de A.; LUCHIARI JUNIOR, A.; ROMANI, LAS (Ed.). Tecnologias da informação e comunicação e suas relações com a agricultura. Brasília, DF: Embrapa, 2014. Cap. 11., 2014.

ZHANG, N.; WANG, M.; WANG, N. Precision agriculture — a worldwide overview. *Computers and electronics in agriculture*, Elsevier, v. 36, n. 2-3, p. 113–132, 2002.

Rede de Sensores Sem Fio Subterrânea: Análise e Teste de Comunicação

Tércio A. S. Filho[1], *Alexandre C. R. Silva*[1], *Marcos Napoleão Rabelo* [2] *Sérgio F. Silva*[2], *Juliano Mendes Ribeiro*[3], *José Carlos C. G. Filho* [2]

[1]Instituto de Biotecnologia
Universidade Federal de Catalão (UFCat)
tercioas@ufg.br, sergio.f.silva@gmail.com

[2]Instituto de Matemática e Tecnologia - IMTec
Universidade Federal de Catalão (UFCat)

[3]Universidade Estadual Paulista - "UNESP"
Departamento de Engenharia Elétrica - DEE
Programa de Pós-Graduação em Engenharia Elétrica - PPGEE
juliano.mendes.r@gmail.com,acrsilva@dee.feis.unesp.br

RESUMO

Redes de sensores sem fio subterrâneas ou WUSNs, (*Wireless Underground Sensor Networks*) são uma categoria de redes de sensores sem fio (WSNs), em que os nós sensores estão enterrados no solo. Os nós de comunicação podem ser posicionados no subsolo profundo, no solo próximo à superfície e ao ar fora do solo. Neste trabalho, foi realizado uma análise da umidade do solo, e realizados modelos analíticos que descrevem a relação entre coeficiente hidráulico e a umidade foram obtidos e comparados com trabalhos realizados na área. Além da análise foi realizado uma rede ponta a ponta do tipo WUSN baseado no padrão IEEE 1451. O projeto da rede WUSN baseado no padrão IEEE 1451 permite uma rede expansível e de baixo custo. Para analisar a viabilidade de tal projeto WUSN foram realizados testes de comunicação entre pares de nós, sendo um subterrâneo e outro na superfície do solo em diferentes distâncias com transmissores de 2.4 GHz. Também foi discutido a teoria matemática relacionada à propagação de ondas no solo e os fatores físicos que interferem em tal propagação, consequentemente, na comunicação.

PALAVRAS-CHAVE: Rede de sensores sem fio subterrânea, IEEE 1451, ZigBee.

ABSTRACT

Wireless sensor networks or WUSNs (Wireless Sensor Networks) are a category of wireless sensor networks (WSNs), where sensors are not buried in the ground. Communication assemblies can be positioned without a deep subsoil, with no soil near the surface and no air outside the soil. In this paper, a moisture analysis of the analytical models describing the relationship between hydraulic coefficients and durability was carried out and compared with the work carried out in the area. In addition to the analyses, a WUSN end-to-end network was implemented based on IEEE 1451 standard. The WUSN network design does not have an IEEE 1451 standard. The objective of this study was to perform peer communication tests, with one soil and another at the soil surface, in some situations with 2,4GHz. Also discussed was a mathematical theory about the propagation of waves in the soil and the factors that interfere in the propagation, consequently, in the communication.

KEYWORDS: Wireless sensor networks underground, IEEE 1451, ZigBee.

INTRODUÇÃO

Pesquisas em *Wireless sensor networks* (WSNs) têm sido fundamental para um mundo conectado, em que diferentes ambientes podem ser monitorados ou controlados de locais remotos. Tal tecnologia tem sido fundamental para o desenvolvimento de casas e edifícios inteligentes, cidades inteligentes, irrigação inteligente, entre outros (SUN; AKYILDIZ; HANCKE, 2011; ALCARAZ et al., 2010).

Embora as implantações de WSNs normalmente sejam feitas em ambientes fechados ou ao ar livre, como por exemplo, casas, edifícios e na agricultura, recentemente as WSNs vêm sendo estudadas e aplicadas para mensurar informações do subsolo, as quais recebem o nome de *wireless underground sensor networks* (WUSNs). As redes de sensores sem fio subterrâneas têm como objetivo monitorar parâmetros físicos do solo, tais como: umidade, ph (potencial hidrogeniônico), deslocamento e temperatura, que possam ser utilizados para auxiliar a tomada de decisão, como por exemplo, no acionamento de sistemas de irrigação ou emissão de alertas sobre riscos de deslizamento de terra, tal como em barragens (ADAMO et al., 2015).

Além dos benefícios da automação de processos, tais como de mão-de-obra, energia elétrica, água em irrigação, entre outros específicos a cada processo, vale destacar que uma automação sem fio é fundamental para várias aplicações. Pode-se citar como exemplo, na agricultura, o uso de fios ligando sensores do subsolo à superfície é um empecilho, pois máquinas, humanos ou animais podem romper, ou até mesmo se acidentar com os cabos. Em barragens, um pequeno deslizamento de terra poderia romper o fio, ocasionando falhas na rede de comunicação. Assim, as WUSNs representam um avanço, pois não há o uso de cabos, sendo a comunicação feita inteiramente através de nós transmissores sem fio (AKYILDIZ; VURAN, 2010).

Neste trabalho é proposto testes com uma rede WUSN baseada no padrão IEEE 1451 para a análise da comunicação entre nós de rede. O padrão aplica o conceito *Tansducer Interface Module* (TIM) e *Network Capable Application Procecessor* (NCAP) conectados por uma interface

específica definida por outro membro da família do padrão IEEE 1451. O TIM é um módulo que possui um circuito de condicionamento de sinal, conversores A/D (Analógico/Digital) e D/A (Digital/Analógico), memória não volátil para armazenamento dos *Transducer Eletronic Data Sheet* (TEDS) e lógica necessária para realizar a comunicação com o bloco NCAP. Quando desenvolvido baseado no subcomitê IEEE 1451.5 é efinido como WTIM (*Wireless Transducer Interface Module*) em que a comunicação é feita sem fio entre o NCAP e o WTIM. O NCAP é um nó de rede com capacidade de processamento local capaz de receber dados tanto de alto ou baixo nível de abstração através de uma rede externa ou localmente, processar esses dados e enviar para o TIMs através de uma interface padronizada baseada em um dos subcomitê da norma IEEE 1451 (ADAMO et al., 2015).

Testes experimentais entre pares de nós, sendo um subterrâneo e outro na superfície do solo, na vertical, mostram que é possível estabelecer uma comunicação com baixa perda de pacotes à uma profundidade máxima de 80cm no solo. Para o desenvolvimento da rede de sensores sem fio foram utilizados módulos transmissão de frequência de 2.4 GHz. Além da análise da comunicação sem fio no solo foi realizado uma análise da modelagem matemática referente a propagação de ondas eletromagnéticas no solo e a influência de variáveis, como, teor de umidade e coeficiente hidráulico de forma a prover um suporte da tendência de comunicação em diferentes condições ambientais. Como a comunicação sem fio foi testado em um solo específico, ilustramos também o processo usado para a análise do solo e, como resultado, sua composição.

Além da introdução, o presente trabalho está dividido da seguinte forma: na Seção 1 apresenta-se uma descrição das redes de sensores subterrâneas e na Seção 2 é descrito o procedimento para análise do solo e o resultado de tal análise. Na Seção 3 será apresentado o desenvolvimento do módulo para os testes de comunicação. Na Seção 4 apresenta-se os testes realizados e análise da rede, e discute-se sobre a tendência de comportamento da comunicação entre nós, conforme o modelo matemático descrito de propagação de ondas no solo.

1 REDES DE SENSORES SEM FIO SUBTERRÂNEAS

Nas redes de sensores sem fio subterrâneas é necessário analisar a influência dos parâmetros do solo sobre a frequência, ou seja, é necessário realizar uma análise do solo para identificar a interferência que o solo exercerá sobre a comunicação dos nós. Exemplo de parâmetros classificadores do solo são: granularidade, a densidade do solo e, a umidade ou conteúdo volumétrico de água. Para que seja calculado esses parâmetros é utilizado o princípio de Peplinski. Peplinski, Ulaby e Dobson (1995), define a constante de propagação da uma onda no solo como: y = $\alpha + j\beta$, compostas pelas Equações 1 e 2, sendo:

$$\alpha = \omega \sqrt{\frac{\mu\varepsilon'}{2}\left[\sqrt{1+\left(\frac{\varepsilon''}{\varepsilon'}\right)^2}-1\right]} \tag{1}$$

$$\beta = \omega \sqrt{\frac{\mu\varepsilon'}{2}\left[\sqrt{1+\left(\frac{\varepsilon''}{\varepsilon'}\right)^2}+1\right]} \tag{2}$$

$\omega = 2\pi f$ é a frequência angular, μ é a permeabilidade magnética, e ε' e ε'' são as partes real e imaginárias da constante dielétrica, respectivamente. De acordo com a equação, a constante de propagação de uma onda no solo varia de acordo com as frações operacionais de frequência, a areia e argila do solo, a densidade a granel e a umidade do solo ou conteúdo volumétrico de água. Consequentemente, a perda do sinal de comunicação também depende desses parâmetros.

O objetivo das redes de sensores sem fio subterrâneo é estabelecer uma eficiente comunicação sem fio entre o solo e a superfície. As aplicações típicas para essas redes incluem a previsão de terremotos, comunicação em minas/túneis, etc. (AKYILDIZ; STUNTEBECK, 2006). Devido às condições de propagação no meio de solo (rocha, areia e água), técnicas de propagação de sinal sem fio tradicionais através de ondas eletromagnéticas só podem ser aplicadas para intervalos de transmissão muito curta, pois podem ocorrer perda de dados durante a transmissão e a vulnerabilidade a alterações das propriedades do solo, tais como umidade (VURAN; SILVA, 2010).

Embora tradicionalmente nós de sensores sem fio se comunicam usando ondas eletromagnéticas através da propagação pelo ar, nas redes de sensores sem fio subterrâneas a propagação é realizada através do solo, rocha e água, em que no modelo de canal para rede de sensores sem fio trata-se de obstáculos para a comunicação. No entanto, há diversas aplicações em diferentes áreas, como: agricultura – redes de sensores sem fio subterrâneo podem ser usadas para monitorar as condições do solo, parâmetros como temperatura, conteúdo mineral e teor de água podem ser mantidos em níveis ideais; Segurança – sensores enterrados a uma profundidade superficial pode ser usado para detectar movimento através de pressão, vibração ou som; Infraestrutura de monitoramento – existe no subsolo uma quantidade significativa de infra-estrutura, incluindo canalizações de água, rede elétrica e de comunicação, em que sensores podem ser usados para monitorar o funcionamento destas auxiliando em tarefas de manutenção (DUNG et al., 2016).

Segundo Yu et al. (2013) a frequência dos sinais eletromagnéticos é influenciada devido as alterações dos parâmetros de solo, profundidade do nó, densidade do solo, frequência do sinal, alterações do ambiente, a água e o conteúdo (a composição) do solo, os quais têm um efeito significativo sobre a atenuação de sinal durante o caminho.

2 PROCEDIMENTO PARA ANÁLISE DO SOLO

O solo foi analisado utilizando uma balança analítica que permite pesar nominalmente 200g com resolução de 0,01g e estufa capaz de manter a temperatura estável entre 105°C e 110°C. Além da balança e da estufa, foram utilizados recipientes tais como: dessecadores que permitam armazenar a amostra sem a variação da umidade, termômetro graduado 0,1°C, relógio com

indicação de segundos, peneiras de 4,8mm, 3,35mm, 2mm, 600μm, 425μm, 250μm, 150μ, 75μ, uma tigela de fundo para coleta do passante na peneira de 75μ, agitador mecânico de peneiras contendo dispositivo de fixação de até 6 peneiras, tampa e fundo.

Os pontos de coleta do solo foi identificado como: A, B, C e D, de acordo com a Figura 1. A profundidade de cada buraco para a coleta dos dados foram de 60cm, após a sua coleta, cada amostra de terra foi seca até a umidade higroscópica por aproximadamente 8 horas. Na Figura 1 apresenta-se o local da coleta do solo em malha quadrática a 4m de distância entre os pontos.



Figura 1: Pontos de coleta do solo.

O solo foi quarteado e reservou-se em média 1Kg de cada uma das amostras, após o quarteamento, passou-se cada solo na peneira de 4,8mm, não ficando material retido. O próximo passo foi passar o solo na peneira com malha de 2mm, os grãos retido na peneira foram lavados e o passante foi para a peneira de 0,075mm no qual foi lavado até que a água saísse limpa em ambas as lavagens. Com o auxílio da pipeta, espátula e agua, colocou-se os solos lavados em cadinhos identificados, removeu-se o excesso de água dos cadinhos e verificou-se os pesos dos mesmos, deixando-os na estufa por 24 horas. Desta forma, foi possível obter os valores de granulometria d_{10}, d_{50} e d_{60}.

2.1 ANÁLISE E EQUACIONAMENTO

A teoria proposta por Mualem (1976), fornece um modelo matemático que expressa a condutividade hidráulica e os valores de umidade como função de carga hidráulica atuante na massa de solo. Para testes laboratoriais, em que é possível fazer medições e determinar valores de carga hidráulica baseados em padrões normativos, as equações deduzidas por GENUCHTEN (1980) podem ser usados para montar rotinas computacionais e validar os modelos teóricos a partir dos dados coletados em campo. A Equação 3 é a primeira, de um conjunto de equações dada em (GENUCHTEN, 1980), que estabelece a condutividade hidráulica como função da umidade e da carga hidráulica que atua na massa de solo.

$$K_r = \theta_{(h)}^{1/2} \left[\int_0^{\theta_{(h)}} \frac{1}{h(x)} dx \Big/ \int_0^1 \frac{1}{h(x)} dx \right]^2, \tag{3}$$

Na Equação 4, os valores da carga hidráulica não foram calculados em laboratório, assim, a Equação 5 foi modificada para que a Equação 4 fosse utilizada. Na Equação 5 é possível

observar os parâmetros m e n. De acordo com GENUCHTEN (1980), os parâmetros m e n são utilizados para ajustar o modelo de forma a fazer com que se adéque o mais próximo dos dados experimentais, que no presente caso, foram os valores de umidade. Os parâmetros de umidade θ foram normalizados de acordo com a Equação 4, sendo:

$$\overline{\theta} = \frac{\theta_h - \theta_r}{\theta_s - \theta_r}, \tag{4}$$

θ_s e θ_r indicam os valores saturado e residual do teor de água do solo, respectivamente. O valor saturado indica o máximo de teor de umidade θ_s no solo e o residual é o valor medido de umidade θ_r como o mínimo, pois é o valor de umidade permanente no solo devido às condições geológicas, meteorológicas, índices de vazios, porosidade e distribuição granulométrica pelos tamanhos e formas dos grãos. A Equação 4 com m=1 foi usada em muitos estudos para descrever dados de umidade (AHUJA L. R.; SWARTZENDRUBER, 1972) e (ENDELMAN, 1974).

$$\theta = \left[\frac{1}{1 + (\alpha h)^n}\right]^m, \tag{5}$$

Na Figura 2 apresenta-se o comportamento padrão da difusividade de fluido no solo com α=0,05. De acordo, o fluxo possui dois padrões bem estabelecidos, sendo, o fluxo saturado e o residual, e entre esses dois pesos tem-se a difusividade do fluído.

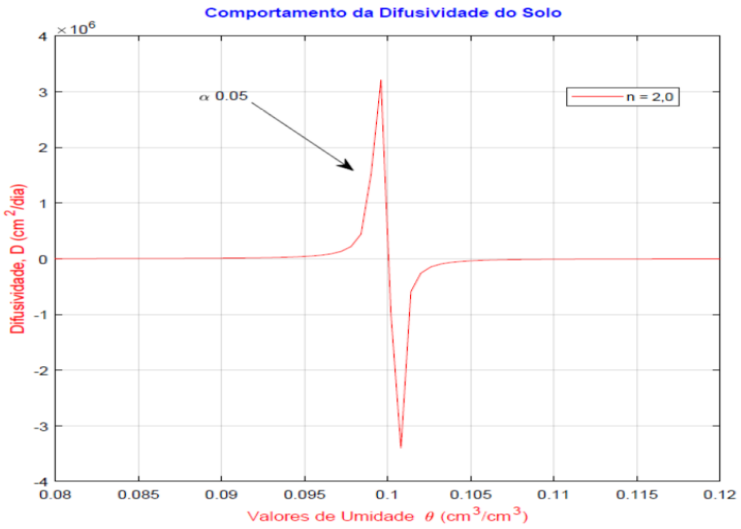
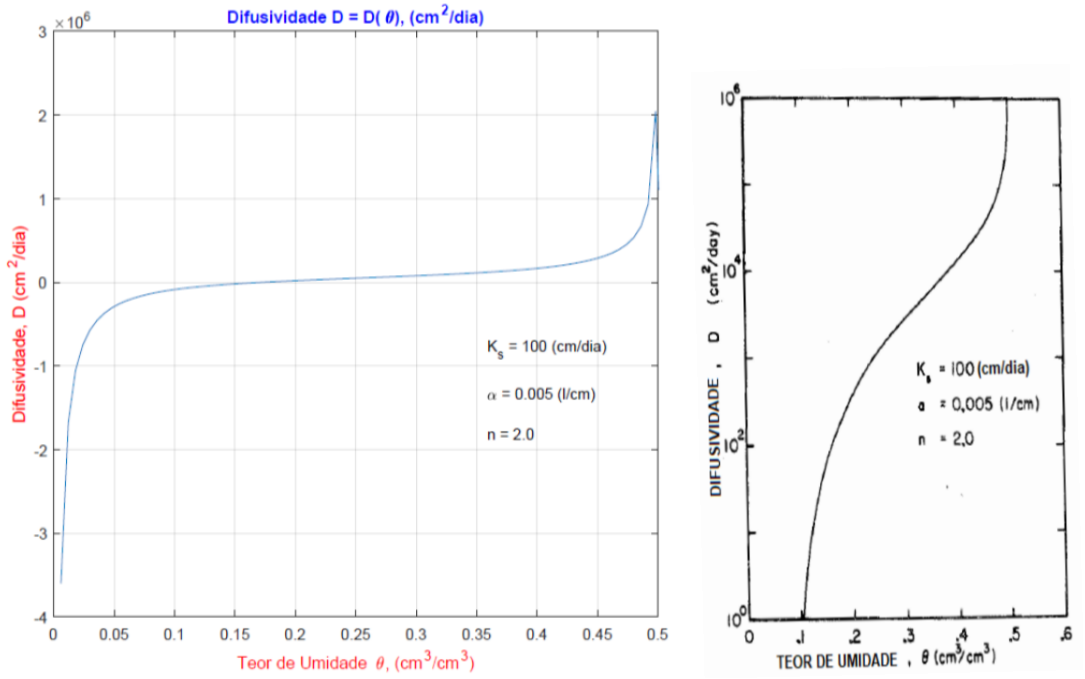


Figura 2: Difundidade de fluído no solo.

Na Figura 3(b), relaciona-se a capacidade de difusividade da água no solo em relação à umidade θ do autor GENUCHTEN (1980). Nota-se um comportamento similar a GENUCHTEN (1980) de acordo com a Figura 3(a). Na Figura 4 representa a variação da condutividade hidráulica como função da umidade, levando em consideração a Equação 6:

$$K_r(\theta) = \theta^{\frac{1}{2}}[1 - (1 - \theta^{\frac{1}{m}})^m]^2, \tag{6}$$

A escala de variação da condutividade hidráulica vai de 0,005 (cm/dia) até 0,030 (cm/dia) o que está dentro da faixa de valores simulados em (GENUCHTEN, 1980).



(a) (b)

Figura 3: a: Comportamento da difusividade D em relação ao teor de umidade θ. b: Comportamento da difusividade D em relação ao teor de umidade θ por (GENUCHTEN, 1980).

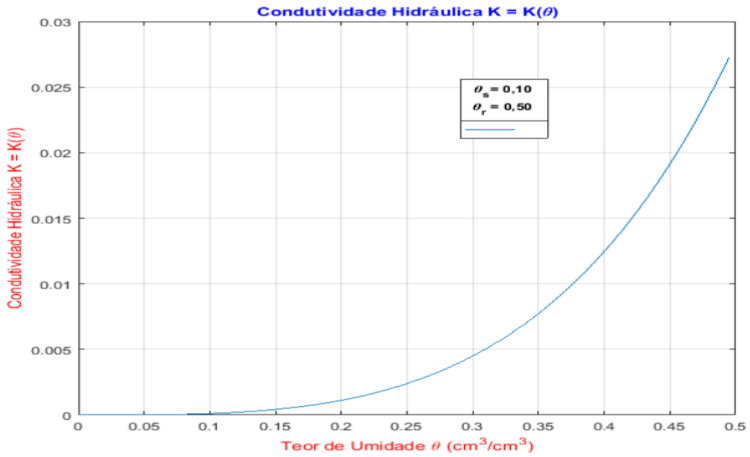


Figura 4: Condutividade hidráulica do solo $K(\theta)$ em relação ao teor de umidade θ

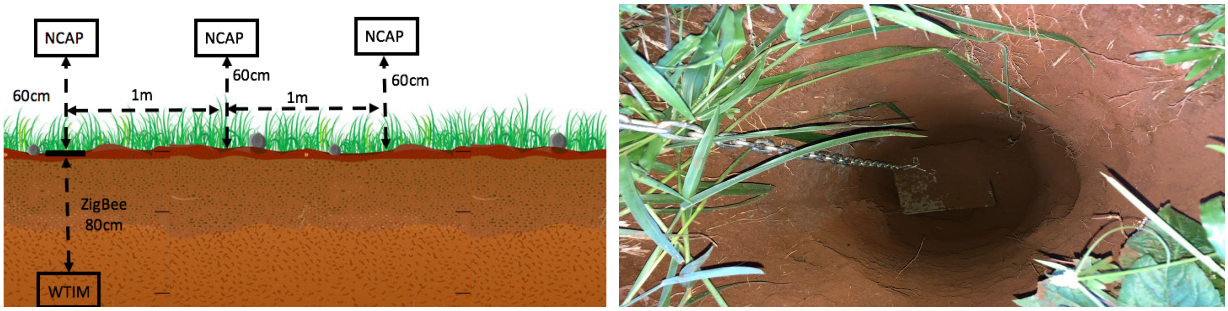
3 DESENVOLVIMENTO DO MÓDULO DE REDE SUBTERRÂNEO

O WTIM desenvolvido neste trabalho possui um microcontrolador ATmega8, circuito de condicionamento de sinais e um transmissor sem fio Xbee PRO S2, com capacidade de roteamento dos dados entre os nós da rede sem fio. Na Figura 5 apresenta o módulo WTIM desenvolvido em laboratório para os testes da comunicação da rede subterrânea. A topologia proposta neste

trabalho é apresentada na Figura 6(a), que tem como objetivo determinar a distância máxima que os módulos são capazes de estabelecer uma eficiente e confiável comunicação subterrânea. A arquitetura para os testes é composta por 2 módulos, um módulo subterrâneo e outro na superfície. Na Figura 6(b) apresenta-se o módulo no sub-solo para a realização dos testes de comunicação.



Figura 5: Módulo WTIM.



(a) (b) Dispositivo no solo.

Figura 6: Experimento realizado em campo. a: Topologia da rede. b: Dispositivo no solo.

A comunicação, a configuração e os frames enviados de um módulo para o outro é realizada por meio do *software XCTU*, criado pela empresa *Digi International*. O *software* possibilita a configuração da característica do transmissor sem fio, podendo ser: Coordenador, Roteador ou *End-Device*. Para este trabalho, os módulos foram definidos da seguinte forma, o módulo da superfície denominado de NCAP foi definido como Coordenador e o WTIM, como *End-Device*.

4 TESTE E ANÁLISE DO MÓDULO SUBTERRÂNEO

Neste trabalho foram realizados testes de comunicação entre os módulos NCAP (superfície do solo) e o WTIM (subterrâneo) baseado na topologia da Figura 5. Uma caixa de metal foi desenvolvida para acoplar o módulo WTIM, com uma saída para a antena e uma bateria de 12V, 9A. Os testes foram realizados em diferentes pontos de acordo com o apresentado na Figura **??**. Na Tabela 1 apresenta-se a média dos bytes transmitidos com sucesso, as perdas e a quantidade recebida entre o NCAP e o WTIM.

Tabela 1: Transmissão dos dados entre o NCAP e o WTIM.

	1 Metro	**2 Metros**
Bytes Enviados	1000	1000
Bytes Recebidos	952	798
Bytes Perdidos	48	202
Média (Recebidos)	95,2%	79,8%
Média (Perdidos)	4,8%	20,2%

O desempenho relacionado a transmissão dos dados pode ser observado no gráfico da Figura 7. De acordo, quanto mais próximo do ponto no qual foi enterrado o módulo WTIM melhor é a comunicação.

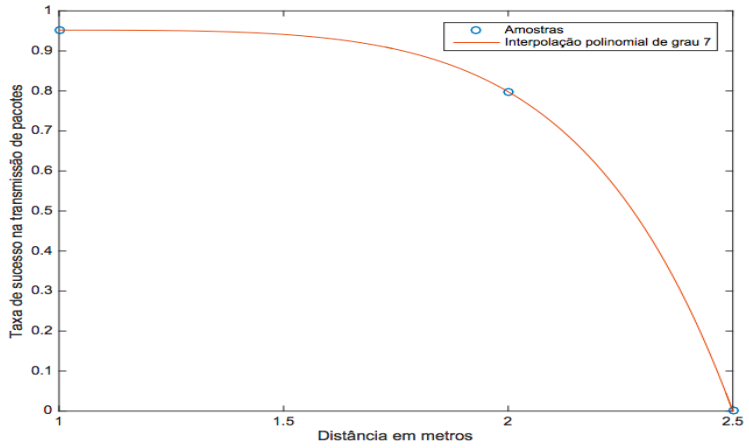


Figura 7: Análise da transmissão dos dados.

5 CONCLUSÕES

As redes de sensores sem fio subterrâneas apresentam diversos desafios relacionadas a comunicação, como por exemplo, compactação do solo e umidade. Tais fatores, interferem na comunicação e devem ser estudados nas redes de transdutores sem fio subterrâneas. De acordo, neste artigo realizou um levantamento bibliográfico apresentando a importância das redes de sensores subterrâneas, suas aplicações e uma estrutura ponto a ponto de comunicação entre dois módulos: NCAP (superfície) e WTIM (subterrâneo). Os testes apresentados demonstram que a comunicação é possível em pequenas distâncias e que é possível desenvolver uma rede subterrânea do tipo mesh intercalando com a superfície. Um outro aspecto é que a rede sendo implementada utilizando o padrão IEEE 1451 a rede se torna expansível e de fácil manutenção. Examinar a comunicação através do solo é significativo para que possa ser analisados os dados provenientes dos transdutores (sensores / atuadores) agregados aos módulos. Os transdutores tem papel relevante, pois, a partir dos dados aferidos por estes é possível controlar sistemas, como por exemplo, sistema de irrigação.

AGRADECIMENTOS

Os autores agradecem ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo fomento da chamada Universal 01/2016, Nº do Processo: 431552/2016-9.

REFERÊNCIAS

ADAMO, F. et al. A smart sensor network for sea water quality monitoring. *IEEE Sensors Journal*, v. 15, n. 5, p. 2514–2522, May 2015.

AHUJA L. R.; SWARTZENDRUBER, D. An improved form of soil-water difusivity funccion. *Soil Sci. Soc.*, v. 36, p. 9–14, 1972.

AKYILDIZ, I. F.; STUNTEBECK, E. P. Wireless underground sensor networks: Research challenges. *Ad Hoc Networks*, v. 4, n. 6, p. 669 – 686, 2006.

AKYILDIZ, I. F.; VURAN, M. C. Wireless sensor networks. Wiley, p. 443–480, 2010.

ALCARAZ, C. et al. Wireless sensor networks and the internet of things: Do we need a complete integration? IEEE, Tokyo (Japan), December 2010.

DUNG, L. T. et al. Impact of soil medium on the path connectivity of sensors in wireless underground sensor networks. p. 60–64, Oct 2016.

ENDELMAN, F. J. e. a. The mathematical modeling of soil-water-nitrogen phenomena. *Ridge National Laboratory*, p. 8, 1974.

GENUCHTEN, M. T. V. A closed-form equation for predicting the hydraulic conductivity of unsaturatedsoils. *Soils Science Society of America Journal*, v. 44, p. 892–898, 1980.

MUALEM, Y. A. New model for predicting the hydraulic conductivity of unsaturated porous media. *WaterResour*, p. 593–622, 1976.

PEPLINSKI, N. R.; ULABY, F. T.; DOBSON, M. C. Dielectric properties of soils in the 0.3-1.3-ghz range. *IEEE Transactions on Geoscience and Remote Sensing*, IEEE, v. 33, n. 3, p. 803–807, 1995.

SUN, Z.; AKYILDIZ, I. F.; HANCKE, G. P. Dynamic connectivity in wireless underground sensor networks. *IEEE Transactions on Wireless Communications*, v. 10, n. 12, p. 4334–4344, December 2011. ISSN 1536-1276.

VURAN, M. C.; SILVA, A. R. Communication through soil in wireless underground sensor networks–theory and practice. Springer, p. 309–347, 2010.

YU, X. et al. A survey on wireless sensor network infrastructure for agriculture. *Computer Standards & Interfaces*, Elsevier, v. 35, n. 1, p. 59–64, 2013.

EaD na Agricultura: análise do perfil discente de um curso de compostagem à distância

Hugo Rogerio Borges de Freitas[1], *Flávio Lages Monteiro Júnior*[1], *Ernani Jardim Reis*[1], *Marco Antônio de Almeida Leal*[1], *Ana Cristina Siewert Garofolo*[1]

[1] Embrapa Agrobiologia, Seropédica, Rio de Janeiro, Brasil, hugo.freitas@embrapa.br, flavio.lages@embrapa.br, ernani.jardim@embrapa.br, marco.leal@embrapa.br, ana.garofolo@embrapa.br

RESUMO

Um dos desafios atuais da transferência de tecnologia agropecuária é a capacidade de processar e disseminar para um contingente expressivo de produtores o conhecimento tecnológico gerado pela pesquisa. Paralelamente a isso, dados do Censo Agropecuário mostram que entre 2006 e 2017 o número de estabelecimentos rurais com acesso à Internet saiu de 75.407 para 1.425.323. Neste contexto, o uso de Ambientes Virtuais de Aprendizagem pode contribuir para o encurtamento do tempo entre a geração e a disseminação destas tecnologias. O objetivo deste trabalho foi elaborar um perfil do público que se interessou pelo curso de Compostagem à distância oferecido pela Embrapa Agrobiologia. Através de técnicas de análise exploratória com base em mineração de textos, foram tratadas as respostas fechadas e abertas dos formulários preenchidos por 4.989 inscritos, do Brasil e de outros seis países. Este público é composto, essencialmente, por adultos jovens, com curso superior, sem vínculo empregatício e somente um terço se considera agente multiplicador. As tecnologias da Embrapa mais usadas por 4,63% deste público são: Fossa Séptica Biodigestora, Adubação verde e ILPF. Dois terços deste público tem interesse em capacitações à distância nos temas “horta”, “manejo”, “solo”, “agrofloresta”, “agricultura orgânica”, “agricultura de precisão”, “controle biológico”, “recuperação de área degradada”, “integração lavoura pecuária”, “controle natural de pragas”, entre outros. Com isso, acredita-se que as informações geradas possam contribuir para futuras tomadas de decisão relacionadas com o tema Transferência de Tecnologia Agropecuária à distância.

PALAVRAS-CHAVE: Educação à distância, Transferência de tecnologia, Transferência de tecnologia agropecuária à distância, Mineração de textos.

ABSTRACT

One of the current challenges of transferring agricultural technology is the ability to process and disseminate the technological knowledge generated by research to a wide contingent of producers. Parallel to this, data from the Agricultural Census show that between 2006 and 2017 the number of rural establishments with Internet access increased from 75,407 to 1,425,323. In this context, the use of Virtual Learning Environments can contribute to the shortening of the time between the generation and the dissemination of these technologies. The objective of this study was to elaborate a profile of the public interested in a Composting course at a distance offered by Embrapa Agrobiology. Through exploratory analysis techniques based on text mining, closed and open responses of the forms filled by 4,989 subscribers from Brazil and six other countries were treated. This audience consists mainly of young adults with a college degree, with no employment relationship, and only one-third are considered multiplier agents. Embrapa technologies used by 4.63% of this public are: “Bio-digester Septic Tank”, “Green Fertilization” and “Integrated Crop/Livestock/Forestry”. Two-thirds of this public is interested in distance learning in the areas of "horta", "management", "soil", "agroforestry", "organic agriculture", "precision agriculture", "biological control", "integrated livestock farming", "natural pest control", among others. With this, it is believed that the information generated can contribute to future decision-making related to the topic of transferring agricultural technology e-learning.

KEYWORDS: Distance education, Technology transfer, Transferring agricultural technology e-learning, Text mining.

INTRODUÇÃO

Um dos grandes desafios para a extensão rural e para a assistência técnica é a decodificação do conhecimento tecnológico gerado pela pesquisa, que tem sido processado em um ritmo cada vez mais rápido, em tecnologias passíveis de adoção, e então disseminá-las para um contingente expressivo de produtores rurais. Atualmente, a adoção de tecnologias atinge um número limitado de produtores e uma inclusão produtiva requer “estratégias inovadoras na criação e na transferência de conhecimentos e de tecnologias para ajudar mais produtores, sobretudo os mais vulneráveis, a participar desse fluxo de crescimento” (EMBRAPA, 2014).

Sendo a “transferência de tecnologia e os serviços de extensão rural, cruciais para a adoção de sistemas de produção modernos e mais sustentáveis, de modo crescente dependerão dos avanços em TIC[1]”, abrindo possibilidades para que os atores da cadeia produtiva sejam treinados, assegurando a sustentabilidade do negócio (EMBRAPA, 2014).

O uso de Ambientes Virtuais de Aprendizagem como mecanismo de transferência de tecnologia pode contribuir para encurtar o tempo entre a geração e a disseminação da tecnologia e também favorecer a massificação devido ao seu caráter abrangente.

Desde 2005 o Centro Regional para o Desenvolvimento da Sociedade da Informação - Cetic.br realiza anualmente a pesquisa TIC Domicílios, com o objetivo de mapear o acesso à infraestrutura de TIC nos domicílios urbanos brasileiros, e desde 2008 incorporou também os domicílios rurais à sua pesquisa. Dados referentes ao indicador A4 (DOMICÍLIOS COM ACESSO À INTERNET) mostram um crescimento do total de domicílios da ordem de 3,4 vezes, saindo de 18% em 2008 para 61% em 2017. Estes mesmos dados agrupados por área (rural e urbana), revelam que o crescimento da proporção de domicílios rurais foi de 4% em 2008 para 34% em 2017 (8,5 vezes) (“Cetic.br - Pesquisa TIC Domicílios”, 2017).

Dados do Censo Agropecuário mostram que existiam 75.407 estabelecimentos agropecuários com acesso à Internet em 2006, e os dados preliminares de 2017 mostram a existência de 1.425.323, um aumento de 19 vezes (“IBGE - Censo Agropecuário”, 2017).

Independente das diferenças de métodos e espaços amostrais, ambas as pesquisas mostraram um aumento relevante de domicílios rurais com acesso à Internet no Brasil. Diante disso não há como desprezar o potencial do uso de Ambientes Virtuais de Aprendizagem como um dos mecanismos para Transferência de Tecnologia Agropecuária à distância.

A Embrapa Agrobiologia, uma unidade de pesquisa da Empresa Brasileira de Pesquisa Agropecuária – Embrapa, localizada em Seropédica/RJ, contribuiu para a vitrine de capacitações online da Embrapa (https://www.embrapa.br/e-campo), lançada em 2018, com o **Curso de Compostagem** (https://www.embrapa.br/e-campo/curso-online-compostagem), com base em uma de suas soluções tecnológicas (“Compostagem de resíduos orgânicos para uso na agricultura”, 2005).

O objetivo deste trabalho foi elaborar, a partir da análise dos dados obtidos, uma base de conhecimento contendo as características pessoais deste público, suas percepções das tecnologias da Embrapa e seus temas de interesse, de modo que tal base possa contribuir nas tomadas de decisão estratégicas relacionadas às ações de Transferência de Tecnologia

[1] TIC: Tecnologia da Informação e Comunicação

Agropecuária à distância para públicos semelhantes ao observado, tendo potencial para atender as necessidades de capacitação com mais precisão.

MATERIAL E MÉTODOS

O universo de estudo considerado nesta pesquisa foi composto por todos os alunos inscritos no curso de Compostagem à distância, das turmas oferecidas no segundo semestre de 2018.

Os dados foram gerados a partir da tabulação dos formulários de inscrição preenchidos, elaborados com questões fechadas, de cunho pessoal (nome, endereço, data de nascimento…), e questões abertas, onde os alunos se expressaram livremente sobre o uso de tecnologias da Embrapa e sugeriram temas de interesse para eventuais futuras capacitações à distância.

Foi realizada uma análise exploratória do conjunto de dados gerados, considerando a execução das seguintes etapas: extração, transformação e carga dos dados; mineração e modelagem dos dados e interpretação dos resultados. Para execução destas etapas foram usadas as ferramentas R-Project (http://www.r-project.org) e Rstudio (http://www.rstudio.com), acrescidos de pacotes específicos para mineração de textos.

RESULTADOS E DISCUSSÃO

Após a exportação dos dados dos formulários de inscrição dos alunos, foi realizado o procedimento de extração, transformação e carga inicial dos dados, identificando 4.979 inscritos, oriundos de 1.357 municípios brasileiros diferentes, distribuídos em todos os estados da federação e no Distrito Federal, e também 10 inscritos de outros seis países, sendo: Argentina (1), Colômbia (1), Guatemala (4), Honduras (1), Paraguai (1) e Portugal (2).

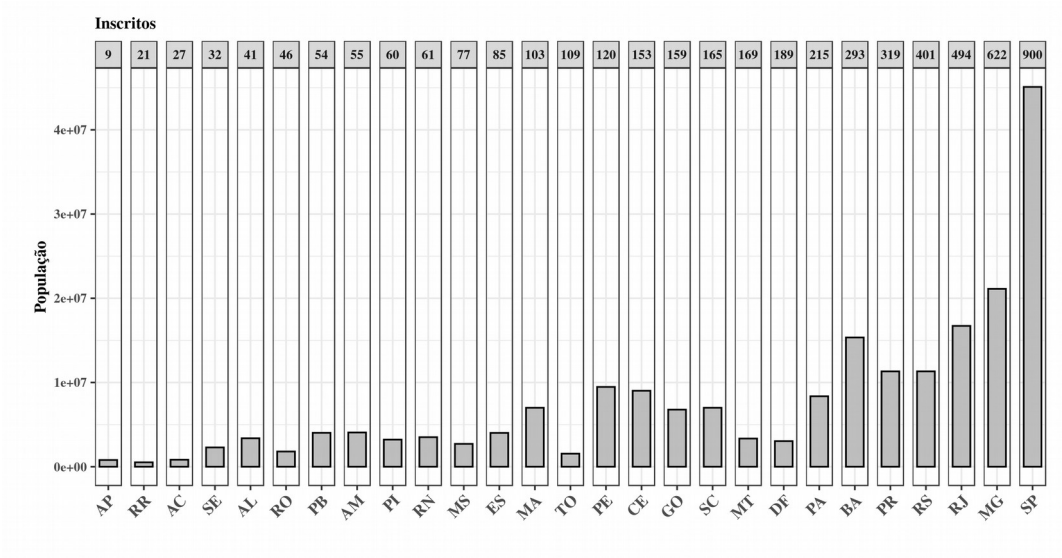
Em seguida à primeira rodada de mineração e modelagem dos dados foi realizada uma análise inicial dos resultados e optou-se por segmentar as respostas em quatro grupos: distribuição geográfica; perfil pessoal; uso de tecnologias Embrapa e temas de interesse.

Distribuição geográfica

A produtividade da agropecuária brasileira apresenta desempenho diferente de acordo com várias características relacionadas a uma determinada região, sendo a difusão do conhecimento e de novas tecnologias uma delas (FELEMA; RAIHER; FERREIRA, 2013). Ou seja, a identificação da localização geográfica do indivíduo que recebeu capacitação em uma determinada tecnologia agropecuária pode contribuir para o entendimento das características que estimularam a procura por tal conhecimento e também, em uma avaliação futura, identificar fatores que contribuíram ou não para a adoção de alguma tecnologia.

Considerando somente os inscritos do Brasil, foi identificada uma distribuição em todo o território nacional, mas com uma variação grande entre os estados. Comparando estes dados com as respectivas populações, percebeu-se que os estados com maior quantidade de inscritos também são os estados mais populosos (Figura 1) e através da estimativa do cálculo do coeficiente de correlação de Spearman (ρ_s=0,8272283) foi identificada uma correlação significativa entre a quantidade de inscritos e a população de um determinado estado.

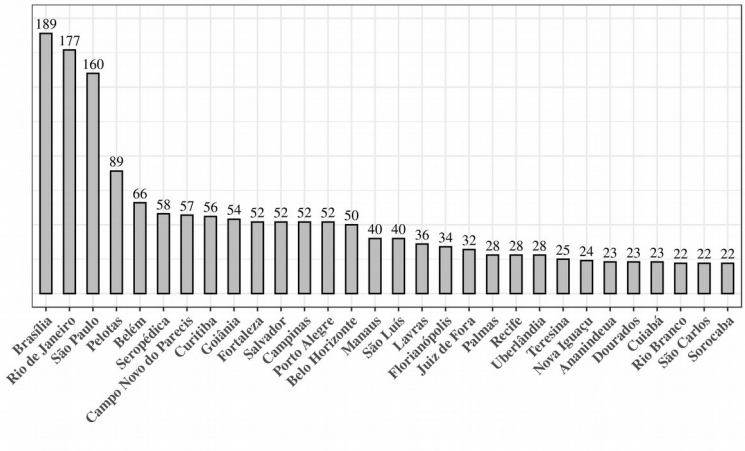
Figura 1 – Inscritos por estado X População por estado



Fonte: dados de pesquisa e IBGE

Na lista dos 30 municípios com maior quantidade de inscritos (Figura 2), há representantes de todas as regiões brasileiras, tanto de capitais (12) quanto do interior (18).

Figura 2 – Municípios com maior quantidade de inscritos



Fonte: dados de pesquisa

Foram representados 1.357 municípios diferentes, ou seja, 24,35% de todos os municípios brasileiros (5.572). Foi identificada grande dispersão quantitativa dos inscritos, sendo 1.275 dos municípios atendidos (93,96%) apresentando 5 ou menos inscrições.

A Figura 3 mostra o resultado da importação dos dados de endereços de todos os inscritos no *Google MyMaps*, o que dá uma visão geral da distribuição geográfica.

Figura 3 – Distribuição geográfica dos inscritos



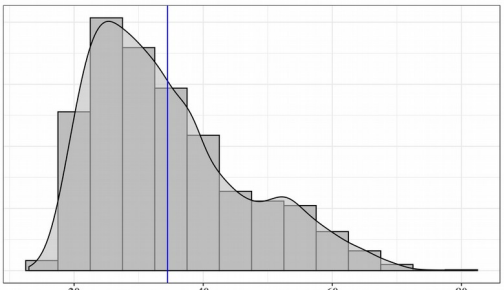
Fonte: dados de pesquisa e Google MyMaps

Perfil pessoal

O conhecimento do perfil do aluno possibilita a implementação de estratégias de desenvolvimento de conteúdo direcionado, permitindo a constante adequação do material didático, contribuindo assim para maior eficiência do processo de ensino-aprendizagem (MOTA; GOMES; LEONARDO, 2014).

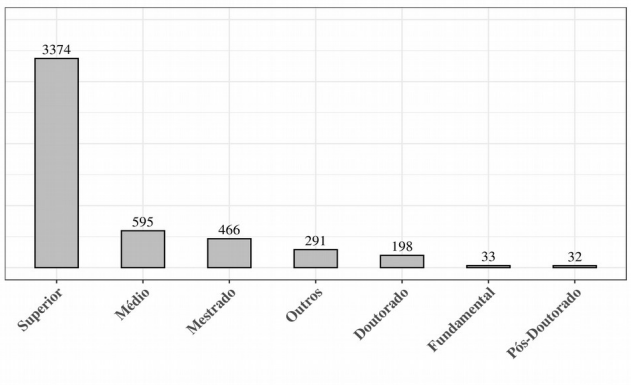
Os dados pessoais revelaram que, essencialmente, é um público adulto jovem, com média de idade de 35 anos (Figura 4) e com curso superior (67,63%) (Figura 5).

Figura 4 – Inscritos por idade



Fonte: dados de pesquisa

Figura 5 – Inscritos por grau de instrução



Fonte: dados de pesquisa

De todos os inscritos, 798 afirmaram ter vínculo empregatício, ou seja, 16,00% do total. Afirmaram não ter vínculo 4.175 (83,68%) e outros 16 (0,32%) não responderam. Outro ponto relevante quando se trata de transferência de tecnologia, é se o capacitado se considera

um agente multiplicador[2]. Dos inscritos, 1.676 (33,59%) afirmaram ser agente multiplicador. Não afirmaram 2.760 (55,32%) e 553 (11.08%) não responderam.

Uso de tecnologias Embrapa

A Embrapa disponibiliza inúmeras soluções tecnológicas[3] que por sua vez são a fonte para as ações de transferência de tecnologia (“Soluções tecnológicas da Embrapa”, 2019).

Durante o planejamento do curso optou-se por criar uma questão que pudesse identificar se o público interessado fazia uso de alguma das tecnologias da Embrapa. Em vez de fornecer uma lista das tecnologias existentes, foi elaborada uma questão aberta para que o interessado pudesse expressar o que ele realmente reconhecia como “tecnologia da Embrapa”.

A Figura 6 mostra uma distribuição de frequência das tecnologias, sendo as mais citadas grafadas com fonte maior, cor mais escura e localizadas mais ao centro, e as menos citadas grafadas com fonte menor, cor mais clara e localizadas mais externamente.

Figura 6: Tecnologias Embrapa mais citadas



Fonte: dados de pesquisa

De todos os inscritos, apenas 231 (4,63%) afirmaram usar alguma tecnologia da Embrapa. A equipe envolvida analisou cada resposta individualmente, uniformizando as grafias de modo a possibilitar o agrupamento de textos que se referiam a mesma tecnologia, mas tinham descrições ligeiramente diferentes. Um ponto identificado foi o fato do público considerar publicações diversas (livros, artigos, páginas *web*…) como tecnologias. Diante disso, estas foram agrupadas em: Material técnico, Livros e Portal. Nos casos que não houve possibilidade de vinculação a alguma solução da Embrapa, as descrições foram mantidas.

[2] **Agentes multiplicadores**: são extensionistas (da iniciativa pública ou privada), consultores autônomos e/ou técnicos de cooperativas, associações, sindicatos, produtores líderes, participantes de eventos de atualização tecnológica e que estão preparados para orientar/repassar o conhecimento e/ou tecnologias gerados pela Embrapa e parceiros para diversos setores da sociedade (EMBRAPA, 2009).

[3] **Solução tecnológica**: conjunto de conhecimentos técnicos testados e validados que assumem o sentido de utilidade para a sociedade e para a Empresa. Abrange o conjunto de Pré-Produtos, Produtos, Processos, Práticas Agropecuárias, Metodologias, Sistemas, Serviços e Ativos de Base Tecnológica da Embrapa em atendimento às demandas produtivas e tecnológicas da sociedade ou público a que se destinam (EMBRAPA, 2009).

Ao final foram identificadas 149 tecnologias diferentes, sendo 40 citadas mais de uma vez, ou seja, 26,85% do total. As três tecnologias mais citadas foram: Fossa Séptica Biodigestora (14), Adubação verde (12) e ILPF (12).

Temas de interesse

Outra questão aberta contida no formulário de inscrição foi: "Quais temas você gostaria de ver em nossos futuros cursos à distância?". O objetivo da equipe foi levantar as necessidades de capacitação à distância almejadas pelos inscritos de modo que esta informação possa ajudar na identificação de cursos já existentes no portfólio que possam atendê-los ou contribuir para uma lista de demandas futuras de desenvolvimento de cursos.

A análise inicial das respostas mostrou que 67,97% dos inscritos sugeriram algum tema, contabilizando 3.391 respostas. Diante da quantidade de dados optou-se por fazer uma análise puramente estatística, usando técnicas de mineração de textos com base no modelo de seleção de n-gramas significativos e não redundantes proposto em MOURA et al. (2010).

Visando melhorar a qualidade dos cálculos de frequência, durante o pré-processamento dos dados foram eliminados todos os caracteres latinos, substituídos os sinais de pontuação por espaços e removidas as palavras com pouco significado (*stopwords*). Em seguida foram extraídos os unigramas (16.063), composto por uma palavra, bigramas (13.168), duplas de palavras, e trigramas (10.941), triplas de palavras.

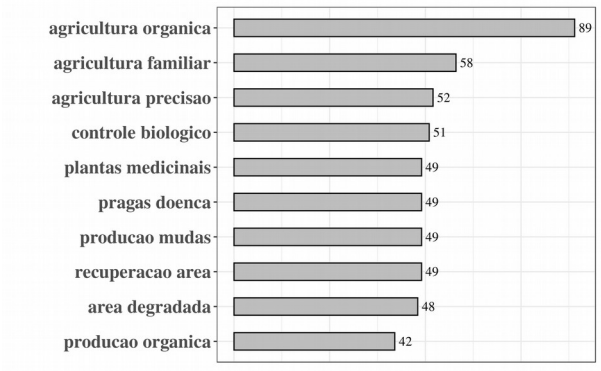
Os 10 unigramas mais citados pelos inscritos foram: "horta", "produção", "agricultura", "manejo", "orgânica", "solo", "cultivo", "compostagem", "agrofloresta" e "curso". Com exceção da palavra "curso", todas as demais identificam um tema na área agrícola, deixando exposto quais são os grandes temas de interesse deste público.

Vários temas relacionados com as atividades agropecuárias são compostos de duas ou mais palavras (Ex.: adubação mineral, calda bordalesa, fibra têxtil natural, impacto ambiental positivo…), e por isso, além das palavras individuais, foram analisadas as frequências de bigramas e trigramas que mais apareceram em uma mesma resposta (ORMOND, 2006).

Observando os bigramas identificados (Figura 7), percebe-se que os três mais citados têm em sua composição a palavra "agricultura", o que ajuda a entender porque este foi o terceiro unigrama mais citado. Os bigramas mostram uma visão mais específica do tema. A palavra "produção"[4] pode estar relacionada com diferentes aspectos na agropecuária mas fica claro que o público tem interesse em "produção de mudas" e "produção orgânica".

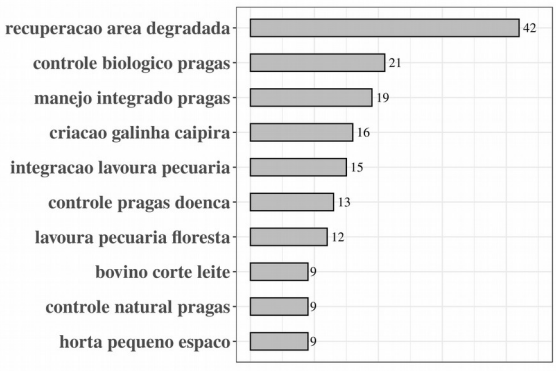
[4] **Produção**: (1) é a acumulação de energia ou biomassa. (2) atividade capaz de produzir, gerar, extrair ou fabricar um bem. (3) o volume produzido. (4) criação de bens e de serviços capazes de suprir as necessidades do homem (ORMOND, 2006).

Figura 7: Temas de interesse – bigramas



Fonte: dados de pesquisa

Figura 8: Temas de interesse – trigramas



Fonte: dados de pesquisa

O levantamento dos trigramas mais citados (Figura 8) também mostrou temas de interesse na área, tais como “recuperação de área degradada” e “integração lavoura pecuária”. Além disso, o aumento do conjunto de palavras possibilita um melhor entendimento das necessidades levantadas. Por exemplo, o bigrama “pragas doença” e o trigrama “controle pragas doença” aparecem na relação dos mais citados, sugerindo que o aspecto “controle” tem mais relevância dentro do tema “pragas doenças”.

CONCLUSÕES

Um público de 4.989 pessoas, distribuídas em todo o território nacional e em alguns outros países, matriculadas em um curso cujo tema é uma das soluções tecnológicas da Embrapa, dá sinais de que o ensino à distância tem potencial para se transformar em uma estratégia inovadora na transferência massiva de tecnologias agropecuárias.

A distribuição geográfica deste público mostra uma dispersão em todos os estados e as características pessoais mostram que é composto por adultos jovens, com curso superior, sem vínculo empregatício, e uma terça parte se considera agente multiplicador.

Somente 4,63% afirmaram usar alguma tecnologia da Embrapa. Parte deste público não demonstrou ter conhecimento efetivo do portfólio de soluções tecnológicas da Embrapa. As mais citadas foram: Fossa Séptica Biodigestora, Adubação verde e ILPF.

Dois terços dos inscritos registraram sugestões de temas para cursos na modalidade à distância, sendo contabilizadas 3.391 respostas. A partir das análises dos unigramas, bigramas e trigramas, foram identificados os temas mais citados por este público: “horta”, “manejo”, “solo” e “agrofloresta”. Temas mais específicos como “agricultura orgânica”, “agricultura de precisão” e “controle biológico” estão entre os mais desejados, assim como “recuperação de área degradada”, “integração lavoura pecuária” e “controle natural de pragas”.

Espera-se que estas informações possam contribuir para tomadas de decisão estratégicas no que se refere a Transferência de Tecnologia Agropecuária à distância, tanto na identificação de temas de interesse quanto no desenvolvimento de conteúdos específicos.

AGRADECIMENTOS

A toda equipe do projeto “Estratégias de ações participativas para inovação na agricultura de base ecológica” que possibilitou a concretização deste trabalho.

REFERÊNCIAS

Cetic.br - Pesquisa TIC Domicílios. Disponível em: <http://cetic.br/pesquisa/domicilios/>. Acesso em: 28 abr. 2019.

Compostagem de resíduos orgânicos para uso na agricultura. Disponível em: <https://www.embrapa.br/busca-de-solucoes-tecnologicas/-/produto-servico/129/compostagem-de-residuos-organicos-para-uso-na-agricultura>. Acesso em: 28 abr. 2019.

EMBRAPA. **Manual do usuário: Sistema de Gestão de Ações de Transferência de Tecnologia (SISGATT)**, 2009. Disponível em: <https://sistemas.sede.embrapa.br/sisgatt/upload/Manual_SISGATT.pdf>. Acesso em: 24 jun. 2019

EMBRAPA. **Visão 2014-2034: o futuro do desenvolvimento tecnológico da agricultura brasileira.** [s.l: s.n.].

FELEMA, J.; RAIHER, A. P.; FERREIRA, C. R. Agropecuária brasileira: desempenho regional e determinantes de produtividade. **Revista de Economia e Sociologia Rural**, v. 51, n. 3, p. 555–573, set. 2013.

IBGE - Censo Agropecuário. Disponível em: <https://sidra.ibge.gov.br/acervo#/S/Q>. Acesso em: 28 abr. 2019.

MOTA, J. B.; GOMES, S. G. S.; LEONARDO, E. D. S. Reflexão sobre o perfil do aluno como determinante para a motivação e aprendizagem em curso de EAD. **Cadernos de Educação, Tecnologia e Sociedade**, v. 7, p. 355, 20 dez. 2014.

MOURA, M. F. et al. Um modelo para a seleção de n-gramas significativos e não redundantes em tarefas de mineração de textos. **Boletim de Pesquisa e Desenvolvimento**, v. 23, out. 2010.

ORMOND, J. G. P. Glossário de termos usados em atividades agropecuárias, florestais e ciências ambientais. 2006.

Soluções tecnológicas da Embrapa. Disponível em: <https://www.embrapa.br/solucoes-tecnologicas>. Acesso em: 29 jun. 2019.

Predição da Produção de Leite Caprino utilizando Redes Neurais Artificiais

Vitória Mendes da Silva[1], Lucas Alberto Teixeira de Rezende[2], João Marcelo Ribeiro[1], Aline Marques Del Valle[1]

[1] Informática, Instituto Federal de Educação, Ciência e Tecnologia do Sul de Minas Gerais, Muzambinho, Minas Gerais, Brasil, vimendesilva@gmail.com, joao.ribeiro@muz.ifsuldeminas.edu.br, aline.valle@muz.ifsuldeminas.edu.br

[2] Agropecuária, Instituto Federal de Educação, Ciência e Tecnologia do Sul de Minas Gerais, Muzambinho, Minas Gerais, Brasil, lucas.teixeira@muz.ifsuldeminas.edu.br

RESUMO

Neste trabalho é apresentada uma ferramenta de apoio à predição da produção de leite caprino e de gerenciamento de produção de setores de caprinocultura. O sistema permite ao produtor e sua equipe gerenciar e armazenar informações referentes aos animais e processos do setor, como produção de leite, reprodução, entre outros, bem como visualizar uma estimativa da produção leiteira, a partir de um conjunto de cabras e de uma data, utilizando Redes Neurais Artificiais. Os resultados encontrados pela rede foram muito próximos aos reais, com desempenho de 99.61% e erro absoluto médio de 0.71%.

PALAVRAS-CHAVE: Caprinocultura, Redes Neurais Artificiais, Predição de Produção de Leite.

ABSTRACT

This work presents a tool to support the prediction of goat milk production and management of goat sector production. The system allows the producer and his team to manage and store information related to the animals and processes in the sector, such as milk production, reproduction, among others, as well as to visualize an estimation of milk production, from a set goats and a date, using Artificial Neural Networks. The results found by the network were very close to the real ones, with performance of 99.61% and average absolute error of 0.71%.

KEYWORDS: Goat breeding, Artificial Neural Networks, Milk Production Prediction.

INTRODUÇÃO

Desde o início das civilizações a criação de cabras esteve presente, assegurando sobrevivência e progresso para a população da época. De acordo com Gonçalves et al. (2002), nas últimas três décadas a caprinocultura leiteira no Brasil vem se tornando uma atividade cada vez mais rentável para o país e despertando o interesse de muitos produtores rurais.

Segundo Fermow (2013), o planejamento e o domínio de produção são importantes, pois isso possibilita um maior controle do que deve ser oferecido ao cliente e do que é esperado para estoque, de forma que não haja surpresas desagradáveis ao produtor. Assim, ter uma estimativa de produção de leite é fundamental para economia/lucro e, além disso, conhecer as variáveis que influenciam a produção pode contribuir para o aumento da produtividade e, em caso de baixa na produção, é possível determinar os fatores relacionados.

Segundo Pardal et al. (2014), a raça, o mês de parto, o tipo de parto e o número de dias de lactação são exemplos de variáveis que podem afetar a produção de leite de cabra. Quando existem dados sobre o domínio de um problema, informações e conhecimentos podem ser obtidos de forma que as decisões tomadas acarretem menor taxa de erros.

Para Guirelli (2006), Sistemas Especialistas, Redes Neurais Artificiais (RNAs), Lógica Fuzzy e Algoritmos Genéticos são exemplos de técnicas da Inteligência Artificial que podem ser utilizadas em aplicações que envolvam predição. Uma possível aplicação seria a predição de produção em diferentes setores, como agricultura, pecuária, indústria, entre outros.

Em alguns trabalhos cujo objetivo era predizer a produção de leite, foi observado que o uso de Redes Neurais Artificiais foi predominante. Segundo Kominakis et al. (2002), a utilização de Redes Neurais apresenta bons resultados. Os autores afirmam que a diferença média entre os rendimentos observados e previstos foi, em sua maioria, não significativa (menor que 0,05) e que quanto maior a quantidade de dados, melhor o resultado.

Assim, este trabalho propõe a realização da predição da produção de leite de cabra utilizando Redes Neurais Artificiais, de forma a determinar e avaliar as variáveis correlacionadas à produção (estação, dias de lactação, raça, entre outras). Sendo a predição uma funcionalidade de um sistema de gerenciamento.

MATERIAL E MÉTODOS

Sistema de Gerenciamento

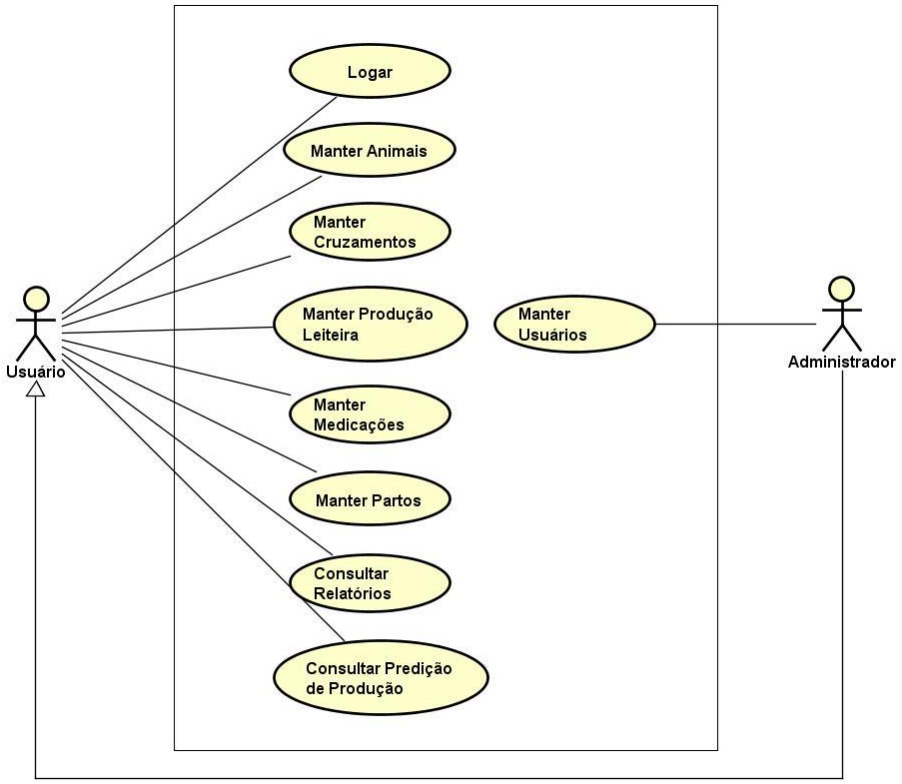
Para o desenvolvimento deste trabalho considerou-se como cliente o Setor de Caprinocultura do IFSULDEMINAS, Campus Muzambinho, representado pelos membros do setor. Foi utilizado o modelo de processo de desenvolvimento de software Iterativo e Incremental.

Iterativo, pois o progresso foi feito com a realização de tentativas de refinamento em versões. E incremental, pois cada versão era composta por funcionalidades novas ou melhoradas.

Primeiramente, foram levantados os requisitos necessários para o desenvolvimento do sistema. Isso foi feito durante reuniões com o cliente e com o auxílio de pesquisas por sistemas de gerenciamento relacionados ao tema. Por meio deste estudo e das reuniões, foram definidas as funcionalidades, as quais foram distribuídas em incrementos no processo de desenvolvimento.

Em seguida, foi modelado o Diagrama de Casos de Uso (Figura 1) utilizando a ferramenta Astah Community. O sistema possui apenas os atores Usuário, que representa os membros do Setor de Caprinocultura, e o Administrador. Ao logar, estão disponíveis as funcionalidades Manter Animais, Manter Cruzamentos, Manter Produção Leiteira, Manter Partos e Manter Medicações. O usuário também pode Consultar Previsão de Produção Leiteira para cabras em uma determinada data e Consultar Relatórios.

Figura 1 – Diagrama de Caso de Uso.



Escolheu-se como linguagem de programação o Python, pois, de acordo com Marques et al. (2011), é uma linguagem de grande simplicidade e poder, que vem ganhando espaço entre os programadores, junto ao framework Django, seguindo o padrão Model, View, Template (MVT). O framework foi escolhido por possuir uma documentação clara e completa e possuir casos de uso de sucesso (PEREIRA, COGO e CHARÃO, 2009).

A cada incremento do sistema, foi realizado um Teste de Usabilidade para que as funcionalidades desenvolvidas fossem validadas pelo cliente, este teste é utilizado para medir o grau de satisfação, eficiência e eficácia alcançadas pelo usuário durante a utilização do sistema. Cada teste durou cerca de 1 hora e todas as modificações e ajustes necessários foram anotados para que, posteriormente, pudessem ser implementados. A partir da implantação do sistema, os usuários puderam utilizar suas funcionalidades e responder um questionário de aceitação[1] para que opiniões acerca da ferramenta desenvolvida pudessem ser colhidas.

Predição da Produção Leiteira

Para realizar a predição da produção leiteira foram seguidas as etapas de preparação de dados e aplicação da rede neural artificial de forma iterativa, conforme descrito abaixo.

- **Preparação dos Dados**

Nesta etapa, os dados armazenados em papéis e coletados desde o ano de 2015 pelo Setor de Caprinocultura do IFSULDEMINAS, Campus Muzambinho, foram recolhidos. Foram selecionados os dados julgados importantes dentre os registros de produção leiteira no decorrer dos anos de 2015, 2016 e de 2017 e registros relacionados a cada animal por meio de fichas cadastrais. Os mesmos foram analisados para que fossem selecionados apenas os dados úteis para compor a base de dados.

Os dados foram fornecidos em tabelas no formato xls e em papéis manuscritos pelos membros do setor. Realizou-se uma transformação para que todos os dados pertencessem a um mesmo repositório. Dados inconsistentes e incompletos foram corrigidos, a fim de se alcançar um conjunto de dados consistente.

Foram criadas tabelas normalizadas na base de dados, a fim de alcançar um conjunto de informações organizado e completo. Os dados foram categorizados de forma adequada, para que houvesse um bom funcionamento na aplicação da Rede Neural Artificial.

Foi feito o cadastro dos dados na base. Foram cadastrados dados específicos de cada animal como identificação, raça, data de nascimento, dentre outros; e referentes à produção leiteira, como pesagem do leite dia após dia no decorrer dos anos supracitados. Foram cadastrados cerca de 5100 registros referentes à produção leiteira.

- **Aplicação da Rede Neural Artificial**

Foi utilizada uma Rede Neural Artificial Perceptron Multicamadas (MLP, do inglês *Multilayer Perceptron*) para realizar a estimativa da produção de leite de cabra do setor.

[1] Disponível em: https://urlzs.com/dvh5Y

Redes do tipo MLP possuem uma ou mais camadas de neurônios que realizam uma função específica, as funções de um neurônio de uma determinada camada são resultados do processamento de neurônios da camada anterior que estão conectados a ele, até uma camada de saída (FACELI et al., 2000). A rede foi integrada ao sistema de gerenciamento, para que o usuário, ao fornecer um conjunto de cabras e uma data, receba a predição da produção leiteira referente àquela data e aos animais.

Para isso foi utilizada a linguagem de programação Python com a biblioteca Sklearn v0.20.0 e o algoritmo MLPRegressor. Este algoritmo possui uma rede neural artificial multicamadas implementada e permite a sua utilização por meio da escolha de parâmetros que sejam da escolha do usuário (PEDREGOSA et al., 2011). Esta etapa foi executada por meio de testes dos parâmetros, para que fosse possível alcançar uma estimativa da produção leiteira próxima da realidade e com um tempo de processamento reduzido ao final do processo.

Utilizou-se como atributos a raça do animal, dias de vida, mês de produção, dias de lactação, mês de parto e média dos últimos cinco dias de produção. Para a realização dos testes foram utilizados 3000 registros para treinamento e 2000 registros para teste. Para cada registro/cabra a saída da rede é a predição de produção de leite do animal. O hardware utilizado foi um Notebook Dell, processador Intel Core i3.

Para cada configuração do algoritmo MLPRegressor foi alcançado um resultado diferente. Por isso, para cada teste foi feita uma variação de parâmetros, aumentando e diminuindo número de camadas e neurônios para cada camada, e quantidade de épocas. A função de ativação hiperbólica (tanh) foi mantida. A cada teste, foram geradas informações como desempenho da rede, erro médio absoluto (em uma escala de zero a um), tempo de processamento (em minutos), total de leite produzido e total de leite encontrado pela rede.

A partir dos resultados obtidos nos testes, foi feita uma avaliação elencando e comparando os resultados alcançados, por meio de tabelas e descrições. Desse modo, a visualização para escolha da melhor configuração obtida para ser implantada no sistema de gerenciamento foi facilitada.

RESULTADOS E DISCUSSÃO

Sistema de Gerenciamento

As Figuras enumeradas de 2 a 5 representam as interfaces do Sistema de Gerenciamento desenvolvido. A Figura 2 se refere à Página de Login onde, com o usuário já cadastrado no sistema, o mesmo pode ter acesso às funcionalidades.

A Figura 3 representa a funcionalidade Manter Produção, onde são cadastradas as pesagens de leite do dia, nos períodos da manhã e da tarde, referente às cabras em produção. A Figura 4 mostra o relatório de produção em formato de gráfico, que representa a produção de uma determinada cabra em um determinado período de tempo informado pelo usuário, representando a funcionalidade Consultar Relatórios.

A Figura 5 representa a funcionalidade Consultar Predição Leiteira, onde ao informar a data e as cabras em produção para àquela data, é trazida a predição da produção leiteira clicando em Gerar Predição. Todas as vezes que a RNA é executada ela também é treinada. O botão nomeado Teste representa os testes realizados para verificação de desempenho da MLP.

Figura 2 – Página de Login.

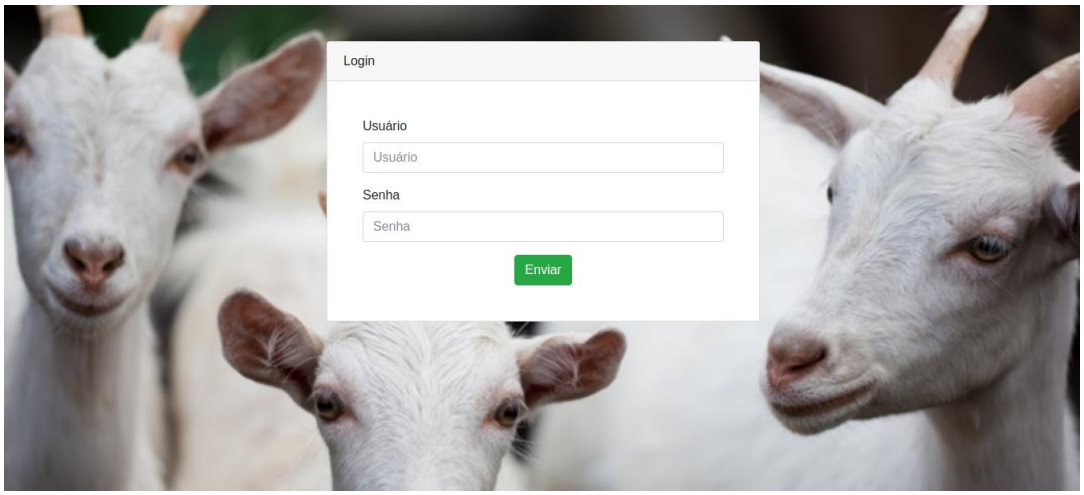


Figura 3 – Listagem Produção Leiteira.

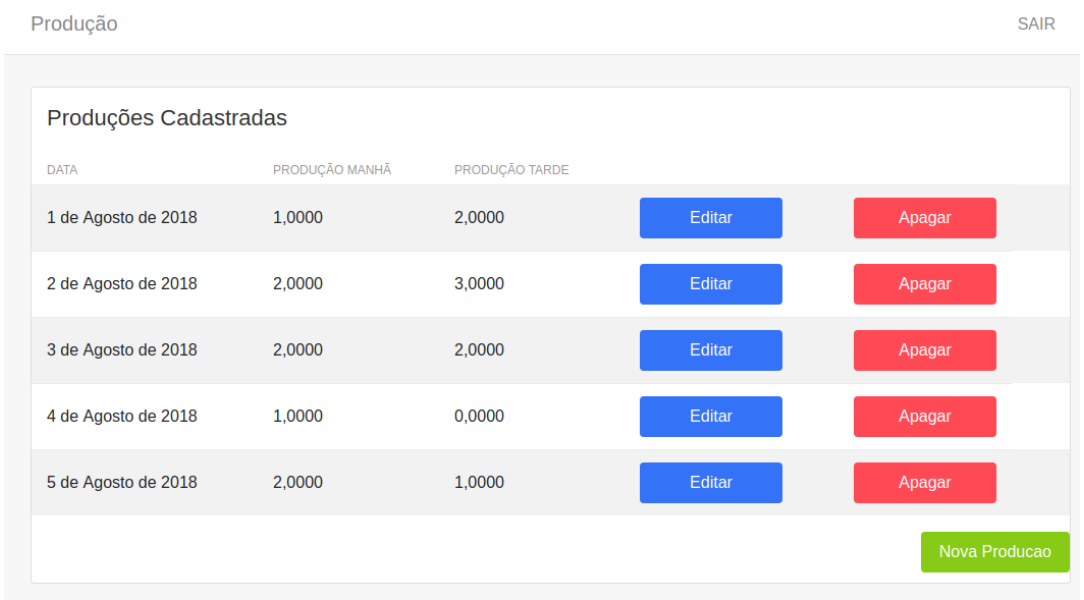


Figura 4 – Relatório Produção da Cabra por período (gráfico).

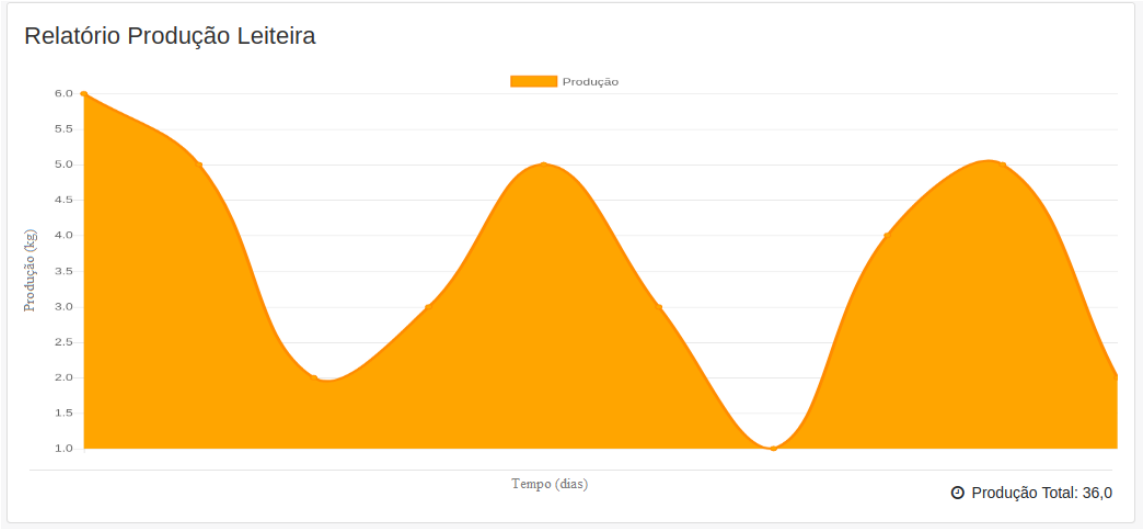
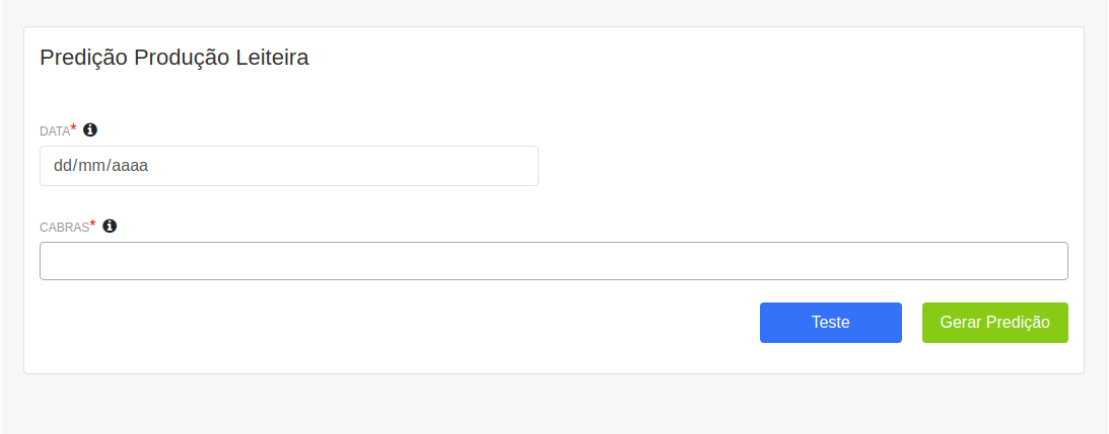


Figura 5 – Consultar Predição Leiteira.



No período de 01 de Novembro de 2018 e 10 de Dezembro de 2018, nove usuários puderam utilizar e avaliar o sistema por meio do questionário, tal questionário foi uma forma de avaliar o grau de satisfação do usuário para com o sistema.

Quanto ao grau de importância do sistema desenvolvido para o Setor de Caprinocultura, em um escala de um a cinco, o sistema foi avaliado pelos respondentes com notas quatro e cinco. Já em relação à satisfação em utilizar o sistema, considerando a mesma escala, as notas também foram quatro e cinco. Sobre a facilidade de uso, 90% dos avaliadores acharam o software fácil de ser utilizado. Quanto às interfaces, foram avaliadas como “Agradável” e “Muito Agradável” por 40% e 60% dos questionados, respectivamente, demonstrando que os usuários gostaram das interfaces. Quando questionados se indicariam o sistema para outros usuários, todos os usuários responderam que recomendariam.

É importante reforçar que todos os avaliadores concordaram em responder o questionário, sendo que nenhum dado pessoal seria repassado em nenhuma situação e que as informações seriam tratadas sempre no coletivo e não individualmente. Apenas os resultados coletivos poderiam ser colocados em imagens para a escrita deste trabalho.

Predição da Produção Leiteira

Com a utilização da Rede Neural Artificial MLP, foi possível chegar a algumas estimativas. Utilizou-se como atributos a raça do animal, dias de vida, mês de produção, dias de lactação, mês de parto e média das últimas cinco produções. Para a realização dos testes que serão apresentados posteriormente foram utilizados 3000 registros para treinamento e 2000 registros para teste. Todos os dados apresentados como resultados dos testes estão normalizados em um intervalo de 0 a 1 (como os da Figura 6).

A Tabela 1 representa um comparativo entre os testes executados no algoritmo MLPRegressor. Como pode ser visto, a RNA pôde admitir resultados diferentes devido à sua configuração (número de camadas, neurônios, número máximo de iterações/épocas e função de ativação) apesar de todos os testes terem sido executados utilizando os mesmos dados.

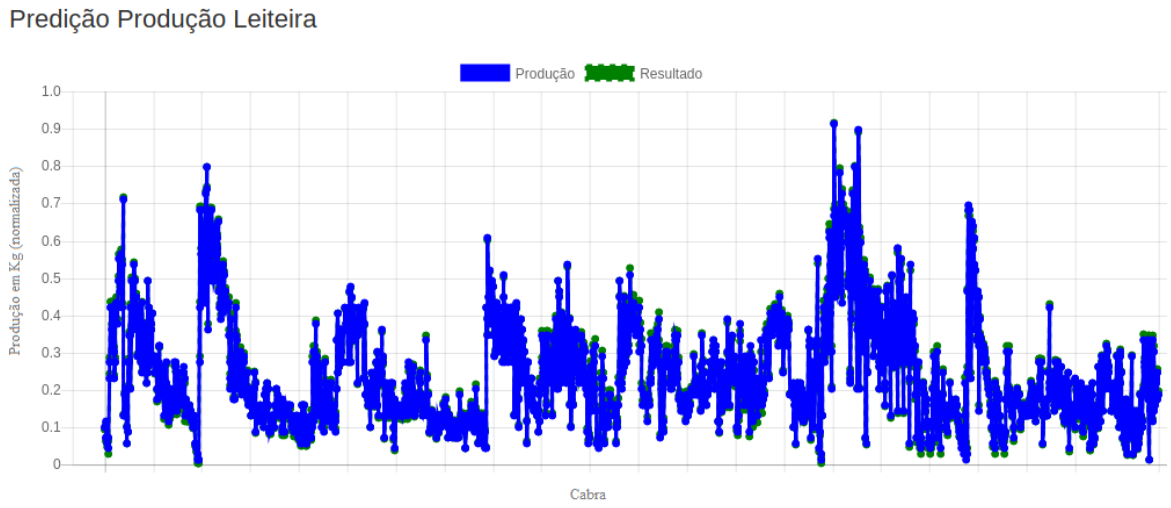
Tabela 1 – Resumo das configurações do algoritmo MLPRegressor de cada teste e seus resultados.

Teste	Camadas Ocultas	Neurônios /Camada Oculta	Iterações /Épocas	Produção Total	Resultado Total	Desempenho	Erro Médio Absoluto	Tempo (min)
1	1	1000	1000	483,1797	483,1845	0,9999	0,00023	70
2	2	100	10000	483,1797	481,5940	0,9737	0,0163	5
3	6	10	10000	483,1797	435,6150	0,5623	0,0747	2
4	4	100	10000	483,1797	488,2066	0,9961	0,0071	2
5	4	10 (1ª camada); 50 (2ª camada); 100 (3ª camada); 100 (3ª camada).	10000	483,1797	524,8867	0,7331	0,0536	2
6	2	50	10000	483,1797	485,0799	0,9592	0,0212	1

De acordo com a Tabela 1, os erros médios para os testes executados variaram de 0.023% a 7.47%, sendo que para a escolha da melhor configuração do algoritmo MLPRegressor também foi avaliado o tempo gasto (treinamento e teste) para convergência da rede e a diferença entre a produção total e o resultado total trazido pela execução da rede. No software, foi implementada a configuração correspondente ao teste 4 (Tabela 1), sendo 4 camadas ocultas com 100 neurônios em cada camada, 10000 como número de iterações/épocas.

O resultado alcançado pode ser visto no gráfico da Figura 6, onde o eixo x representa os 2000 registros/cabras e o eixo y representa a produção normalizada, para a produção real é exibida a curva de cor azul e para o resultado obtido pela rede a curva de cor verde. Pode-se perceber que a produção de leite real e os resultados encontrados pela rede foram muito próximos, com desempenho de 99.61% e erro absoluto médio de 0.71%.

Figura 6 – Resultado Quarto Teste MLP.



CONCLUSÕES

Sistema de Gerenciamento

Com o desenvolvimento do Sistema de Gerenciamento conclui-se que esse é de extrema importância para o cliente, já que todos os dados produzidos por eles eram armazenados em papéis e não possuía nenhum tipo de informatização, o que tornava muitas tarefas mais difíceis e demoradas de se fazer.

A ferramenta desenvolvida em formato web foi feita pensando na facilidade de uso dos usuários, por meio de uma interface limpa e objetiva. A maior parte dos usuários que avaliou a ferramenta demonstrou satisfação ao utilizá-la e afirmou que a interface é agradável e que o sistema é de fácil utilização.

Predição da Produção de Leite

Com a realização do trabalho, conclui-se que é possível alcançar uma boa estimativa de produção de leite caprino utilizando Redes Neurais Artificiais e os atributos raça do animal, dias de vida, mês de produção, dias de lactação, mês de parto e média dos últimos cinco dias de produção. Entretanto, é necessário possuir uma boa quantidade de dados e, no caso do algoritmo MLPRegressor, é preciso conhecer seus parâmetros, pois verificou-se que para

conjunto de parâmetros são admitidos diferentes resultados, o que implica no desempenho final da RNA.

Os erros médios das saídas das RNAs se mostraram baixos na maior parte dos testes realizados. E, na maioria dos testes o tempo de execução foi pequeno, apenas um caso alcançou um tempo maior que 60 minutos, que pode ser justificado devido à quantidade de camadas ocultas.

Como trabalho futuro, poderia ser realizada uma análise de sensibilidade, onde seriam verificadas quais variáveis mais contribuem na predição da produção leiteira. Poderiam ser adicionadas outras variáveis que interferem na produção leiteira de cada animal à base de dados, como dados sobre alimentação, genéticos (parentesco, relacionado ao pai) e a ordem de partos, pois o terceiro e quarto partos possuem maior produção. Além disso, outros algoritmos de predição poderiam ser desenvolvidos para fins de validação do resultado alcançado pela RNA e de comparação.

REFERÊNCIAS

FACELI, K. et al. Inteligência Artificial - Uma Abordagem de Aprendizado de Máquina. Grupo Gen-LTC, 2000.

FERMOW, E. A. C. Planejamento e controle da produção: um estudo de caso em uma indústria de laticínio no município de Ministro Andreazza-RO. 2013.

GONÇALVES, H. C. et al. Parâmetros e tendência genética da produção de leite de cabra no Brasil. Revista Brasileira de Zootecnia, p. 2204-2208, 2002.

KOMINAKIS, A. P. et al. A preliminary study of the application of artificial neural networks to prediction of milk yield in dairy sheep. Computers and electronics in agriculture, v. 35, n. 1, p. 35-48, 2002.

MARQUES, D. L. et al. Atraindo alunos do ensino médio para a computação: Uma Experiência Prática de Introdução à Programação utilizando Jogos e Python. In: Anais do Workshop de Informática na Escola. 2011. p. 1138-1147.

PARDAL, P. et al. Estudos dos factores ambientais influenciam a produção de leite de cabras das raças Saanen e Alpina. In: Revista da Unidade de Investigação do Instituto Politécnico de Santarém, Número especial do Congresso Investigação, inovação e tecnologia: novos desafios. Unidade de Investigação do Instituto Politécnico de Santarém (UIIPS), 2014. p. 30-46.

PEDREGOSA, F. et al. Scikit-learn: Machine learning in python. Journal of machine learning research, v. 12, n. Oct, p. 2825–2830, 2011.

PEREIRA, A.; COGO, V. V.; CHARÃO, A. S. Frameworks para Desenvolvimento Rápido de Aplicações Web: um Estudo de Caso com CakePHP e Django. 2009.

Identificação de pontos de estabilidade e flexibilidade no domínio de formação de preço de venda de produtos da horticultura aplicado à agricultura familiar

Giovane Galvão[1], Simone Nasser Matos[2]

[1] Departamento de Informática, Universidade Estadual de Ponta Grossa, Ponta Grossa, Paraná, Brasil, giovanegalvao@gmail.com

[2] Departamento de Informática, Universidade Tecnológica Federal do Paraná, Ponta Grossa, Paraná, Brasil, snasser@utfpr.edu.br

RESUMO

A gestão de custo na agricultura permite ao agricultor precificar seu produto considerando fatores naturais, como a condição do clima, o solo e a localização da propriedade. No cenário da horticultura aplicada à agricultura familiar, a formação de preço de venda tem passado por mudanças, por isso é necessário desenvolver inovações e estratégias para que o pequeno agricultor familiar possa enfrentar a concorrência de mercado. Este artigo apresenta a identificação de pontos de estabilidade e flexibilidade no domínio de custos de produtos hortícolas na agricultura familiar. Os métodos de custeio utilizados como fonte para extração desses pontos foram: Curva ABC, Custeio por Absorção e Custeio Variável. Como resultado, obteve-se um grafo de fluxo com os pontos de estabilidade e flexibilidade do domínio, além de uma lista com a descrição dos padrões que foram considerados pontos de variação no contexto analisado, os quais podem ser usados no desenvolvimento de um framework de domínio na área de precificação.

PALAVRAS-CHAVE: Agricultura, Gestão de Custos, *Hot spots, Frozen spots*.

ABSTRACT

Cost management in agriculture allows farmers to price their produce considering natural factors such as climate condition, soil and property location. In the horticultural scenario applied to family farming, selling price formation has been changing, so it is necessary to

develop innovations and strategies so that the small family farmer can face market competition. This paper presents the identification of stability and flexibility points in the field of cost of vegetables in family farming. The costing methods used as source for extraction of these points were: ABC Curve, Absorption Costing and Variable Costing. As a result, we obtained a flow graph with the stability and flexibility points of the domain, as well as a list with the description of the patterns that were considered variation points in the analyzed context, which can be used in the development of a framework. domain in the pricing area.
KEYWORDS: Agriculture, Costs Management, Hot Spots, Frozen Spots.

INTRODUÇÃO

A agricultura familiar é uma modalidade da agricultura considerada importante para o desenvolvimento do Brasil. No total, são aproximadamente 4,4 milhões de famílias que praticam esse segmento e este valor representa cerca de 84% dos estabelecimentos rurais do país (PLANO SAFRA, 2017).

Na produção de horticulturas pela agricultura familiar, a definição e o cálculo dos custos dos produtos agrícolas, ao contrário de outros setores da economia, são influenciados por fatores resultantes do caráter da produção agrícola. Entre os mais relevantes estão: os fatores naturais, que incluem as condições do solo, as condições climáticas e a localização da área agrícola. Estes fatores determinam a qualidade da terra e o rendimento das culturas (STAŠOVÁ; BAJUS, 2017).

O desenvolvimento de tecnologias relacionadas a área de precificação na agricultura familiar torna-se importante porque a maioria dos pequenos agricultores não se utilizam de algum tipo de prática de gestão de custo, geralmente, fazem o uso de anotações em papeis (CREPALDI, 2012), ou seja, fazem o uso do método dedutivo para formar o preço de venda de seu produto.

Uma das práticas de gestão de custos está relacionada a formação de preço de venda, pois por meio dela é possível administrar os custos de produção e definir o preço do produto para a venda (CREPALDI, 2012). Alguns softwares de formação de preço de venda já existentes na área agrícola, destinam-se a precificação de equipamentos agrícolas, como é o caso do PRAPAG (MERCANTE et al., 2010), MAQCONTROL (PIACENTINI et al., 2012), entre outros.

Na literatura existem trabalhos de formação de preço de venda da área agrícola que utilizam apenas um *software* específico ou apenas uma metodologia de custeio. Por isto, este

trabalho propõe a identificação de pontos de estabilidade e flexibilidade para o desenvolvimento de um *framework* de domínio que considere mais de uma forma de custeio e os tipos de custos de produção, para isso serão utilizados os dados extraídos da literatura com base nos resultados do mapeamento sistemático realizado para essa pesquisa (GALVÃO; MATOS, 2019), da Secretaria da Agricultura e Abastecimento do Estado do Paraná (SEAB), Companhia Nacional de Abastecimento (CONAB) e Embrapa, com o objetivo de identificar pontos em comum e específicos entre os métodos de custeio: Curva ABC, Custeio por Absorção e Custeio Variável.

GESTÃO DE CUSTOS NA AGRICULTURA

A gestão de custos é um fator relevante devido a globalização econômica, o aumento da concorrência e a diminuição das margens de lucros (CREPALDI, 1998, p. 21).

Os proprietários e/ou agrônomos de fazendas utilizam o conhecimento sobre custos para avaliar economicamente suas atividades e verificar os fatores de produção, tais como: terra, mão de obra e capital, de maneira eficiente, completa e econômica (LOPES et al., 2015).

Na formação dos custos de uma determinada produção se necessita realizar o planejamento e definir as estratégias. Os métodos de custeios são utilizados para realizar a precificação de uma determinada produção. Os métodos de Custeio Baseado em Atividades (ABC), Custeio por Absorção e Custeio Variável, são exemplos de práticas utilizadas pelos pequenos agricultores para a gestão de custo (SENAR, 2015).

O Custeio Baseado em Atividades (ABC) é uma técnica de controle e alocação de custos que permite a identificação dos processos e atividades existentes; análise e o controle dos custos envolvidos em um processo ou atividade; atribuição dos custos aos produtos, tendo como parâmetro a utilização de geradores de custos (SCHIER, 2011).

O método de Custeio por Absorção considera a apropriação de todos os custos de produção de um determinado produto, assim como os demais gastos relativos ao esforço aplicado na produção (MARTINS, 2003).

O método de Custeio Variável divide os custos de produção de um produto em custos fixos e variáveis. Os custos fixos não são apropriados aos produtos, pois estão relacionados a estrutura da organização e não da produção, enquanto, os custos variáveis são aqueles decorrentes da produção (MEGLIORINE, 2011; LEONE, 1997).

FRAMEWORK DE DOMÍNIO

Uma definição para *framework*, segundo Gamma et al. (1994), é um projeto reusável para uma categoria específica de *software* composto por classes abstratas e concretas. Johnson (1997) propôs uma nova definição para *framework*, a qual diz o seguinte: "Projeto reutilizável de uma parte ou todo de um sistema, que é representado por um conjunto e classes abstratas e pelo modo que elas interagem". Com relação ao domínio os *frameworks* podem ser classificados em *frameworks* de aplicação ou de infraestrutura, de suporte ou de integração de middleware, e de domínio.

No processo de implementação de frameworks de domínio torna-se importante a identificação de pontos de estabilidade (*frozen spots*) e flexibilidade (*hot spots*) (PREE, 1995), sendo que: Pontos de estabilidade (*frozen spots*): aspectos comuns a todas aplicações do domínio e Pontos de flexibilidade (*hot spots*): pontos que são específicos a cada aplicação-exemplo. No caso deste trabalho, as aplicações-exemplo são os métodos de Custeio Baseado em Atividades (ABC), Custeio por Absorção e Custeio Variável.

TRABALHOS RELACIONADOS

Com o intuito de verificar o panorama atual de desenvolvimento de *frameworks* de formação de preço de venda aplicado a área agronômica realizou-se um mapeamento sistemático. Foi observado que dos 23 trabalhos selecionados no mapeamento, apenas quatro utilizam um modelo de precificação automatizado, logo, a maioria dos métodos utilizados são de natureza manual (GALVÃO; MATOS, 2019).

Os modelos automatizados presentes nos trabalhos de Nasira e Hemageetha (2012), Ye Lu et al. (2015), Raja et al. (2017) e Leshed et al. (2018), nas quatro ocorrências, são aplicações específicas e não *frameworks*, estes sistemas utilizam atributos relacionados aos custos gerados no processo de produção de uma determinada cultura.

Os autores Nasira e Hemageetha (2012) desenvolveram um sistema de previsão de preços de vegetais, com foco em tomates, utilizando o método de *Backpropagation*. Utilizou-se a ferramenta MatLab para a automatização. No trabalho de Ye Lu et al. (2015) teve como finalidade a previsão de preços da produção de pimentões verdes, por meio do Modelo de Previsão PSO-BP.

Em Raja et al. (2017), utilizando a regressão não-linear, foi apresentado uma tentativa de prever o rendimento e o preço da safra que um agricultor pode obter de sua terra, analisando padrões em dados passados. Os produtos analisados foram: amendoim, batata, carne, cebola, óleo, tomate e trigo. Leshed et al. (2018) desenvolveu uma ferramenta para

cafeicultores calcularem o custo de sua produção, utilizando um modelo de custo de produção subjacente.

Nota-se que existem trabalhos na literatura que precificam culturas na área agrícola, porém, a maioria dos trabalhos utilizam-se de apenas uma metodologia de custeio para tal fim. Pensando-se nisso, a proposta deste trabalho é identificar os pontos de estabilidade e flexibilidade de mais de um método de custeio para que seja possível desenvolver um framework de precificação no domínio agrícola.

MATERIAL E MÉTODOS

Para realizar a identificação de pontos de estabilidade e flexibilidade para o domínio de formação de preço de venda de produtos da horticultura aplicado à agricultura familiar utilizou-se a metodologia de Braga (2003), na qual, esse processo ocorre no início da construção do *framework.*

A determinação dos pontos de estabilidade e flexibilidade é realizada por meio de um grafo da linguagem de padrões, caso exista, ou, a partir da seção "Próximos Padrões" – uma das seções da documentação do padrão de uma linguagem de padrões.

O grafo da linguagem de padrões indica a interação entre os padrões e a ordem em que devem ser aplicados (BRAGA, 2003). A figura 1 ilustra um exemplo de um grafo genérico para a linguagem de padrão.

Na figura 1, o nome do padrão encontra-se no formato "Nome do Padrão", a relação de dependência entre os padrões da linguagem é representada por meio de setas de direcionamento e a ordem de aplicação dos padrões é indicada usando o número que se encontra entre parênteses. Os grupos, representam o agrupamento dos padrões de acordo com sua finalidade e seu propósito (MATOS, 2008).

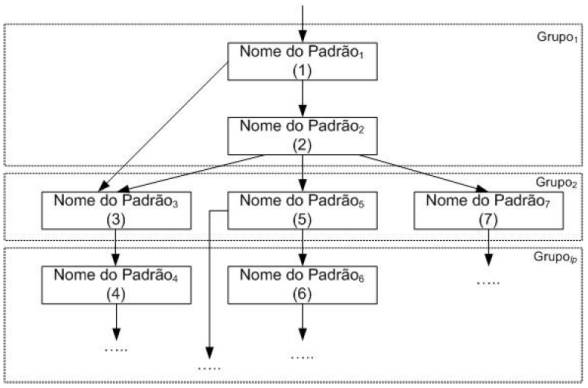
A identificação por meio da documentação indica a sequência de aplicação dos padrões, além disso, tem-se a descrição de quais padrões devem ser verificados após o padrão selecionado ter sido aplicado (BRAGA, 2003).

Ao final da análise do grafo da linguagem de padrões e da documentação, obtêm-se a primeira lista dos pontos de flexibilidade, que é composta dos seguintes campos: nome do padrão, sua descrição, tipo, a seção fonte correspondente na linguagem de padrões (opcional) e por fim seu respectivo número (BRAGA, 2003; MATOS, 2008). Os tipos possíveis dos pontos de flexibilidade podem ser os seguintes:

- PARTIC_ESCOLHA – Estabelece que haverá escolha entre participantes de um padrão.

- PADRÃO_OPCIONAL – Determina que todo padrão pode ser aplicado opcionalmente em aplicações específicas.
- PARTIC_OPCIONAL – Permite que um participante de um padrão não seja utilizado em instanciações particulares.

Figura 1 - Exemplo genérico de um grafo de uma linguagem de padrão.



Fonte: Braga (2003)

Em seguida, estuda-se cada padrão que compõe a linguagem de padrões de forma individual com o objetivo de identificar outros pontos de flexibilidade do *framework*. Assim, tem-se a segunda lista dos pontos de flexibilidade, que pode ser refinada identificando-se outros pontos de flexibilidade não explícitos na linguagem de padrões. Após isso, deve-se analisar os aspectos não funcionais da aplicação, pois pode-se gerar novos pontos (BRAGA, 2003; MATOS, 2008).

RESULTADOS E DISCUSSÕES

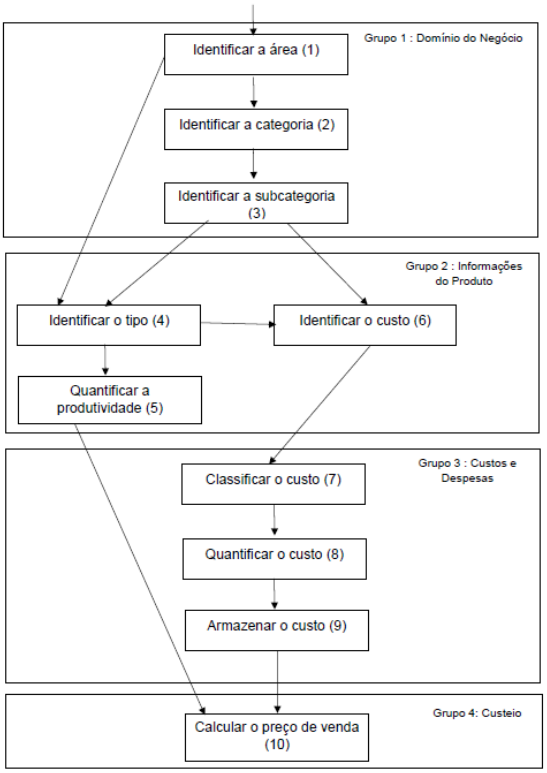
A partir da linguagem de padrões definida para o domínio de formação de preço de venda de produtos hortícolas aplicado à agricultura familiar construiu-se um diagrama de fluxo de aplicações dos padrões com o objetivo de determinar os pontos de estabilidade e flexibilidade. O grafo está ilustrado na figura 2.

Com base na linguagem de padrões criada, nota-se que ela possui quatro grupos: Grupo 1 – "Domínio do Negócio", Grupo 2 – "Informações do Produto", Grupo 3 – "Custos e Despesas" e Grupo 4 – "Custeio". O Grupo 1 possui como nome de padrões os seguintes: *Identificar a área* (1), *Identificar a categoria* (2), *Identificar a subcategoria* (3).

Para o Grupo 2 existem os seguintes nomes de padrões: *Identificar o tipo* (4), *Quantificar a produtividade* (5) e *Identificar o custo* (6). O Grupo 3 é constituído dos seguintes padrões: *Classificar o custo* (7), *Quantificar o custo* (8), *Armazenar o custo* (9). Por

fim, o Grupo 4 é composto pelo padrão *Calcular o preço de venda* (10). Nesta linguagem, um exemplo de relacionamento de dependência entre os padrões é a seguinte (1) – (2) – (3) – (6) – (7) – (8) – (9) – (10), sendo que os números representam a ordem em que esses padrões serão aplicados.

Figura 2 - Grafo desenvolvido com base na linguagem de padrões do domínio da pesquisa.



Fonte: (Os autores, 2019)

Após a análise do grafo de aplicação de padrões, elaborou-se uma tabela com os pontos que podem sofrer variação no domínio. A lista de pontos de flexibilidade foi baseada no modelo proposto por Braga (2003) e pode ser vista na tabela 1. Os outros pontos presentes na figura 2 que não constam na tabela 1, foram considerados os pontos de estabilidade, ou seja, aqueles que não sofrem variação entre os métodos de custeio analisados, são eles: (1) – (2) – (3) – (4) e (5).

Com base na análise do grafo representado na figura 2, identificou-se os seguintes pontos de flexibilidade: *Identificar o custo* (6), *Classificar o custo* (7), *Quantificar o custo* (8), *Armazenar o custo* (9) e *Calcular o preço de venda* (10).

Por exemplo, o padrão *Classificar o custo* foi considerado um ponto de flexibilidade pois dependendo da metodologia do domínio utilizado na formação de preço de venda, essa classificação pode apresentar as seguintes denominações: custo direto, custo indireto, custo

unitário, custo variável, despesas, entre outros. De forma parecida acontece com o padrão *Calcular o preço de venda*, pois conforme o método de custeio (Curva ABC, Custo por Absorção e Custo Variável) escolhido seu funcionamento será distinto.

Tabela 1 – Lista de pontos de flexibilidade.

Nome do Padrão	Descrição	Tipo	Nº Padrão
Identificar o custo	Um custo pode ou não ter um tipo a ele relacionado. É possível ter múltiplos tipos ou tipos aninhados.	PARTIC_ESCOLHA	6
Classificar o custo	Um custo pode ter diversas classificações dependendo do método de custeio utilizado.	PARTIC_ESCOLHA	7
Quantificar o custo	Um custo pode ser único, pode ter múltiplas instâncias, pode ser gerenciado em diferentes unidades.	PARTIC_ESCOLHA	8
Armazenar o custo	Pode ser que a aplicação não necessite gerenciar a armazenagem de custo.	PADRÃO_OPCIONAL	9
Calcular o preço de venda	Existem diferentes formas de custeio de uma produção, cada uma delas tem sua metodologia específica.	PARTIC_ESCOLHA	10

Fonte: (Os autores)

Após a identificação dos pontos de flexibilidade e estabilidade, será possível realizar a implementação do projeto de framework de domínio para formação de preço de venda de horticulturas aplicado à agricultura familiar.

CONSIDERAÇÕES FINAIS

Este artigo apresentou a identificação de pontos de estabilidade e flexibilidade no domínio de formação de preço de venda de produtos da horticultura aplicado à agricultura familiar.

Para que fosse possível o levantamento dos requisitos para a identificação dos pontos, analisou-se dados utilizados em três métodos de custeio: Curva ABC, Custo por Absorção e Custo Variável, extraídos da literatura, da Secretaria da Agricultura e Abastecimento do Estado do Paraná (SEAB), Companhia Nacional de Abastecimento (CONAB) e Embrapa.

Com o grafo de fluxo construído e a lista de pontos de variação do domínio, será possível realizar, como trabalho futuro, o projeto de *framework* que precifique determinada produção, assim contribuindo com a formação de preço de venda de produtos agrícolas, bem

como, aplicar a gestão de custo na agricultura familiar tendo em vista que é uma modalidade que necessita da aplicação de inovações tecnológicas.

REFERÊNCIAS

BRAGA, R. T. V. Um processo para construção e instanciação de frameworks baseados em uma linguagem de padrões para um domínio específico. 2003. 232f. Tese (Doutorado em Ciência da Computação) – Universidade de São Paulo, São Carlos.

CREPALDI, S. A. Contabilidade gerencial, teoria e prática. Atlas, 1998.

CREPALDI, S. A. Contabilidade rural: Uma Abordagem Decisorial. 7 ed. São Paulo: Atlas, 2012.

FAYAD, M. E.; SCHMIDT, D. Object-oriented application frameworks. Communications of the ACM, v. 40, n. 10, p. 32-38, 1997.

GALVÃO, G., MATOS, S. N., GUEIBER, E. Precificação no Agronegócio: Um mapeamento sistemático. Revista de Extensão e Estudos Rurais (REVER), 2019.

GAMMA, E. et al. Design patterns: elements of reusable object-oriented software. Reading: Addison-Wesley, 1994. 424p.

GOMES, A.T.; CEREDA, M.P.; VILPOUX, O. Desidratação Osmótica: uma tecnologia de baixo custo para o desenvolvimento da agricultura familiar. Revista Brasileira de Gestão e Desenvolvimento Regional, Taubaté, v.3, p.212-226, 2007.

GURA, A. Gestão de custos: práticas utilizadas em propriedades rurais familiares. 2018. 104 f. Dissertação (Mestrado em Engenharia de Produção) - Programa de Pós-Graduação em Engenharia de Produção, Universidade Tecnológica Federal do Paraná, Ponta Grossa, 2018.

JOHNSON, R. E.; FOOTE, B. Designing reusable classes. Journal of the Object-Oriented Programming, v.1, n.2, p.22-35, 1988.

JOHNSON, R. E. Frameworks = (Components + Patterns). **Communications of the ACM**, v. 40, n.10, p.39-43, 1997.

LEONE, G. S. G. Custos: planejamento, implementação e controle. São Paulo: Atlas, 1997.

LESHED, G.; HUANG, M.; Mansbach, L. Designing for transparency of cofee production costs, CHI EA ’18 Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems, 2016.

LOPES, M. A.; MORAES, F. d.; CARVALHO, F. d. M.; PERES, A. A. d. C.; BRUHN, F. R. P.; REIS, E. M. B. The effect of technological levels on profts of milk production systems participating in the “full bucket” program: a multicase study, Semina: Ciências Agrárias (4): 2909–2922, 2015.

MARTINS, E. Contabilidade de custos. São Paulo: Atlas, 2003.

MATOS, S. N. Um Método Dirigido por Responsabilidades para Obtenção Antecipada de Pontos de Estabilidade e de Flexibilidade no Desenvolvimento de Frameworks de Domínio. 2008. 274f. Tese (Doutorado em Informática) – Instituto Tecnológico de Aeronáutica, São José dos Campos.

MEGLIORINE, E. Custos análise e gestão. 3. ed. São Paulo: Person, 2011.

MERCANTE, E. et al. Praprag - software para planejamento racional de máquinas agrícolas. Engenharia Agrícola, scielo, v. 30, p. 322 – 333, 04 2010. ISSN 0100-6916.

NASIRA, G. M.; HEMAGEETHA, N. Vegetable price prediction using datamining classification technique, Proceedings of the International Conference on Pattern Recognition, Informatics and Medical Engineering., 2012.

PIACENTINI, L. et al. Software para estimativa do custo operacional de máquinas agrícolas - maqcontrol. Engenharia Agrícola, scielo, v. 32, p. 609–623, 2012. ISSN 0100-6916.

PLANO SAFRA DA AGRICULTURA FAMILIAR 2017/2020 (2017). Disponível em <http://www.mda.gov.br/sitemda/plano-safra-da-agricultura-familiar-20172020> acesso em 22 de junho de 2019.

PREE, W. Design patterns for object-oriented software development. New York: Addison-Wesley, 1995. 268p.

RAJA, S. K. S.; RISHI, R.; SUNDARESAN, E.; SRIJIT, V. Demand based crop recommender system for farmers, 2017 IEEE International Conference on Technological Innovations in ICT For Agriculture and Rural Development (TIAR 2017).

SENAR (2015). Gestão de Custos. Disponível em <http://senar-es.org.br/doc/uc/UC %2013%20-%20Gesta%CC%83o%20de%20Custos.pdf> acesso em 11-09-2019.

SCHIER, C. U. da C. Gestão de Custos. 2. ed. Paraná: Ibpex, 2011.

STAŠOVÁ, L. H.; BAJUS, R. Costs of growing wheat and oilseed rape in slovakia and other v4 countries, Potravinarstvo Slovak Journal of Food Sciences (1): 685–695, 2017.

YASSIN, A; FAYAD, M. E. Application frameworks: a survey. In: FAYAD, M. E.; JOHNSON, R. E. Domain-specific application frameworks: frameworks experience by industry. New York: John Wiley & Sons, 2000. Cap. 29, p.615-632.

YE, L., LI, Y.; LIANG, W.; SONG, Q.; LIU, Y.; QIN, X. Vegetable price prediction based on pso-bp neural network, 2015 8th International Conference on Intelligent Computation Technology and Automation, 2015.



Comunicação modular para uma estação meteorológica

Thiago Piovesan[1], *Marcelo Felipe Chan Yu*[2], *José Alberto dos Santos*[2], *Antonio Marcos Massao Hachisuca*[1]

[1]Centro de Engenharias e Ciências Exatas, Universidade Estadual do Oeste do Paraná
Foz do Iguaçu, Paraná, Brasil
thiagoppiovesan@gmail.com, antonio.hachisuca@unioeste.br

[2]Parque Tecnológico Itaipu
Foz do Iguaçu, Paraná, Brasil
marcelo.fcy@bolsista.pti.org.br, jose.santos@pti.org.br

RESUMO

Com o avanço das indústrias de tecnologia e de redes surgiu o conceito de Internet das Coisas onde passou a não somente conectar pessoas, mas também objetos ao mundo aumentando a eficiência de todas as áreas e serviços. O aumento populacional e a limitação das áreas de plantio são dois grandes problemas do mundo, que por sua vez, requerem uma solução que aumente a produtividade. Um outro agravante é a falta de cobertura de redes de telefonia e a extensão das lavouras, tornando necessário a utilização de equipamentos e infraestrutura que atenda o requisito de cobertura e de baixa manutenção. Assim, a aplicação da tecnologia LoRa e do protocolo LoRaWAN, ambos amplamente disseminados no conceito de Internet das Coisas, é uma das soluções encontradas para atender os requisitos existentes. Então, empregando-se de *softwares* e *hardwares* necessários foram desenvolvidos aparatos para validar a aplicação da tecnologia e do protocolo no campo. Por fim, foi constatado que o resultado obtido com a adoção da proposta foi satisfatória, facilitando a tomada de decisão do agricultor.

PALAVRAS-CHAVE: Internet das Coisas, LoRa, LoRaWAN.

ABSTRACT

With the development of technology and networking industries, the concept of the Internet of Things came into being, which not only connected people, but also objects to the world, increasing the efficiency of all areas and services. Population growth and the limitation of plantation areas are two major problems in the world, which in turn require a solution that increases productivity. Another aggravating factor is the lack of coverage of the cellular networks and the extension of the cutlery, making it necessary to use equipment and infrastructure that meets

the requirements of energy consumption and coverage. Therefore, the application of radio frequency technology, called LoRa, and one of its most widespread protocols called LoRaWAN, was one of the best outputs found to meet all existing requirements. Then, using the softwares and necessary hardwares, devices were developed to validate the application of data technology and protocol in the field. Finally, it was found that the result obtained with the adoption of the proposal was satisfactory, facilitating the decision making of the farmer.
KEYWORDS: Internet of Things, LoRa, LoRaWAN.

INTRODUÇÃO

Internet das Coisas, do inglês *Internet of Things* (IoT), surgiu do avanço das indústrias de tecnologia e de redes de computadores com base em elementos do cotidiano, comumente denominada de "coisas", e a Internet (SANTOS et al., 2016; GALEGALE et al., 2016). O resultado é a criação de uma rede interconectada com a característica de não somente conectar humanos, mas também, humano-coisa e coisa-coisa (DINIZ, 2006).

Dada a grande quantidade de coisas conectadas à Internet, aproximadamente mais de seis coisas por pessoa, novas variáveis devem ser consideradas para as aplicações de IoT como a demanda de uma alta ou baixa taxa de transmissão de dados, alta ou baixa cobertura e a facilidade ou não em realizar manutenções (LEMOS, 2013; SANT'ANA, 2017).

Uma outra variável importante a se considerar é o local onde a aplicação será instalada. Áreas rurais possuem problemas na conexão dos sensores e demais aparelhos já que naturalmente são mais afastadas em relação aos centros urbanos, fato esse que resulta na falta de monitoramento em tempo real comprometendo a segurança e a integridade (GOULART, 2018).

Diante disso, o protocolo de comunicação sem fio mais propensa para as soluções de IoT na agricultura são os baseados em *Long Power Wide Area Network* (LPWAN), em especial a rede *Long Range*® (LoRa) e o protocolo LoRaWAN, uma vez que apresenta baixo consumo energético bem como supre as aplicações para o monitoramento de variáveis e coleta de dados que influenciam de maneira direta ou indireta na produção agrícola (TEIXEIRA; ALMEIDA, 2017).

Um exemplo de aplicação da tecnologia LoRa, na produção agrícola, é explicitado em Santos, Righi e Costa (2018) onde foi possível obter um ganho de 17,94% no desenvolvimento das folhas e 14,29% no peso de um cultivo de rúcula. Dessa maneira, fica evidente a revolução em que IoT acarreta na facilidade e melhoria na produção, desenvolvimento de produtos e até mesmo no aprendizado.

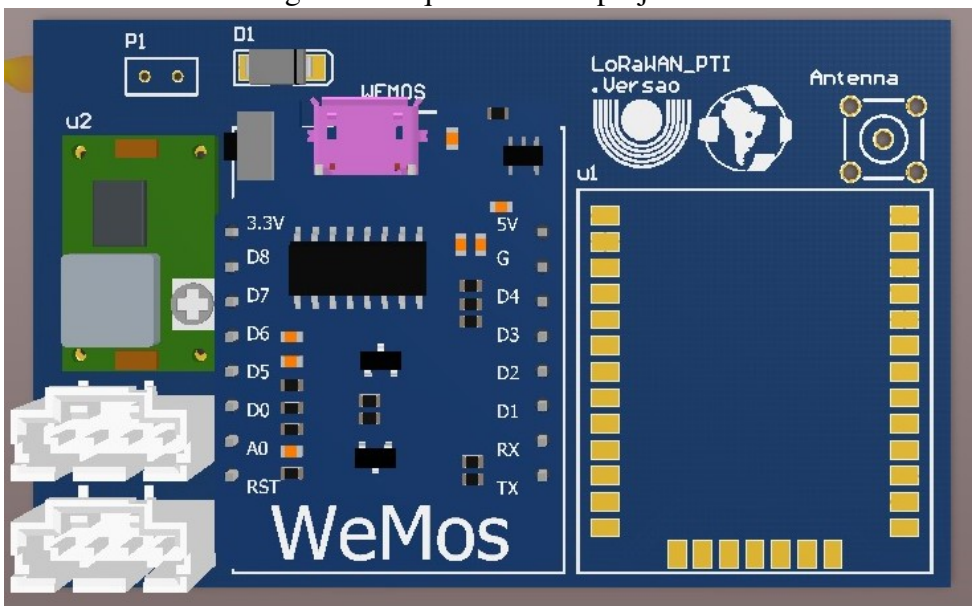
Assim, este artigo apresenta a instalação de uma infraestrutura LoRaWAN, com o desenvolvimento tanto do *hardware* quanto do *software*, para monitorar estações meteorológicas com o objetivo de validar a tecnologia e possibilitar o uso da mesma em futuras aplicações.

MATERIAL E MÉTODOS

O experimento visa demonstrar a aplicação de conceitos de IoT em campo, empregando a tecnologia LoRa e o protocolo LoRaWAN, para coletar dados meteorológicos. Para isso, foi utilizado o módulo LoRa RHF76-052, da empresa RisingHF, e o *software* Autodesk® EAGLE para projetar uma Placa de Circuito Impresso, do inglês *Printed Circuit Board* (PCB), capaz de coletar os dados dos sensores e enviá-los por LoRa com a finalidade de mitigar os erros de leitura comumente realizadas de maneira manual.

A escolha do módulo RHF76-052 se deu pela variedade de tipos de comunicação suportados por ele (UART, I2C, SPI, entre outros) e a possibilidade de programá-lo por comandos ATs. Ilustrada na Figura 1, a PCB projetada emprega ESP8266 como microcontrolador, o protocolo I2C para coletar os dados dos sensores meteorológicos e o protocolo UART para realizar o envio dos mesmos por LoRa.

Figura 1: Esquemático do projetado



Fonte (Próprio Autor)

Os ensaios foram realizados em parceria com a Universidade Estadual do Oeste do Paraná, campus Foz do Iguaçu, nas dependências do Parque Tecnológico Itaipu (PTI) e Itaipu Binacional com a montagem de duas estações meteorológicas composta pela PCB projetada, sensores de temperatura e umidade, placa solar, bateria e uma antena de 915 MHz. Vale ressaltar que a antena receptora de dados, do inglês *gateway* (GW), está instalada numa altura de aproximadamente 35 metros do solo.

LoRa

A tecnologia LoRa tem por objetivo transmitir dados a longas distâncias, com baixo consumo energético, em troca da redução do tamanho dos pacotes de dados (KNIGHT; SEEBER, 2016).

Para tal, emprega-se a técnica de modulação de sinal denominada de Espalhamento Espectral, do inglês *Chirp Spread Spectrum* (SEMTECH, 2015).

Nesta técnica, o sinal a ser transmitido é espalhado no campo das frequências aumentando a resistência contra interferências externas. Na especificação da tecnologia LoRa, o espalhamento espectral está relacionado com o parâmetro Fator de Espalhamento, do inglês *Spreading Factor* (SF), na qual pode ser configurado em seis valores distintos (de 7 a 12) obtendo um aumento no tempo de propagação do sinal no ar, e consequentemente uma maior cobertura, na ordem crescente de SF (ALLIANCE, 2015).

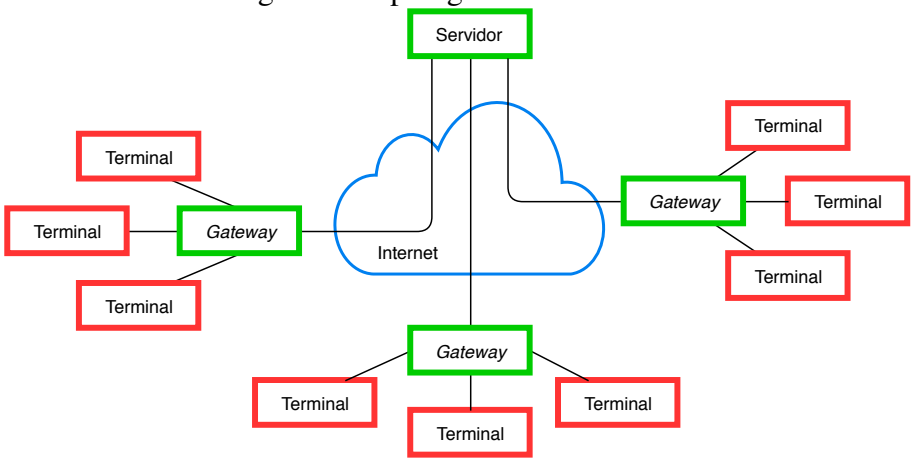
Entretanto, uma desvantagem dessa tecnologia é a comunicação entre dispositivos ser ponto a ponto, impossibilitando a existência de agentes intermediários (servidor ou outros dispositivos), conectados nesta comunicação.

LoRaWAN

A rede LoRaWAN é composta por terminais, GW e servidores. Terminais, dispositivos ou "coisas", são sensores ou atuadores com restrições de alimentação. O GW é responsável por intermediar a rede LoRaWAN com as demais redes e o servidor o encarregado por receber e analisar os dados enviados pelos terminais (BANKOV; KHOROV; LYAKHOV, 2016; RAZA; KULKARNI; SOORIYABANDARA, 2017).

Denominada de "estrela de estrelas", a topologia da rede LoRaWAN é ilustrada na Figura 2 evidenciando as comunicações entre os componentes existentes. As redes dessa topologia são favoráveis no quesito tráfego de rede e roteamento de informações, contudo, necessitam de um GW com capacidade elevada para receber uma grande quantidade de dados de inúmeros terminais (TEIXEIRA; ALMEIDA, 2017).

Figura 2: Topologia em LoRaWAN



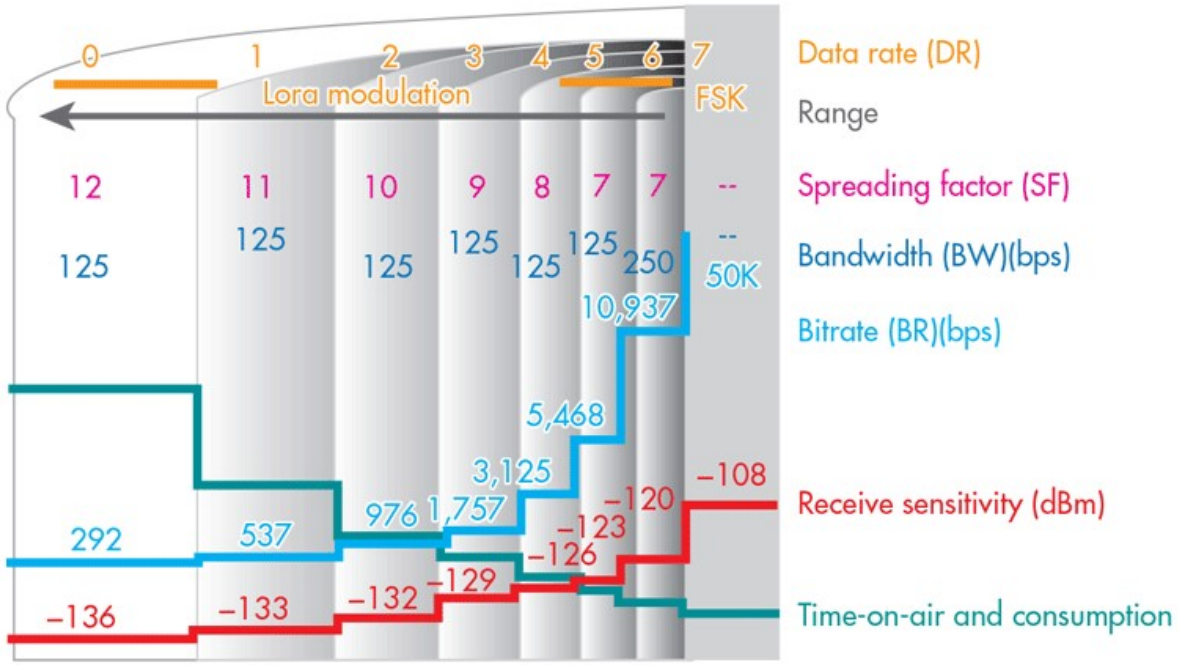
Fonte (ORTIZ et al., 2018)

Uma outra característica do protocolo LoRaWAN é a taxa de dados adaptativa, do inglês *Adaptive Data Rate* (ADR), responsável por gerenciar e alterar a taxa de dados com a finalidade

de garantir o recebimento dos dados no GW empregando-se a menor quantidade de energia e de perdas possível. A Figura 3 ilustra a sensibilidade do protocolo LoRaWAN em relação aos parâmetros existentes.

Diante do exposto, é comum empregar tanto a tecnologia LoRa quanto o protocolo LoRaWAN em soluções de IoT para aproveitar os benefícios de ambos. Baixo consumo energético e longo alcance da tecnologia LoRa e a grande quantidade de dispositivos conectados na rede em virtude da topologia "estrela de estrelas" do protocolo LoRaWAN (ORTIZ et al., 2018; TEIXEIRA; ALMEIDA, 2017).

Figura 3: Taxa de dados adaptativa



Fonte (TEIXEIRA; ALMEIDA, 2017)

RESULTADOS E DISCUSSÃO

Empregando os componentes de *software* e *hardware* listados na seção anterior, juntamente com peças remanescentes de estações meteorológicas, foi possível adaptar duas estações para coletar dados meteorológicos e enviá-los por meio de LoRaWAN. Ilustrados na Figura 4, as estações foram instaladas em dois locais distintos, sendo a primeira distando em torno de 350 metros do GW e a segunda a aproximadamente 3 quilômetros.

Figura 4: Estações meteorológicas

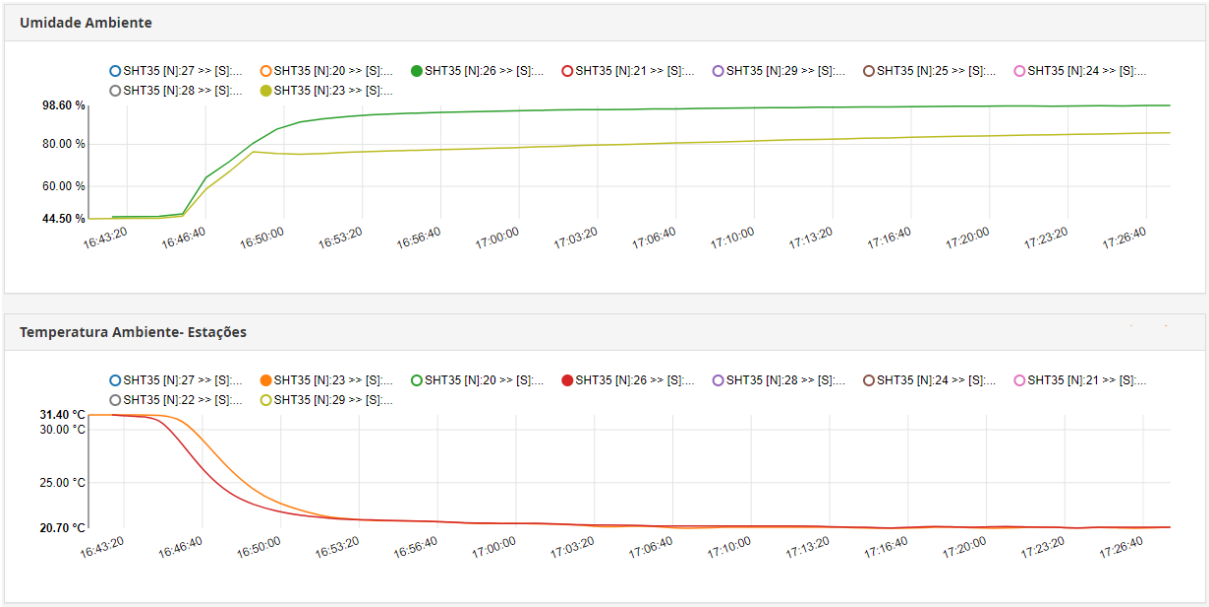


Fonte (Próprio autor)

Com o recebimento dos dados numa frequência média de aproximadamente um minuto, e a utilização de um sistema de armazenamento e visualização de dados, torna-se rápida e prática a identificação do estado meteorológico do local e, consequentemente, a tomada de decisão por parte do agricultor. Uma exemplo disso é ilustrado na Figura 5, onde as estações enviam os dados para a plataforma denominada "Sistema de Monitoramento de Estações Climáticas" (SMEC2), que os expressa graficamente.

Na ocasião em que a figura foi gerada, a temperatura dos dois locais era a mesma, porém a umidade de uma era superior que a outra, o que, dependendo do cultivo, pode significar a necessidade de umedecer o local menos úmido ou a necessidade de secar o local mais úmido. Uma outra aplicação para a coleta extensiva de dados é a possibilidade de prever comportamentos através dos dados históricos, permitindo-se assim, a estimativa precisa da quantidade de chuva num dado período do ano.

Figura 5: Dados de temperatura e umidade mensurados pelas estações



Fonte (Próprio autor)

CONCLUSÕES

Mediante o crescimento populacional desenfreado e limitação das áreas de plantio, o emprego da tecnologia na agricultura é visto como uma solução para aumentar a produtividade, a sustentabilidade e a economia de recursos naturais. Aplicando os conceitos de IoT é possível coletar uma grande quantidade de dados, a curta ou a longas distâncias, e disponibilizá-los de maneira acessível de forma a agilizar o processo de tomada de decisão dos interessados.

A zona rural é caracterizada pelas grandes distâncias e/ou ambientes de pouca circulação, fatos esses que propiciam o emprego da tecnologia LoRa e o protocolo LoRaWAN pois disponibiliza uma comunicação a longa distância, com baixo consumo energético e por consequência um baixo custo de manutenção.

Elaborado de maneira conjunta entre o Centro Latino-americano de Tecnologias Abertas (Celtab), Universidade Estadual do Oeste do Paraná, PTI e Itaipu Binacional, foi possível adaptar duas estações meteorológicas para transmitir os dados mensurados por LoRaWAN a distâncias de aproximadamente 350 metros e 3 quilômetros do GW. Atualmente, além das duas estações mencionadas acima, existem outras duas estações adaptadas e instaladas numa propriedade rural em Toledo/PR transmitindo os dados mensurados por LoRaWAN para a plataforma SMEC2.

Logo, analisando as aplicações citadas no decorrer deste trabalho e considerando o sucesso nos testes realizados, fica aparente a versatilidade da tecnologia LoRa e do protocolo LoRaWAN, além de serem um possível substituto em relação às tecnologias convencionais de comunicação (3G e satélite) para aplicações na agricultura.

AGRADECIMENTOS

O presente artigo foi realizado com apoio do Parque Tecnológico Itaipu e da Itaipu Binacional. O trabalho desenvolvido faz parte do projeto “Infraestrutura LoRaWAN” do Centro Latino-americano de Tecnologias Abertas em conjunto com a Universidade Estadual do Oeste do Paraná campus Foz do Iguaçu.

REFERÊNCIAS

ALLIANCE, L. Lorawan specification. Tech. Rep., 2015.

BANKOV, D.; KHOROV, E.; LYAKHOV, A. On the limits of lorawan channel access. In: IEEE. 2016 International Conference on Engineering and Telecommunication (EnT). [S.l.], 2016. p. 10–14.

DINIZ, E. H. Internet das coisas. GV-executivo, v. 5, n. 1, p. 59, 2006.

GALEGALE, G. P. et al. Internet das coisas aplicada a negócios: um estudo bibliométrico. JISTEM-Journal of Information Systems and Technology Management, Universidade de São Paulo-USP, v. 13, n. 3, p. 423–438, 2016.

GOULART, V. Sistema iot para monitoramento de porteiras utilizando lora e lorawan. CIMATech, v. 1, n. 5, 2018.

KNIGHT, M.; SEEBER, B. Decoding lora: Realizing a modern lpwan with sdr. In: Proceedings of the GNU Radio Conference. [S.l.: s.n.], 2016. v. 1, n. 1.

LEMOS, A. A comunicação das coisas. internet das coisas e teoria ator-rede. Seminários Internacionais Museu Vale, 2013.

ORTIZ, F. M. et al. Caracterização de uma rede sem-fio de baixa potência e longo alcance para internet das coisas. In: Anais do XXXVI Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos. Porto Alegre, RS, Brasil: SBC, 2018. ISSN 2177-9384. Disponível em: <http://portaldeconteudo.sbc.org.br/index.php/sbrc/article/view/2485>.

RAZA, U.; KULKARNI, P.; SOORIYABANDARA, M. Low power wide area networks: An overview. IEEE Communications Surveys & Tutorials, IEEE, v. 19, n. 2, p. 855–873, 2017.

SANTOS, B. P. et al. Internet das coisas: da teoria à prática. Minicursos SBRC-Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuıdos, 2016.

SANTOS, U. J. L.; RIGHI, R. da R.; COSTA, C. A. da. Modelo baseado em rede de sensores sem fio para predição de disfunções em cultivos agrícolas. In: 10º Simpósio Brasileirão Ubíqua e Pervasiva (SBCUP 2018). Porto Alegre, RS, Brasil: SBC, 2018. v. 10. ISSN 2595-6183. Disponível em: <http://portaldeconteudo.sbc.org.br/index.php/sbcup/article/view/3287>.

SANT’ANA, J. M. d. S. Redes lorawan: implantação e desenvolvimento de aplicações. 2017.

SEMTECH, A. 120022. LoRa Modulation Basics, 2015.

TEIXEIRA, G. B.; ALMEIDA, J. V. P. d. Rede LoRa® e protocolo LoRaWAN® aplicados na agricultura de precisão no Brasil. Dissertação (B.S. thesis) — Universidade Tecnológica Federal do Paraná, 2017.

SEGMENTAÇÃO DE IMAGENS DE VANTS PARA DETECÇÃO DE FERRUGEM DA SOJA

Gercina Gonçalves da Silva[1], Ariadne Barbosa Gonçalves[2], Junior Silva Souza[1], Diogo Soares da Silva[1], Nícolas Alessandro de Souza Belete[2,3], Denilson de Oliveria Guilherme[2], Marcelo Theophilo Folhes[4], Hemerson Pistori[1,2]

[1] Universidade Federal de Mato Grosso do Sul, Aquidauana, Mato Grosso do Sul, Brasil, gercina.goncalves@gmail.com, estrutura.pilha@gmail.com, diogo.ec.2013@gmail.com

[2] Universidade Católica Dom Bosco, Campo Grande, Mato Grosso do Sul, Brasil, ariadne.gon@gmail.com, nicolas.belete@unir.br, denilsond@gmail.com, pistori@ucdb.br

[3] Universidade Federal de Rondônia, Cacoal, Rondônia, Brasil, nicolas.belete@unir.br

[4] IbiGeo Geociência Aplicada, Campo Grande, Mato Grosso do Sul, Brasil, marcelo@ibigeo.com.br

RESUMO

Neste artigo são apresentados os resultados da avaliação de um programa de computador desenvolvido para reconhecimento da ferrugem asiática na soja, empregando visão computacional, utilizando o algoritmo SLIC para segmentação e geração de superpixels que compõem dois bancos de imagens gerados a partir de dois VANTS e distribuídas nas classes: Palha, Solo, Folha Sã e Ferrugem. Foram extraídos dessas imagens atributos de texturas, cor e forma, que foram utilizados para medir o desempenho dos classificadores SVM, KNN, J48 e Florestas Aleatórias no ambiente Weka através das métricas: coeficientes de similaridade de Jaccard e de Yule, PCC, Medida-F, área sob a Curva ROC e Kappa. Com o coeficientes de similaridade o banco de imagens A obteve melhores resultados com o classificador SVM enquanto que o banco de Imagens B com o classificador J48. Com a geração da matriz de confusão e posterior cálculo de métricas para cada classe, verificou-se que a classe 'ferrugem' obteve melhores resultados quando avaliada através do coeficiente de Jaccard, observada no banco de imagens B, com o classificador SVM (0,953) e J48 (0,932). Entre os classificadores o melhor desempenho obtido foi verificado com o algoritmo Florestas Aleatórias através da métrica Área sob a Curva ROC em ambos os bancos de imagens. Os resultados indicam que a atuação do programa desenvolvido para reconhecimento da patologia que foi implementado e avaliado permite um bom desempenho no reconhecimento da ferrugem asiática.

PALAVRAS-CHAVE: doenças foliares, superpixel, VANT.

ABSTRACT

In this paper we present the results of the evaluation of a computer program developed for the recognition of Asian rust in soybean using computational vision using the SLIC algorithm for segmentation and generation of superpixels that compose two banks of images generated from two VANTS and distributed in groups: Straw, Soil, Healthy Leaf and Rust. From these

images, we extract the attribute texture, color and shape. These data were used to measure the performance of the machine learning algorithms of the Weka: Support Vector Machines, KNearest Neighbor, J48, Random Forest. The metrics used were: Jaccard and Yule similarity coefficients, Correct Classification Percentage, F-Measure, area under the ROC curve and Kappa. The results obtained with the similarity coefficients showed that image dataset1 obtained better results with the SVM classifier while image dataset2 with the classifier J48. These classifiers were used to generate the confusion matrix and subsequent calculation of metrics for each class. The Rust class obtained better results when evaluated through the Jaccard coefficient, observed in the image dataset2, with the SVM classifier (0.953) and J48 (0.932). Among the classifiers, the best performance was obtained with the Random Forests algorithm, with the metric Area under the ROC curve in both image datasets. The results indicate that the program allows a satisfactory performance in recognition of Asian Rust.
KEYWORDS: leaf diseases, superpixel, UAV.

INTRODUÇÃO

A soja é uma das principais commodities produzidas mundialmente, e é cada vez mais importante no contexto econômico do Brasil, integrando o segmento do agronegócio que possui uma representatividade de 23% no Produto Interno Bruto – PIB nacional. A planta é facilmente cultivada e pode proporcionar um rendimento superior a 4.000 kg por hectare. Todavia diversos fatores podem provocar a redução do potencial de rendimento, entre os quais se destacam as doenças (YORINORI et al.,1993; EMBRAPA, 2013).

Estima-se que no Brasil as perdas anuais na produção causadas por doenças da soja variam entre 15 a 20%, e algumas doenças podem ocasionar 100% de perda da safra (EMBRAPA, 2013). O monitoramento frequente da plantação favorece a identificação dos sintomas iniciais das principais doenças. A vistoria da lavoura deve ocorrer durante o crescimento vegetativo e principalmente antes do início do florescimento da soja, independente da ocorrência ou não de doenças (YORINORI et al., 2003). Usualmente, o monitoramento da soja é feito por amostragem (MORAES, 2007) e, comumente, são realizadas pelos próprios produtores e/ou técnicos, sendo uma tarefa que demanda tempo e recursos financeiros.

Os veículos aéreos não tripulados (VANTs) surgem como ferramenta de auxílio às atividades agrícolas entre as quais se destaca o monitoramento de lavouras de soja (TORRESSÁNCHEZ et al., 2015), possibilitando a coleta de dados com resoluções espaciais e temporais desejadas (HONKAVAARA et al., 2013). A análise de imagens aéreas obtidas por VANTs requer a utilização de programas especialistas. Assim, pesquisadores vêm trabalhando no desenvolvimento de programas aplicados à resolução de problemas agrícolas, a exemplo da pesquisa de Rumpf et al. (2010) que propôs um sistema para a detecção e

diferenciação de doenças na cultura da beterraba sacarina. O programa utiliza Máquinas de Vetores de Suporte (SVM) e, na classificação múltipla de folhas de beterraba sadias e doentes (cercosporiose, ferrugem de beterraba sacarina e oídio), o programa alcançou uma precisão superior a 86%.

Considerando o potencial produtivo da soja no Brasil, bem como a importância da utilização de mecanismos que inovem o processo de monitoramento da fitossanidade da produção agrícola, justifica-se o desenvolvimento de um software que possibilite a detecção rápida e precoce de doenças através do processamento de imagens aéreas obtidas por VANTs, podendo incidir em economia e/ou redução de prejuízo através da automatização da detecção de doenças em plantações agrícolas (JESUS et al., 2015).

Propõe-se nesse artigo a validação do programa DIAGSOJA[1] que utiliza conceitos de visão computacional e reconhecimento de padrões. Tem-se por objetivo reconhecer doenças em plantações de soja a partir das imagens capturadas por VANTs. As imagens utilizadas na validação do programa foram obtidas em um ensaio experimental de plantio de soja. O programa realiza a segmentação das imagens, utilizando o algoritmo Simple Linear Iterative Clustering – SLIC - para agrupar pixels em regiões coesas, com características comuns de cor, textura e brilho. Esses superpixels foram utilizados para criar o banco de imagens para testes de classificadores.

MATERIAL E MÉTODOS

Foi necessária a condução de um ensaio/experimento de plantio de soja numa fazenda em uma área de 1 ha. A variedade utilizada foi a BMX Potencia RR, plantada no dia 25 de novembro de 2014, quando foram semeadas 16 sementes por metro linear de plantio. As coletas de imagens aéreas foram realizadas ao longo do ciclo produtivo da cultura, entre os dias 01/12/2014 e 27/03/2015, através de voos com altura média de 5 metros. As imagens deram origem a dois bancos de imagens digitais de plantas de soja sadias e doentes para teste do software o que se justifica, pois, as imagens foram capturadas por dois VANTs: 1) VANT multirotor, octacoptero com câmera Canon 70d e; 2) VANT DJI Phantom com uma câmera GoPro HERO3 Black Edition acoplada.

Com o primeiro equipamento gerou-se o banco de imagens intitulado aboveR3C4 (above: nome da empresa responsável pelo equipamento, R3: estádio da soja no período em

[1] O DIAGSOJA é um programa desenvolvido pelo INOVISÃO, que é um grupo de pesquisa de desenvolvimento e inovação em visão computacional e tem como principal objetivo a integração entre pesquisa, desenvolvimento e inovação para contribuir com o desenvolvimento do estado do Mato Grosso do Sul.

que as imagens foram coletadas e; C4: quantidade de classes do banco de imagens), e a partir do segundo equipamento gerou-se o banco de imagens phantomR6C4 (phantom: equipamento, R6: estádio da soja no período em que as imagens foram coletadas e C4: quantidade de classes do banco de imagens). Neste artigo o banco aboveR3C4 é referenciado como banco de imagens A e o banco phantomR6C4 é referenciado como banco de imagens B.

O software para reconhecimento de doenças na soja utilizou conceitos de visão computacional (SZELISKI, 2010), onde a aquisição de imagem é o primeiro passo, seguidos do pré-processamento, segmentação e processamento de alto nível (MILANO; HONORATO, 2010). A classificação é obtida com a utilização de algoritmos de reconhecimento de padrões. Para o reconhecimento automático dos sintomas de doenças na soja, optou-se por segmentar as imagens em regiões que compartilham propriedades comuns, tais como brilho, textura e forma. Para tanto, utilizou-se o programa DIAGSOJA, desenvolvido com o auxílio do Play Framework 2.3.7, concebido para a criação de partições com o algoritmo SLIC que ao agrupar vários pixels da imagem em uma única região (superpixel) reduz o custo computacional do agrupamento (ACHANTA et. al., 2010) conforme verifica-se na Figura 1. Assim, procedeu-se a segmentação das imagens com o DIAGSOJA, utilizando para tanto 5 iterações e 40 segmentos. A partir da segmentação de imagens, foram gerados superpixels que compuseram dois bancos de imagens, pois as duas aeronaves geraram imagens com padrões diferentes entre si.

Figura 1 - Exemplo de uma imagem original (à esquerda) segmentada com o extrator de superpixels SLIC (à direita).

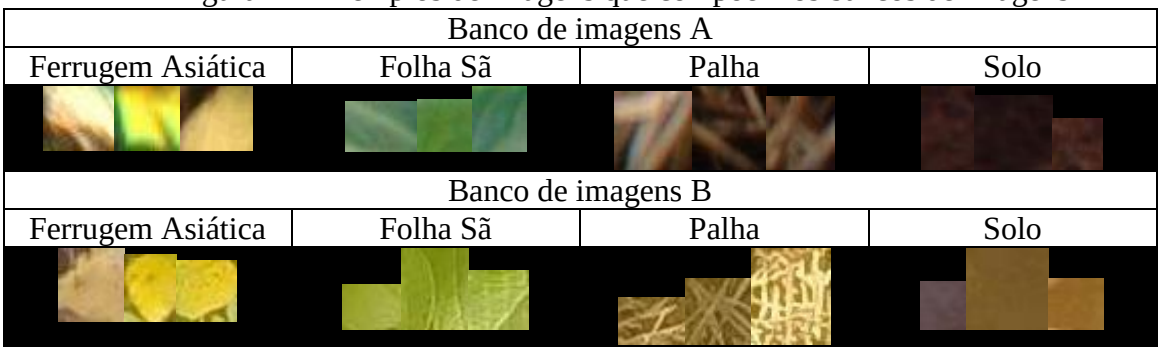


Fonte: Autores, 2019.

Na extração dos atributos de texturas foram realizados cálculos de correlação e entropia com a matriz de co-ocorrência. Ainda na descrição da textura, utilizou-se também de Padrões Binários Locais e, a extração de atributos de cor ocorre através de cálculos de média, desvio padrão, valor máximo e mínimo dos canais dos espaços de cores: RGB, HSV e CieLab. Também foram extraídos histogramas de gradiente nas imagens em níveis de cinza, bem como atributos de forma: formfactor, roundness. A máquina utilizada para o desenvolvimento do trabalho foi um Notebook da marca Samsung NP-RV411-AD3 c/ Intel Core i3, 3GB.

Destaca-se que nesse experimento, apenas uma doença afetou a plantação utilizada para o estudo: ferrugem asiática. Assim, em ambos os bancos, as imagens foram distribuídas nas classes: Ferrugem Asiática, Folha Sã, Palha e Solo, conforme Figura 2. Para o banco de imagens A foram gerados 1125 superpixels e para o banco de imagens B foram gerados 702 superpixels. Com esses superpixels obtiveram-se informações fornecidas pelos atributos das imagens de cada classe, em ambos os bancos de imagens.

Figura 2 – Exemplos de imagens que compõem os bancos de imagens



Fonte: Autores, 2019

A partir de então, utilizou-se da aprendizagem supervisionada empregando-se alguns métodos de classificação (SVM, KNN, J48 e Florestas Aleatórias) com os quais foram realizados os testes no ambiente Weka versão 3.7. Foi utilizada a validação cruzada (Santos et al., 2009) que distribuiu o conjunto de dados em 10 partições, onde 9 partes foram utilizadas para treinamento dos classificadores e uma parte foi utilizada para validação.

Para verificar qual classificador obteve melhor desempenho utilizou-se dos coeficientes de similaridade: coeficientes de Jaccard (LARA, 2007) e de Yule (MEYER, 2002), onde os resultados com valores próximos a 0 são considerados de pior resultado e valores próximos a 1 são considerados de melhor resultado. Com os resultados obtidos junto aos coeficientes de similaridade, foram geradas as matrizes de confusão para os classificadores com melhor desempenho para ambos os bancos de imagens (aboveR3C4 e phantomR6C4). Os valores exibidos nas matrizes permitiram calcular os coeficientes de similaridade para cada classe

abordada, verificando assim em qual classe o programa obteve melhor desempenho. Para avaliar os classificadores foram utilizadas as seguintes métricas: PCC, Medida-F, área sob a Curva ROC e Análise de concordância Kappa. Para verificar se houve diferença significativa entre os classificadores foi aplicada a Análise de Variância.

RESULTADOS E DISCUSSÃO

O fungo Phakopsora pachyrhizi, responsável por provocar a ferrugem asiática na soja, foi verificado no ensaio experimental ainda no estádio vegetativo da planta, em período que foi caracterizado por incidência de muita chuva na região. As primeiras características da doença percebidas durante o monitoramento do ensaio experimental e coleta de imagens foi a visualização de lesões na cor verde claro. Posteriormente a coloração alterou-se, evoluindo para a cor amarelo ferrugem, havendo também alterações no tamanho das lesões na face inferior dos folíolos.

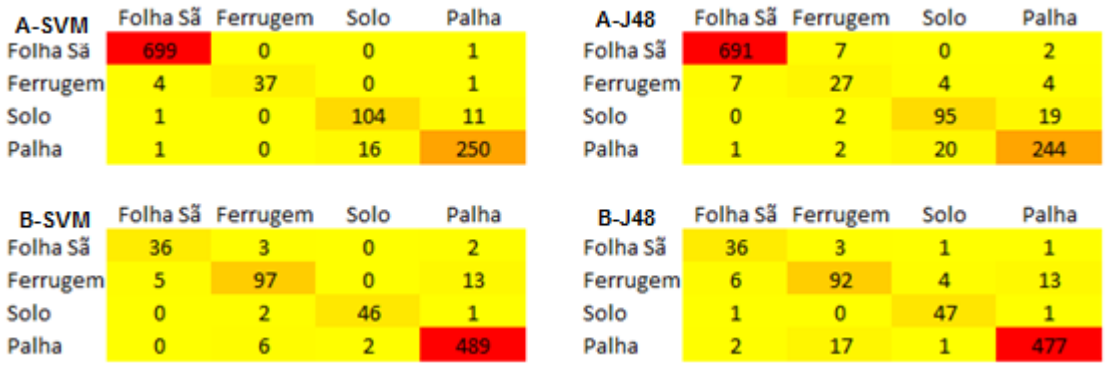
Após a segmentação das imagens através do DIAGSOJA, procedeu-se as classificações dos dois bancos de imagens (aboveR3C4 – banco de imagens A, e phantomR6C4 – banco de imagens B), utilizando os algoritmos de classificação SVM, KNN, J48 e Florestas Aleatórias (FA), e a técnica de amostragem validação cruzada com 10 dobras. Os resultados das classificações foram avaliados através dos coeficientes de similaridade de Jaccard e de Yule. Verificou-se que, tanto o coeficiente de Jaccard como o coeficiente de Yule, indicaram que as imagens do banco de imagens A (aboveR3C4) apresentaram melhores resultados com o classificador SVM, e as imagens do banco de imagens B (phantomR6C4), com o classificador J48, conforme Tabela 1.

Tabela 1 - Avaliação dos classificadores considerando os índices de similaridade.

Métricas	SVM		KNN		J48		FA	
	A	B	A	B	A	B	A	B
Coeficiente de Jaccard	**0,990**	0,842	0,935	0,624	0,971	**0,861**	0,971	0,780
Coeficiente de Yule	**0,990**	0,857	0,935	0,712	0,970	**0,873**	0,971	0,820

A Figura 3 apresenta as matrizes de confusão para os bancos de imagens A e B, para as quais foram utilizados, respectivamente, os classificadores SVM e J48, pois obtiveram melhores desempenhos com ambos os coeficientes, objetivando avaliar qual classe - ferrugem asiática, folha sã, palha ou solo- foi melhor classificada. Nas matrizes de confusão, as imagens em tom vermelho representam os valores mais altos. Destaca-se que quanto mais próximos os valores das diagonais principais da matriz de confusão sejam do valor total de imagens contidas na classe, menor foi a confusão entre as classes.

Figura 3 - Matriz de confusão - Banco de imagens A e B, Classificadores SVM e J48



Fonte: Autores, 2019.

Utilizando uma matriz de confusão, Rezende (2003) demonstra como calcular medidas quantitativas através da classificação de pixels, considerando problemas em que a classificação envolve apenas duas classes. Conforme observado na Figura 3, este artigo trabalha com 4 classes e, assim, para calcular as medidas quantitativas, a classe de interesse (classe analisada) foi considerada como a classe positiva, enquanto todas as outras foram consideradas como negativas (LANTZ, 2015). A partir das matrizes, foram calculados os coeficientes de Jaccard e Yule para cada classe - ferrugem asiática, folha sã, palha e solo, em cada banco de imagens, conforme resultados apresentados na Tabela 2.

Tabela 2 - Classe melhor classificada considerando os coeficientes de Jaccard e Yule.

Classes	Coeficiente de Jaccard				Coeficiente de Yule			
	SVM		J48		SVM		J48	
	A	B	A	B	A	B	A	B
Palha	**0,990**	0,783	**0,976**	0,720	0,989	0,870	**0,967**	0,792
Solo	0,881	0,770	0,509	0,681	**0,995**	0,868	0,697	0,782
Folha Sã	0,788	0,902	0,679	0,855	0,855	**0,954**	0,777	**0,884**
Ferrugem	0,893	**0,953**	0,836	**0,932**	0,931	0,928	0,880	0,874

Os resultados fornecidos através do coeficiente de Jaccard indicaram que tanto o classificador SVM quanto o J48 obtiveram melhores resultados com a classe Palha para o banco de imagens A, e com a classe Ferrugem, para o banco de imagens B. Com os resultados fornecidos através do coeficiente de Yule verificou-se que para o banco de imagens A, os melhores resultados foram com a classe Solo (classificador SVM), e com a classe Palha (classificador J48), enquanto que com o banco de imagens B, o melhor resultado foi obtido na classe Folha Sã, tanto com o classificador SVM quanto com o J48.

O desempenho de cada classificador foi testado através das métricas PCC, Medida-F, área sob a Curva ROC e Análise de concordância Kappa, e o resultado está demonstrado na Tabela 3. Com a métrica PCC, verificou-se que a maior porcentagem de imagens corretamente classificadas como parte das classes ocorreu com o classificador SVM no que se refere ao banco de imagens A e, com o classificador Florestas Aleatórias (FA) no banco de imagens B.

Na utilização da métrica Medida-F, optou-se pelo uso da macro Medida-F, por proporcionar um peso igual a cada classe (ASCH, 2013). Assim, verificou-se que com o banco de imagens A, o melhor desempenho foi obtido com o classificador SVM, enquanto para o banco de imagens B, os classificadores SVM e Florestas Aleatórias obtiverem igual desempenho.

As Curvas de Características de Operação do Receptor (Curvas ROC Receiver Operating Characteristic) é uma representação gráfica da taxa de verdadeiros positivos versus taxa de falsos positivos. O desempenho entre os dois bancos de imagens, avaliado através Área sob a Curva ROC foi o mesmo utilizando o classificador Florestas Aleatórias. Na avaliação com a métrica Kappa verificou-se que tanto no banco de imagens A quanto no B a concordância é quase perfeita.

Tabela 3 - Desempenho dos classificadores com o DIAGSOJA

Métricas	Banco de Imagens	Classificadores			
		SVM	KNN	J48	FA
PCC	A	**97,06±1,29**	91,17±2,30	94,12±1,90	95,87±1,90
	B	95,07±2,43	91,51±2,93	94,53±2,94	**95,38±2,06**
Medida-F	A	**97,00±0,01**	91,00±0,02	94,00±0,02	95,00±0,02
	B	**95,00±0,02**	91,00±0,03	94,00±0,03	**95,00±0,02**
Área sob a Curva ROC	A	99,00±0,01	97,00±0,02	99,00±0,01	**100±0,00**
	B	97,00±0,05	81,00±0,11	94,00±0,07	**100±0,01**
Kappa	A	**95,00±0,02**	84,00±0,04	89,00±0,03	92,00±0,03
	B	**89,00±0,05**	82,00±0,06	88,00±0,06	**89,00±0,05**

Para verificar se houve diferença significativa entre os classificadores foi aplicada a Análise de Variância para os dois bancos de imagens. Considerando os resultados valor p=0,31542 e 0,33745, a um nível de significância de 0,05%, pode-se afirmar que não existem evidências estatísticas de que os desempenhos sejam diferentes entre os classificadores. Automatizar tanto quanto possível o processo de reconhecimento de doenças é uma

necessidade para os produtores já que, a depender da patologia que atinge a plantação, uma demora no diagnóstico pode comprometer o ganho do produtor (AFRIDI et al., 2014). /

CONCLUSÕES

Considerando a importância da soja, bem como o potencial produtivo que não é alcançado devido, entre outros fatores, ao impacto de doenças, verificou-se com essa pesquisa que a utilização de um sensor para aquisição de imagens, acoplado aos VANTs, permite agilidade na aquisição de imagens e monitoramento de lavouras produtivas. Esse fato, somado a atuação do programa desenvolvido para reconhecimento das patologias que foi implementado e avaliados por meio de várias métricas, permite um bom desempenho no reconhecimento da ferrugem asiática. Os resultados obtidos junto aos classificadores demonstraram que a proposta, bem como sua utilização, pode atuar como nova alternativa para o reconhecimento dos sintomas da doença nas folhas de soja.

O emprego de aprendizagem automática e visão computacional como base para programas de reconhecimento tem se mostrado de grande valia na identificação de focos de pragas ou doenças. Nesse sentido o experimento demonstrou que é possível detectar sintomas de doenças foliares em plantações de soja utilizando imagens obtidas por VANTs e, os testes realizados utilizando o programa de computador para reconhecimento de doenças demonstraram realizar distinção entre as classes conforme os resultados estatísticos já apresentados.

O software encontra-se em fase de teste e deve contar com ajustes. Outros fatores poderão proporcionar melhores resultados, como testes com os parâmetros a serem estabelecidos pelo usuário na geração de arquivo para análise. Além disso, é possível que a utilização de um banco de imagens mais robusto possa proporcionar melhores desempenhos. Outro fator limitante para a análise foi o surgimento de apenas uma doença no experimento, quando doenças com características diferentes são relevantes para avaliar o desempenho do programa. Além disso, o fato de a doença ter como uma de suas características seu início na parte baixa da planta pode dificultar o reconhecimento da doença no seu estágio inicial. Nesse sentido, a busca por uma forma de obtenção de imagens da parte baixa da planta pode favorecer o reconhecimento dos sintomas das doenças no estágio inicial, onde é possível ao produtor entrar com o uso de defensivos e reduzir o prejuízo.

AGRADECIMENTOS

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001, da NVIDIA Corporation pela doação da placa gráfica e da Fundação de Apoio ao Desenvolvimento do Ensino, Ciência e Tecnologia do Estado de Mato Grosso do Sul (FUNDECT).

REFERÊNCIAS

ACHANTA, R.; SMITH, K.; LUCCHI, A.; FUA, P.; and SUSSTRUNK, S.. SLIC superpixels. Technical report, EPFL, Tech.Rep. 149300, 2010.

AFRIDI, M. J.; LIU, X.; MITCHELL MCGRATH, J.. "An Automated System for Plant-level Disease Rating in Real Fields," in Proceedings of the 22nd International Conference on Pattern Recognition (ICPR 2014), Stockholm, Sweden, August 24-28, 2014.

ASCH, V. V.. “Macro- and micro-averaged evaluation measures,” Tech. Rep., University of Antwerp, 2013.

EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA - EMBRAPA SOJA. Tecnologias de produção de soja – região central do Brasil 2014. - Londrina: Embrapa Soja, 2013.

HONKAVAARA, E.; SAARI, H.; KAIVOSOJA, J.; PÖLÖNEN, I.; HAKALA, T.; LITKEY, P.; MÄKYNEN, J.; PESONEN, L. Processing and Assessment of Spectrometric, Stereoscopic Imagery Collected Using a Lightweight UAV Spectral Camera for Precision Agriculture. Remote Sens. 2013, 5, 5006-5039.

JESUS, L.F.; SILVA, V.B.; ROCHA, F.G.. Uso de software para detecção de doenças na cultura da soja com o auxílio de um drone autônomo. Computer on the Beach 2015 -

LANTZ, B.. Machine Learning with R. Packt, 2° Edição, 2015.

LARA, A. C.. Segmentação de movimento usando morfologia matemática. Dissertação. Universidade de São Paulo, 2007

MEYER A. S.. Comparação de coeficientes de similaridade usados em análises de agrupamento com dados de marcadores moleculares dominantes. Dissertação de Mestrado. ESALQ, Universidade de São Paulo. Piracicaba SP. 2002.

MILANO, D. D.; HONORATO, L. B. Visão computacional. – Universidade Estadual de Campinas– Faculdade de Tecnologia. 2010.

MORAES, S.A. de Quantificação de doenças de plantas. 2007. Artigo em Hypertexto. Disponível em: <http://www.infobibos.com/Artigos/2007_1/doencas/index.htm>. Acesso em: 22/9/2016.

REZENDE, S. O.. Sistemas inteligentes: fundamentos e aplicações. Barueri, SP: Manole, 2003.

RUMPF, T.; MAHLEIN, A.K.; STEINER, U.; OERKE, E.C.; DEHNE, H.W.; PLÜMER, L.. Early detection and classification of plant diseases with Support Vector Machines based on hyperspectral reflectance. Comput. Electron. Agric., 74 (1) (2010), pp. 91–99.

SZELISKI, R..Computer Vision: Algorithms and Applications. Springer, 2010.

TORRES-SÁNCHEZ, J., LÓPEZ-GRANADOS, F., PEÑA, J. M. (2015). An automatic object-based method for optimal thresholding in UAV images: Application for vegetation detection in herbaceous crops. Computers and Electronics in Agriculture , 114, 43-52.

UGALE ,V.; GUPTA, D.. A Comprehensive Survey on Agricultural Image Processing. International Journal of Science and Research (IJSR). Volume 5 Issue 1, January 2016.

YORINORI, J. T.; CHARCHAR, M.J.D'A.; NASSER; L.C.B.; HENNING, A.A...Doenças da soja e seu controle. In: ARANTES, N.E.; SOUZA, P.I. de M. de., ed. Cultura de soja nos cerrados. Piracicaba: POTAFOS, 1993. p.333-397.

YORINORI, J. T.; GODOY, C. V.; PAIVA, W. M.; FREDERICK, R. D.; COSTAMILAN, L. N.; BERTAGNOLLI, P. F.; NUNES JUNIOR, J.. Evolução da ferrugem da soja (Phakopsora pachyrhizi) no Brasil, de 2001 a 2003. Fitopatologia Brasileira, Brasília, v. 28, p. S210, ago. 2003.



Espacialização de dados de temperatura do ar e soma térmica para o Rio Grande do Sul no contexto do manejo do Capim-Sudão BRS Estribo

Allan Sampaio Pires[1], *Gustavo Trentin*[1,2], *Alexandro Gularte Schäfer*[1]

[1]Programa de Pós-Graduação em Computação Aplicada, Universidade Federal do Pampa
Bagé, Rio Grande do Sul, Brasil
allanpires@unipampa.edu.br, alexandroschafer@unipampa.edu.br

[2]Embrapa Pecuária Sul
Bagé, Rio Grande do Sul, Brasil
gustavo.trentin@embrapa.br

RESUMO

A partir de uma parceria entre Embrapa e Sulpasto, foi desenvolvida uma espécie de forrageira anual de verão denominada Capim-Sudão BRS Estribo. Devido à origem recente, esta planta ainda carece de informações que relacionem sua produção com parâmetros agroclimáticas como a temperatura do ar e a soma térmica. O presente estudo teve como objetivo espacializar dados de temperatura do ar e de soma térmica relacionada ao período de produção do Capim-Sudão BRS Estribo, para o estado do Rio Grande do Sul. A metodologia tem como base a operação de bases digitais espacializadas com recursos de Sistemas de Informação Geográfica (SIG). Foram utilizados dados altimétricos em formato matricial oriundos da Shuttle Radar Topography Mission (SRTM) com resolução espacial de 90 metros e médias mensais de temperatura do ar compiladas para um período de trinta anos. A espacialização da temperatura foi feita com modelos de regressão, calculados mensalmente em relação à latitude e à altitude. Mais especificamente, utilizou-se a interpolação por tendência para ajuste dos dados da superfície terrestre e o método de mínimos quadrados ordinários para a determinação dos coeficientes de regressão das equações de regressão linear. A partir dos modelos de regressão, foram gerados mapas de média mensal de temperatura do ar e de soma térmica para todos os meses do ano. Por fim, foram gerados os mapas das somas térmicas acumuladas referentes ao período de produção do Capim-Sudão BRS Estribo no estado, que se dá entre os meses de outubro a abril.

PALAVRAS-CHAVE: Soma termica, Estatistica espacial, Sistemas de Informação Geográfica, Capim-Sudão.

ABSTRACT

From a partnership between Embrapa and Sulpasto, a species of annual summer forage called Capim-Sudão BRS Estribo has been developed. Due to the recent origin, this plant still lacks information that relates its production with agroclimatic parameters such as the air temperature and the thermal sum. The present study aimed to spatialize air temperature and thermal sum data related to the production period of the Capim-Sudão BRS Estribo for the state of Rio Grande do Sul. The methodology is based on the operation of digital bases with resources of Geographic Information Systems (GIS). It was used data in raster format from Shuttle Radar Topography Mission (SRTM) with spatial resolution of 90 meters and monthly averages of air temperature compiled for a period of thirty years. The spatialization of the temperature was done with regression models, calculated monthly in relation to latitude and altitude. More specifically, was used the interpolation by tendency to adjust the land surface data and the ordinary least squares method for the determination of the regression coefficients of the linear regression equations. From the regression models, were generated maps of monthly average air temperature and thermal sum for all months of the year. Finally, the maps of accumulated thermal sums were generated for the period of production of Capim-Sudão BRS Estribo in Rio Grande do Sul, which occurs between October and April.

KEYWORDS: Thermal sum, Spatial statistics, Geographic Information Systems, Capim-Sudão.

INTRODUÇÃO

A partir de uma parceria entre Embrapa e Sulpasto, foi desenvolvido uma espécie de forrageira anual de verão denominada Capim-Sudão BRS Estribo (SILVEIRA et al., 2015). Estudos mostram que esta forrageira proporciona um impacto bioeconômico positivo na produção (EMBRAPA, 2017).

O desenvolvimento do Capim-Sudão BRS Estribo está diretamente relacionado a fatores agroclimáticos, dentre eles a temperatura do ar. A forrageira desenvolve-se com temperaturas do ar entre 11°C e 34°C, considerando-se 30°C como o ideal para o seu desenvolvimento e crescimento. Quando exposta a temperaturas críticas, próximas ao ponto de congelamento, ocorrerão danos foliares letais e, de acordo com a idade dos tecidos, tempo de exposição e condições nutricionais, o ciclo da planta poderá ser finalizado. O encerramento do ciclo também ocorrerá em casos de temperaturas inferiores a 8°C e rebaixamento do pasto inferiores a 15 cm. Em plantas com mais de 15cm, entretanto, que essa temperatura é letal para as folhas novas, porém a base da planta e as folhas antigas sobrevivem e mantêm crescimento razoável. Persistindo a temperatura crítica por dias consecutivos, observa-se grande comprometimento da planta, ocasionando assim, o fim do ciclo de utilização (SILVEIRA et al., 2015). Embora as temperaturas não correspondam a um fator limitante para o crescimento da cultivar (na ocorrência de boa disponibilidade hídrica) para as condições nacionais, a temperatura constitui-se fator de grande relevância no que tange à sua época de plantio e o seu manejo (RUELAS, 2018). Devido a origem

recente, esta planta ainda carece de informações que relacionem sua produção com parâmetros agroclimáticas como a temperatura e a soma térmica, que pode ser definida como a quantidade mínima de radiação solar necessária para o desenvolvimento do vegetal (BRUNINI et al., 1976).

Atualmente, os registros de temperatura do ar são pontuais, disponíveis apenas para os locais em que se realiza a sua medição, ou seja, nos locais em que existem estações climatológicas. Sabe-se que a temperatura do ar varia de acordo com a latitude e longitude, sendo influenciada ainda pela altitude, podendo sofrer um decréscimo de até 1 ° C a cada 180 metros de elevação na altitude (FRITZSONS; MANTOVANI; AGUIAR, 2008). Utilizando técnicas de estatística espacial, é possível constituir um modelo capaz de predizer valores de temperatura do ar para áreas carentes de dados. Um dos métodos utilizados faz uso de equações de regressão linear (ESRI, 2016), juntamente com o método de mínimos quadrados ordinários (*ordinary least squares*) para a determinação dos coeficientes e da constante de regressão (KALKHAN, 2011). No presente estudo, fez-se uso do método dos mínimos quadrados para espacializar as médias mensais de temperatura do ar e da soma térmica no estado do Rio Grande do Sul, visando delimitar regiões de maior e menor aptidão para o cultivo do Capim-Sudão BRS Estribo.

MATERIAL E MÉTODOS

Para a espacialização das médias mensais de temperaturas do ar foram obtidos dados altimétricos em formato matricial do estado do Rio Grande do Sul, disponibilizados pela Embrapa Monitoramento por Satélite (MIRANDA, 2005). Realizou-se o download das cenas do satélite SRTM que abrangem o estado e logo após confeccionou-se o mosaico de imagens, que posteriormente foi convertido para SIRGAS 2000 (Sistema de Referência Geocêntrico para as Américas), que é o referencial geodésico adotado para as Américas (FORTES, 2003).

Tratamento dos dados

Os dados meteorológicos das médias mensais de temperatura foram obtidos junto ao Atlas Climático da Região Sul do Brasil (WREGE et al., 2012), calculados entre os anos de 1976 e 2005. Estes valores foram obtidos a partir do registro de estações meteorológicas completas, instaladas no Rio Grande do Sul, Santa Catarina e Uruguai. Os dados foram tratados, colocando-os em arquivos, de forma que ficassem dispostos em tabelas contendo o nome das localidades, a latitude, a longitude e a altitude de cada estação, além dos valores das médias mensais das temperaturas de janeiro a dezembro. A partir da tabela, foi criado um arquivo, com estes dados, em formato shapefile para a utilização nos cálculos de estimativa dos coeficientes de regressão que posteriormente foram utilizados nas equações que dão origem aos mapas de médias mensais das temperaturas do ar e de soma térmica para o Rio Grande do Sul.

Espacialização dos dados

Para a espacialização dos dados foi utilizada a ferramenta de análise espacial (*Spatial Analyst Tools* do software ArcGIS. Este software pode ser considerado um software de Sistemas de Informações Geográficas (*Geographic Information System – GIS*), que possibilita a análise de relações espaciais, a detecção de padrões, a predição de variáveis, além de facilitar a criação de mapas com os resultados do processo de modelagem.

Foi utilizado o método de interpolação por tendência para a espacialização dos dados. Esta técnica utiliza uma interpolação polinomial global que se ajusta a uma superfície definida por uma função matemática polinomial para os pontos amostrais. Esta superfície muda gradualmente e captura padrões de escala nos dados. Este processo resulta em uma superfície suave que representa tendências graduais na superfície sobre a área de interesse. É utilizada geralmente para gerar uma superfície a partir dos pontos amostrais quando esta varia gradualmente de região para região na área de interesse. Este processo serve para criar uma superfície gradualmente variável usando polinômios de baixa ordem para descrição de um processo físico (ESRI, 2016).

Para a criação das imagens de espacialização da temperatura do ar e da soma térmica, montou-se as equações de regressão linear múltipla, utilizando os coeficientes de regressão estimados através do método de mínimos quadrados ordinários (*ordinary least squares* ou OLS). Esta é uma técnica de otimização matemática que busca a aproximação máxima dos valores preditos com os reais, tentando minimizar a soma dos quadrados das diferenças das ordenadas (resíduos) entre os pontos gerados pela função e os pontos correspondentes nos dados (KALKHAN, 2011). Foram definidas a constante de regressão (β_0), os coeficientes de regressão relativos à latitude (β_1), à longitude (β_2) e à altitude (β_3), assim como o coeficiente de determinação (R^2). Esse coeficiente é a razão entre a variação explicada e a variação total, ou seja, é interpretado como a fração da variação da amostra em y (variável dependente) que é explicada por x (variável independente). (WOOLDRIDGE, 2015). Para a utilização das equações de regressão linear foi utilizado a ferramenta de estatística espacial (*Spatial Statistics Tools*) do ArcMAP. Com isso montou-se a seguinte fórmula:

$$T = \beta_0 + \beta_1 * Latitude + \beta_2 * Longitude + \beta_3 * Altitude$$

Em que:

- T: temperatura média do ar em graus (° C), sendo esta a variável dependente;
- Latitude: variável independente;
- Longitude: variável independente;
- Altitude: variável independente;
- β_0: constante de regressão;
- β_1, β_2 e β_3: coeficientes de regressão.

A soma térmica média de cada estação foi multiplicada pelo número de dias de cada

mês para a obtenção dos totais acumulados nestes períodos. A soma térmica é quantidade de graus/dia necessário para o desenvolvimento de uma determinada planta, e no caso do capim-sudão BRS Estribo, este cálculo foi feito da seguinte forma:

$$ST = T_{\text{med}} - T_{\text{b}}$$

Onde:

- ST: Soma térmica;
- T_{med}: Temperatura média do ar diária;
- T_b: Temperatura base, que é o mínimo de temperatura que a planta necessita para seu desenvolvimento, sendo neste caso considerado 10° C.

As imagens das temperaturas médias mensais e da soma térmica mensal foram calculados através de álgebra de mapas. A aplicação desta técnica foi feita na calculadora Raster (*Raster Calculator*) do ArcGIS. Esta ferramenta executa uma expressão algébrica de linha única usando operadores matemáticos simples, semelhante a de uma calculadora. As 12 equações de regressão, referentes aos 12 meses do ano, foram inseridas na calculadora raster. A partir daí foram calculadas as temperaturas médias do ar para o estado do Rio Grande do Sul, gerando imagens em formato matricial. As equações de regressão linear das temperaturas médias mensais e das somas térmicas mensais podem ser observadas respectivamente nas tabelas 1 e 2.

Tabela 1: Equações de regressão linear para temperaturas médias mensais

Mês	Equação de regressão linear
Jan	15.651072 + 0.452775*LAT - 0.425876*LON - 0.00458*ALT
Fev	22.203453 + 0.382232*LAT - 0.253723*LON - 0.004495*ALT
Mar	24.107957 + 0.402506*LAT - 0.206114*LON - 0.004493*ALT
Abr	28.687055 + 0.456181*LAT - 0.091106*LON - 0.004402*ALT
Mai	30.113801 + 0.47312*LAT - 0.013358*LON - 0.004407*ALT
Jun	30.777881 + 0.547631*LAT + 0.00184*LON - 0.003694*ALT
Jul	29.661391 + 0.598978*LAT - 0.041948*LON - 0.003649*ALT
Ago	26.288945 + 0.674324*LAT - 0.172104*LON - 0.003388*ALT
Set	27.917449 + 0.701167*LAT - 0.17986*LON - 0.003748*ALT
Out	21.735562 + 0.648047*LAT - 0.317165*LON - 0.003871*ALT
Nov	19.899617 + 0.650309 * LAT - 0.392551*LON - 0.004039*ALT
Dez	15.680551 + 0.58329*LAT - 0.475674*LON - 0.004261*ALT

Os período entre os meses de outubro e fevereiro são considerados os melhores para o plantio do capim-sudão BRS Estribo no estado do Rio Grande do Sul (SILVEIRA et al., 2015), por isso foram calculadas as somas térmicas acumuladas nos períodos entre outubro a abril, novembro a abril, dezembro a abril, janeiro a abril e fevereiro abril. Para a obtenção dos coeficientes de regressão, da constante de regressão, a elaboração da equação de regressão linear e a obtenção dos mapas, foram utilizados os mesmos procedimentos utilizados para as médias mensais de temperatura do ar. As equações de regressão linear podem ser observadas na tabela 3.

Tabela 2: Equações de regressão linear para soma térmica mensal

Mês	Equação de regressão linear
Jan	175.183231 + 14.036027*LAT - 13.202154*LON - 0.141982*ALT
Fev	341.696685 + 10.702502*LAT - 7.104234*LON - 0.125865*ALT
Mar	437.346681 + 12.477688*LAT - 6.389543*LON - 0.139278*ALT
Abr	560.611655 + 13.685418*LAT - 2.733191*LON - 0.132058*ALT
Mai	623.527832 + 14.666723*LAT - 0.414109*LON - 0.136606*ALT
Jun	617.447551 + 15.729195*LAT + 0.32672*LON - 0.109604*ALT
Jul	600.484753 + 17.43216*LAT - 0.844009*LON - 0.110536*ALT
Ago	503.884956 + 20.799818*LAT - 5.30017*LON - 0.105057*ALT
Set	537.523476 + 21.03502*LAT- 5.395804* - 0.11244*ALT
Out	363.80242 + 20.089448*LAT - 9.832119*LON - 0.120014*ALT
Nov	296.988512 + 19.509278*LAT - 11.776523*LON - 0.121173*ALT
Dez	176.097081 + 18.081983*LAT - 14.745884*LON - 0.132078*ALT

Tabela 3: Equações de regressão linear para soma térmica acumulada por períodos

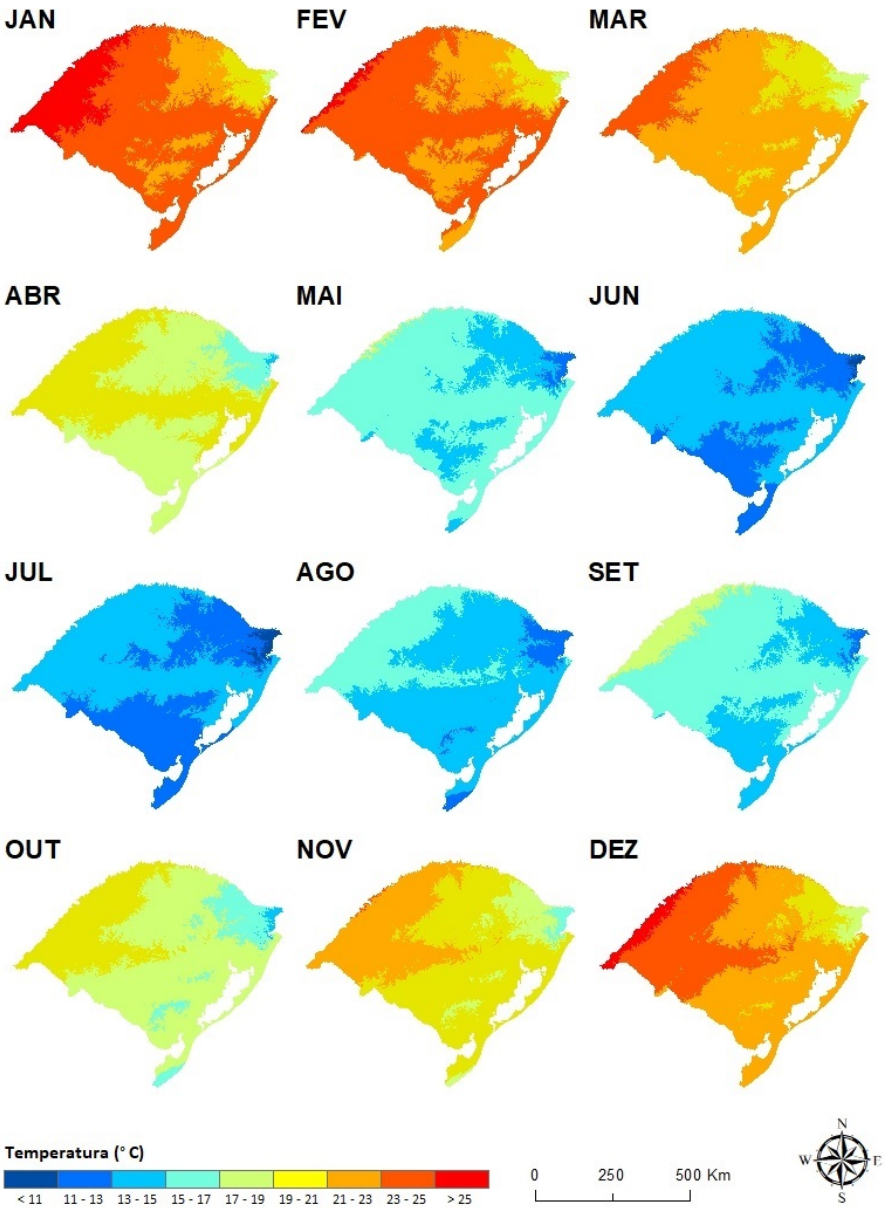
Mês	Equação de regressão linear
Out - Abr	2351.726265 + 108.582344*LAT -65.783647*LON - 0.912448*ALT
Nov - Abr	1987.923845 + 88.492897*LAT - 55.951529*LON - 0.792434*ALT
Dez - Abr	1690.935333 + 68.983618*LAT - 44.175006*LON - 0.671261*ALT
Jan - Abr	1514.838252 + 50.901635*LAT - 29.429122*LON - 0.539183*ALT
Fev - Abr	1339.655021 + 36.865608*LAT - 16.226969*LON - 0.397201*ALT

RESULTADOS E DISCUSSÃO

As médias mensais de temperatura do ar espacializadas para o estado do Rio Grande do Sul são apresentadas na Figura 1.A menor média de temperatura foi 9,55° C observada no mês de julho e a maior média foi de 26,48° C no mês de janeiro. Nota-se que as regiões mais quentes no verão encontram-se principalmente no noroeste do estado, na divisa com a Argentina, enquanto as médias mais baixas do inverno localizavam-se em regiões como a campanha gaúcha, extremo sul e localidades de altitude elevada localizadas próximas ao estado de Santa Catarina. Constatou-se também que o mês de dezembro foi o que apresentou maior variação de temperatura entre as regiões.

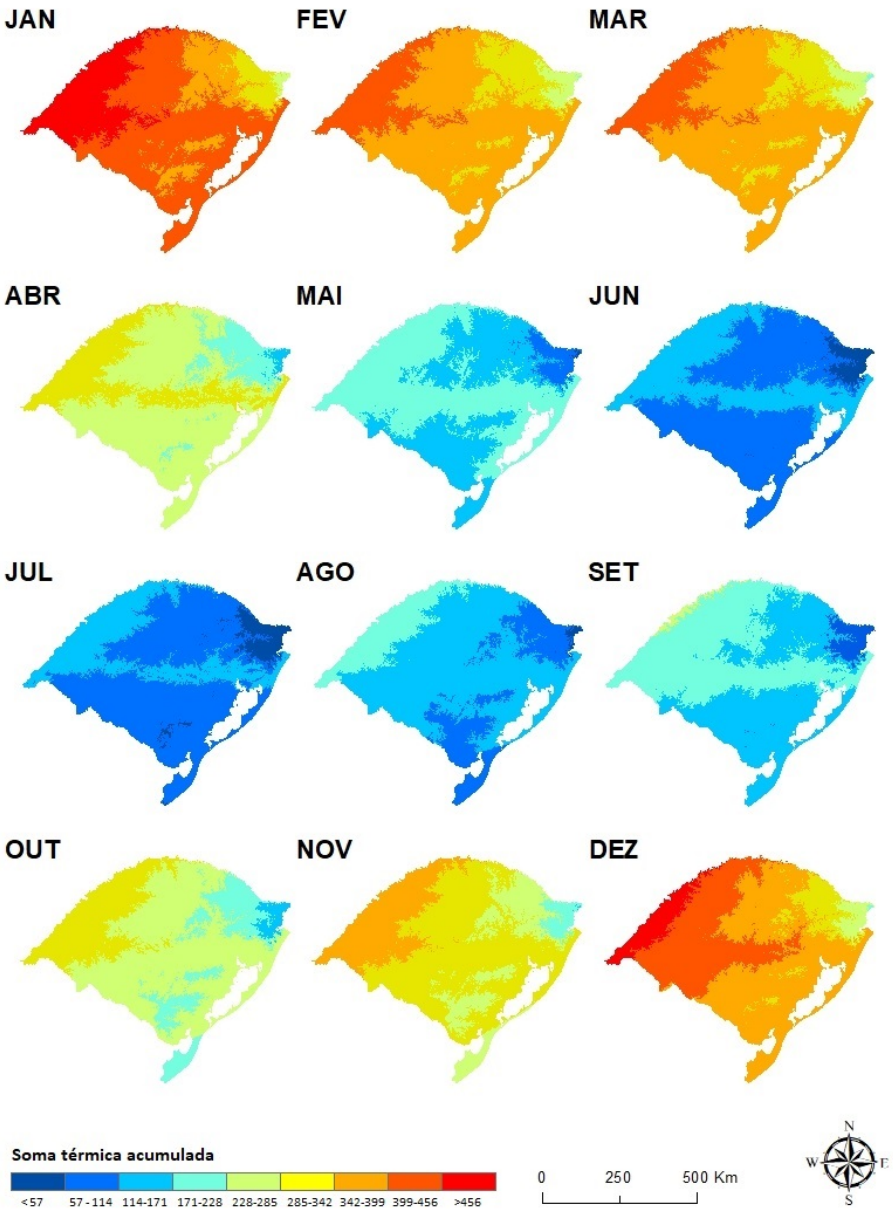
Apresenta-se, na Figura 2, os resultados da aplicação da álgebra de mapas e da reclassificação da espacialização dos dados da soma térmica acumulada de cada mês. O valor máximo de 511,15 graus-dia no mês de janeiro, na região noroeste, na fronteira com a Argentina. No mês de julho, algumas regiões apresentaram temperatura base abaixo do mínimo para o desenvolvimento da planta e a quantidade de graus-dia foi considerado como nulo.

Figura 1: Temperatura média mensal do ar no RS



Fonte (O autor)

Figura 2: Soma térmica mensal no RS

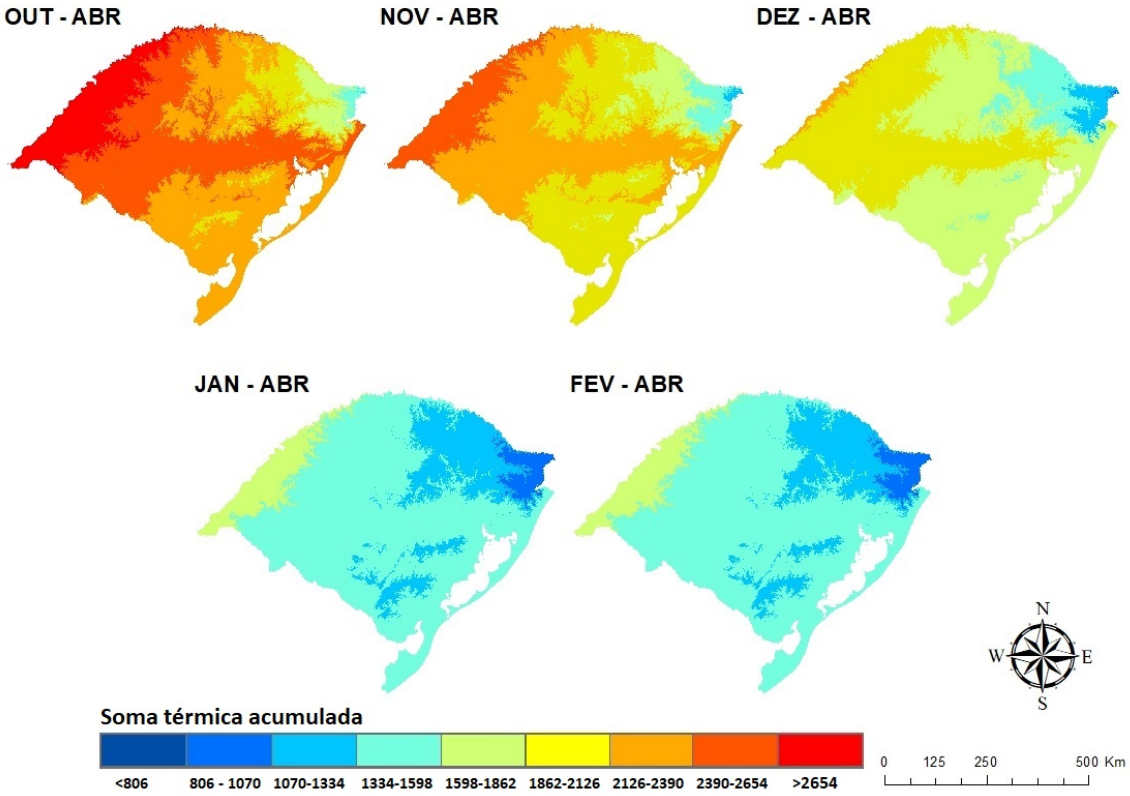


Fonte (O autor)

Por fim, apresenta-se na figura Figura 3 as imagens resultantes da álgebra de mapas e da reclassificação das somas térmicas acumuladas nos períodos de produção do capim-sudão BRS Estribo na região, que se dá entre os meses de outubro a abril.

Foram somados os valores nos períodos compreendidos entre outubro e abril, novembro e abril, dezembro e abril, janeiro e abril e fevereiro e abril. No maior período, compreendido entre os meses de outubro a abril, a soma térmica acumulada foi de 2921,11 graus-dia, na região noroeste, na divisa com a Argentina, e o menor valor acumulado foi de 1256,55 graus-dia, na região próxima a Torres. No menor período considerado, entre os meses de fevereiro a abril, as máximas e mínimas foram respectivamente 1190 e 542,59 graus-dia, nas mesmas regiões citadas anteriormente.

Figura 3: Soma térmica acumulada



Fonte (O autor)

CONCLUSÕES

Apresentou-se neste artigo os procedimentos e resultados da espacialização dos registros de médias mensais de temperatura do ar e de soma térmica mensal para o estado do Rio Grande do Sul. Levando em consideração o período de produção do Capím-Sudão BRS Estribo, foram gerados mapas da soma térmica acumulada para períodos entre outubro e abril. Este trabalho faz parte de um estudo mais amplo, em que nas proximas etapas se investigará o cultivo do capim-sudão BRS Estribo com relação à temperatura do ar para toda a região sul do Brasil com o objetivo de contribuir com a definição de zonas de melhor aptidão para o seu cultivo e de técnicas de manejo apropriadas. A utilização de sistemas de informação geográficas foi de grande relevância, possibilitando a aplicação de técnicas de estatística espacial e a geração de mapas com os dados espacializados.

REFERÊNCIAS

BRUNINI, O. et al. Temperatura-base para alface cultivar"white boston", em um sistema de unidades térmicas. *Bragantia*, SciELO Brasil, v. 35, n. 1, p. 213–219, 1976.

EMBRAPA. *Impacto bioeconômico do capim-sudão BRS Estribo na produção de bovinos de corte*. [S.l.], 2017. (18m31s). Disponível em: ¡https://youtu.be/noXoQ1flRyM¿. Acesso em: 18 mai. 2019.

ESRI. *ArcGIS Desktop*. [S.l.], 2016. (Disponível em: ¡http://desktop.arcgis.com/en/arcmap/¿. Acesso em: 18 jun. 2019.

FORTES, L. P. S. O sistema de referência sirgas2000. In: *XXI Congresso Brasileiro de Cartografia, Belo Horizonte, MG*. [S.l.: s.n.], 2003.

FRITZSONS, E.; MANTOVANI, L. E.; AGUIAR, A. V. de. Relação entre altitude e temperatura: uma contribuição ao zoneamento climático no estado do paraná. *Revista de Estudos Ambientais*, v. 10, n. 1, p. 49–64, 2008.

KALKHAN, M. A. *Spatial statistics: geospatial information modeling and thematic mapping*. [S.l.]: CRC press, 2011.

MIRANDA, E. d. Brasil em relevo. *Campinas: Embrapa Monitoramento por Satélite*, 2005. Disponível em: ¡http://www.relevobr.cnpm.embrapa.br¿. Acesso em: 19 Mai. 2019.

RUELAS, J. L. J. Modelagem do parâmetro de irrigação: um estudo de caso para o capim sudão brs estribo. Universidade Federal do Pampa, 2018.

SILVEIRA, M. da et al. Aspectos relativos à implantação e manejo de capim-sudão brs estribo. *Embrapa Pecuária Sul-Documentos (INFOTECA-E)*, Bagé: Embrapa Pecuária Sul, 2015., 2015.

WOOLDRIDGE, J. M. *Introductory econometrics: A modern approach*. [S.l.]: Nelson Education, 2015.

WREGE, M. S. et al. *Atlas climático da região sul do Brasil: estados do Paraná, Santa Catarina e Rio Grande do Sul.* [S.l.]: Pelotas: Embrapa Clima Temperado; Colombo: Embrapa Florestas, 2012., 2012.

Modeling technology diffusion of beef cattle in Brazil using a cellular automata

Gustavo Gianotto Bratfisch[1], Luis Gustavo Barioni[2], Rafael de Oliveira Silva[3],Célia Regina Grego[4],Bruno Henrique Pereira[5],Jair Bertolucci Junior[6]

[1] Universidade Estadual de Campinas, Campinas, São Paulo, Brasil, gustavo@bratfisch.br

[2] Departamento de Modelagem Agroambiental, EMBRAPA, Campinas, Brasil, luis.barioni@embrapa.br

[3] Global Academy of Agriculture and Food Securit, The university of Edinburgh, Edinburgh, Scotland, mathros@gmail.com

[4] EMBRAPA, Campinas, Brasil, celia.grego@embrapa.br

[5] Departamento de Modelagem Agroambiental, EMBRAPA, Campinas, Brasil, zehwallace@gmail.com

[6] Departamento de Modelagem Agroambiental, EMBRAPA, Campinas, Brasil, jairbortolucci@hotmail.com

RESUMO

Introduzimos um modelo de autômato celular representando a difusão da intensificação do gado de corte considerando a preferencia do fazendeiro. Nos também analisamos as propriedades teóricas do modelo e a sensibilidade dos principais parâmetros. O modelo representa a dinâmica de adoção de dois sistemas de manejos, intensivo (ϕ) e extensivo(ξ). Foi calculado a probabilidade do estado mudar como uma função de proporção de vizinhos em cada estado e a preferencia relativa por ϕ ou ξ. O modelo inclui o efeito da inercia (I), isso é, a resistência (ou dificuldade) inerente do produtor de mudar o seu estado dado um pedaço de terra, e três tipos diferentes de comportamento de adoção (AB). O primeiro (AB_1) esta associado aos primeiros que adotam que são exclusivamente guiados pela sua percepção de vantagem a mudança tecnológica; o segundo (AB_2), esta relacionado com a aversão a riscos, por isso provável que a mudança de estado é voltado para os sistemas com maior frequência na vizinhança

(influenciadores), i.e "comportamento de manada" e o terceiro (AB_3) é o comportamento que considera o efeito interativo da percepção de vantagem e dos influenciadores.

Para fornecer informações para o uso e calibração, executamos uma análise de sensitiva dos parâmetros de velocidade de adoção, o estado de equilíbrio e padrões espaciais. A análise revelou que o nível de agregação espacial esta relacionado com os valores relativos do conjunto de parâmetros.

ABSTRACT

We introduce a cellular automata model representing the diffusion of technological intensification of beef cattle considering farmer's preferences. We also analyze the theoretical properties of the model and the sensitivity of key parameters. The model represents the dynamics of adoption of two types of management systems, intensive (ϕ) and extensive (ξ). We compute the probability of state change as a function of the proportion of neighbors in each state and the relative preference for ϕ or ξ. The model includes the effect of inertia (I), that is, the inherent resistance (or difficulty) of a producer to change the state of a given plot of land, and three different types of adoption behaviors (AB). The first (AB_1) is associated to "early adopters" that are exclusively guided by their perception of advantages of technological change; the second (AB_2), is related to high risk aversion, so most probable change of state is towards the systems with the highest frequency in the neighborhood (influencers), i.e. "herd behavior"; and the third (AB_3) is a behavior that considers the interactive effect of advantage perception and influencers.

To give insights for use and calibration, we ran sensitivity analysis of parameter values of speed of adoption, equilibrium state, and spatial state patterns.. The analysis revealed the level of spatial aggregation is associated to the relative values of the parameters set.

INTRODUCTION

Brazil is a major global agricultural player and has an ambition to balance growing production with environmental costs (de Oliveira Silva et al. 2018). Particularly beef cattle production has long been linked the loss of biodiversity and greenhouse gas (GHG) emissions, mainly because of methane (CH_4) emitted by the around 200 million heads (M hd) via enteric fermentation process and associated natural conversion for grazing pastures. In the past, 1975 to 1996, beef cattle production was based on extensive grazing, characterized by low input management and low-quality pasture with limited carrying capacities. Along that period, increases in production were highly correlated with pasture expansion (Martha Jr 2012).

However, at least since 1996 the production pattern has changed and beef cattle has embarked on an intensification era, and production growth is now explained by gains in productivity of both pasture and animal performance. Animal growth rates are due to improved genetics through breeding and the use of on-pasture supplements and feedlot systems. Pasture productivity gains are due to better management, via pasture improvement/restoration by chemical and mechanical treatment in the soil. As opposed to extensive systems, these productivity measures characterize intensive systems. Despite Brazil has already embarked on an intensification era extensive cattle ranching, and underutilized pastures remain representative in many regions of the country, meaning current average stocking rates are far relatively low, around 1.1 head/ha (IBGE, 2015).

To increase beef cattle productivity, the government supports the adoption of intensification measures. Those include the restoration of degraded pasture (de Oliveira Silva et al. 2015, 2017; Negra et al. 2014). Research show that intensified system is more profitable and more sustainable, as GHG emissions per kg of beef is reduced (de Oliveira Silva et al. 2017; De Oliveira Silva et al. 2018, 2016). However, there are several barriers of technology adoption, including the lack of technical information and support, skilled labour, access to rural credit, land registration issues and a risk aversion by behaviour by famers (Strassburg et al. 2014). Optimization approaches usually assume a rational decision maker, e.g., adoption of intensification technology is a function of economic return and/or risk alone. These methods have been applied to understand economic return and environmental impacts of intensification of livestock systems in Brazil (Cohn et al. 2014) but they are not able to explore factors other than profit maximization inherent to the decision process of technology adoption. Thus, understanding the influence of geographical specificities, farmers preferences and networking on the adoption of intensification measures is essential to complement traditional approach of maximization of economic returns. Addressing the barriers of adoption and technology diffusion is essential to support governmental policies targeting reduction of deforestation via the sustainable intensification of livestock in Brazil (MOZZER et al 2011, 2015). Unfortunately, applying technology diffusion models to the intensification of livestock production in Brazil is a highly underexplored, most of the time is used with deforestation (KAIMOWITZ ,2008) or with agriculture (Gil et at 2015, 2016; CORTNER et at 2019). This work develops a spatially explicit model of technology diffusion using a cellular automata (CA) approach that introduces farmer's preferences in addition to the network effect. We use historical data to calibrate the model and apply it to explore the barriers and the intensification of beef cattle systems in Brazil.

METHODS

Cellular Automata modelling and implementation

A cellular automata (CA) is a discrete dynamical system, space, time, and the states of the system are discrete values. Each point in a regular spatial lattice, or cell is associated with a state within the a domain called alphabet. The states of the cells in the lattice are updated according to a local rule. That is, the state of a cell at a given time depends only on the previous of the cell and its neighbors.

A cellular automata is characterized by some fundamental properties**. The lattice of cells**, representing the structure or geometry of the grid of cells. The dimension of the lattice of cells can be one, two, or three dimensions in Euclidean space. **Discrete States,** a cell can assume only a finite number of possible states**. Local interactions,** a state of **a** cell can only be influenced by its neighbors**. Discrete dynamics**, every cell update at the same discrete time, according to a transition rule at the model.

In this paper, we use a two-dimensional lattice and the Moore-neighborhood (KRETZ, 2005), consisting of the central cell and the eight adjacent cells

We assume each cell is a plot of land which can assume three different states. The "inactive state" coded as "0" are cells not used for grazing, e.g. a mountain, a river, a city, cropland which are unavailable for state change. State "1" represents an extensive livestock plot; and state "2" represents an intensive livestock plot, where pastures are well managed, animals are supplemented and/or fedloted. The cells in the borders of the latice are assumed as inactives.

We use a transition function f to compute the probability of state change of a plot in time step t, (C_t) to the next one $C_{(t+1)}$:

$$C_{(t+1)} = f\left(C_t(i) \middle| i \in N(r)\right), \quad (1)$$

Where $N(r)$ is the set of neighbor cells of the cell r,

Changing a plot from extensive to intensive state is a "big decision" which demands significant investments, more demanding management and long-term consequences for the individual and the land. For that reason, individuals facing changes in type of management tend to be reluctant to change both because of giving up a possibly acceptable current state and the fear of failure (i.e. low perception of control). To account for the trend of the farmer to remain in the current

state, we used the consept of “inertia” for the transition rule. The transition rule is then given by:

$$f(N_\phi, E_\phi) = I_{min} + k\left(1 - e^{\left(-\lambda(\alpha * N_\phi + \beta * E_\phi + \gamma * N_\phi * E_\phi)\right)}\right) \quad (2)$$

$$\alpha + \beta + \gamma = 1 \quad (3)$$

$$k = \frac{I_{max} - I_{min}}{e^{-\lambda}} \quad (4)$$

Where, N_ϕ is the number of neighbors in the system ϕ $E_\phi$is the preference for the system ϕ, $I_{min}$is the minimal inertial probability; I_{max} is the maximum probability of keeping current state (1 is the default value); λ is a parameter used to calibrate the speed of convergence to the steady-state. The parameters $\alpha, \beta \; and \; \gamma$ (Eq. 3) are the associated weights of the neighborhood, preference and the crossproduct term, respectively. Thus the terms in the exponential represents the weight of each type of behavior in a given scenario.

In the case a cell has no neighbors in the same system, and the preference for that system is zero, $f(0,0) = I_{min}$. In the opporsite case, all the neighbors are in the same system and with maximum preference for the same system, $f(1,1) = I_{max}$. The parameter a (Eq. 4) is derived by combinging f(0,0) and f(1,1,).

Steady-state conditions

A deterministic cellular automata follows a pattern akin to the second law of thermodynamics: starting from a partially disordered state, the system evolves towards a state of equilibrium (TAATI , 2018) In our case, this happens when the flow of cells changing from state ϕ to ξ equals the flow of cells changing from ξ to ϕ. The flows are given by the number of cells in a given state multiplied by their probability of change; i .e.$1 - f(N_\phi, E_\phi)$. Therefore, we can find the equilibrium point through Equation 5.

$$N_\phi * \left(1 - f(N_\phi, E_\phi)\right) = N_\xi * \left(1 - f(N_\xi, E_\xi)\right) \quad (5)$$

$$N_\phi + N_\xi = 1 \quad (6)$$

$$E_\phi + E_\xi = 1 \quad (7)$$

Solving equation (5) analytically is impractible so we solve it numerically.

Evaluation of spatial aggregation

Spatial patterns are observed when a CA is simulated for an extended time (WOLFRAN, 1982), with different parameters. We analyse how each parameter affect the spatial dependence

patterns through a semivariogram, which shows the variance from data for separated by specific distance between points (ISAAKS; SRIVASTAVA 1989, VIERIA et al. 1983, VIERIA 2002). After making the curve adjustment in the semivariogram, we calculate the degree of aggregation (gc), given by:

$$gc = \left(\frac{C_1}{C_0 + C_1}\right) * 100 \quad (8)$$

Where C_0 is the nugget of the function, C_1 is the semivariogram effect and $C_0 + C_1$ is the asymptotic value of the function. According to Zimback et al. (2001), $gc > 75\%$ indicate strong spatial dependence; $25\% \geq gc \geq 75\%$ moderate spatial dependence and $gc < 25\%$ weak spatial dependence.

Sensitivity analysis

We evaluate the sensitivity of model outputs to the parameters $\gamma, \alpha, \beta \; and \; \lambda$ in in a 50 years model runs. We assumed fixed arbitrary values for the preference for the intensive system in $E_\phi = 0.7$, and the preference for the extensive system in $E_\xi = 1 - E_\phi$. We also assumed $I_{max} = 1$ and $I_{min} = 0.8$. We also compared steady-state values simulated and calculated with the equation (5).

RESULTS

Model results

Figure 2 shows the comparison of proportion in different set of parameters. When β is maximum, the model converges slightly faster as the model becomes closer to deterministic behavior and have the highest value of intensive proportion. We noticed λ is related with the proportion in each system, so higher λ higher the difference between the proportion of the management system, furthermore, shows that bigger λ longer it takes to reach the equilibrium

Figure 1 - a) Proportion of intensive System b) Proportion of extensive System

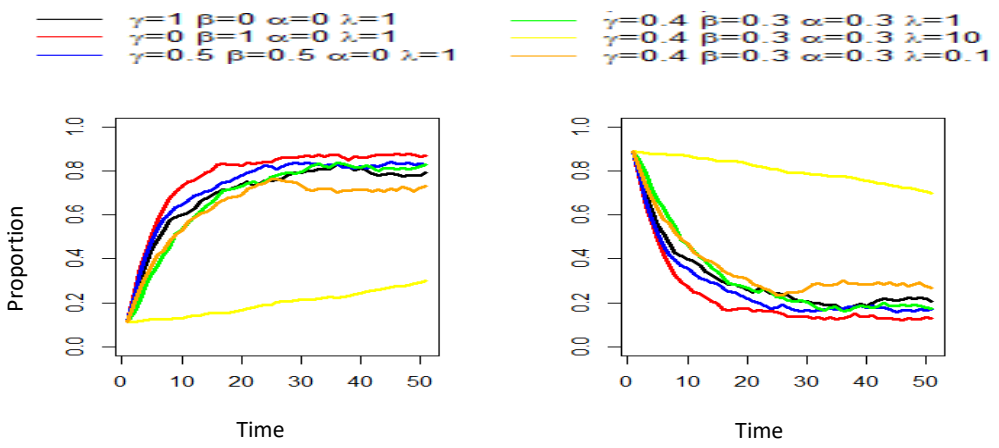


Figure 2 shows the dynamics of extensive and intensive management systems' adoption over time. The simulated value when reaches the equilibrium is approaching the calculated value, with this result we can use the equation (5) to calculate the value of equilibrium without doing the simulations. Also, equilibrium values agree with Figure 1 simulation, i.e., values around 0.2 and 0.8 for extensive an intensive. However, the smallest λ do not reaches equilibrium faster than the others.

Figure 2 - Simulation of -a)γ=1,β=0,α=0 e λ=1 b)γ=0,β=1,α=0 e λ=1 c)γ=0.5,β=0.5,α=0.0 e λ=1 d)γ=0.4,β=0.3,α=0.3 e λ=1 e)γ=0.4,β=0.3,α=0.3 e λ=3 f) e)γ=0.4,β=0.3,α=0.3 e λ=0.1

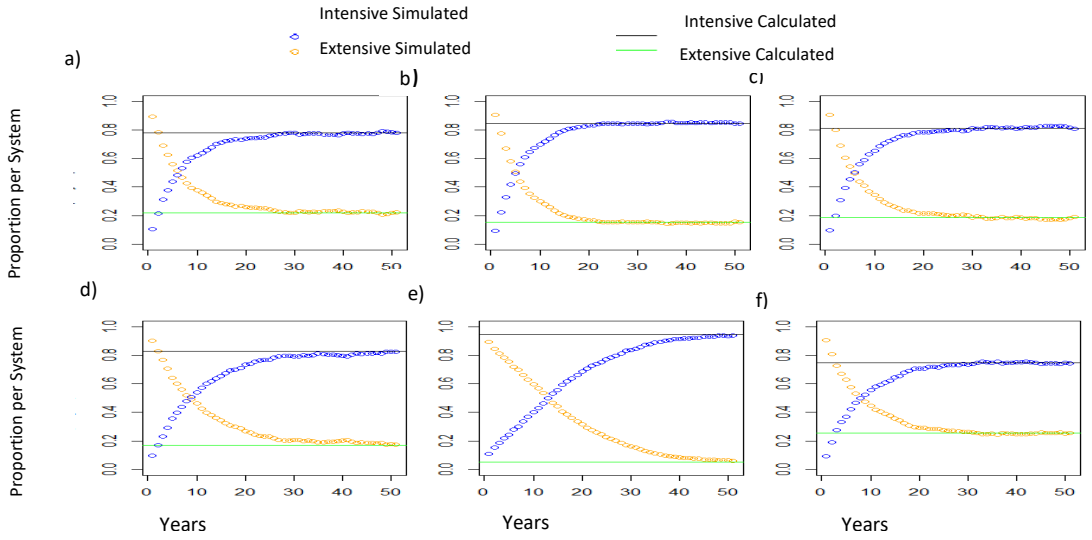
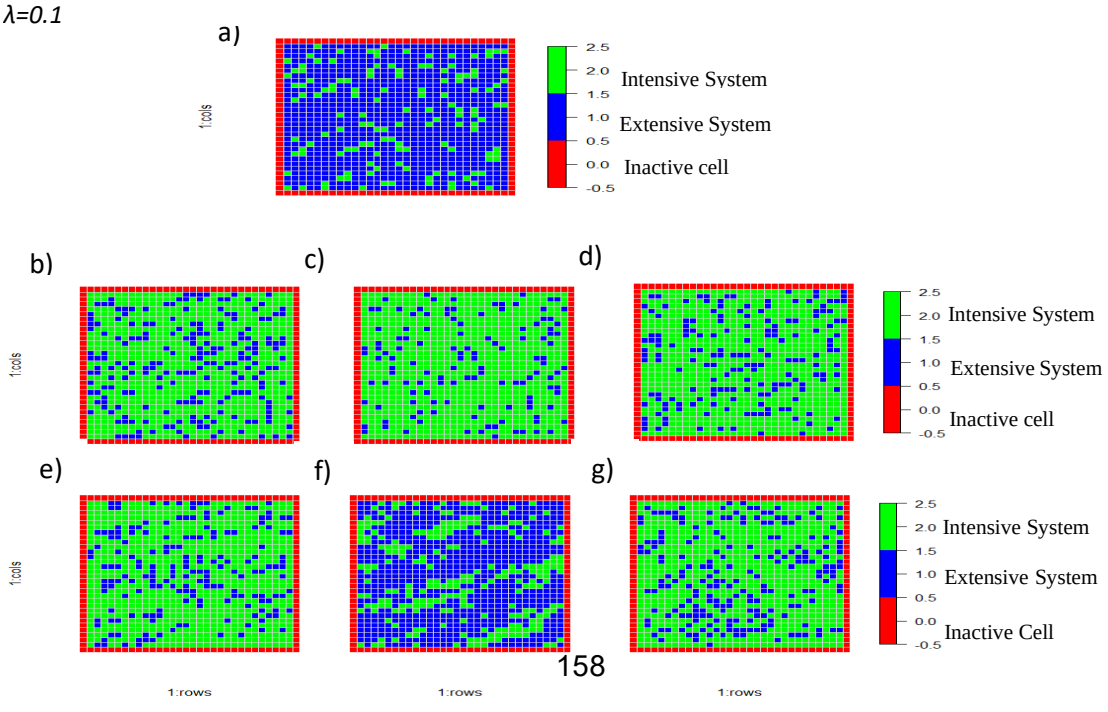


Figure 3 shows the spatial deposition of the plot of the land. We noticed that when β is high, the states are more randomly spread in space while a high α or γ tend to produce secure spatial aggregation as the influence of the neighborhood are taken into account in the automata. We can notice that λ is not related with spatial aggregation.

Figure 3– a) Initial Grid – Final Grid with the parameters b)γ=1,β=0,α=0 e λ=1 c)γ=0,β=1,α=0 e λ=1 d)γ=0.5,β=0.5,α=0.0 e λ=1 e)γ=0.4,β=0.3,α=0.3 e λ=1 f)γ=0.4,β=0.3,α=0.3 e λ=3 g))γ=0.4,β=0.3,α=0.3 e λ=0.1

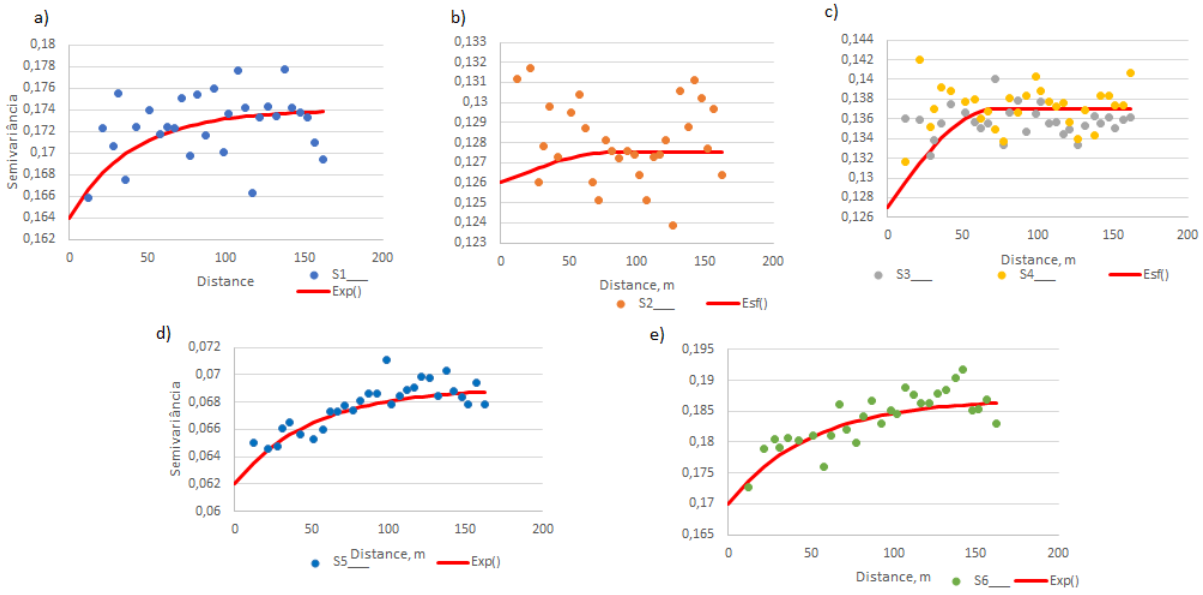


Geostatistics

The spatial aggregation differences shown in Figure 4 ware analyzed through the semivariogram technique for 6 different sets of parameters, $S1 = (\gamma = 1, \beta = 0, \alpha = 0 \text{ e } \lambda = 1)$, $S2 = (\gamma = 0, \beta = 1, \alpha = 0 \text{ e } \lambda = 1)$, $S3 = (\gamma = 0.5, \beta = 0.5, \alpha = 0 \text{ e } \lambda = 1)$, $S4 = (\gamma = 0.4, \beta = 0.3, \alpha = 0.3 \text{ e } \lambda = 1)$, $S5 = (\gamma = 0.4, \beta = 0.3, \alpha = 0.3 \text{ e } \lambda = 10)$,$S6 = (\gamma = 0.4, \beta = 0.3, \alpha = 0.3 \text{ e } \lambda = 0.1)$. For $S1, S4 \; and \; S5$ the exponential curve presented best fit while for S2 and S3 the spherical curve was best. As the dataset for $S3 \; and \; S4$ were similar, we combined the two in only one regression was made for them.

With the semivariogram plotted and using the equation (11), we have the following levels of aggregation degree: $gc(S1) = 5.74$, $gc(S2) = 1.17$, $gc(S3\_S4) = 7.29$, $gc(S5) = 10.11$, $gc(S6) = 9.09$, so $gc(S2)$ is the lowest value because β it will not carry any neighborhood influence, after that $gd(S1)$ do not have a high value because it is working only with γ that have some of the preference and neighborhood influence at the same time, besides this two value all the other stay close this happens when we added to the model some value to α.

Figure 4- a)Semivariogram S1 b)Semivariogram S2 c)Semivariogram S3 and S4 d)Semivariogram S5 e)Semivariogram S6



Conclusion

The cellular automata model presented herein contributes to the understanding of farmer behavior beyond profit maximization assumptions. By accounting for geographical proximity

of influencers it is able to produce different patterns of spatial dependence, depending on the parameters set. The model can also be calibrated for speed of adoption. Future modeling work will focus on empirically calibrate the model to different regions of Brazil and on determining the economic drivers and psychological influencers of the behavior defined by the parameter values.

AGRADECIMENTOS

Agradeço a Embrapa Informática Agropecuária e a Universidade de Edimburgo, pelo apoio a pesquisa.

REFERÊNCIAS

ILACHINSKI, A. Cellular Automata. A discrete universe. World Scientific; 2001

ADAMATZKY, A., 1994. Identification of Cellular Automata. Taylor & Francis, London.

ISAAKS, E.H. and Srivastava, R.M., 1989: An introduction to applied geostatistics. Oxford Univ. Press, New York, Oxford.

BRAZIL. 2015. **Intended Nationaly Determined Contribution (INDCs).** *Library of Congress*.http://www4.unfccc.int/submissions/INDC/Published Documents/Brazil/1/BRAZIL iNDC english FINAL.pdf.

MOZZER, G B. Agriculture and Cattle Raising in the Context of a Low Carbon Economy. *In* Motta, R.S. (Ed.) ***CLIMATE CHANGE IN BRAZIL: Economic, Social and Regulatory Aspects***. Brasília, DF: IPEA. 2011, Cap 4, p. 107 – 122. https://www.ipea.gov.br/agencia/images/stories/PDFs/livros/livros/livro_climatechange_ingles.pdf#page=108.

VIEIRA, S.R.; HATFIELD, J.L.; NIELSEN, D.R. & BIGGAR, J.W. Geostatistical theory and application to variability of some agronomical properties. Hilgardia, 51:1-75, 1983.

VIEIRA, S.R,; MILLETE, J.; TOPP, G.C.; REYNOLDS, W.D. Handbook for geostatistical analysis of variability in soil and climate data. In: ALVAREZZ, V.H.; SCHAEFER, C.R.; BARROS, N.F.; MELLO, J.W.V; COSTA, L.M. (Ed.). Tópicos em Ciência do solo. Viçosa: Sociedade Brasileira de Ciência do solo, 2002. vol.2, p.1-45, 2002.

ZIMBACK, C. R. L. Análise espacial de atributos químicos de solos para fins de mapeamento da fertilidade do solo. 2001. 114 f. Tese (Livre-Docência) - Faculdade de Ciências Agronômicas, Universidade Estadual Paulista, Botucatu, 2001

WOLFRAM, S. Cellular automata as simple-organizing systems. Technical report. Caltech preprint CALT. Passadena California, 1982, p 68-938.

IBGE (2015) *Censo Agropecuário 2006* (2006 Agricultural Census), accessed 5 June 2015; http://www.sidra.ibge.gov.br/bda/acervo/acervo2.asp.

Negra, C., S. Vermeulen, L. G. Barioni, T. Mamo, P. Melville, and M. Tadesse, 2014: Brazil, Ethiopia, and New Zealand lead the way on climate-smart agriculture. *Agric. Food Secur.*, **3**, doi:10.1186/s40066-014-0019-8.

de Oliveira Silva, R., L. G. Barioni, T. Z. Albertini, V. Eory, C. F. E. Topp, F. A. Fernandes, and D. Moran, 2015: Developing a nationally appropriate mitigation measure from the greenhouse gas GHG abatement potential from livestock production in the Brazilian Cerrado. *Agric. Syst.*, **140**, doi:10.1016/j.agsy.2015.08.011.

——, ——, J. A. J. Hall, A. C. Moretti, R. Fonseca Veloso, P. Alexander, M. Crespolini, and D. Moran, 2017: Sustainable intensification of Brazilian livestock production through optimized pasture restoration. *Agric. Syst.*, **153**, doi:10.1016/j.agsy.2017.02.001.

De Oliveira Silva, R., L. G. Barioni, J. A. J. Hall, M. Folegatti Matsuura, T. Zanett Albertini, F. A. Fernandes, and D. Moran, 2016: Increasing beef production could lower greenhouse gas emissions in Brazil if decoupled from deforestation. *Nat. Clim. Chang.*, **6**, 493–497, doi:10.1038/nclimate2916.

——, L. G. Barioni, G. Queiroz Pellegrino, and D. Moran, 2018: The role of agricultural intensification in Brazil's Nationally Determined Contribution on emissions mitigation. *Agric. Syst.*, **161**, doi:10.1016/j.agsy.2018.01.003.

Strassburg, B. B., Latawiec, A. E., Barioni, L. G., Nobre, C. A., Da Silva, V. P., Valentim, J. F., ... & Assad, E. D. (2014). When enough should be enough: Improving the use of current agricultural lands could meet production demands and spare natural habitats in Brazil. *Global Environmental Change*, *28*, 84-97.

Cohn, A. S., Mosnier, A., Havlík, P., Valin, H., Herrero, M., Schmid, E., ... & Obersteiner, M. (2014). Cattle ranching intensification in Brazil can reduce global greenhouse gas emissions by sparing land from deforestation. *Proceedings of the National Academy of Sciences*, *111*(20), 7236-7241.

MOZZER, Gustavo Barbosa; VIEIRA, Simone Aparecida. The Role of Agriculture on the UNFCCC Negotiation Process. In: **Impact of Climate Change on Water Resources in Agriculture**. CRC Press, 2015. p. 9-28.

KRETZ, Tobias; SCHRECKENBERG, Michael. Moore and more and symmetry. In: **Pedestrian and evacuation dynamics 2005**. Springer, Berlin, Heidelberg, 2007. p. 297-308.

TAATI, Siamak. Statistical equilibrium in deterministic cellular automata. In: **Probabilistic Cellular Automata**. Springer, Cham, 2018. p. 145-164.

KAIMOWITZ, David; ANGELSEN, Arild. Will livestock intensification help save Latin America's tropical forests?. **Journal of Sustainable Forestry**, v. 27, n. 1-2, p. 6-24, 2008.

GIL, Juliana; SIEBOLD, Matthias; BERGER, Thomas. Adoption and development of integrated crop–livestock–forestry systems in Mato Grosso, Brazil. **Agriculture, Ecosystems & Environment**, v. 199, p. 394-406, 2015.

GIL, Juliana Dias Bernardes; GARRETT, Rachael; BERGER, Thomas. Determinants of crop-livestock integration in Brazil: Evidence from the household and regional levels. **Land Use Policy**, v. 59, p. 557-568, 2016.

CORTNER, O. et al. Perceptions of integrated crop-livestock systems for sustainable intensification in the Brazilian Amazon. **Land use policy**, v. 82, p. 841-853, 2019.

OpenSoils: Uma Plataforma de Apoio à Ciência do Solo

Sérgio Manuel Serra da Cruz[1], *Filipe Klinger*[1], *Alexia Duarte*[1], *Pedro Vieira Cruz*[1], *Elton Carneiro Marinho*[2], *Eber Assis Schmitz*[2]

[1] Departamento de Computação, Universidade Federal Rural do Rio de Janeiro, Seropédica, Rio de Janeiro, Brasil, serra@pet-si.ufrrj.br

[2] Programa de Pós-Graduação em Informática, Universidade Federal do Rio de Janeiro, Ilha do Fundão, Rio de Janeiro, Brasil, elton@ufrj.br

RESUMO

Solos são recursos não renováveis e parte vital do meio ambiente; os usos racionais bem como a gestão dos dados de solos são problemas globais que afetam a todos os países do mundo. O objetivo deste artigo é apresentar a plataforma *OpenSoils* que armazena, conecta e compartilha grandes quantidades de dados curados de solos brasileiros. *OpenSoils* é uma plataforma aberta, elástica, multiusuária que descreve, organiza e harmoniza grandes conjuntos de dados de perfis e tradagens de solos. Também oferece dados curados e geração de relatórios e mapas permitem que os usuários consultem os dados de solo da sua região. *OpenSoils* é uma das primeiras infraestruturas voltados para segurança de solos baseada na tríade de conceitos de *e-Science* e *Open Science*.

PALAVRAS-CHAVE: Agricultura 4.0, Aplicativos móveis, Segurança de Solos.

ABSTRACT

Soils are nonrenewable resources; they are an essential asset of the environment. Rational use of the soils and data management are global, growing and critical problems that affect every country in the world. The goal of *OpenSoils* is to connect and share large amounts of cured soil data at the Brazilian level aiding researchers to developed soils researchers. *OpenSoils* is an open, elastic, provenance-oriented framework that collects, stores, describes, organizes, shares and harmonizes large data sets of soil profiles and boreholes. It also offers high-quality data and maps allowing users to navigate through the data. *OpenSoils* is one of the first soils security-based infrastructures based on e-Science and Open Science concepts.

KEYWORDS: Agriculture 4.0, Mobile apps, Soils Security.

INTRODUÇÃO

A Ciência do Solo é um domínio complexo tanto do ponto de vista científico como também da integração e gestão de dados. Ela incorpora em suas práticas produtoras de grandes volumes de dados que variam desde as análises físico-químicas de solos à geolocalização e mapas de perfis apoiada por satélites. Portanto, a capacidade de integrar e processar inteligentemente os *datasets* é uma questão crucial para enfrentar as novas demandas ambientais, agrícolas, econômicas, científicas e sociais vivenciados pela sociedade atual (FECHER e FRIESIKE, 2014) e (WOLFERT et al., 2017).

De acordo com Koch et al. (2013) e Wolfert et al. 2017, o solo é um dos recursos naturais não renováveis mais importantes que sustentam a vida terrestre e humana. O solo contribui para o sequestro de carbono, a sustentabilidade energética e climática além de ajudar a manter a biodiversidade e a proteção global do ecossistema. A segurança de solos está diretamente relacionada com a manutenção e melhoria global desses recursos para produção de alimentos, fibras e águas. Do ponto de vista da computação, a segurança de solos é uma área de pesquisa pouco explorada e que pode ser classificada como intensiva em processamento de dados e que demanda a adoção imediata de inúmeras técnicas de computação comuns às áreas de *big data*, inteligência artificial, processamento distribuído, gestão do conhecimento e governança de dados. Até o momento, ainda não se verificam muitas plataformas integradoras de dados voltadas para esse contexto. Isto é, os levantamentos de solos são realizados de modo desconectado e sem apoio sistemático de repositórios de dados compartilháveis, exacerbando o problema de "silos de dados", onde cada pesquisador armazena subconjuntos de dados formatados heterogeneamente, sob a forma de planilhas, sem semântica explicitamente compartilhada (CRUZ et al., *2018*).

O ciclo de vida das pesquisas em solos assemelha-se com o ciclo de vida de experimentos científicos descritos por Mattoso et al. (2010). Ele é amplo e começa com o planejamento das coletas a serem realizadas no campo, perpassa por diferentes tipos de procedimentos experimentais (em campo ou em laboratórios) e continua na estação de trabalho dos pesquisadores com análises, interpretações e visualizações de dados e finaliza com a produção de *datasets* isolados, artigos e mapas que podem ser disponibilizados para tomadores de decisão e demais interessados.

Motivação e objetivo

As principais motivações para o desenvolvimento da plataforma *OpenSoils* foram contribuir com os objetivos de desenvolvimento sustentável enunciados na agenda 2030 da

ONU[1] e oferecer apoio sistemático aos pesquisadores brasileiros na área de Ciência do solo, fornecendo soluções inovadoras e de baixo custo para atender os problemas descritos no relatório 011.713/2015-1 do Tribunal de Conta da União (TCU). Por exemplo, o Brasil dispõe apenas de levantamentos de solo de caráter geral, com mapas de pequena escala, sendo que menos de 5% do território nacional conta com mapas de solos em escala 1:100.000 ou maior e carece de políticas de governança de dados de solos (PRONASOLOS, 2019). A plataforma *OpenSoils* foi concebida para ser um ativo computacional distribuído, modular, colaborativo, aberto, de acesso público voltado para a integração, organização, sistematização e operacionalização de dados decorrentes de estudos e levantamentos pedológicos multi-escalares, podendo ser utilizado em um amplo arco de usuários, finalidades e projetos que variam desde simples mapeamentos de atributos de solos até recomendação e educação em solos.

Neste artigo, defendemos a necessidade de conduzir pesquisas interdisciplinares considerando os papéis da Ciência da Computação e da Agricultura 4.0 (MASSRUHÁ e LEITE, 2017) para enfrentar os desafios da sustentabilidade dos solos. Vislumbramos que várias abordagens computacionais podem ser incorporadas, dentre elas: workflows científicos, proveniência de dados (FREIRE et al., 2008), bancos de dados abertos, ontologias, inteligência artificial, aprendizado de máquina, computação em nuvem, princípios FAIR (WILKINSON et al, 2016), entre outros. Tais abordagens visam apoiar a comunidade de pesquisadores de solos a realizar investigações mais extensas e oferecer novos conhecimentos e soluções para a sociedade, contribuindo para a preservação do solo.

O objetivo deste artigo é divulgar a plataforma *OpenSoils* para cientistas de solos, agrônomos e tomadores de decisão. A plataforma foi concebida para apoiar projetos de naturezas distintas, desde grandes projetos multi-institucionais de cobertura nacional tais como o PRONASOLOS até pequenos projetos locais de mapeamento visando a aptidão, recomendação de uso de solos em propriedades agrícolas ou em cidades.

MATERIAL E MÉTODOS

O desenvolvimento da plataforma *OpenSoils* teve seu início em 2018 e, desde então, vem sendo ampliada, diferindo das demais presentes na literatura (por exemplo, FeBR[2]). *OpenSoils* tem como fundamentação teórica a *Open Science* (ciência aberta) (FECHER e FRIESIKE, 2014) e *e-Science* (e-Ciência) (HEY et al., 2009) aplicados a área de Ciências do solo.

[1] https://nacoesunidas.org/pos2015/agenda2030/

[2] http://coral.ufsm.br/febr/

Open Science é um movimento global abrangente que engloba um arcabouço de saberes aplicados a geração e ampla difusão do conhecimento científico para diversos estratos sociais. É um conceito estimulado por governos e órgão de fomento com vistas para tornar a pesquisa científica mais acessível, ampliando a reprodutibilidade e o compartilhamento de dados científicos, tornando o conhecimento mais inteligível em todos os níveis da sociedade. Atualmente, existem algumas infraestruturas científicas abertas (por exemplo, OpenAIRE, OSF, EOSC), no entanto, ainda não têm contemplam os desafios de segurança dos solos.

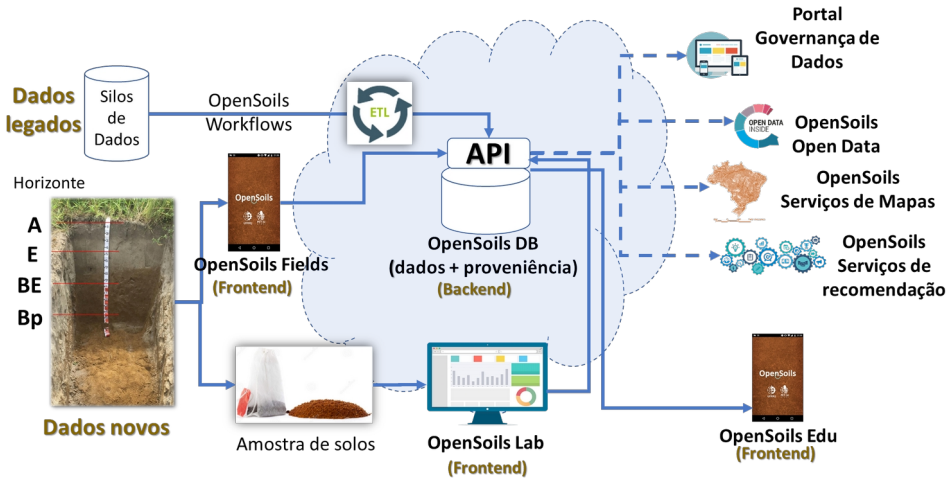
e-Science representa uma nova forma de se fazer ciência no Século XXI (HEY et al., 2009). O conceito diz respeito a um movimento que apregoa a concepção de uma ciência fundamentada no uso intensivo de processamento de dados e na colaboração. A incorporação da *e-Science* no âmbito dos estudos agronômicos mundiais se amplia a cada dia, no entanto, ainda é incipiente no Brasil (BAMBINI et al., 2018). No contexto brasileiro, a adoção da *e-Science* no setor agropecuário e de solos pode se beneficiar da evolução das redes de telefonia móvel, do barateamento dos sensores e drones e da computação em nuvem permitindo que os agricultores, agrônomos e tomadores de decisão se conectem a micro serviços distribuídos através de dispositivos móveis, auxiliando no processo de gestão da cadeia produtiva com conveniência e custos acessíveis. O detalhamento das metodologias adotadas neste trabalho pode ser resumido da seguinte forma.

Os paradigmas adotados no desenvolvimento do *backend* da plataforma estão centrados na teoria de sistemas agrícolas de próxima geração (SAPG) proposta de Antle et al., (2017). Os paradigmas adotados no desenvolvimento das interfaces (*frontend*) dos aplicativos móveis são "centrados no usuário", desenvolvimento ágil e *design thinking* apresentados por Krippendorff (2000).

RESULTADOS E DISCUSSÃO

OpenSoils é a primeira plataforma brasileira de e-ciência e ciência aberta que é elástica, multiusuária, multicamada e orientada a micros serviços capazes de manipular grandes volumes de dados primários de projetos multi-escalares de levantamento de solos. Ela é capaz de processar dados dos projetos e equipes e armazenar dados geolocalizados de cada camada dos perfis e tradagens de solos de um estudo pedológicos. Além disso, harmoniza e mantém dados curados etiquetados com metadados de proveniência retrospectiva relativos a todo o ciclo de atividades/experimentos realizadas diretamente no campo ou em laboratórios que executam experimentos úmidos ou digitais. A plataforma é compatível com o sistema de classificação preconizada pelo Sistema Brasileiro de Classificação de Solos (SIBCS, 2019).

Figura 1. Visão conceitual da plataforma *OpenSoils*, ilustração baseado em Cruz et al., (2018).



Fonte: Elaborado pelos autores.

Resumidamente, o *backend* da plataforma é composto por um conjunto de APIs, banco de dados e camadas de serviço (PASS) concebidos para serem distribuídos, seguros, ter alta disponibilidade e operar em nuvens de computadores. O *frontend* da plataforma é composto por muitos aplicativos móveis e ferramentas Web. A Figura 1 resume o ciclo de vida de pesquisa em solos segundo a plataforma e ilustra de modo conceitual os principais componentes da arquitetura originalmente proposta por Cruz et al. (2018). Tecnicamente, a plataforma é composta por diversas camadas, a saber:

Camada I - É a camada das interfaces voltadas para os usuários (*e.g.* especialistas em solos, pesquisadores, formuladores de políticas públicas, agricultores, profissionais da extensão rural, assistência técnica e estudantes) que utilizam os aplicativos móveis ou portal da Web para acessar os dados ou serviços digitais ofertados pelo *backend*. Além disso, a camada oferece as funcionalidades de governança através da central de controle de projetos e operações que são utilizados por curadores, administradores de dados e gestores de projetos.

As aplicações cliente podem inserir/consultar dados ou administrar serviços no banco de dados *OpenSoils DB* através da API *OpenSoils*, microsserviços ou aplicações (por exemplo, o portal *OpenSoils Lab* e os aplicativos *OpenSoils Field* e *OpenSoils Edu*). O aplicativo *OpenSoils Field*s permite coletar dados diretamente no campo e rastrear a rota de cada amostra enviada aos laboratórios de análise de solos. Normalmente, cada amostra de solo tem sua análise morfológica realizada no local da coleta, esta pode ser complementada posteriormente em laboratórios. O aplicativo *OpenSoils Field* submete, através de conexão segura, dados de solos brutos para o *backend* que opera em nuvem (Figura 1). Depois disso, cada amostra de solo pode ser etiquetada eletronicamente e enviada para laboratórios onde os

pesquisadores realizam experimentos úmidos ou de natureza computacional registrados através de ferramentas digitais, tais como o *SisGExp* (CRUZ e NASCIMENTO, 2019).

Camada II – É uma camada de serviços internos, trata-se do *backend* da plataforma que oferece as APIs e microsserviços utilizadas pelos demais componentes. Utiliza modelos científicos e matemáticos que são representados por workflows científicos centrados em dados (analisam a consistência dos dados legados recebidos da *camada I* e inserem descritores de proveniência retrospectiva segundo a especificação PROV-DM da W3C (MOREAU e MISSIER, 2013). Proveniência são metadados que explicam as origens de um item de dado (FREIRE et al., 2008). A camada também pode suportar conexões com ferramentas de terceiros, tais como o RFlow (CRUZ e NASCIMENTO, 2019) que encapsula workflows científicos baseados em *scripts* estatísticos descritos em linguagem R, permitindo a captura da proveniência retrospectiva dos scripts associados aos experimentos. Essa camada também permite a ingestão de dados legados de solos armazenados em silos de dados dispersos na Internet. Estes podem ser incorporados por meio de workflows científicos do tipo ETL (Extração-Transformação-Carga) desenvolvidos em ferramentas de manipulação de dados (por exemplo, Pentaho/Kettle, entre outros).

Camada III - É a camada mais interna do *OpenSoils*, armazena, classifica e organiza os dados primários e secundários de solos. A estrutura interna suporta um grau diversificado de granularidade de dados e se apoia no *schema* denominado *OpenSoilsDB* (CRUZ et al., *2018*). Esta camada pode armazenar tanto os dados recém coletados no campo quanto os dados legados incorporados a partir dos silos de dados pré-existentes, sendo anotados com metadados proveniência. Resumidamente, o *schema* é capaz de armazenar grandes volumes de dados de solos e de governança. Os dados de solos são georreferenciados e representam as camadas dos perfis e tradagens, são dados de alta qualidade que são voltados para as comunidades brasileira e internacional interessadas em padronização e harmonização de dados de solos. Os dados de solos da plataforma utilizam os padrões de classificação do Sistema Brasileiro de Classificação de Solos (SIBCS, 2019). Cada descrição de perfil de solo possui mais 250 atributos para registrar as propriedades mineralógicos, morfológicos, químicos, físicos, sensores proximais, fertilidade, contaminações por metais pesados e ambientais do solo, entre outros. Além disso, o *OpenSoilsDB* suporta o versionamento de dados abertos e dados georreferenciados, imagens de perfis e dados físicos de experimentos (*e.g.* curva de retenção e dados de espectrometria de solos).

Os dados e metadados de governança, são as informações relativas aos projetos, protocolos de pesquisa/trabalho, usuários, grupos de trabalho, permissões, entre outros. Eles

são tão importantes quanto os científicos e são principalmente utilizados na *camada IV*. Grande parte das ações necessárias para garantir perenidade e a qualidade dos dados se apoiam na capacidade de permitir que pesquisadores executem e gerenciem *datasets* e experimentos com captura sistemática de metadados de proveniência. A proveniência dos dados produzidos pelos workflows da *camada II* consiste no registro da derivação de um resultado por um processo computacional (por exemplo, a ingestão de perfil do solo, coleta de uma imagem, geração de um mapa de atributos de solos). O *OpenSoils DB* adota a recomendação PROV-DM para armazenar a proveniência retrospectiva das execuções dos workflows científicos e scripts.

O banco de dados também permite o armazenamento de grandes volumes de dados legados de solos que podem ser importados por meio de workflows ETL da *camada II*. Além disso, para suportar o paradigma dos dados abertos, oferecem conexão com outras plataformas que podem compartilhar *datasets* harmonizados e curados para toda a comunidade de usuários da plataforma.

Camada IV – A camada de governança de dados do *OpenSoils*, que ainda está em desenvolvimento, permite que especialistas desenvolvam estratégias de governança de dados nos níveis estratégico, tático e operacional. A camada amplia a acessibilidade e a reutilização de dados de solos e permite difundir os conhecimentos sobre levantamentos e segurança de solos tanto para a comunidade científica quanto para cidadãos, agricultores e gestores públicos e privados. A camada suportará os princípios FAIR para a gestão e compartilhamento de dados (WILKINSON et al, 2016).

Adotar os princípios de governança de dados é vantajoso porque não só alinha a plataforma *OpenSoils* com as principais demandas dos órgãos nacionais tais como o TCU, Infraestrutura Nacional de Dados Abertos (INDA) e a Infraestrutura Nacional de Dados Espaciais (INDE), como também facilita o compartilhamento de dados para órgãos internacionais de gestão de dados agrícolas (por exemplo, GODAN (Global Open Data for Agriculture and Nutrition), CGIAR (Consultative Group for International Agricultural Research), Research Data Alliance (RDA), World Soil Information (ISRIC), World Data System (ICSU-WDS), entre outros).

A *camada IV* disponibiliza serviços datacêntricos e processos padronizados que asseguram a transparência das atividades de coleta e distribuição de dados. A camada incorpora regras, políticas e padrões de segurança, direitos e responsabilidades de usos de dados, permitindo o gerenciamento e licenciamento de *datasets* por parte de profissionais da área de curadoria. Embora a governança e a curadoria de dados de solos de longo prazo

tenham recebido pouca atenção dos governos e das comunidades de pesquisa em Ciência do solo, acreditamos que essa camada seja fundamental para contribuir com as políticas de gestão de dados científicos.

Protótipo da plataforma

Um protótipo da plataforma *OpenSoils* já se encontra operacional e armazena dados de mais de 13.000 perfis/tradagens de solos de todo o Brasil; seus aplicativos móveis são gratuitos e podem ser encontrados na loja no Google Play[3]. O portal Web e a ferramenta *OpenSoils Lab* estão disponíveis em *www.opensoils.org*. Testes de campo sobre a usabilidade da plataforma estão sendo realizados com especialistas em projetos de mapeamentos digitais de solos.

Figura 2. Exemplos de telas descrevendo o ciclo de inclusão de dados, relatório de perfil de solo e visualização de dados produzidos pela plataforma *OpenSoils*.



Fonte: Elaborado pelos autores.

A Figura 2 exemplifica um caso de operação integrada da plataforma através de algumas telas dos seus módulos principais. No detalhe da Figura 2A, exemplifica-se o processo de coleta de dados diretamente no campo através do aplicativo *OpenSoils Fields* (*camada I*) e posterior armazenamento no banco de dados na nuvem (*camada III*). A seguir, há o tratamento e harmonização de dados no portal *OpenSoils* Lab (Figura 2B) e a geração de relatórios digitais com descrição de um perfil contendo dados curados e imagens dos

[3] https://play.google.com/store/apps/details?id=br.ufrrj.petsi.opensoils&hl=en

horizontes daquele perfil (Figura 2C) (*camadas II e III*). Por fim, a visualização espacial dos perfis ilustrados através do aplicativo *OpenSoils Edu* (Figura 2D), onde os *pins* (cor vermelha) representam a geolocalização de todos os perfis de um projeto de levantamento de dados de solos.

CONCLUSÕES

Ao mesmo tempo em que os solos fornecem um espaço vital para os seres humanos, eles estão sob pressão, devido ao aumento da população e crescentes demandas por alimentos ou aos usos competitivos da terra.

Neste artigo, apresentamos o *OpenSoils*, uma plataforma centrada em dados de solos que fornece apoio ao ciclo de vida de estudos na Ciência do solo. A plataforma, está em atualização constante, mas já pode ser utilizada em projetos de levantamentos pedológicos multi-escalares, na conservação e o uso sustentável do solo ou em atividades ligadas à Agricultura 4.0. A plataforma está em constante evolução, porém já é capaz de armazenar e compartilhar *datasets* curados e permite a coleta de dados em diversos pontos do ciclo de vida, sendo capaz de auxiliar nas tomadas de decisão.

Como trabalhos futuros, planejamos continuar a implementação da plataforma, em especial no mapeamento da API da *camada II* que facilitarão conectividade com redes de sensores de solos e ferramentas de terceiros. Adicionalmente, investiremos mais esforços nos módulos da *camada IV* que permitirão investigar a ligação semântica entre os *datasets* de solos do *OpenSoils* com *datasets* produzidos por organismos internacionais, tais como *WoSIS Soil Profile Database*. Além disso, planejamos incorporar recursos do modelo de ciência cidadã e envolver uma rede de voluntários para auxiliar na coleta de novos dados de solos a serem utilizados nas pesquisas em segurança de solos.

AGRADECIMENTOS

O presente trabalho foi parcialmente realizado com apoio da Programa de Educação Tutorial (PET-SI, da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001 e do Conselho Nacional de Pesquisas (CNPq) – Código de Financiamento produtividade DT-II (315399/2018-0). Agradecimentos ao professor Marcos Cedia pelo apoio na área de solos.

REFERÊNCIAS

ANTLE J.M., JONES J.W., ROSENZWEIG C.E., Next generation agricultural system data, models and knowledge products: Introduction. Agricultural Systems.; 2017, pp. 155:186-190.

BAMBINI, M.D., BONACELLI, M. B. M., HIHA, R. H., Pesquisa Agropecuária No Contexto Da E-Science: Monitoramento de Temas e Plataformas de Data Science. 56 Congresso SOBER, Campinas, SP. 2018, 20pp.

CRUZ, S. M. S et al. OpenSoils: e-Science em Segurança de Solos II. Encuentro Latinoamericano de e-Ciencia (TICAL), Cartagena da Índias. Anais da Octava Conferência de Directores de Tecnología de Información y Comunicación en Instituciones de Educación Superior, y II Encuentro Latinoamericano de e-Ciencia 2018. 8p.

CRUZ, S. M. S e NASCIMENTO, J. A. P., Towards integration of data-driven agronomic experiments with data provenance. Computers and Electronics in Agriculture, v. 161, 2019, pp. 14-28.

CARACCIOLO C. et al., The AGROVOC linked dataset, Semant. Web, vol. 4, no. 3, 2013, pp. 341–348.

DEELMAN, E., GANNON, D., SHIELDS, M., TAYLOR, I., Workflows and e-Science: An overview of workflow system features and capabilities, Futur. Gener. Comput. Syst., vol. 25, no. 5, 2009, pp. 528–540.

FECHER B., FRIESIKE S., Open Science: One Term, Five Schools of Thought, Open. Sci., 2014, pp. 17–47.

FREIRE, J., KOOP, D., SANTOS, E., SILVA, C. T., Provenance for computational tasks: A survey, Comput. Sci. Eng., vol. 10, no. 3, 2008, pp. 11–21.

HEY T., TANSLEY S., TOLLE K., The Fourth Paradigm: Data-Intesive Scientific Discovery. 2009.

KOCH A. et al., Soil Security: Solving the Global Soil Crisis, Glob. Policy, vol. 4, no. 4, Nov. 2013, pp. 434–441.

KRIPPENDORFF, K. Propositions Of Human-Centeredness: A Philosophy For Design. In: Durling, D. Friedman, K. (Eds.). Doctoral Education in Design: Foundations for the Future. Staffordshire (UK): Staffordshire Univ.Press. 2000, pp.55-63.

MASSRUHÁ, S. M. F. S.; LEITE, M. A. de A. AGRO 4.0 no meio rural. Futuro das TICs na AGRO 4.0. 2017.

MATTOSO M. et al., Towards supporting the life cycle of large scale scientific experiments, Int. J. Bus. Process Integr. Manag., vol. 5, no. 1, 2010, pp. 79.

MOREAU, L., MISSIER, P., PROV-DM: The PROV Data Model, W3C Recommendation, 2013. [Online]. http://www.w3.org/TR/prov-dm/.

PRONASOLOS. 2019. [Online]. https://www.embrapa.br/pronasolos

SiBCS. 2019. [Online]. https://www.embrapa.br/solos/sibcs

WILKINSON M. D. et al., The FAIR Guiding Principles for scientific data management and stewardship, Sci. Data, vol. 3, Mar. 2016. pp. 160018.

WOLFERT S., GE L., VERDOUW C., BOGAARDT M. J., Big Data in Smart Farming – A review, Agric. Syst., vol. 153, 2017, pp. 69–80.

Visão Computacional Aplicada ao Monitoramento Bovino: Uma Análise dos Biomas Nacionais

Matheus Serrão Marinato[1], *Jayne Ribeiro da Silva Machado*[1], *Wanderlany Fialho de Abreu*[1], *Márcio José Moutinho da Ponte*[1]

[1] Instituto de Engenharia e Geociências, Universidade Federal do Oeste do Pará, Santarém, Pará, Brasil, marinatostm21@gmail.com, jaynesilva661@gmail.com, lanyfialho@hotmail.commjmoutinho@yahoo.com.br

RESUMO

A necessidade de monitorar a quantidade de gado existente na propriedade do criador é imprescindível. Na maioria dos casos a quantidade é obtida de forma aproximada, sendo verificada visualmente e em situações de necessidade real, realizada através de contagem por separação ou confinamento, no qual o indivíduo distrai os bovinos para serem levados a um local fechado, com intuito de realizar a contagem. Porém é um grande problema, pois esse meio de avaliação demanda muito tempo e pode ocasionar comprometimento do rebanho. Neste contexto, este trabalho tem o objetivo de desenvolver um protótipo baseado no modelo de mistura gaussiana (GMM) utilizando técnicas de visão computacional em diferentes biomas, no intuito de monitorar, contar, o rebanho bovino evitando perdas e facilitando o manejo dos animais.

PALAVRAS-CHAVE: Gado, Mistura Gaussiana, Manejo.

ABSTRACT

The need to monitor the amount of livestock existing in the property of the creator is indispensable. In most cases the quantity is obtained in an approximate way, being visually verified and in situations of real necessity, performed by counting by separation or confinement, in which the individual distracts the cattle to be taken to a closed place , in order to perform the counting. But it is a big problem, because this means of evaluation demands a lot of time and can cause commitment of the herd. In this context, this work aims to develop a

prototype based on the Gaussian mixing model (GMM) using computational vision techniques in different biome, in order to monitor, count, the cattle herd avoiding losses and facilitating Animal management.
KEYWORDS: Cattle, Gaussian Mixture, Management.

INTRODUÇÃO

De acordo com o Ministério da Agricultura, pecuária e abastecimento (2019), a pecuária é um dos grandes elementos contribuintes para a economia do Brasil, com uma representatividade de 50% nas atividades agropecuárias, proporcionando qualidade nas atividades de manejo do rebanho bovino. A aplicação de tecnologias no processo de produção, pode estimular a competitividade e qualidade dos produtos no negócio. (Lomba et al., 2017)

A Implantação de sistemas computacionais é um investimento no negócio que facilita o manuseio, automação, agilidade e prevenção de perdas. A pecuária não está fora deste contexto, sendo uma área de extrema importância na economia brasileira e no mercado internacional, existindo a necessidade de implementar tecnologias e métodos avançados na atividade de monitoramento bovino, pois as suas práticas demandam desenvoltura e qualidade do rebanho.

Apesar dos avanços tecnológicos, alguns processos pertencentes ao monitoramento bovino ainda são realizados manualmente, como a contagem do rebanho. Este procedimento geralmente é complexo e exige um grande esforço, dado que os animais são conduzidos à um local, no intuito de determinar o quantitativo existente. (Handling and Housing, 2002)

Este processo é arriscado, visto que durante a etapa de preparação da contagem manual pode ocorrer acidentes, como lesões na pele ou até mesmo a perda do animal, dessa forma impactando na produtividade. Além disso, o bovino passa bastante tempo confinado sem alimento, causando prejuízos econômicos ao produtor. O estresse também reduz o ganho de peso e deixa o animal com baixa imunidade, podendo ficar vulnerável a doenças. Desta forma à implantação de tecnologias no manejo bovino poderá minimizar o risco de acidentes e agilizar o processo de contagem. (Handling and Housing, 2002)

Neste contexto, este trabalho tem o objetivo de detectar e contar automaticamente, utilizando imagens de vídeo, à fim de facilitar a quantificação e monitoramento da espécie, utilizando técnicas de visão computacional baseado no modelo de mistura gaussiana.

MATERIAL E MÉTODOS

Trabalhos Relacionados

Ribeiro (2019) descreve em seu estudo, o processo de desenvolvimento de um programa que aplica algoritmos de Visão Computacional na contagem do gado, utilizando imagens aéreas coletadas em bancos de imagens públicas. O processamento das imagens foi executado utilizando técnicas de visão computacional e processamento de imagens, por meio da linguagem de programação Python. O autor discute que a relevância do trabalho, pela análise estatística dos dados coletados por pecuaristas aplicados na administração dos diversos processos de manejo, vacinação e controle de doenças em gados, especialmente devido ao constante crescimento da exportação de produtos gerados pela pecuária no Brasil.

Sensolo e Buhr (2015), apresentam o desenvolvimento de um sistema de visão computacional, que utiliza a segmentação com a finalidade de contar os troncos de madeira organizados em pilhas, a partir da análise de imagens digitais. Esses troncos expõem semelhanças de formato e textura, no qual permite à diferenciação do restante da imagem. Entretanto, variações de iluminação na cena e posicionamento dos troncos na pilha, impede a detecção automática pelo computador. Com o auxílio do framework OpenCV, criou-se uma aplicação semi automatizada, em que delimita a área da pilha e em seguida realiza a separação dos troncos e os quantifica.

Lucarini et al. (2004), projeta e implementa um sistema semi-automático de contagem de microrganismos, aplicando técnicas de Processamento de Imagens. O procedimento metodológico, se baseia na técnica tradicionalmente empregada em microbiologia da contagem de células em hemacitômetro. Para isto, foi implementado um sistema que captura imagens de um microscópio óptico através de uma câmera de vídeo e encaminha a imagem para um computador. No qual, se emprega a digitalização e limiarização da imagem e a quantidade de microrganismos existentes em uma amostra, em que é calculado pelo meio da manipulação do algoritmo de rotulação de áreas conectadas.

Procedimentos Metodológicos

Para implementação dos métodos será utilizada a IDE Matlab versão R2018b, que dispõem de bibliotecas com funções específicas de visão computacional. Também será aplicado o modelo de mistura gaussiana (GMM) com a finalidade de realizar testes em vídeos dos biomas da Amazônia, cerrado, pampa, pantanal e caatinga, visando detectar e contar o gado. O Matlab é essencial para execução dos objetivos e análise dos resultados.

Esse protótipo baseia-se na análise do processamento de imagens usando vídeos retirados de telejornais, todos no formato 480x640. Os resultados gerados serão avaliados,

mediante matriz de confusão, avaliando os diferentes biomas quanto a eficiência do modelo proposto.

Modelo de Mistura Gaussiana

O modelo desenvolvido aplica misturas Gaussian a cada frame e transforma imagens rgb em imagens binárias. Os pixels correspondentes que não sofrem variações de estado, o valor 1 (preto) é atribuído e para pixels que sofrem mudanças drásticas no estado, o valor 0 (branco) é atribuído. (SANTOSH et al., 2013)

Em termos matemáticos, temos:

$$1) \quad P(x_t) = \sum_{i=1}^{k} \omega_{i,t}\, \eta(x_t; \mu_{i,t}, \sum_{i,t})$$

Em que, evidentemente,

$$2) \quad \sum_{i=1}^{k} \omega_{i,t} = 1$$

A média da mistura, equivale,

$$3) \quad \mu_t = \sum_{i=1}^{k} \omega_{i,t}\, \mu_{i,t,}$$

Isto é, a soma ponderada dos meios das densidades dos componentes. Onde a variável que representa o pixel atual no frame, K é o número de distribuições, e t representa o tempo (o índice do frame), é uma estimativa do peso do iésimo Gaussiano na mistura no tempo t, é o valor médio do iésimo Gaussiano na mistura no tempo t, é a matriz de covariância do iéssico Gaussiano na mistura no tempo t.

Implementação do Modelo de Mistura Gaussiana

O protótipo implementado usando o modelo de mistura gaussiana no monitoramento bovino, executa algumas etapas para detecção. As etapas utilizadas são exemplificadas na figura 1 a seguir.

Figura 1 –Exemplificação das etapas executadas pelo protótipo implementado

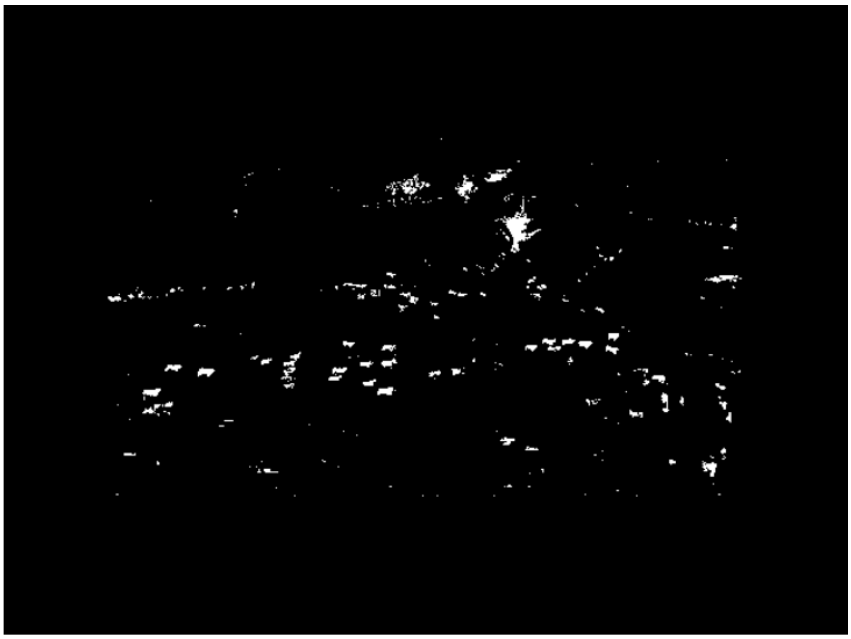


Fonte: (Autor, 2019)

Inicialmente o protótipo realiza a separação do vídeo em frames, no qual consiste na quebra do vídeo em vários quadros de imagem. Após essa separação, é iniciado um treinamento, onde é aplicado na detecção no primeiro plano do vídeo.

Dessa forma é processado o frame inicial e com isso sendo segmentado a partir do plano de fundo. O uso da detecção de primeiro plano, necessita que à quantidade frames utilizados seja informado, com a finalidade de aplicar o modelo de mistura gaussiana. Neste trabalho foi empregado o uso de 350 quadros para inicialização do modelo, após a inicialização, o detector começa a produzir resultados de segmentação mais confiáveis, mediante um dos quadros de vídeo e a utilização da máscara de primeiro plano. A figura 2 apresenta o resultado da segmentação de primeiro plano para inicialização do modelo de mistura gaussiana.

Figura 2 – Aplicação da segmentação de primeiro plano para inicialização do modelo de mistura gaussiana



Fonte: (Autor, 2019)

Posteriormente a utilização da segmentação de primeiro plano, se faz necessário o uso da aplicação do filtro morfológico de dilatação com a finalidade de suavizar os ruídos apresentados no frame e dessa forma preencher as lacunas expostas na detecção. Esse filtro trata da junção de dois elementos (imagem e seu elemento estruturante) empregando adição vetorial para expansão da imagem. A figura 3 apresenta a utilização do filtro de dilatação para suavização dos ruídos.

Figura 3 – Utilização do filtro morfológico de dilatação para suavização dos ruídos



Fonte: (Autor, 2019)

Depois da aplicação do filtro morfológico, caixas delimitadoras são desenhadas em cada componente correspondente a um objeto identificado, no qual é filtrado novamente no primeiro plano detectado, rejeitando regiões de interesse que apresentam, menos de 90 pixels na detecção.

Na contagem do gado é utilizado o número de regiões encontradas no vídeo e apresentada no canto esquerdo da imagem. A figura 4 mostra a detecção e contagem do gado.

Figura 4 – Detecção e contagem do gado, após a aplicação do modelo de mistura gaussiana



Fonte: (Autor, 2019)

Finalizado a etapa de treinamento, o processo anterior é efetuado novamente para o restante dos frames, exibindo a detecção e o número total do gado em todo o vídeo.

RESULTADOS E DISCUSSÃO

Na análise dos resultados do protótipo foi aplicado a matriz de confusão (Brilhador,2011), onde verificou-se a quantidade de:

- Verdadeiro positivo (VP): Referem-se as entradas positivas que foram corretamente detectadas.
- Falso positivo (FP): Referem-se as entradas negativas que foram incorretamente detectadas como positivas.

Após isto, foi calculado a taxa de eficiência e de erro, manuseando os valores do verdadeiro positivo (VP) e do falso positivo (FP), as fórmulas empregadas para esse cálculo se encontram na figura 5 abaixo.

Figura 5 – Fórmulas usadas para o cálculo dos Resultados

$$Eficiencia = \frac{(VP*100)}{(VP+FP)} \qquad Taxa\ Erro = \frac{(FP)}{(VP+FP)}$$

Fonte: (Brilhador, 2011)

A aplicação das fórmulas, gerou como resultado os dados que se encontram na tabela 1.

Tabela 1 – Resultados do uso dos filtros para detecção do gado no bioma

Resultado	Valor Real	VP	FP	Eficiência em %	Taxa de erro
Amazônia	45	39	10	79,59	20,40
Cerrado	10	6	6	50	50
Pampa	11	5	42	10,63	89,36
Pantanal	80	60	33	64,51	35,48
Caatinga	48	17	42	28,81	71,18

Avaliando os dados gerados na utilização do modelo mistura gaussiana para detecção do gado, expôs as melhores taxas de eficiência os biomas Amazônia e Pantanal. Porém os biomas do cerrado, pampa e caatinga exibiram índice baixo respectivamente 50%, 10,63% e 28,81%, pelo fato dos vídeos encontrados apresentarem filmagem descalibrada e muitos ruídos, situações como estas que dificultam uma detecção eficiente do gado. O protótipo mostrou ser eficiente na implementação dos biomas, realizando a detecção do gado em até uma determinada área. Além disso, a baixa qualidade dos vídeos empregados e a longa distância do objeto de interesse influenciaram nos resultados.

CONCLUSÕES E TRABALHOS FUTUROS

O uso do processamento digital de imagens para detecção dos gados é desafiador. A aplicação de técnicas em visão computacional, empregando o modelo de mistura gaussiana, mostrou ser eficiente nos diversos tipos de biomas.

A qualidade dos vídeos influenciaram nos resultados, causando interferências na detecção do gado, principalmente nos biomas do pampa e da caatinga. A questão das oscilações e da diversidade em cada bioma afetou na detecção e contagem do gado.

Apesar desses fatores, foi possível criar uma solução para o monitoramento do gado, recorrendo a uma simulação do problema desenvolvida na ferramenta Matlab. Como trabalhos futuros, pretende-se criar uma base de dados utilizando na coleta de campo um drone, aprimorando o modelo de mistura gaussiana e aplicando outras técnicas de visão computacional.

REFERÊNCIAS

BRASIL. Agropecuária Brasileira em Números. fev 2019. Disponível em: <http://www.agricultura.gov.br/assuntos/politica-agricola/agropecuaria-brasileira-em-numeros>. Acesso em: 28 set. 2019.

BRILHADOR, A. Análise Semi-Automática do Arranjo Espacial de Plantas de Milho utilizando Visão Computacional. 111 f. Dissertação – Mestrado – Programa de Pós-graduação em Informática, Universidade Tecnológica Federal do Paraná. Cornélio Procópio, 2015.

Health and Safety Executive. Handling and housing cattle, Agriculture Information Sheet No 35, 05/02.

LOMBA, Luíz Fernando Delboni. Título: Sistema para monitoramento da movimentação bovina e aferição dos comportamentos. Anais do XI Congresso Brasileiro de Agroinformática (SBIAgro 2017), Campinas-SP, 2017.

LUCARINI, C.; SILVA, L. A. da; BIANCH, R. A. C. Um Sistema para a Contagem Semi-Automática de Microorganismos. Pesquisa & Tecnologia FEI - Nº 26, 2004.

RIBEIRO, Nathan Gustavo Vieira; GUEDES, Gustavo Bartz; BARBIERI, Tamires Tessarolli. Aplicação de algoritmos de Visão Computacional na contagem de gado por meio de processamento de imagens aéreas. Revista Eletrônica Argentina-Brasil de Tecnologias da Informação e da Comunicação, [S.l.], v. 1, n. 10, jun. 2019. ISSN 2446-7634. Disponível em: <https://revistas.setrem.com.br/index.php/reabtic/article/view/343>. Acesso em: 01 out. 2019. doi: http://dx.doi.org/10.5281/zenodo.3338098.

SANTOSH, D. H. H.; POORNESH, P. V.; RAO, L. N.; KUMAR, N. A. Tracking Multiple Moving Objects Using Gaussian Mixture Model. International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-3, Issue-2, May 2013.

SENSOLO, A. B.; BUHR, K. P. Desenvolvimento de um Sistema de Visão computacional para Identificação e Contagem de Troncos de Madeiras Armazenados em Pilhas. Unoesc & Ciência - ACET Joaçaba, v. 6, n. 2, p. 175-184, Dezembro 2015.



Diagnóstico Automático de Ceratoconjuntivite Infecciosa Bovina por meio de Imagens Termográficas e *Deep Learning*

Dhyonatan Santos de Freitas[1], *Sandro da Silva Camargo*[1], *Helena Brocardo Comin*[1], *Robert Domingues*[2], *Emanuelle Gaspar Baldo*[2], *Fernando Flores Cardoso*[1,2]

[1]Programa de Pós-Graduação em Computação Aplicada,
Universidade Federal do Pampa & EMBRAPA Pecuária Sul
Bagé, Rio Grande do Sul, Brasil
dhyonatanfreitas.aluno@unipampa.edu.br, sandro.camargo@unipampa.edu.br,

[2]Embrapa Pecuária Sul, Empresa Brasileira de Pesquisa Agropecuária
Bagé, Rio Grande do Sul, Brasil
robertdomingues@embrapa.br, emanuellebaldo@embrapa.br, fernandocardoso@embrapa.br

RESUMO

A Ceratoconjuntivite Infecciosa Bovina (CIB) é a doença ocular com maior incidência em bovinos no mundo, causando perdas no rebanho e prejuízos ao produtor. Atualmente, o diagnóstico para CIB é realizado por meio da avaliação dos sinais clínicos por um especialista e confirmados por exames laboratoriais, o que geralmente é uma tarefa custosa. Nesse sentido, é notória a necessidade em aprimorar as formas de diagnóstico da doença. Neste trabalho, é discutido o uso de Redes Neurais Convolucionais (RNC) para diagnóstico automático da CIB. A abordagem proposta foi treinada e testada em uma base de imagens termográficas da região ocular bovina de animais sadios e experimentalmente infectados. Foram desenvolvidos seis modelos de RNC com arquiteturas distintas e avaliados utilizando validação cruzada. Os melhores resultados indicam que as RNCs são capazes de identificar os sinais clínicos da CIB em imagens termográficas com uma acurácia próxima a 87%.

PALAVRAS-CHAVE: Redes Neurais Convolucionais, Doenças Oculares, Conjuntivite, *Moraxella bovis*, Inteligência Artificial.

ABSTRACT

Infectious Bovine Keratoconjunctivitis (IBK) is the ocular disease with the highest incidence in cattle herds in the world, causing losses in herds and, consequently, for farmers. Currently, the diagnosis for IBK is performed by evaluating clinical signs by an expert and confirmed by laboratory exams, which is usually a costly task. In this sense, it is perceived the need to improve the forms of identification of the disease. In this paper, the use of Convolutional Neural

Networks (CNN) for IBK automatic diagnosis is discussed. The proposed approach was trained and tested on a thermographic images dataset of the bovine ocular region of healthy and experimentally infected animals. Six CNN models were developed with different architectures and evaluated using cross-validation. The best results indicate that CNNs are able to identify the clinical signs of IBK in thermographic images with a accuracy rate close to 87%.
KEYWORDS: Convolutional Neural Networks, Ocular Diseases, Conjuntivitis, *Moraxella bovis*, Artificial Intelligence.

INTRODUÇÃO

A criação de bovinos é uma das principais atividades do agronegócio nacional, representando cerca de 5,5% do Produto Interno Bruto brasileiro (CEPEA, 2018). Além disso, o Brasil está inserido em posição de destaque no competitivo mercado internacional da carne. Apesar da posição favorável, melhorias nos aspectos gerenciais, zootécnicos e econômicos se fazem necessárias para manter a competitividade da pecuária brasileira (FILHO; EUCLIDES, 2010).

Do ponto de vista dos processos zootécnicos, percebe-se um grande esforço para integrar novas tecnologias que possam colaborar com o incremento da produtividade e redução de custos nas fazendas. Uma alternativa promissora é o desenvolvimento de ferramentas que auxiliem no diagnóstico precoce e na prevenção de doenças como a CIB. A CIB é considerada a doença ocular de maior impacto na criação de bovinos, sendo responsável por perdas econômicas e produtivas consideráveis em rebanhos de todo o mundo (POSTMA; CARFAGNINI; MINATEL, 2008; ALEXANDER, 2010). Trata-se de uma doença contagiosa causada pela bactéria denominada *Moraxella bovis*. Animais infectados apresentam manifestações clínicas que iniciam com lacrimejamento, fotofobia, edema e opacidade de córnea, podendo evoluir para úlcera de córnea e, nos casos mais graves, cegueira irreversível (CHANDLER; BAPTISTA; TURFREY, 1979). Embora não seja uma doença geralmente fatal, tem altas taxas de morbidade, chegando a afetar cerca de 80% do rebanho após 3 a 4 semanas do início do surto (POSTMA; CARFAGNINI; MINATEL, 2008).

As lesões características da CIB, que se originam na área central da córnea e progridem de forma centrífuga, facilitam o diagnóstico pelos sinais clínicos. Porém, o diagnóstico conclusivo depende do isolamento e identificação da *Moraxella bovis* em laboratório, por meio da coleta do líquido conjuntival de animais infectados (ANGELOS et al., 2007). Por vezes, o surto de infecção já se encontra em fase avançada quando identificado pelo produtor, podendo agregar-se a isso a indisponibilidade imediata dos meios de diagnóstico adequados.

Com o advento da tecnologia no campo, ferramentas inovadoras e biotecnologias são uma alternativa para o avanço em pesquisa animal, principalmente na identificação de doenças e promoção do bem-estar. A utilização de métodos não destrutivos nem invasivos tem sido uma alternativa preferencial para obtenção de dados. Neste sentido, a Termografia Infravermelha (TIV) é um método viável na medicina veterinária como forma de obtenção de dados para subsidiar o diagnóstico precoce de doenças (LEÃO et al., 2015).

Neste contexto, visando a necessidade de se ter um diagnóstico mais rápido e eficiente

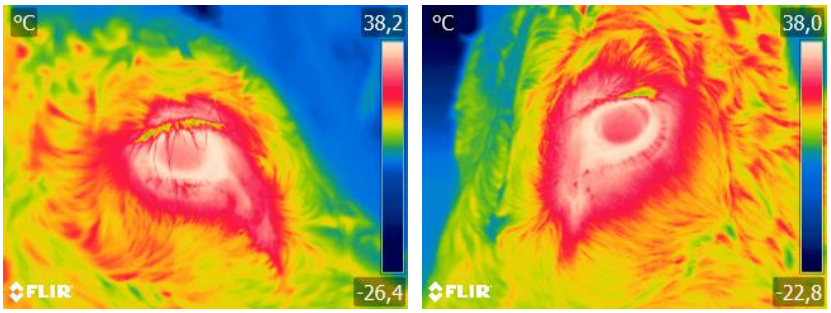
em casos de CIB, este trabalho tem como objetivo desenvolver um método computacional, baseado em *Deep Learning*, capaz de identificar autonomamente os sinais clínicos da CIB em imagens termográficas da região ocular bovina e fornecer o diagnóstico, simulando o trabalho do especialista.

MATERIAL E MÉTODOS

Base de Dados e Ferramentas Computacionais

A base de imagens disponível foi desenvolvida pela Embrapa Pecuária Sul onde os pesquisadores observaram o desenvolvimento da CIB em 11 animais da raça Hereford. Os registros termográficos foram realizados em cinco dias (dia 0, 2, 4, 6 e 8 pós infecção), sendo que no dia zero os animais foram observados sadios e nos dias seguintes já com a manifestação da doença em um dos olhos, que foi experimentalmente infectado (procedimento aprovado pela CEUA/CPPSul nº 10/2016). A base é composta por 186 imagens termográficas da região ocular bovina, compactadas no formato JPEG, com dimensões de 320×240 pixels e três canais do espaço de cores RGB. As imagens foram produzidas utilizando uma câmera térmica da marca FLIR® modelo T300. A Figura 1 mostra imagens termográficas da região ocular de um bovino infectado (a) e de um sadio (b). As avaliações do progresso da doença foram classificadas pelo

Figura 1: Imagens termográficas da região ocular bovina.



(a) Animal infectado pela CIB (b) Animal sadio

especialista por meio de escores, conforme o Guia de Coleta de Dados de Ceratoconjuntivite Bovina Infecciosa (CBI) (GASPAR et al., 2016). A Tabela 1 apresenta a classificação de cada escore de acordo com os sinais clínicos do animal no momento da avaliação. Para a construção

Tabela 1: Relação entre escores e sinal clínico identificado no momento da avaliação.

Escore	Sinal Clínico
0	Nenhum sinal clínico
1	Lacrimejamento (ausência de lesão na córnea)
2	Úlceras na córnea (de até 4mm de diâmetro)
3	Úlceras na córnea maiores que 4mm (pode haver opacidade da córnea)
4	Ruptura corneal (rompimento da úlcera)

dos métodos, foi utilizada a linguagem de programação Python (versão 2.7) e as bibliotecas

para aprendizado de máquina *Tensorflow* (versão 1.5) e *Keras* (versão 2.2.4). A escolha dessas ferramentas foi baseada em sua popularidade e eficiência quando comparado a outras estruturas de aprendizado profundo com distribuição em código aberto (RAMPASEK; GOLDENBERG, 2016; KOVALEV; KALINOVSKY; KOVALEV, 2016).

Pré-processamento dos dados

As imagens originais continham análises de temperatura em seu conteúdo. Com isso, o primeiro passo do processamento dos dados foi remover essas análises de temperatura, as quais poderiam prejudicar a extração de características pelas RNCs. Nesta tarefa, o software disponibilizado pela fabricante da câmera, o FLIR Tools®, foi utilizado. Em seguida, com base nas informações contidas na planilha de avaliação dos especialistas, apenas as imagens classificadas com os escores 0 e 1 foram selecionadas e separadas em duas classes. A primeira classe contém 33 imagens identificadas com escore 1 e a segunda classe contém 111 imagens com escore 0.

Além disso, as classes supramencionadas foram divididas em três agrupamentos: treinamento, teste e validação. Essa divisão foi baseada na técnica da validação cruzada utilizando o método *holdout*, o qual garante a exclusividade dos dados que serão empregados na avaliação dos modelos.

Como a base de imagens original possui poucas amostras para cada classe e redes neurais necessitam de um volume grande de dados, foi utilizada uma técnica conhecida como aumento de dados nos agrupamentos de treino e teste. Essa técnica tem sido empregada frequentemente em trabalhos de reconhecimento de imagens por meio de *deep learning* (PEREZ; WANG, 2017; HAN; LIU; FAN, 2018).

Desta forma, a partir das imagens originais, foi gerada uma nova base de imagens, como mostra a Tabela 2.

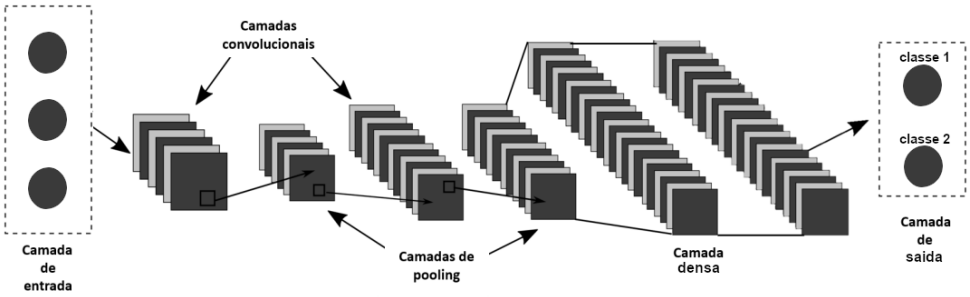
Tabela 2: Base de imagens gerada a partir do processo de aumento de dados.

Agrupamento	**Antes do processo de aumento de dados**		**Após o processo de aumento de dados**	
	Imagens de olhos infectados	**Imagens de olhos sadios**	**Imagens de olhos infectados**	**Imagens de olhos sadios**
Treinamento	19	65	938	949
Teste	7	23	322	327
Validação	7	23	-	-
Total de imagens	33	111	1260	1276

Desenvolvimento das Redes Neurais Convolucionais

RNC é um tipo de rede neural especializada no processamento de dados de entrada que possuem um formato de matriz, usualmente utilizadas para o processamento de informações visuais (GOODFELLOW; BENGIO; COURVILLE, 2016). São uma espécie de extrator de características automático que antecede uma rede neural do tipo *Multilayer Feedforward* para classificação do

Figura 2: Arquitetura de uma RNC.



Fonte: Adapdata de (ARAÚJO et al., 2017)

resultado em classes distintas. A arquitetura de uma RNC é apresentada na Figura 2. A arquitetura de uma RNC é constituída pela camada de entrada juntamente com a combinação de blocos de camadas convolucionais seguidas de funções de ativação que geram mapas de características. Esses mapas de características são processados por uma operação de *pooling*, a qual é utilizada para reduzir a dimensionalidade dos dados e refinar o mapa de características gerado na camada anterior. Por fim, um conjunto de camadas completamente conectadas (camada densa), também acompanhadas por funções de ativação, são responsáveis por gerar o resultado final.

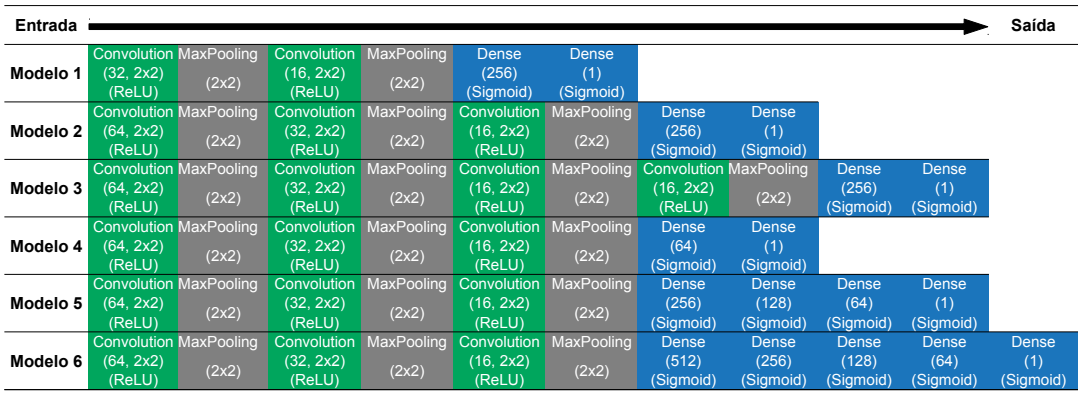
A configuração de uma RNC é definida pelos seguintes hiperparâmetros: i) Camadas convolucionais, onde para cada camada deve-se definir o número de filtros, seu tamanho e o tamanho do passo dado durante a convolução; ii) Camadas de *pooling*, onde são definidos o tamanho e o tipo do filtro utilizado; e iii) Camadas densas, onde são definidos a quantidade de camadas ocultas e o número de neurônios contidos em cada uma delas.

O processo de desenvolvimento de uma rede neural para a solução de novos problemas é um desafio, principalmente no que tange a busca pela arquitetura com melhor capacidade de abstração e generalização dos dados. Nesse sentido, buscou-se desenvolver modelos de RNC com arquiteturas distintas e avaliar sua capacidade de identificar os sintomas da CIB em imagens termográficas da região ocular bovina. Essas arquiteturas foram criadas utilizando as bibliotecas Keras e Tensorflow. A Figura 3 apresenta a arquitetura de cada um dos modelos de RNCs desenvolvidos. A variação dos hiperparâmetros observou a composição de dois tipos de arquiteturas: a primeira com maior número de camadas convolutivas e um número fixo de camadas totalmente conectadas, como pode ser visto nos modelos 1, 2 e 3; a segunda com um número fixo de camadas convolutivas e uma variação crescente de camadas totalmente conectadas, como é visto nos modelos 4, 5 e 6. Essa proposta foi elaborada com o intuito de observar se modelos com maior capacidade de extração de características e também de classificação teriam vantagem em relação a modelos com uma arquitetura mais simples.

Métricas de avaliação

Técnicas de aprendizado de máquina necessitam de uma avaliação confiável dos resultados. Medidas que avaliam a qualidade de uma tarefa de classificação são construídas a partir de uma

Figura 3: Arquitetura dos modelos de RNCs desenvolvidas.



matriz de confusão, contabilizando exemplos reconhecidos corretamente e incorretamente para cada classe (SOKOLOVA; JAPKOWICZ; SZPAKOWICZ, 2006). Em classificações binárias, essas matrizes são compostas por quatro elementos: Verdadeiros Positivos (VP); Falsos Positivos (FP); Verdadeiros Negativos (VN) e Falsos Negativos (FN), como mostra a Tabela 3. As métricas

		Classe Real	
		Positivo	**Negativo**
Classe Prevista	**Positivo**	VP	FP
	Negativo	FN	VN

Tabela 3: Matriz de confusão.

mais utilizadas para avaliar algoritmos de aprendizado de máquina têm sido a taxa de acurácia e a taxa de erro, as quais são definidas pelas Equações 1 e 2. Porém, a avaliação utilizando apenas essas métricas é frágil, pois as mesmas não atendem plenamente às necessidades de problemas de aprendizagem nos quais as classes são igualmente importantes (SOKOLOVA; JAPKOWICZ; SZPAKOWICZ, 2006).

$$Acurácia = \frac{VP + VN}{VP + FP + VN + FN} \tag{1}$$

$$Erro = 1 - Acurácia \tag{2}$$

Por outro lado, duas medidas que estimam separadamente o desempenho de um classificador em diferentes classes são sensibilidade e especificidade. A sensibilidade explica a capacidade do sistema em predizer corretamente a condição para casos que realmente a possui e a especificidade explica a capacidade do sistema em predizer corretamente a ausência da condição para casos que realmente não a possui (KOHAVI, 1998). A sensibilidade e a especificidade são definidas pelas Equações 3 e 4.

$$Sensibilidade = \frac{VP}{VP + FN} \tag{3}$$

$$Especificidade = \frac{VN}{VN + FP} \tag{4}$$

Neste trabalho, as métricas citadas serão utilizadas para avaliar os modelos de RNC em dois momentos: i) etapa de treinamento das RNCs, utilizando a acurácia e a taxa de erro; ii) etapa de validação, utilizando todas as métricas descritas.

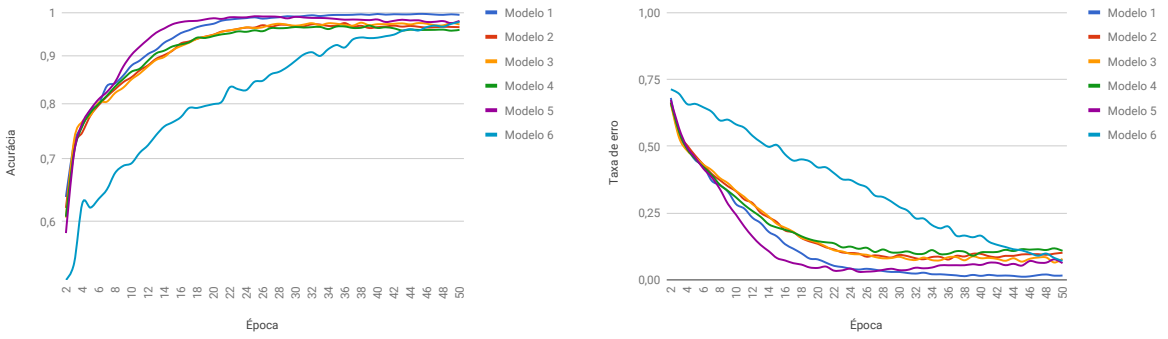
RESULTADOS E DISCUSSÃO

A utilização da TIV como método de obtenção de dados mostrou-se ser uma alternativa confiável e não invasiva para auxiliar no diagnóstico de animais acometidos pela CIB, conforme apontado em (LOPES et al., 2018). Diante disso, foi proposto neste trabalho o uso de *deep learning* para a construção de modelos de RNCs com o objetivo de reconhecer os sinais clínicos da CIB em imagens termográficas da região ocular bovina. Os modelos de RNCs desenvolvidos foram submetidos a uma sessão de treinamento na base de dados disponível e, posteriormente, validados.

Treinamento dos modelos

Foram utilizadas um total de 2536 imagens termográficas da região ocular bovina (Tabela 2) para treinar seis modelos de RNCs com arquiteturas distintas. Os modelos de RNCs desenvolvidos foram submetidos a uma rotina de treinamento de 50 épocas cada. O número de épocas é determinado como uma condição de parada para o algoritmo de treinamento. A Figura 4(a) apresenta os resultados da acurácia e a Figura 4(b) os resultados da taxa de erro obtidos por cada modelo na etapa de treinamento. Os resultados alcançados na etapa de treinamento mostram

Figura 4: Resultados da etapa de treinamento dos modelos de RNC.



(a) Acurácia

(b) Taxa de erro

que os modelos apresentaram uma acurácia entre 95% e 99%, apresentando uma boa capacidade de generalização no conjunto de teste, validando a técnica de aumento de dados utilizada para construir a base de dados de treinamento.

Nesta etapa, modelos com mais camadas responsáveis pela extração de características apresentaram maior capacidade de aprendizado em decorrência do número de épocas de treinamento, contudo são mais suscetíveis a ocorrência de *overfitting*, por isso o equilíbrio entre

o número de camadas convolutivas e totalmente conectadas é desejável. Observou-se ainda que modelos com arquiteturas mais profundas indicam a necessidade de mais épocas de treinamento, como pode ser visto na curva de aprendizado do modelo 6 na Figura 4(a).

Validação dos modelos

Após a etapa de treinamento, os modelos foram avaliados em relação à sua capacidade de classificar novas amostras. Para isso, foi utilizada uma parte do conjunto de imagens exclusivamente destinadas a esta etapa, sendo 30 imagens térmicas da região ocular bovina (Tabela 2), classificadas em duas classes pelos especialistas. Os modelos de RNC foram submetidos a análise da base de imagens citada e os resultados computados por meio de uma matriz de confusão (Tabela 3) e as Equações 1, 2, 3 e 4, como mostra a Tabela 4.

Tabela 4: Resultados obtidos pelos modelos de RNC na etapa de validação.

Métricas	**Modelos de RNCs**					
	1	**2**	**3**	**4**	**5**	**6**
VP	3	3	3	4	2	6
FP	0	0	2	1	0	3
VN	23	23	21	22	23	20
FN	4	4	4	3	5	1
Acurácia	0,8667	0,8667	0,8000	0,8667	0,8333	0,8667
Taxa de erro	0,1333	0,1333	0,2000	0,1333	0,1667	0,1333
Sensibilidade	0,4286	0,4286	0,4286	0,5414	0,2857	0,8571
Especificidade	1,0000	1,0000	0,9130	0,9565	1,0000	0,8696

Os resultados da etapa de validação demonstram que os modelos tiveram uma acurácia média de 85% $\pm$ 2,79 e um erro médio de 25%. Isso demonstra a viabilidade das RNCs em detectar os sinais clínicos da CIB em imagens termográficas. Além disso, os valores da sensibilidade e especificidade evidenciam a capacidade dos modelos em predizer corretamente valores de VP e VN, respectivamente. Os modelos 1, 2, 3 e 5 apresentaram menor capacidade de predizer valor de VP, o que não é recomendado em uma ferramenta para diagnóstico médico precoce. Os modelos 4 e 6 apresentaram um equilíbrio entre a capacidade de predizer corretamente valor de VP e VN, o que em diagnóstico de doenças é desejável (DAVIS; GOADRICH, 2006).

O desbalanceamento da base de dados original pode ser um fator que influenciou nos resultados deste trabalho, uma vez que haviam mais imagens classificadas como escore 0 do que imagens classificadas como escore 1. Nesse sentido, a coleta de novas imagens classificadas com escores diferentes de 0 é uma alternativa a ser explorada.

CONCLUSÕES

A utilização da TIV como método não invasivo de obtenção de dados mostrou-se uma alternativa eficiente e promissora para o auxílio na detecção dos sinais clínicos da CIB, conforme relatado em (LOPES et al., 2018) e neste trabalho.

O uso de técnicas de aumento de dados mostrou-se uma alternativa eficiente dada a quantidade reduzida de imagens disponíveis, o que é uma realidade comum em termos de diagnóstico por imagem na pecuária. Os melhores resultados obtidos neste trabalho indicam que as RNCs são capazes de classificar corretamente os sinais clínicos da CIB em imagens termográficas com uma acurácia próxima de 87%. Contudo, aprimoramentos são necessários para encontrar a arquitetura que melhor se adapta aos dados disponíveis. Além disso, um treinamento com maior número de épocas se mostrou desejável ao passo que a complexidade dos modelos aumenta. Desta forma, este trabalho proporcionou um avanço no que tange a forma de diagnosticar a CIB. A utilização de imagens termográficas é uma forma não invasiva de coletar dados dos animais e, por meio das RNCs, a interpretação desses dados se torna uma ferramenta capaz de apoiar o médico veterinário no momento de seu diagnóstico definitivo.

Como possíveis trabalhos futuros, estima-se que o desenvolvimento de novas arquiteturas de RNCs, juntamente com uma base de treinamento maior, podem fornecer resultados mais acurados. Além disso, a construção de bases de dados maiores para este fim é desejável para o aprimoramento de técnicas de reconhecimento de padrões e apoio ao diagnóstico.

AGRADECIMENTOS

Os autores agradecem a Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul pela bolsa de mestrado, a Embrapa-SEG pelo financiamento número 02.13.02.13.10.002 e a Prof [a]. Concepta Margaret McManus Pimentel pelo fornecimento do equipamento de termografia.

REFERÊNCIAS

ALEXANDER, D. Infectious bovine keratoconjunctivitis: a review of cases in clinical practice. *The Veterinary clinics of North America. Food animal practice*, Elsevier, v. 26, n. 3, p. 487–503, 2010.

ANGELOS, J. A. et al. Moraxella bovoculi sp. nov., isolated from calves with infectious bovine keratoconjunctivitis. *International Journal of Systematic and Evolutionary Microbiology*, Microbiology Society, v. 57, n. 4, p. 789–795, 2007.

ARAÚJO, F. H. et al. Redes neurais convolucionais com tensorflow: Teoria and prática. *SOCIEDADE BRASILEIRA DE COMPUTAÇÃO. III Escola Regional de Informática do Piauí. Livro Anais-Artigos and Minicursos*, Sociedade Brasileira de Computação, v. 1, n. 1, p. 382–406, jun 2017.

CEPEA. *Centro de Estudos Avançados em Economia Aplicada: PIB do Agronegócio Brasileiro*. 2018. Disponível em: `https://www.cepea.esalq.usp.br/br/pib-do-agronegocio-brasileiro.aspx`. Acesso em: 15/05/2019.

CHANDLER, R.; BAPTISTA, P.; TURFREY, B. Studies on the pathogenicity of moraxella bovis in relation to infectious bovine keratoconjunctivitis. *Journal of comparative pathology*, Elsevier, v. 89, n. 3, p. 441–448, 1979.

DAVIS, J.; GOADRICH, M. The relationship between precision-recall and ROC curves. In: *Proceedings of the 23rd International Conference on Machine Learning*. New York, NY, USA: ACM, 2006. (ICML '06), p. 233–240. ISBN 1-59593-383-2.

FILHO, K. E.; EUCLIDES, V. Desenvolvimento recente da pecuária de corte brasileira e suas perspectivas. *PIRES, AV Bovinocultura de corte. Piracicaba: FEALQ*, v. 1, p. 11–40, 2010.

GASPAR, E. et al. Guia de coleta de dados de ceratoconjutivite bovina infecciosa (CBI). *Embrapa Pecuária Sul-Documentos (INFOTECA-E)*, Bagé: Embrapa Pecuária Sul, 2016., 2016.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. *Deep learning*. London, England: MIT Press Cambridge, 2016.

HAN, D.; LIU, Q.; FAN, W. A new image classification method using cnn transfer learning and web data augmentation. *Expert Systems with Applications*, Elsevier, v. 95, p. 43–56, 2018.

KOHAVI, R. Glossary of terms. *Special issue on applications of machine learning and the knowledge discovery process*, v. 30, n. 271, p. 127–132, 1998.

KOVALEV, V.; KALINOVSKY, A.; KOVALEV, S. Deep learning with theano, torch, caffe, tensorflow, and deeplearning4j: Which one is the best in speed and accuracy? Minsk: Publishing Center of BSU, 2016.

LEÃO, J. M. et al. Uso da termografia infravermelha na pecuária de precisão. *Embrapa Gado de Leite-Artigo em periódico indexado (ALICE)*, Cadernos Técnicos de Veterinária e Zootecnia, n. 79, p. 97-109, 2015., 2015.

LOPES, T. et al. Utilização de imagens termográficas para diagnóstico de ceratoconjuntivite infecciosa bovina em animais da raça hereford. *Simpósio de Iniciação Cientifica da Embrapa Pecuária Sul*, v. 8, 2018.

PEREZ, L.; WANG, J. The effectiveness of data augmentation in image classification using deep learning. *arXiv preprint arXiv:1712.04621*, 2017.

POSTMA, G. C.; CARFAGNINI, J. C.; MINATEL, L. *Moraxella bovis* pathogenicity: an update. *Comparative immunology, microbiology and infectious diseases*, Elsevier, v. 31, n. 6, p. 449–458, 2008.

RAMPASEK, L.; GOLDENBERG, A. Tensorflow: Biology's gateway to deep learning? *Cell systems*, Elsevier, v. 2, n. 1, p. 12–14, 2016.

SOKOLOVA, M.; JAPKOWICZ, N.; SZPAKOWICZ, S. Beyond accuracy, F-score and ROC: A family of discriminant measures for performance evaluation. In: *Proceedings of the 19th Australian Joint Conference on Artificial Intelligence: Advances in Artificial Intelligence*. Berlin, Heidelberg: Springer-Verlag, 2006. (AI'06), p. 1015–1021.

Quantificação da severidade de doenças foliares em plantas e manutenção de um banco de imagens digitais

João Batista Tolentino Júnior[1], *Amanda do Prado Matto*[2], *Ana Gabriela Mabe Martins*[3], *Adriana Terumi Itako*[2]

[1] Departamento de Agricultura, Biodiversidade e Florestas, Campus de Curitibanos, Universidade Federal de Santa Catarina, Curitibanos, Santa Catarina, Brasil, joao.tolentino@ufsc.br, adriana.itako@ufsc.br

[2] Programa de Pós-Graduação em Agronomia, , Universidade Estadual de Maringá, Maringá, Paraná, Brasil, pradomattosa@gmail.com

[3] Curso de Graduação em Agronomia, Campus de Curitibanos, Universidade Federal de Santa Catarina, Curitibanos, Santa Catarina, Brasil, anamabe83@gmail.com

RESUMO

Tarefas como a determinação da severidade de doenças em plantas pode ser facilitada e ter uma maior precisão com a utilização de softwares que quantifiquem a área foliar da planta afetada a partir de imagens digitais. O objetivo deste trabalho é propor uma metodologia semiautomática para quantificação da severidade de doenças foliares de plantas baseada nas propriedades estatísticas da imagem e manter um banco de imagens de fotografias digitais, devidamente catalogadas. Folhas de plantas de goiaba-serrana (*Acca sellowiana*) com ausência e presença de lesões de mancha de Phyllachora (*Phyllachora* spp.) foram coletadas e fotografadas. As imagens digitais foram editadas para destacar as lesões causadas pela doença. Posteriormente, foram analisadas no software R pelo pacote EBImage pelo processo de limiarização entre as lesões e o restante da folha. Propriedades estatísticas relacionadas a cor foram calculadas e correlacionadas com a severidade da doença. As imagens foram disponibilizadas em um banco de imagens digitais no website *fitopatologia-ccr.ufsc.br* junto com um arquivo com metadados das imagens. A metodologia proposta foi capaz de quantificar com precisão a severidade de doenças foliares de plantas. A disponibilização em um banco de imagens dos resultados da quantificação da severidade e das propriedades estatísticas das imagens poderá ser utilizado na aplicação de modelos mais complexos que sejam capazes de estimar com precisão a severidades a partir destas medidas.

PALAVRAS-CHAVE: Processamento de imagens, Imagem RGB, Metadados.

ABSTRACT

Determining the severity of plant diseases can be easier with softwares that quantify the leaf area affected from digital images. The objective of this work is to propose a methodology for quantification of the severity of foliar diseases of plants based on the statistical properties of the image and to maintain a database of digital photographs. Leaves of feijoa (*Acca sellowiana*) plants with absence and presence of lesions of Phyllachora spot (*Phyllachora* spp.) were collected and photographed. Digital images were edited to highlight lesions. Afterwards, they were analyzed in software R by the EBImage package through the process of thresholding between the lesions and the rest of the leaf in order to quantify the severity. Statistical properties related to color were calculated and correlated with the severity. The images were made available in a database of digital images in the site *fitopatologia-ccr.ufsc.br* along with metadata. The proposed methodology was able to accurately quantify the severity of foliar plant diseases. The availability in a bank of images of the results can be used in development of more complex models that are able to accurately estimate the severity from these measurements.

KEYWORDS: Image processing, RGB image, metadata.

INTRODUÇÃO

A determinação da severidade de uma doença pela análise visual é demorada e sujeita a erros, necessitando de grande conhecimento sobre a doença avaliada. O olho humano tem capacidade limitada em interpretar pequenas diferenças na severidade, devido à subjetividade da percepção da quantidade da doença pelo avaliador, com variações na avaliação entre diferentes avaliadores ou pelo mesmo avaliador(MADDEN *et al*., 2007).

A agricultura vem sendo cada vez mais auxiliada pelas tecnologias. Tarefas como a determinação da severidade de doenças em plantas pode ser facilitada e ter uma maior precisão com a utilização de softwares que quantifiquem a área foliar da planta afetada a partir de imagens digitais (MARTINELLI *et al*., 2015). Os equipamentos e softwares vem se tornando cada vez mais complexos, mas também há uma preocupação com a facilidade de uso destes (BOCK e NUTTER, 2011).

Although the equipment and software continue to become more sophisticated, these technologies are also becoming easier to use. As a result, image analysis to measure plant disease is becoming increasingly widely used, and has now been applied in the study of numerous plant diseases

A limiarização é uma técnica muito utilizada de segmentação de imagem, consistindo em separar regiões em duas classes: o fundo e o objeto. A maneira mais simples de limiarização baseia-se na bipartição do histograma da imagem, convertendo os pixels com tom de cinza maior ou igual a um certo valor de limiar em brancos e os demais em pretos. Pode-se utilizar técnicas de cálculo do valor ótimo de limiar com base nas propriedades estatísticas da imagem (OTSU, 1970).

Vários trabalhos buscam a identificação e/ou quantificação de doenças em plantas a partir da análise de imagens digitais, utilizando processos de limiarização e segmentação de imagens, histogramas de cor e/ou obtenção de características inerentes à imagem (ALI *et al.*, 2017; BARBEDO, 2014; BARBEDO *et al*., 2016; CAMARGO, SMITH, 2009; IQBALA *et al*., 2018; KRUSE *et al*., 2014; SENGAR *et al*., 2018). Ainda assim, há grandes desafios a serem superados na quantificação de doenças, relacionados com as características das doenças e ao processo de obtenção das imagens (BARBEDO, 2016; HAMUDA *et al*., 2016).

Segundo Conci *et al.* (2008), as imagens digitais podem ser caracterizadas por algumas propriedades estatísticas da imagem, como média e desvio padrão dos tons de cinza da região de interesse, média e desvio padrão dos tons de cinza da região de fundo. Estas propriedades podem ser altamente relacionadas com a severidade da doença foliar (SINGH e MISRA, 2017).

O objetivo deste trabalho é propor uma metodologia semi-automática para quantificação da severidade de doenças foliares de plantas baseada nas propriedades estatísticas da imagem e manter um banco de imagens de fotografias digitais, devidamente catalogadas.

MATERIAL E MÉTODOS

Este trabalho compreende as etapas de quantificação da severidade de doenças de plantas, em imagens digitais, pela metodologia computacional proposta e a manutenção de um banco de imagens digitais que poderá ser utilizado para a quantificação da severidade de doenças de plantas.

Para a quantificação da severidade, foram coletadas 13 folhas de plantas de goiaba-serrana (*Acca sellowiana*) com a ausência e a presença de lesões de mancha de Phyllachora (*Phyllachora* spp.). As doenças foram identificadas a partir da análise de sua sintomatologia conforme Amorim *et al*. (2018).

Todas as folhas foram fotografadas em um ambiente com fundo claro e boa iluminação. Foi utilizada uma câmera fotográfica digital modelo Canon T5 na face adaxial e abaxial. As

imagens foram salvas no formato RAW e convertidas em formato TIF para análise digital utilizando o software RawTherapee versão 5.5.

Posteriormente à obtenção das imagens digitais, estas foram submetidas à edição no software GIMP versão 2.10 para delimitação das áreas lesionadas. Para isso, foram destacadas as áreas lesionadas pela doença utilizando a ferramenta pincel na cor preta.

As imagens editadas foram analisadas no software R com o pacote EBImage (PAU *et al*., 2010). Inicialmente, as imagens foram decompostas em seus três níveis de cor RGB: vermelho (red), verde (green) e azul (blue), utilizando o comando *channel.* A camada verde (green) foi escolhida para as etapas seguintes de limiarização, visto a coloração verde predominante da folha. A correção *gamma* foi utilizada com o intuito de alterar o contraste entre o objeto e o fundo da imagem. Em um primeiro momento este comando foi empregado para tornar a folha mais escura em relação ao fundo, e em um segundo momento tornando as lesões mais evidentes com o clareamento da folha. O expoente da correção *gamma* foi escolhido após análise preliminar de um pequeno grupo de imagens.

Por fim, foi realizado o processo de limiarização, a partir da determinação de um limiar que permite a separação entre o fundo e o objeto de interesse. Este limiar foi calculado segundo a metodologia de Otsu (1979). Após a segmentação da imagem, foi utilizado o comando *ComputeFeatures*, que analisa as regiões em destaque da imagem e realiza cálculos das suas dimensões, dentre elas, a área em pixels. A porcentagem da área foliar lesionada foi, então, calculada pela divisão da área ocupada pelas lesões pela área foliar total.

A partir das imagens originais, foram calculadas as seguintes propriedades estatísticas da imagem: média dos tons das cores RGB e seus respectivos desvios padrão, média dos tons das cores RGB acima e abaixo do limiar estabelecido seus respectivos desvios padrão e razão do brilho médio das cores RGB. Foram calculados os coeficientes de correlação linear destas propriedades com a porcentagem de área foliar lesionada.

As imagens originais, as imagens editadas com destaque para as lesões e os resultados da análise das imagens foram disponibilizadas em um banco de imagens digitais no website *fitopatologia-ccr.ufsc.br*. Junto ao banco de imagens, também foi disponibilizado um arquivo com metadados das imagens, contendo informações a respeito da obtenção das imagens e os resultados da análise computacional: porcentagem de área foliar lesionada e propriedades estatística das imagens.

RESULTADOS E DISCUSSÃO

A utilização da metodologia proposta foi capaz de quantificar a área foliar lesionada. As imagens foram editadas para destacar as lesões com a cor preta (Figura 1). Isso permitiu uma perfeita limiarização das lesões e, consequentemente, a quantificação da área foliar lesionada.

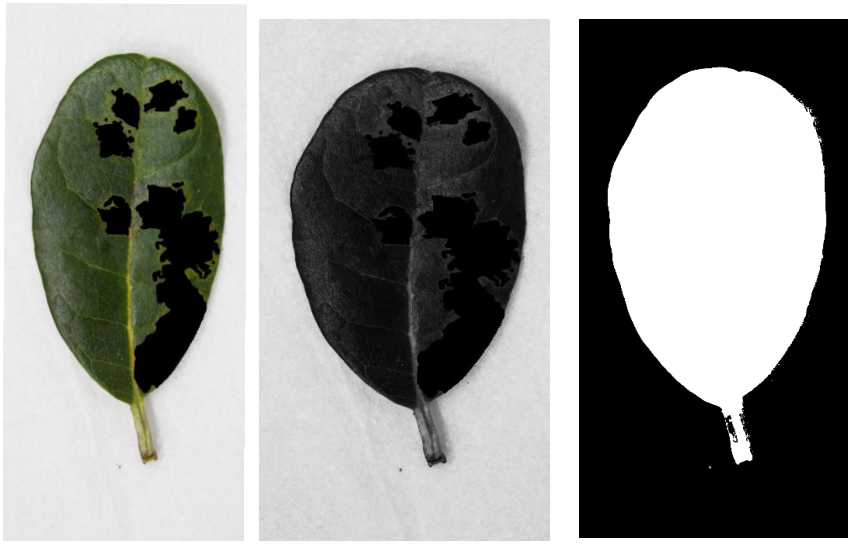
Figura 1 –Imagem original (esquerda) e imagem editada (direta) com destaque para as lesões de mancha-de-Phyllachora em folhas de goiaba-serrana. Exemplo ilustrativo para apenas uma imagem obtida.



Fonte: Os autores

A escolha do canal verde e a correção *gamma* resultaram em uma imagem adequada para o processo de limiarização. Ao utilizar um expoente maior que 1,0 para a correção *gamma* (no exemplo foi utilizado 2,0), mais escura a folha em relação ao fundo (Figura 2), evidenciando e facilitando a limiarização, com consequente segmentação da imagem em folha/fundo da imagem.

Figura 2 – Correção gamma na imagem do centro tornando a folha mais escura em relação ao fundo da imagem, permitindo a separação da folha do restante da imagem (imagem da direita). Exemplo ilustrativo para apenas uma imagem obtida.



Fonte: Os autores

Com o expoente de correção *gamma* menor que 1,0 (no exemplo foi utilizado 0,2), a folha apresentou menor contraste, ficando mais clara, evidenciando as lesões (Figura 3). Seguindo o mesmo processo de limiarização, o resultado foi uma ótima segmentação de lesão/fundo da imagem.

Figura 3 – Correção *gamma* na imagem do centro tornando as lesões escuras em relação à folha, permitindo a separação das lesões do restante da imagem (imagem da direita). Exemplo ilustrativo para apenas uma imagem obtida.



Fonte: Os autores

Com os objetos de interesse devidamente delimitadas, obtiveram-se as dimensões e a área das folhas e das lesões, como consta na Tabela 1, um exemplo de uma imagem de folha de goiaba-serrana e de uma imagem de videira. As propriedades estatísticas das imagens originais estão apresentadas na Tabela 2, para as mesmas imagens do exemplo anterior.

Tabela 1 – Dimensões dos objetos de interesse limiarizados em folhas com lesões de mancha-de-Phyllachora em folhas de goiaba-serrana obtidos pelo comando *computeFeatures.shape*. Exemplo ilustrativo para apenas uma imagem obtida.

Imagem de folha de goiaba serrana com lesões de mancha-de-Phyllachora	
	Área (pixel)
Lesões	473773
Folha	1930509
%área lesionada	24,5%

Tabela 2 – Propriedades estatísticas extraídas das imagens.

Propriedades Estatísticas	
Média dos tons de vermelho (mR)	0,2860
Média dos tons de verde (mG)	0,3187
Média dos tons de azul (mB)	0,1923
Desvio-padrão dos tons de vermelho (sR)	0,1143
Desvio-padrão dos tons de verde (sG)	0,1169
Desvio-padrão dos tons de azul (sB)	0,0999
Média dos tons de vermelho acima do limiar (mR+)	0,3999
Média dos tons de verde acima do limiar (mG+)	0,4418
Média dos tons de azul acima do limiar (mB+)	0,2306
Desvio-padrão dos tons de vermelho acima do limiar (sR+)	0,0847
Desvio-padrão dos tons de verde acima do limiar (sG+)	0,0809
Desvio-padrão dos tons de azul acima do limiar (sB+)	0,1019
Média dos tons de vermelho abaixo do limiar (mR-)	0,2122
Média dos tons de verde abaixo do limiar (mG-)	0,2598
Média dos tons de azul abaixo do limiar (mB-)	0,1206
Desvio-padrão dos tons de vermelho abaixo do limiar (sR-)	0,0553
Desvio-padrão dos tons de verde abaixo do limiar (sG-)	0,0796
Desvio-padrão dos tons de azul abaixo do limiar (sB-)	0,0365
Brilho médio dos tons de vermelho (bmR)	0,3588
Brilho médio dos tons de verde (bmG)	0,3999
Brilho médio dos tons de azul (bmB)	0,2412

A análise de correlação realizada a partir desses dados, resultou em coeficientes de correlação linear com amplitude de |0,0054| a |0,8303|, apresentado tanto correlações positivas quanto negativas (Tabela 3). Mesmo as correlações lineares apresentando uma tendência de baixos valores, ainda é possível que sejam encontradas boas relações não lineares entre as propriedades da imagem avaliadas.

Tabela 3 – Coeficientes de correlação linear entre a porcentagem de área lesionada e as propriedades estatísticas das imagens.

Propriedades Estatísticas	
Média dos tons de vermelho (mR)	0,2032
Média dos tons de verde (mG)	0,2505
Média dos tons de azul (mB)	- 0,4990
Desvio-padrão dos tons de vermelho (sR)	-0,3460
Desvio-padrão dos tons de verde (sG)	-0,5363
Desvio-padrão dos tons de azul (sB)	-0,1112
Média dos tons de vermelho acima do limiar (mR+)	0,0847
Média dos tons de verde acima do limiar (mG+)	0,1202
Média dos tons de azul acima do limiar (mB+)	-0,4267
Desvio-padrão dos tons de vermelho acima do limiar (sR+)	0,1554
Desvio-padrão dos tons de verde acima do limiar (sG+)	-0,0567
Desvio-padrão dos tons de azul acima do limiar (sB+)	0,0054
Média dos tons de vermelho abaixo do limiar (mR-)	0,5412
Média dos tons de verde abaixo do limiar (mG-)	0,4123
Média dos tons de azul abaixo do limiar (mB-)	-0,3403
Desvio-padrão dos tons de vermelho abaixo do limiar (sR-)	-0,2590
Desvio-padrão dos tons de verde abaixo do limiar (sG-)	-0,4737
Desvio-padrão dos tons de azul abaixo do limiar (sB-)	-0,8303
Brilho médio dos tons de vermelho (bmR)	0,5470
Brilho médio dos tons de verde (bmG)	0,3802
Brilho médio dos tons de azul (bmB)	- 0,4411

Após todas as análises e cálculos realizados com as imagens, os resultados foram disponibilizados em um banco de imagens digitais no website *fitopatologia-ccr.ufsc.br*. Buscou-se organizar as informações das imagens de uma forma intuitiva para que possam ser facilmente utilizadas por outros pesquisadores. As imagens estão no formato RAW em alta resolução, mais adequadas para a análise digital. Soluções semelhantes podem ser encontradas, como o repositório digipathos (www.digipathos-rep.cnptia.embrapa.br) da Embrapa (BARBEDO *et al.*, 2018). Este repositório foi criado com o objetivo de suportar estudos em métodos de detecção e reconhecimento de doenças de plantas. Fujita *et al*. (2018) criaram um banco de imagens com mais de nove mil imagens da cultura do pepino cuidadosamente capturadas com o objetivo de identificar sintomas do míldio e de viroses. No entanto, enfatiza-se que os objetivos do nosso trabalho são diferentes, ao enfocar na quantificação da severidade das doenças de plantas.

A obtenção e organização de bancos de imagens é laboriosa, pois envolve várias etapas, além de um conhecimento profundo das doenças em estudo para a correta delimitação das áreas lesionadas. No estudo aqui apresentado, foi possível obter mensurações precisas da severidade utilizando a metodologia proposta de destaque das áreas lesionadas. Esse é um primeiro passo que permite a utilização de modelos mais complexos, que já demonstraram ser bastante promissores para a quantificação automática da severidade em imagens digitais.

Pethybridge e Nelson (2015) desenvolveram uma aplicação para smartphone, denominada Leaf Doctor, que permite a quantificação da severidade de doenças de plantas a

partir de imagens digitais de modo semiautomático. Os autores afirmam que as estimativas da severidade são feitas com alta acurácia. Ainda é possível citar a utilização de outros métodos de determinação da severidade, como a utilização de lógica *fuzzy* (SANNAKKI *et al*., 2011; SIBIYA e SUMBWANYAMB, 2019), support vector machine (ARIVAZHAGAN *et al*., 2013; HALLAU *et al*., 2018) ou *machine learning* (MOHANTY *et al*., 2016; SLADOJEVIC *et al*., 2016). Todas estas técnicas avançadas necessitam como dado de entrada medidas acuradas da severidade. É na obtenção destas medições que está a principal contribuição desta pesquisa. A utilização de modelos mais complexos fica além dos objetivos propostos para este trabalho, mas com a disponibilização do banco de imagens, espera-se a contribuição de outros pesquisadores no desenvolvimento de modelos que possam estimar com o menor erro a severidade de doenças foliares de plantas.

As relações entre as propriedades estatísticas das imagens e a percentagem de área foliar lesionada podem ser bastante altas, como pôde ser demonstrado pela análise de correlação linear simples. Resultados similares já foram encontrados por Gavhale e Gawande (2014) e Mainkar *et al*. (2015). Como citado anteriormente, espera-se que a utilização de técnicas mais avançadas permita uma melhor estimativa da severidade das doenças de plantas a partir de imagens digitais.

CONCLUSÕES

A metodologia proposta foi capaz de quantificar com precisão a severidade da mancha de Phyllachora em folhas de goiaba-serrana, nas condições estudadas.

A disponibilização em um banco de imagens dos resultados da quantificação da severidade e das propriedades estatísticas das imagens poderá ser utilizado na aplicação de modelos mais complexos que sejam capazes de estimar com precisão a severidades a partir destas medidas.

REFERÊNCIAS

ALI, H. *et al*. Symptom based automated detection of citrus diseases using color histogram and textural descriptors. Computers and Electronics in Agriculture. V. 138, n. 1. P. 92-104, 2017.

BARBEDO, J. G. A. An automatic method to detect and measure leaf disease symptoms using digital image processing. Plant Disease. v 98, n 12, 2014. p. 1709-1716

AMORIM, L. *et al.* Manual de Fitopatologia, v 1. 5 ed. Editora Ceres, 573 p. 2018.

ARIVAZHAGAN, S *et al*. Detection of unhealthy region of plant leaves and classification of plant leaf diseases using texture features. Agricultural Engineering International: CIGR Journal, v. 15, n. 1, p. 211-217, 2013.

BARBEDO, J. G. A. A review on the main challenges in automatic plant disease identification based on visible range images. Biosystems engineering 144, 2016. p 52-60

BARBEDO, J. G. A. *et al*. Annotated Plant Pathology Databases for Image-Based Detection and Recognition of Diseases. IEEE Latin America Transactions, v. 16, n. 6, p. 1749-1757, 2018.

BARBEDO, J. G. A. *et al*. Identifying multiple plant diseases using digital image processing. Biosystems engineering, 147, 2016. p. 104-116

BOCK, C. H., NUTTER JR, F. W. Detection and measurement of plant disease symptoms using visible-wavelength photography and image analysis. CAB Reviews: Perspectives in Agriculture, Veterinary Science, Nutrition and Natural Resources. v. 6: p. 1-15, 2011.

CAMARGO, A.; SMITH. J. S. An image-processing based algorithm to automatically identify plant disease visual symptoms. Biosystems engineering 102, 2009. p 9-21

CONCI, A. *et al* Computação Gráfica: Processamento e Análise de Imagens Digitais, v. 2 Campus/Elsevier, 2008. 528 p.

FUJITA, E. *et al*. A Practical Plant Diagnosis System for Field Leaf Images and Feature Visualization. International Journal of Engineering & Technology, v. 7, n. 4.11, p. 49-54, 2018.

GAVHALE, K. R.; GAWANDE, U. An overview of the research on plant leaves disease detection using image processing techniques. IOSR Journal of Computer Engineering, v. 16, n. 1, p. 10-16, 2014.

HALLAU, L. *et al*. Automated identification of sugar beet diseases using smartphones. Plant Pathology. v. 67, p. 399-410, 2018.

HAMUDA, E. *et al*. Um levantamento de técnicas de processamento de imagens para extração e segmentação de plantas no campo. Computers and Electronics in Agriculture, v 125, 2016. p 184-199

IQBALA, Z. *et al*. An automated detection and classification of citrus plant diseases using image processing techniques: a review. Computers and Electronics in Agriculture, 153, 2018. p. 12–32.

KRUSE, O. M. O. *et al*. Pixel classification methods for identifying and quantifying leaf surface injury from digital images. Computers and Electronics in Agriculture, v 108, 2014. p. 155-165

MADDEN, L. V. *et al*. The study of plant disease epidemics. St Paul: APS Press, 2007. 432 p

MAINKAR, P. M. *et al*. Plant leaf disease detection and classification using image processing techniques. International Journal of Innovative and Emerging Research in Engineering, v. 2, n. 4, p. 139-144, 2015.

MARTINELLI, F. *et al*. Advanced methods of plant disease detection. A review. Agronomy for Sustainable Development. v. 35, n. 1, p. 1-25, 2015

MOHANTY, S. P. *et al*. Using deep learning for image-based plant disease detection. Frontiers in plant science, v. 7, p. 1419, 2016.

OTSU, N. A Threshold Selection Method from Gray-Level Histograms. IEEE Transactions on Systems, Man, and Cybernetics, v. 9, n. 1, p. 62-66. 1979.

PAU, G. *et al*. EBImage—an R package for image processing with applications to cellular phenotypes. Bioinformatics, v. 26, n. 7, 2010. p. 979-981

PETHYBRIDGE, S.J., NELSON, S.C. Leaf Doctor: A New Portable Application for Quantifying Plant Disease Severity. Plant Disease, v. 99, n. 10, p. 1310-1316, 2015.

s and Electronics in Agriculture, 138, 2017, p. 92–104

SANNAKKI, S.S. *et al*. Leaf Disease Grading by Machine Vision and Fuzzy Logic. International Journal of Computer Technology and Applications. v. 2, p. 1709–1716, 2011.

SENGAR, N. *et al*. Técnica de visão computacional para identificação e quantificação da doença do oídio em folhas de cerejeira. Computing 100, 2018. p. 1189-1201

SIBIYA, M., SUMBWANYAMBE, M. An Algorithm for Severity Estimation of Plant Leaf Diseases by the Use of Colour Threshold Image Segmentation and Fuzzy Logic Inference: A Proposed Algorithm to Update a "Leaf Doctor" Application. AgriEngineering, v. 1, p. 205–219, 2019

SINGH, V. MISRA, A.K. Detection of plant leaf diseases using image segmentation and soft computing techniques, Information Processing in Agriculture, v. 4, n. 1, p. 41-49, 2017.

SLADOJEVIC, S. *et al*. Deep neural networks based recognition of plant diseases by leaf image classification. Computational intelligence and neuroscience, v. 2016, 2016.

Proposta e avaliação de uma solução baseada em Fog Computing para a Pecuária Leiteira de Precisão

Márcio Spenst[1,3], Renata W. Suñé[2], Marcos J. Yokoo[1,2], Leonardo B. Pinho[1]

[1] Programa de Pós-Graduação em Computação Aplicada – UNIPAMPA – Bagé – RS, leonardopinho@unipampa.edu.br

[2] Unidade Pecuária Sul – EMBRAPA – Bagé – RS, renata.sune@embrapa.br, marcos.yokoo@embrapa.edu.br

[3] Campus Bagé – IFSUL – Bagé – RS, marciospenst@ifsul.edu.br

RESUMO

O trabalho apresenta os desafios preliminares da concepção de um modelo de sistema computacional escalável e distribuído, aplicado à gestão e qualificação de processos existentes na cadeia do leite. O modelo proposto é baseado em uma arquitetura *cloud-fog-thing*, prevendo diferentes tipos de dispositivos finais microcontrolados, de baixo custo e baixa capacidade de processamento, instalados no sistema de ordenha (sala de ordenha e sala de máquinas), compostos por sensores e atuadores, capazes de se comunicar, de forma oportunística, com dispositivos móveis de capacidade intermediária presentes na camada "*fog*", bem como explorando mecanismos de comunicação veicular e de redes tolerantes a atrasos para comunicação nas diferentes camadas. Resultados preliminares, amparados por metodologia exploratória-experimental, indicam que as tecnologias podem auxiliar os produtores e a indústria de laticínios a produzir seus produtos com qualidade de acordo com a exigência da legislação em vigor. Estes resultados reforçam o potencial da solução proposta, evidenciando a relevância de uma pesquisa explicativa que permita quantificar o desempenho da solução em cenários realísticos.

PALAVRAS-CHAVE: Cadeia do Leite, Sensores e Atuadores, *Cloud-fog-thing*, Redes Veiculares.

ABSTRACT

The paper presents the preliminary challenges of the conception of a scalable, dis*tributed computational system model applied to the management and qualification of existing processes in the milk supply chain. The proposed model is based on a cloud-fog-thing*

architecture, providing different types of low-cost, low-capacity, microcontrollable final devices installed in the milking system (milking room and machine room), composed of sensors and actuators, capable of communicating opportunistically with mobile devices of intermediate capacity present in the fog layer, as well as exploiting mechanisms of vehicular communication and delay tolerant networks for communication in the different layers. Preliminary results, supported by exploratory-experimental methodology, indicate that the proposed system can help producers and the dairy industry to produce their products with quality as required by current legislation. These results reinforce the potential of the proposed solution, evidencing the relevance of an explanatory research that will allows to quantify its performance in realistic scenarios.
KEYWORDS: Milk supply chain, Sensors and Actuators, Cloud-fog-thing, Vehicular networks.

INTRODUÇÃO

O Rio Grande do Sul é o terceiro maior produtor de leite do país, produzindo em torno de 4,5 bilhões de litros de leite anualmente. Nos últimos anos, a produção de leite no estado vem sofrendo uma redução de números de produtores, devido ao baixo preço pago e os altos custos de produção. Aqueles que permanecem na atividade investem cada vez mais em tecnologias para garantir a qualidade do produto e um menor custo de produção (RIES, 2019).

Recentemente a Instrução Normativa (IN) 76/2018, a qual regulamenta a qualidade de leite cru refrigerado, determinou que no momento do recebimento do leite no laticínio a temperatura do mesmo deve estar em 7 ºC, sendo que anteriormente a temperatura aceita era de 10 ºC, excepcionalmente admitindo o recebimento com até 9 ºC (MAPA, 2018, a). Essa nova exigência trouxe preocupação para muitos produtores de leite, devido ao fato das propriedades produtoras de leite estarem localizadas distantes dos laticínios. Geralmente as estradas até as propriedades não estão em boas condições, na grande maioria das vezes sendo estradas de chão precárias, de modo que existe uma grande preocupação de que os laticínios precisarão investir em mais caminhões para realizar a coleta do leite ou ainda reformular as rotas de coleta, devido ao fato da Instrução Normativa estar mais exigente quanto à temperatura do leite armazenado no tanque do caminhão na chegada ao laticínio. Como resultado, vislumbra-se uma tendência excludente, isto é, um cenário no qual pequenos produtores pode acabar sendo excluídos da cadeia produtiva.

Por outro lado a Instrução Normativa 77/2018, regulamenta que o leite ordenhado e armazenado em tanque de expansão direta, deve atingir a temperatura de 4 ºC em até 3 horas

após a ordenha. Além disso, regulamenta que o leite estocado nos tanques de refrigeração deve ser amostrado pelo responsável pela coleta do leite, para análises em laboratório da Rede Brasileira de Laboratórios de Controle de Qualidade do Leite (RBQL), com uma frequência mínima mensal para avaliar parâmetros tais como: teor de gordura, contagem de células somáticas, contagem padrão em placas e resíduos de produtos de uso veterinário. Caso apresentem resultados fora do padrão por 3 meses consecutivos, deve ser interrompido o recolhimento do leite da propriedades (MAPA, 2018, b).

A partir deste contexto normativo, nota-se que uma das dificuldades na fase inicial da cadeia produtiva do leite é a necessidade de refrigeração, que tem dois objetivos principais: a inibição da multiplicação bacteriana e o aumento do tempo de armazenamento no local de ordenha, de modo a reduzir os custos de transporte até a indústria de laticínios. Depois de seguir todos os procedimentos de ordenha e higiene, o rápido resfriamento do leite para 4 ºC é a melhor maneira de evitar o crescimento de micro-organismos e alterações químicas. O tempo de resfriamento para atingir a temperatura de armazenamento, normalmente em 4 ºC, também é um fator crítico. Os tanques de resfriamento de leite a granel foram especialmente projetados para resfriar o leite a 4 ºC em um período máximo de 3 horas (DELAVAL, 2018b). Portanto, conclui-se que há necessidade de investimentos contínuos em boas práticas para prevenção da contaminação e do crescimento microbiano na cadeia produtiva do leite para reduzir problemas tecnológicos e econômicos na indústria de laticínios. Para Batalha (2007), a cadeia de produção pode ser entendida como “uma sucessão de operações de transformação dissociáveis, capazes de ser separadas e ligadas entre si por um encadeamento técnico”. Logo, há necessidade e oportunidade para qualificar diferentes etapas do processo da cadeia do leite, em especial as etapas compreendidas entre a ordenha e o recebimento do leite no laticínio, por meio do desenvolvimento de tecnologias no contexto de Pecuária Leiteira de Precisão, também reconhecida pelo termo em inglês *Precision Dairy Farming* (PDF).

De acordo com Pereira et al. (2015), PDF é definida como uma tecnologia para melhorar o controle da variabilidade animal e espacial, visando principalmente a economia social e ambiental. Tecnologias de precisão podem utilizar sensores para mensurar variáveis de algum situação em especial, gerando, assim, informações que visam um aconselhamento, uma tomada de decisão. Além disso o uso da pecuária de precisão contribui para uma produção de qualidade e a segurança do produto entregue ao laticínio.

A falta de conexão ou uma conexão não tão eficiente proporcionada pelas operadoras de telecomunicação é um problema que está se tornando crítico no meio rural, sobretudo porque vários equipamentos que chegam ao mercado possuem diversos recursos que somente podem

ser utilizados com uma conectividade de qualidade. Segundo o IBGE (2018), nas áreas rurais da região Sul do Brasil, apenas 53,1% dos domicílios pesquisados utilizam a Internet. Na maioria das vezes o acesso à Internet funciona somente na sede da propriedade rural, de modo que no campo ou mesmo na sala de ordenha esse recurso frequentemente não está disponível. Portanto, representa um cenário com significativos desafios de conectividade, caracterizada como a "habilidade de conectar ou comunicar com um outro computador ou sistema computacional" (MERRIAM-WEBSTER, 2019).

No cenário exposto, emerge de forma crescente a discussão sobre o potencial de emprego das Tecnologias da Informação e Comunicação (TIC) em pesquisas aplicadas à agropecuária, foco deste trabalho. Em especial, destaca-se a oportunidade de utilização da arquitetura *cloud-fog-thing*, ou simplesmente *fog/edge*, a qual busca explorar a concepção de escalabilidade com baixo custo de implantação, inclusive por meio dos avanços das pesquisas e padrões relacionados à comunicação veicular e redes tolerantes a atraso para mitigar a falta de conectividade, permitindo viabilizar a comunicação nas suas diferentes camadas.

Considerando os princípios de PDF e a evolução das TIC aplicadas à agropecuária, o principal problema passa a ser sua integração com o objetivo de melhorar o sistema produtivo para garantir que o leite, produzido por diversos produtores, tenha seus parâmetros de qualidade monitorados desde a ordenha, sendo resfriado de maneira correta e que o mesmo chegue até a indústria com a temperatura regulamentada pela IN 76/2018. Deste problema, surgem duas questões relevantes de pesquisa, uma principal e uma secundária: (i) É possível viabilizar o monitoramento da cadeia do leite, desde o processo de ordenha, passando pelo armazenamento em tanques de resfriamento e depois por um sistema de transporte coletivo, até a chegada a indústria, por meio de um sistema baseado na arquitetura fog/edge, associado a modelos de comunicação aplicáveis a cenários com baixa conectividade?; e (ii) É possível agregar elementos atuadores ao sistema proposto de modo a identificar precocemente anomalias no processo produtivo, evitando a contaminação do leite armazenado no tanque de resfriamento de um produtor e/ou no meio de transporte coletivo?

Partindo dos problemas enfrentados pelos produtores, pela indústria e pelos órgãos governamentais responsáveis por garantir a qualidade do leite e seus derivados, este trabalho, norteado pelas questões de pesquisa, tem como objetivo principal identificar alternativas tecnológicas relevantes para o contexto e, a partir destas, propor e avaliar a eficácia de um modelo de sistema computacional, com escalabilidade suficiente para atender um conjunto grande e crescente de produtores, capaz de monitorar (e eventualmente atuar sobre) variáveis relevantes da cadeia produtiva de leite em áreas rurais com baixa conectividade de rede.

MATERIAIS E MÉTODOS

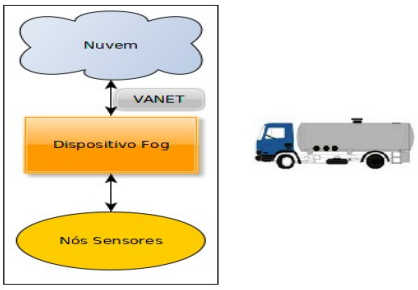
Foi realizada uma pesquisa exploratória-experimental, visando conhecer mais detalhes e possíveis soluções aplicáveis ao problema, partindo de uma revisão inicial da literatura identificando aspectos computacionais passíveis de aplicação na solução do problema, servindo de requisitos para a solução a ser concebida. A partir do melhor entendimento das variáveis do problema, foi concebido um modelo de solução amparado por aspectos aplicáveis identificados na fase de revisão da literatura, destacando os materiais e métodos usados na implementação de um protótipo funcional. Na sequência, desenvolveu-se uma avaliação experimental preliminar do protótipo de nó sensor, incluindo a instalação do dispositivo implementado no tanque resfriador de uma propriedade rural real. A avaliação permitiu identificar o comportamento potencial do dispositivo, evidenciando principalmente suas virtudes e limitações em atender os objetivos do trabalho.

Cabe destacar que, para a construção do modelo proposto, foram realizadas visitas a uma propriedade rural, com o objetivo de conhecer as instalações e a rotina do processo de ordenha e recolhimento do leite. O processo de ordenha desenvolvido na propriedade selecionada para o estudo de caso, localizada no município de Hulha Negra-RS, é realizado 2 vezes ao dia, na parte da manhã, iniciando por volta das 5 horas e 30 minutos, e na parte da tarde, com início por volta das 16 horas e 30 minutos. Em média é realizada a ordenha de 30 vacas lactantes, da raça holandesa, produzindo em torno de 400 a 500 litros de leite por dia, dependendo da época do ano. O sistema de ordenha utilizado é o conhecido como “espinha de peixe 4x2”, onde na sala de ordenha há espaço para quatro vacas de cada lado, totalizando oito vacas na sala simultaneamente. O leite ordenhado é enviado para um pré tanque de armazenamento, com capacidade de aproximadamente 10 litros, tendo como função de armazenar uma certa quantidade de leite para então ser enviado para o tanque resfriador de leite. A propriedade possui um tanque de expansão/resfriamento de leite com capacidade de armazenamento de 1.000 litros. A coleta do leite pelo caminhão tanque é realizada a cada dois dias, sempre pela parte da manhã, entre 7 e 8 horas. Portanto, em regra geral, cada coleta corresponde a quatro ordenhas, duas na parte da manhã e duas na parte da tarde.

MODELO PROPOSTO

Considerando os requisitos identificados, é proposto um modelo de sistema que visa a coleta automatizada de variáveis na cadeia produtiva de leite, além da disseminação das dados coletadas de cada produtor, envolvendo uma série de produtores de uma determinada região. Este é dividido em quatro elementos (Figura 1), detalhados a seguir.

Figura 1 – Modelo Proposto.



Fonte: (Autor, 2019)

Nós Sensores

Pretende-se coletar a temperatura ambiente, umidade ambiente, a temperatura do leite armazenado no tanque de resfriamento, a quantidade de leite armazenado, o consumo de energia elétrica e o pH do leite enquanto o mesmo se encontra dentro do pré tanque.

Nuvem

Em uma arquitetura normal os dados coletados pelos sensores são enviados diretamente para a nuvem, e somente depois processados. Neste processo uma grande quantidade de dados é enviada, exigindo assim que se tenha uma boa conectividade disponível, o que nem sempre é garantida. Ao utilizar *Fog Computing*, uma nova camada é introduzida, onde os dados são processados e assim reduzindo a quantidade de informação a ser enviada para a nuvem. (BONOMI et al., 2012).

Dispositivo Fog

A ideia baseia-se no fato do caminhão que faz a coleta do leite na propriedade rural, se conectar ao dispositivo *fog*, onde serão avaliadas diferentes tecnologias como *Wireless Personal Area Network* (WPAN) e *Wireless Local Area Network* (WLAN) para determinar aquela que possua as características adequadas para a comunicação e coleta dos dados armazenados no *fog*. Dados que já estariam validados através de um pré-processamento realizado na própria plataforma Arduino. A vantagem do processamento ser realizado na borda da rede, é que possui suporte a mobilidade e baixa latência de rede. *Fog Computing* suporta a aplicação da Internet das Coisas (ARIF; WANG; BALAS, 2018).

Comunicação entre Camadas - VANET/DTN

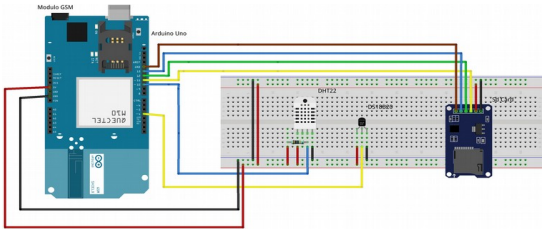
Redes veiculares Ad Hoc são unidades de armazenamento e processamento que podem se comunicar através de redes sem fio, cooperando entre si criando a rede. Permite a

comunicação de uma unidade para outra unidade ou de uma unidade para uma infraestrutura (WANG et al., 2015). Os dispositivos *fog* irão atuar como *data smarts* agrupando dados de produtores de uma determinada região, bem como servirem de interface de comunicação, por meio de princípios de *Vehicular Ad hoc Networks* (VANET) e *Delay Tolerant Networks* (DTN), com uma nuvem computacional na qual são armazenados dados consolidados de todos os produtores, na forma de um *data warehouse*, permitindo uma visão gerencial global de indicadores de qualidade do sistema de produção.

IMPLEMENTAÇÃO DO PROTÓTIPO

Foi desenvolvido de um protótipo utilizando Arduino, plataforma de prototipagem eletrônica que torna o processo de automação mais acessível a todos (ARDUINO, 2018). Na proposta do protótipo (Figura 2), foram utilizados alguns sensores e módulos próprios para Arduino.

Figura 2 – Protótipo de monitoramento de temperatura do resfriador de leite.



Fonte: (Autor, 2018)

Optou-se por utilizar o sensor DHT22 para monitorar a temperatura e umidade ambiente, tendo em vista que é um sensor com capacidade de medir a temperatura na faixa de -40 ºC a 80 ºC, juntamente com a umidade relativa do ar na faixa de 0 a 100%. Este foi escolhido por ser mais estável que o sensor DHT11, já que testes preliminares com este outro modelo mostraram uma grande variação da temperatura em pouco tempo. Concluiu-se que, mesmo possuindo um custo mais elevado, o DHT22 mostrou-se mais eficiente em relação ao DHT11. Para o monitoramento da temperatura do leite armazenado no tanque de resfriamento, foi utilizado o sensor a prova d'agua DS18B20. Sua faixa de medição é de -55 ºC até 125 ºC, com uma tolerância de 0,5 ºC. A escolha deste sensor se deu pelo fato do mesmo ser a prova d'agua, suportando locais mais úmidos para mensurar a temperatura de dentro do tanque de resfriamento. Um *Data Logger Shield* para Arduino foi utilizado para armazenar as informações coletadas pelo protótipo. São armazenadas as informações em um Cartão SD com capacidade de 8 GB, juntamente com a data e hora, obtidas através relógio de tempo real RTC DS1307 integrado com a placa. Para que o relógio funcione, é necessária uma pilha CR1220. Um *display* TM1637 foi adotado para mostrar informações como hora,

temperatura do tanque de resfriamento, temperatura e umidade do ambiente, atualizadas a cada quinze segundos. Dentre os vários *displays* disponíveis para Arduino a escolha pelo TM1637 se deu pelo fato de ser um visor com 4 dígitos, satisfatórios para a visualização das informações deste modelo, bem como ser de fácil prototipação e pelo seu reduzido custo.

Os sensores para verificação da quantidade de leite armazenada no tanque de resfriamento de leite, medir o pH e o consumo de energia não foram implementados nesta etapa do projeto. O modelo completo proposto prevê a coleta automatizada de dados tanto do pré tanque como do tanque de resfriamento. Realizando um pré-processamento destes dados coletados, espera-se incorporar elementos atuadores nos nós sensores, por exemplo, para acionar válvulas quando identificado o leite com alguma anomalia através do sensor de pH instalado no pré tanque.

A Tabela 1 mostra o custo do protótipo, com valores expressos em reais e dólar com cotação de R$ 4,10 na data de 10/09/2019.

Tabela 1 – Placas e sensores.

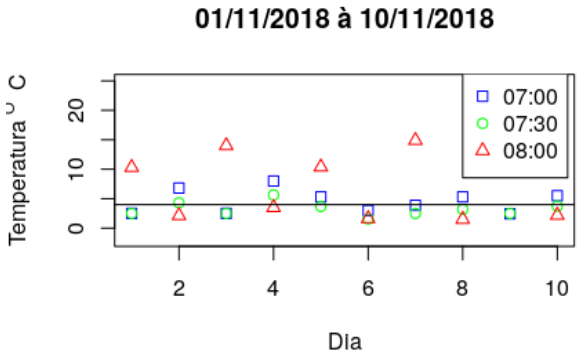
Descrição	**Real**	**Dólar**
Placa Arduino Uno	R$ 54,90	U$ 13,40
Sensor DS18B20	R$ 16,90	U$ 4,12
Sensor DHT22	R$ 49,90	U$ 12,17
Display TM1637	R$ 14,90	U$ 3,64
Data Logger	R$ 27,90	U$ 6,80
Fonte de Alimentação	R$ 39,90	U$ 9,73
Pilha CR1220	R$ 15,00	U$ 3,65
Cartão de Memória de 8 GB	R$ 25,00	U$ 6,09
Total	**R$ 244,40**	**U$ 59,60**

Fonte: (Autor, 2019)

Foram realizados alguns testes preliminares com o protótipo, com o intuito de verificar seu funcionamento. Em particular colocou-se o sensor de temperatura DS18B20 dentro de uma geladeira, aferindo após como os dados ficaram armazenados no cartão de memória. Após foram realizados alguns testes reais, em uma propriedade rural. O protótipo foi instalado em cima do tanque de resfriamento e o sensor de temperatura foi acondicionado na parte interna do tanque, não entrando em contato com o leite armazenado.

Foram realizadas coletas de dados preliminares entre os dias 01/11/2018 à 10/11/2018, sendo efetuada a leitura da temperatura interna do tanque de resfriamento, temperatura e umidade ambiente, a cada minuto. Neste período, a ordenha na parte da manhã se encerrava por volta das 7 horas, sendo realizada a coleta do leite pelo caminhão entre 7 horas e 8 horas. Registrou-se que as coletas foram realizadas nos dias 1, 3, 5, 7, 9 de novembro, estando o tanque de resfriamento com aproximadamente 1.000 litros em cada uma delas.

Figura 3 – Dados Coletados.



Fonte: (Autor, 2019)

A linha horizontal no gráfico (Figura 3) representa a temperatura de 4 ºC. Observou-se que nos dias em que não houve a coleta do leite (dias 2, 4, 6, 10), a temperatura sempre ficou dentro do padrão até às 8 horas da manhã. Nos dias em que houve a coleta do leite pelo caminhão (dias 1, 3, 5, 7 e 9), a temperatura entre 7 horas e 7 horas e 30 minutos se manteve dentro do padrão exigido pela IN, ou seja em 4 ºC.

CONCLUSÕES

A pesquisa exploratória-experimental relatada neste trabalho serviu para adquirir maiores conhecimentos sobre o domínio do problema e tecnologias potenciais para a concepção de um sistema eficaz para aumentar a eficácia da cadeia do leite. Em particular, identificou-se a necessidade da utilização de mais alguns sensores para a construção do modelo aqui proposto.

Após efetuados os testes preliminares, foi possível concluir que uma solução de baixo custo é viável de ser instalada em uma propriedade produtora de leite. Em particular, notou-se a necessidade de monitorar se o tanque de resfriamento está ligado ou desligado, pois ao ser realizada a coleta do leite pelo caminhão, o tanque é desligado, não havendo assim a necessidade de ser feito o monitoramento. Também foi verificado que existe a necessidade de ser coletado o momento exato do início da ordenha, para verificar com maior acurácia o tempo que o tanque de resfriamento está levando para chegar até a temperatura exigida na legislação.

O modelo proposto apresenta um instrumento de auxílio, tanto para o produtor como para a indústria, para que sejam monitoradas de forma mais eficiente variáveis relevantes da produção de leite, desde a coleta do leite no tanque de resfriamento localizado no produtor até a chegada do caminhão tanque na indústria. Com o desenvolvimento e validação deste

modelo, espera-se contribuir significativamente para a eficientização de diferentes etapas da cadeia, sobretudo dos processos de ordenha e de logística de coleta de leite cru, respeitando as normas nacionais e internacionais de qualidade que a cada ano se tornam mais rigorosas.

REFERÊNCIAS

ARDUINO, 2018. Site Oficial. Disponivel em: <https://www.arduino.cc/>. Acessado em: 01 dez. 2018.

ARIF, M.; WANG, G.; BALAS, V. Secure vanets: Trusted communication scheme between vehicles and infrastructure based on fog computing. p. 235–246, 2018.

BATALHA, M. Gestão agroindustrial. 3. ed. Sãp Paulo: Atlas, 2007.

BONOMI, F. et al. Fog computing and its role in the internet of things. In: Proceedings of the First Edition of the MCC Workshop on Mobile Cloud Computing. New York, NY, USA: ACM, 2012. (MCC '12), p. 13–16. ISBN 978-1-4503-1519-7. Disponivel em: <http://doi.acm.org/10.1145/2342509.2342513>.

DELAVAL. Porque resfriar o leite. Informações técnicas da DeLaval. 2018. Disponível em: <http://www.delaval.com.br>. Acessado em: 01 set. 2018.

IBGE. Pesquisa Nacional por Amostras de Domicílios Contínua. 2018. Disponível em: <https://biblioteca.ibge.gov.br/visualizacao/livros/liv101631_informativo.pdf>. Acessado em: 28 mai. 2019.

MAPA, 2018. Instrução Normativa N o 76, DE 30 de Novembro de 2018. Disponível em: <http://www.in.gov.br/materia/-/asset_publisher/Kujrw0TZC2Mb/content/id/52750137>. Acessado em: 02 jan. 2019.

MAPA, 2018. Instrução Normativa N o 77, DE 26 de Novembro de 2018. Disponível em: <http://www.in.gov.br/materia/-/asset_publisher/Kujrw0TZC2Mb/content/id/52750141/do1-2018-11-30-instrucao-normativa-n-77-de-26-de-novembro-de-2018-52749887>. Acessado em: 02 jan. 2019.

MERRIAM-WEBMASTER. **Connectivity.** Disponível em: <https://www.merriam-webster.-com/dictionary/connectivity>. Acessado em 20 set. 2019.

PEREIRA, L. et al. Pecuária leiteira de precisão: conceitos e tecnologias disponíveis. São Paulo: Embrapa Gado de Leite, 2015. 114 p.

RIES, J. Bovinocultura de Leite. 2019. Disponivel em: <http://www.emater.tche.br/site/area-tecnica/sistema-de-producao-animal/bovinos-de-leite.php#.XDODtMtKg-Y>. Acessado em: 02 jan. 2019.

WANG, M. et al. Real-time path planning based on hybrid-vanet-enhanced transportation system. IEEE Transactions on Vehicular Technology, v. 64, p. 1664–1678, 2015.

Start Green: Uma proposta de ferramenta pedagógica às crianças para promover interações com a natureza

Ana Elisandra Campos da Silva[1], *Leandro da Silva Camargo*[1], *Marcel da Silva Camargo*[1]

[1] Instituto Federal de Educação, Ciência e Tecnologia Sul-rio-grandense (IFSUL)

Bagé - RS - Brasil, anaecdasilva@gmail.com, leandrocamargo@ifsul.com.br, marcel@sigasistemas.com.br

RESUMO

As questões que envolvem o meio ambiente precisam ser trabalhadas como temática transversal na educação dos estudantes das séries iniciais, pois o futuro da humanidade depende dessa relação salutar estabelecida entre a natureza e homem, através do uso adequado dos recursos naturais disponíveis. Nesse viés, o presente trabalho tem como objetivo o desenvolvimento de uma aplicação para auxiliar professores da rede pública de ensino a despertar o interesse dos alunos quanto a produção do próprio alimento, mostrando que é possível o cultivo de alimentos sem a utilização de agrotóxicos e respeitando as características da área produtiva. A metodologia para concepção desse produto de software foi iniciada com uma conversa com educadores para avaliar a infraestrutura disponível nas escolas, bem como conhecer o espaço ocupado pelas hortas. Passando para a prototipagem e validação da proposta. Em seguida, foi iniciada a implementação do software e convergência das tecnologias adotas. Como resultado é esperado que o software permita observar as condições ambientais, como disponibilidade hídrica, temperatura e umidade do solo através de dados coletados por sensores conectados ao solo e apresentar esses dados em um avatar com ilustrações dinâmicas e com reações humanoides, tornando o processo mais divertido e útil permeando toda prática educacional. O sistema permite que tipos diferentes de culturas sejam pré-cadastradas, cada qual preservando suas características produtivas e identificação do avatar que ilustra o respectivo vegetal.

PALAVRAS-CHAVE: Sensores, Condições ambientais, Educação ambiental.

ABSTRACT

The issues surrounding the environment need to be addressed as a transversal theme in the students education in the initial grades, since the humanity future depends on the healthy relationship established between nature and man, appropriate through use natural resources available. In this bias, the present work aims to develop an application to help public school teachers network to arouse the students' interest in the production of their own food, showing that it is possible to grow food without using pesticides and respecting productive area characteristics. The methodology for designing this software product was initiated by a conversation with educators infrastructure available in schools, as well as space know gardens occupied. Prototyping on moving and proposal validation. Then, the software implementation and convergence technologies adopted was started. As a result, it is expected that the software will allow observing the environmental conditions, such as water availability, soil temperature and soil moisture through data collected by soil sensors connected and presenting this data in an avatar with dynamic illustrations and with humanoid reactions, making the process more fun and useful pervading all educational practice. The system allows different types of cultures to be pre-registered, each preserving its productive characteristics and identifying the avatar that illustrates the respective plant.

KEYWORDS: Sensors, Environmental conditions, Environmental education.

INTRODUÇÃO

Um projeto chamado horta escolar como eixo gerador de dinâmicas comunitárias, educação ambiental e alimentação saudável e sustentável, foi desenvolvido em 2005, na cidade de Bagé/RS, com recursos do Fundo Nacional de Desenvolvimento da Educação (FNDE/MEC), e apoio da Organização das Nações Unidas para a Agricultura e Alimentação (FAO). A ideia daquele projeto foi de proporcionar aos alunos a aquisição de conhecimentos em disciplinas escolares e educação ambiental. A colheita de verduras e legumes também enriquecerá a merenda escolar (FARIA, 2007).

Para tornar aquele projeto mais interativo e, consequentemente, despertar a atenção das crianças, a inserção de alguma tecnologia pode potencializar os objetos previstos anteriormente. Nesse sentido, se propõe o desenvolvimento de uma aplicação de software que permita o acesso a partir de qualquer dispositivo móvel, visando a aproximação do estudante com o meio ambiente, através de interações em tempo real das mudanças percebidas pelos sensores e apresentadas no dispositivo computacional.

A demanda crescente pela produção de alimentos em grande escala, mostra um aumento na utilização de produtos químicos para dirimir as perdas nas culturas. Contudo, a utilização desses produtos químicos acaba afetando de forma negativa a saúde das pessoas (CONTRERAS, 2011).

O uso generalizado dos agrotóxicos nas mais diferentes condições ambientais, muitos problemas começaram a ser percebidos e diagnosticados, tais como a ocorrência de resíduos em alimentos, a contaminação de solos e águas, o efeito em organismos não-visados e a intoxicação de trabalhadores rurais (CAMPAGNOLLA; BETTIOL, 2003). Nesse contexto, a produção de alimentos orgânicos nas escolas pode ser potencializada com o envolvimento dos estudantes desde a etapa de plantio até a colheita acompanhando todo o ciclo produtivo.

Segundo Aguiar e Passos (2014), a educação transdisciplinar pressupõe inclusão, sinônimo de cidadania, aliada a tecnologia é decisiva na promoção do desenvolvimento do ser humano, entendido como o acréscimo de possibilidades e potencial de cada um, pois colabora com a expansão da inteligência individual e coletiva, permitindo o pensar e o agir de forma dinâmica.

A tecnologia quando utilizada de forma correta, traz benefícios para as crianças à longo prazo, despertando a curiosidade para descobrir coisas novas. Favorecendo o processo formativo dentro e fora da escola, pois, a tecnologia utilizada de forma correta favorece o relacionamento interpessoal entre os alunos e mantém o foco das atividades escolares (PAIVA; COSTA, 2015). A ferramenta de software proposta nesse trabalho, visa auxiliar os docentes às práticas interdisciplinares e buscando o envolvimento e o despertar do interesse dos estudantes com a produção de alimentos de forma consciente, respeitando os limites ecossistêmicos e hídricos.

Trabalhos Correlatos

Existem diversos trabalhos relacionados à automatização de hortas, um desses trabalhos prevê a utilização de tecnologia para agricultura de precisão, que propõe o desenvolvimento de um protótipo de um sistema de irrigação preciso automatizado, de baixo custo e que possa ser utilizado por pequenos agricultores. Nesse projeto Barbosa (2013) pretende controlar a umidade de estufas agrícolas baseado na tecnologia Arduino juntamente com sensores.

A proposta apresentada por Lino *et al.* (2017) se refere a um projeto desenvolvido no Departamento de Engenharia Agrícola da Universidade Federal do Ceará, tem como objetivo otimizar o processo de irrigação, de forma a suprir a necessidade de água das culturas e

também a redução do consumo de água no campo, como também potencializar o tempo do trabalhador rural de maneira acessível aos pequenos produtores.

O trabalho de Lordani e Arcoverde (2013) utiliza software aplicativo gratuito para registro de dados na horta em ambiente escolar, através da ferramenta QR-code, um aplicativo para ler estas informações e outro para gerar, a ideia é codificar todas as Horticultura. Os estudantes por meio dos celulares com o aplicativo de leitura podem fazer o levantamento de informações *in loco* na horta.

Já a proposta de Valpreda e Zonda (2016) é a que mais se assemelha ao trabalho proposto neste artigo, pois traz um kit de sensor de jardinagem para crianças, uma das partes consiste em um mini computador composto por sensores de umidade do solo, temperatura e luminosidade que é colocado diretamente no solo. Esses sensores enviam as informações para um aplicativo que por sua vez apresenta um avatar interpretando emoções conforme o estado da planta. A Tabela 1 apresenta um comparativo das funções existentes em cada projeto.

Tabela 1 - Comparativo entre os trabalhos correlatos.

Identificação do trabalho	App/Sistema	Controla consumo de água	Sugestão de colheita	Identifica a cultura	*Ranking* de turmas
Irrigação Automatizada (BARBOSA, 2013)	X	X	X		
Horta Didática (LINO *et al*., 2017)		X		X	
Registro de Dados na Horta (LORDANI; ARCOVERDE, 2013)	X			X	
A Gardening Sensor Kit For Children (VALPREDA; ZONDA, 2016)	X	X		X	X
Autonomous Green (Autor, 2019)	X	X	X	X	

Fonte: Autor (2019)

Os trabalhos correlatos pesquisados reforçam a importância da proposição desse trabalho, ao inserir a temática do meio ambiente de forma mais pedagógica e orientada às crianças das séries iniciais, embora existam algumas semelhanças entre tais trabalhos, em nenhum deles foi observado a utilização de um avatar que esboça as reações da cultura conforme a necessidade de insumos e características percebidas no ambiente.

MATERIAL E MÉTODOS

O ponto de partida para a concepção desse produto de software foi uma conversa com educadores para avaliar a infraestrutura disponível nas escolas, bem como conhecer o espaço

ocupado pelas hortas. Ao iniciar o processo de criação do produto foi definido a utilização do ciclo de vida do software seguindo o modelo espiral, já que esse modelo visa pequenas entregas com produtos que possam ser utilizados mesmo que ainda não tenha sido finalizado o projeto. Neste contexto foram desenvolvidos protótipos até chegar em um produto final, a estrutura do projeto inicia-se com o planejamento, na sequência é realizado a modelagem, logo após, a construção do projeto e então é feita a pequena entrega com um produto funcionando. Por fim o processo recomeça até chegar à um produto final.

Comunicação

Para realização do trabalho, foi realizado uma reunião com a Secretaria Municipal de Educação visando apresentar a proposta de criação de uma solução de software para incentivar os alunos das séries iniciais, a despertar o interesse pelo cuidado com a natureza através do cultivo de hortas.

Tal incentivo ocorrerá através de um aplicativo que demonstra o status do cultivar com reações ilustrativas. A partir daí foram prototipadas as telas da aplicação para validação junto aos docentes e alguns estudantes selecionados para participar do projeto. Após a validação inicial foi realizado o planejamento de custos, cronograma e a análise dos riscos.

Materiais

A arquitetura do projeto foi dividida em dois níveis, um correspondendo à toda a parte do software, que não haverá custos no início e outro nível representando a parte do hardware onde estão conectados os sensores e placas. Este nível apresenta um custo médio de R$ 287,00 (duzentos e oitenta e sete reais) como estimativa apresentada na Tabela 2.

Tabela 2 - Estimativa de custos dos componentes de hardware para a montagem do protótipo.

Hardware	**Preço em R$**
Arduino Uno R3	55,00
Raspberry PI	180,00
Sensor de luminosidade	3,00
ESP8266	24,00
Higrômetro de Umidade do Solo	15,00
Sensor de temperatura	10,00

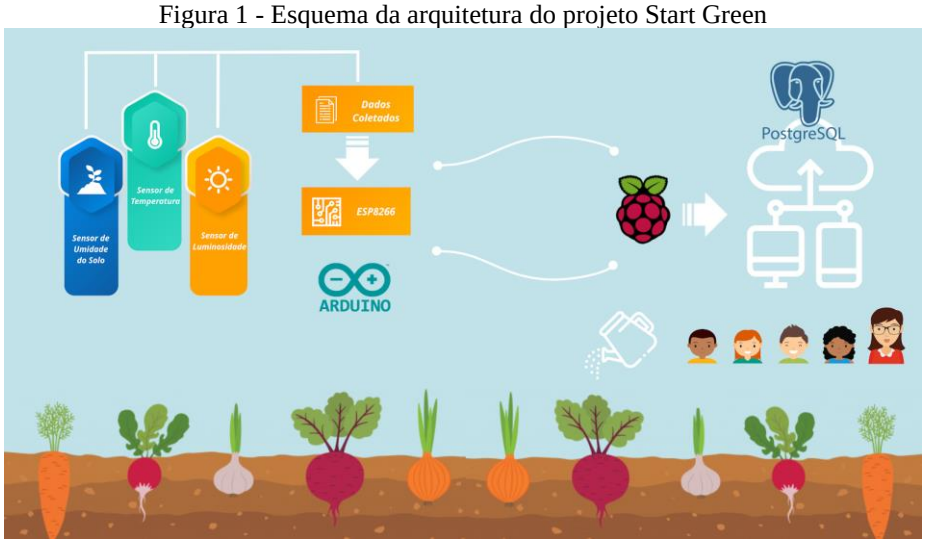
Fonte: Autor (2019)

O custo total do projeto depende, em partes, da infraestrutura de rede disponível nas escolas, bem como das dimensões da área produtiva. Contudo, os valores são acessíveis, frente aos benefícios e a possibilidade de replicação do projeto em várias escolas. O custo mais elevado que trata da aquisição do Raspberry PI, como uma estrutura remota de persistência dos dados pode suportar o compartilhamento por vários usuários, diluindo o custo do investimento inicial.

RESULTADOS E DISCUSSÃO

A Figura 1 apresenta a arquitetura concebida neste projeto, com o conjunto de sensores que foram instalados na placa Arduino para a coleta dos parâmetros ambientais. Nessa arquitetura os dados são enviados para o servidor na nuvem, e inseridos no banco de dados PostgreSQL, essa coleta é suportada por um microcomputador Raspberry PI. E após a inserção no banco, os dados serão resgatados pela aplicação que mostra aos usuários de forma intuitiva.

Figura 1 - Esquema da arquitetura do projeto Start Green



Fonte: Autor (2019)

A parte da aplicação web do projeto proposto começou a ser criado em março de 2019, vem sendo desenvolvido na linguagem PHP que segundo Dall'Oglio (2018) é atualmente uma das linguagens mais utilizadas no desenvolvimento de aplicações web. Para o desenvolvimento do layout está sendo utilizado o framework Materialize que tem uma estrutura front-end responsiva e moderna que é baseada na ferramenta Material Design.

Uma API (*Application Programming Interface*) é um conjunto de rotinas, padrões e instruções de programação que permitem que desenvolvedores criem aplicações que possam acessar determinado serviço na internet (DOS SANTOS, 2013). Foi utilizado uma API de previsão do tempo para que fosse possível a previsão de chuva, frio ou dias ensolarados,

assim como a possibilidade do disparo de notificações referente à previsão de chuva para o dia seguinte.

A aplicação poderá ser acessada por qualquer navegador web, para otimizar a utilização pelo usuário, a aplicação poderá ser salva no celular ou computador como se fosse um aplicativo. Para que isso seja possível o sistema será convertido em PWA (*Progressive Web Apps*) aplicativo da web, que pode ser instalado e distribuído sem *marketplaces*, funciona sem conectividade com a internet e se parecem com aplicativos comuns (BIORN-HANSEN; MAJCHRZAK; GRONLI, 2017).

O sistema recebe os dados da planta através de sensores conectados à uma placa Arduino, em que pese existem diversos modelos de placa Arduino, neste projeto foi utilizado o Arduino UNO R3, por ter o melhor custo-benefício. O modelo UNO R3 possui presença de conectores centrais, terminais que permitem mais conexões como os sensores e apresenta conector micro-USB para alimentação da placa.

Foram utilizados três sensores, um sensor de temperatura LM35 que é capaz de medir a temperatura em graus centígrados, este sensor tem boa precisão e sensibilidade, podendo ser utilizado como termômetro, além de possuir um baixo custo e ser fácil de ser encontrado no Brasil (LUNARDELLI, 2014).

Um sensor de umidade do solo possui duas hastes de prova, que vão diretamente no solo, o funcionamento é baseado na resistividade elétrica do solo, ou seja, à medida que o solo é umedecido, a sua resistência elétrica diminui, permitindo a corrente fluir da haste energizada, para a haste de medição, conectada ao controlador principal (DE MELO; SOUSA; SARAIVA; MACEDO-FILHO, 2016).

O sensor de luminosidade que é um resistor dependente de luz ou fotoresistência (LDR) é um componente eletrônico passivo do tipo resistor variável, ou seja, sua resistência varia conforme a intensidade da luz que incide sobre ele. Este sensor é comumente utilizado na iluminação pública (LIMA; NOBRE; DE ALENCAR, 2016).

Para gravar os dados coletados no banco de dados foi necessário a integração do Arduino com o microcomputador Raspberry PI. A utilização das duas placas gera um desempenho melhor, já que a placa Arduino é melhor quando utilizada para conexões de hardware e o microcomputador Raspberry responde melhor às integrações de software. O banco de dados em questão é o PostgreSQL, foi escolhido por ser um software de código aberto e possuir alta performance.

O sistema foi submetido à testes para garantir a qualidade do projeto, primeiramente foi realizado um teste de segurança que garantirá que os dados sejam armazenados e

disponibilizados de forma segura para o usuário. O teste funcional garante se a aplicação proposta cumpra o que foi sugerido. Também foi efetuado um teste de usabilidade que é um teste focada na experiência do usuário, consistência da interface, layout e acesso às funcionalidades.

O primeiro protótipo entregue foi uma parte do sistema, que abrangia login, e esse primeiro protótipo possibilita realizar inserção de culturas e de professores no sistema. A segunda entrega foi o cadastro de plantio e a criação de um avatar ainda estático, sem as reações propostas, logo a próxima entrega o avatar já exibia as reações de forma dinâmica. O último protótipo entregue foi o case com os sensores e as placas e a integração entre sistema e os sensores de coleta dos dados do solo.

O usuário pode acessar o sistema através do navegador, sem perder memória do dispositivo que estiver sendo utilizado. A interface do sistema ficou bem intuitiva com um aspecto de clareza e objetividade, além disso foi alcançado o objetivo de apresentar diretamente no navegador um sistema equivalente à um aplicativo.

A ilustração da planta possui diversas interações, como dicas de plantio e até reações ao receber toques na tela, a fim de não deixar o sistema cair em desuso e desinteresse. O sistema está integrado com API de previsão do tempo, propiciando a previsão em tempo real e com ilustrações dinâmicas no cenário do respectivo avatar. Na Figura 6 é exibido algumas telas com reações do avatar que variam de acordo com os dados coletados para acompanhamento em tempo real.

Figura 1 - Telas da aplicação (A) parâmetros ambientais, (B) notificações e (C) cadastro de nova cultivar.

(A) (B) (C)



Fonte: Autor (2019)

As placas e sensores foram introduzidas de forma organizada dentro da estrutura de um recipiente reciclado, com intuito de suportar a exposição às intempéries no ambiente externo. Tal estrutura foi totalmente vedada para que não houvesse infiltração e consequentemente produzisse avarias nos dispositivos. A Figura 2 mostra estrutura para a coleta dos dados.

Figura 2 - Protótipo com os sensores na (A) estrutura fechada e (B) parte interna do protótipo.

(A) (B)



Fonte: Autor (2019)

A estrutura é bastante compacta e gera pouca emissão de calor, dessa forma permite a operação por um período de tempo, sem a necessidade de manutenção ou preocupação com questões climáticas quando exposta em ambientes externos.

CONCLUSÕES

O sistema Start Green atende aos requisitos elencados junto à profissionais da educação como um instrumento que auxilie os professores no estímulo de estudantes para o envolvimento com as questões ambientais, ao incluir em sala de aula esse tema de forma transversal e, ainda, como incentivo à produção do próprio alimento, de maneira orgânica. Evidenciando o uso dos recursos naturais e a forma correta de tratar cada cultivar. O sistema possui a capacidade de apresentar aos alunos de forma divertida os conceitos e as tarefas demandadas em uma plantação.

Ao analisar o projeto em funcionamento, se verificou uma possível expansão futura, que trata da criação de um mecanismo com tarefas a serem realizadas na horta e um *ranking* de quais as turmas obtiveram o melhor desempenho e consequentemente conseguiram realizar a colheita de produtos de boa qualidade. A ideia com isso é de realizar uma premiação ao final de cada ano letivo, ampliando assim a motivação para cuidar da melhor forma possível dos recursos ambientais, bem como a capacidade de produção do próprio alimento, de forma saudável, sustentável e respeitando as características de cada cultura.

REFERÊNCIAS

AGUIAR, I. A.; PASSOS, E. A tecnologia como caminho para uma educação cidadã. p.1-24, USP (FAU), 2014.

BARBOSA J. W. Sistema de Irrigação Automatizado utilizando a plataforma Arduino. Fundação Educacional do Município de Assis – FEMA - Assis, 2013, 57 p.

BIORN-HANSEN, A.; MAJCHRZAK, T. A.; GRANLI, T. M. Progressive Web Apps: The Possible Web-native Unifier for Mobile Development. In WEBIST, p. 344-351, 2017.

CAMPAGNOLLA, C.; BETTIOL, W. Métodos alternativos de controle fitossanitário. Jaguariúna: Embrapa Meio Ambiente - Livro científico (ALICE). 2003. 279 p.

CONTRERAS, J. A modernidade alimentar: entre a superabundância e a insegurança. História: Questões & Debates, [s.l.], v. 54, n. 1, p.19-45, 30 jun. 2011.

DALL'OGLIO, P. PHP Programando com orientação a Objetos. Novatec Editora. 2018.

DE MELO, T. R. Sistema automático de irrigação de baixo custo. Universidade Estadual do Piauí, UESPI. 2016.

DOS SANTOS, M. C. Conversando com uma API: um estudo exploratório sobre TV social a partir da relação entre o twitter e a programação da televisão. Revista Geminis, v. 4(1), p. 89-107. 2013.

FARIA, S. FNDE começa Horta Escolar em Bagé. Educação Básica. maio, 2007. Disponível em: < http://portal.mec.gov.br/ultimas-noticias/211-218175739/3510-sp-1876026392> Acesso em: 31 de março de 2019.

LIMA, E. M. S.; NOBRE A. Y. M; DE ALENCAR, R. A. E. Automação Residencial de baixo custo com Arduino Mega e Ethernet Shield. Departamento de Eletrônica, UTFPR, Paraná, 2016, 44 p.

LINO, D. R.; NETO, B. P. L.; TEIXEIRA, A. D. S. Irrigação automatizada com plataforma de desenvolvimento Arduino na horta didática da Universidade Federal do Ceará, Irriga - Edição Especial IRRIGA & INOVAGRI, v. 1, n. 1 2017.

LORDANI; ARCOVERDE. Utilização de software aplicativo gratuitos para registro de dados na horta. OS DESAFIOS DA ESCOLA PÚBLICA PARANAENSE NA PERSPECTIVA DO PROFESSOR PDE, v. 1, p. 1-18, 2013.

LUNARDELLI, M. V. P. Sensor de temperatura utilizando plataforma Arduino. Universidade do Vale da Paraíba, 2014.

PAIVA, N. M. N.; COSTA, J. S. A influência da tecnologia na infância: desenvolvimento ou ameaça, p.8-16, 2015.

VALPREDA, F.; ZONDA, I. Grüt: A Gardening Sensor Kit for Children. Sensors, [s.l.], v. 16, n. 2, p.231-245, 16 fev. 2016.



Modelagem e avaliação de desempenho de banco de dados espaço-temporal aplicado a experimento de manejo de pastagem

Ricardo Robaina[1], *Alexandro Schaffer*[1], *Leonardo Pinho*[1], *Gustavo Trentin*[2]

[1]Universidade Federal do Pampa - Bagé, Rio Grande do Sul, Brasil
ricardorobaina11@gmail.com, {alexandroschafer, leonardopinho}@unipampa.edu.br

[2]Embrapa Pecuária Sul - Bagé, Rio Grande do Sul, Brasil
gustavo.trentin@embrapa.br

RESUMO

A alta demanda por produção de alimentos intensifica a necessidade de aumento de produtividade em sistemas de agricultura e de pecuária. A Agropecuária de Precisão leva em consideração a variabilidade espaço-temporal presente nos sistemas de produção em busca deste objetivo. O mapeamento dessa variabilidade gera um grande volume de dados, o que torna os bancos de dados tradicionais (padrão SQL) ineficientes. O presente artigo apresenta a modelagem, a implementação e a avaliação de desempenho de um banco de dados espaço-temporal padrão NoSQL aplicado a um experimento de manejo de pastagem.

PALAVRAS-CHAVE: Banco de Dados, Dados Georreferenciados, MongoDB, Modelagem Espacial.

ABSTRACT

The high demand for food production intensifies the need for productivity. Precision Agriculture uses temporal and space variability present in the production systems in search of this demand. Mapping this variability generates a large volume of data, making inefficient traditional SQL databases (SQL standard). The present paper presents the modeling, implementation, and performance evaluation of a NoSQL-based spatiotemporal database applied to the data collection of a pasture experiment.

KEYWORDS: Data Base, Geospatial Data, MongoDB, Spatial Data Design.

INTRODUÇÃO

Nos últimos 50 anos, houve a duplicação da população do planeta, fato que intensificou a demanda pela produção de alimentos em geral. O Estado do Rio Grande do Sul é um dos maiores

produtores de alimentos do Brasil, atuando fortemente na produção agropecuária. Em relação à pecuária, segundo dados do censo agropecuário de 2017, o estado possui o sexto maior rebanho de gado do país, cerca de onze milhões e meio de animais distribuídos em mais de duzentas e sessenta mil propriedades rurais (IBGE, 2017). Além disso, no ano de 2018, o valor bruto da produção (VBP) de bovinos no Brasil foi de mais setenta e seis bilhões de reais, o que representa cerca de treze por cento do VBP total do país (MAPA, 2018). Nesse contexto, a quantidade finita de recursos naturais em nosso planeta em conjunto com a alta demanda por produção alimentar resulta em uma necessidade de maximização da produtividade em sistemas de produção de alimentos (SAATH; FACHINELLO, 2018).

A Agropecuária de Precisão (AP) pode contribuir para o aumento da produtividade através do mapeamento das variabilidades espaço-temporais do sistema de produção, através do emprego de dispositivos computacionais. No processo produtivo, tais dispositivos são empregados nas tarefas de coleta, armazenamento, processamento e análise de grandes quantidades de dados, com o objetivo de produzir conhecimento útil e auxiliar os processos de tomada de decisão do produtor rural (KAMILARIS; KARTAKOULLIS; PRENAFETA-BOLDÚ, 2017).

Na pecuária, a produtividade está diretamente relacionada com o tipo de manejo de pastagem utilizado. Atualmente, as técnicas de manejo de pastagem mais utilizadas levam em consideração a altura da vegetação (COSTA; QUEIROZ, 2013). O crescimento vegetal, por sua vez, está relacionado com a temperatura do ar à qual a planta está exposta. Esta relação é dada devido ao efeito que a temperatura exerce na velocidade das reações e dos processos internos do vegetal. Tais processos ocorrem adequadamente somente entre certos limites térmicos, que variam dependendo da espécie (PEREIRA; ANGELOCCI; SENTELHAS, 2007).

A medição da altura da pastagem pode ser realizada de forma empírica ou através da utilização de instrumentos de medida específicos, como a régua de manejo apresentada em (COSTA; QUEIROZ, 2013). Tanto a medição empírica quanto a utilização de instrumentos de medida possuem um custo operacional alto, isto é, é necessário o deslocamento até a área de interesse para a realização da medição, muitas vezes essas áreas sendo de difícil acesso. Uma forma alternativa de mensurar a taxa de crescimento de um vegetal é através da análise de dados meteorológicos, com base no sistema de unidades térmicas ou graus-dia. Esse método pressupõe que a planta necessita de uma certa quantidade de energia para o seu desenvolvimento, representada pela soma térmica do período de cultivo (FILHO et al., 2008).

A obtenção de dados meteorológicos em determinado local no espaço pode ser realizada por meio de sensores. A coleta automatizada de dados gera continuamente um grande volume de dados heterogêneos que necessitam ser armazenados, levando a um cenário de *Big Data* (WOLFERT et al., 2017). Karmas, Tzotsos e Karantzalos (2016) apontam que nesses casos os bancos de dados padrão *Not only SQL* (NoSQL) são mais adequados em comparação com os bancos de dados relacionais (SQL). Dentre estes, o banco de dados MongoDB é um dos mais utilizados atualmente, segundo (KAMILARIS; KARTAKOULLIS; PRENAFETA-BOLDÚ, 2017). Perondi et al. (2019) utilizam esse banco de dados em uma ferramenta de suporte à decisão para

escolha de época de semeadura e que integra o projeto AgroClimate (AGROCLIMATE, 2019). Baker et al. (2014) fazem uso da mesma tecnologia de banco de dados em um sistema de predição da ocorrência de doenças no cultivo de batata. O banco de dados MongoDB foi utilizado em um estudo que avalia o efeito de ondas eletromagnéticas, utilizadas para comunicação sem fio, em colônias de abelhas (HENRY et al., 2019). Em Robaina e Camargo (2019) foi realizado uma análise acerca da utilização de dados georreferenciados em bancos de dados NoSQL na Agropecuária de Precisão. Para isso, os autores fizeram uma revisão bibliográfica no periódico *Computers and Electronics in Agriculture*. Tal trabalho conclui que os bancos de dados NoSQL são mais adequados para cenários de AP pois possuírem uma maior flexibilidade e escalabilidade em relação aos bancos de dados relacionais. Além disso, foi identificada a sub-utilização do banco de dados MongoDB nesse contexto, pois o banco de dados possui funcionalidades que se enquadram em demandas de aplicaçações de AP como: suporte a dados e a consultas georreferenciadas; ferramenta para visualização de dados (MongoDB Compass) e ferramenta voltada a dispositivos móveis (MongoDB Mobile) (MONGODB, 2019b), concebida para utilização em dispositivos de *Internet of Things* (IoT). Com o objetivo de identificar a tecnologia de bancos de dados mais adequada nestes cenários, Fatima e Wasnik (2016) realizam uma comparação entre as tecnologias de bancos de dados Relacional (SQL) e NoSQL, utilizando os bancos de dados MySQL e MongoDB, respectivamente. A partir da simulação de uma rede de sensores, os autores concluem que o banco de dados MongoDB possui uma melhor performance e escalabilidade em relação ao banco de dados MySQL. Além disso, os autores ressaltam que a propriedade ACID não é presente nos bancos de dados do padrão NoSQL que em contrapartida possuem a estrutura dos dados flexível.

Para lidar com entidades que possuem localização no espaço e que podem sofrer alteração ao longo do tempo, é possível utilizar os conceitos e técnicas de bancos de dados espaço-temporais. Um banco de dados espaço-temporal é um banco de dados que incorpora conceitos de bancos de dados espaciais, temporais e espaço-temporais, possibilitando o registro e a manipulação simultânea dos aspectos espaciais e temporais dos dados (CHEN, 2017). Entre os bancos de dados do padrão NoSQL, o MongoDB destaca-se por possuir suporte a dados georreferenciados.

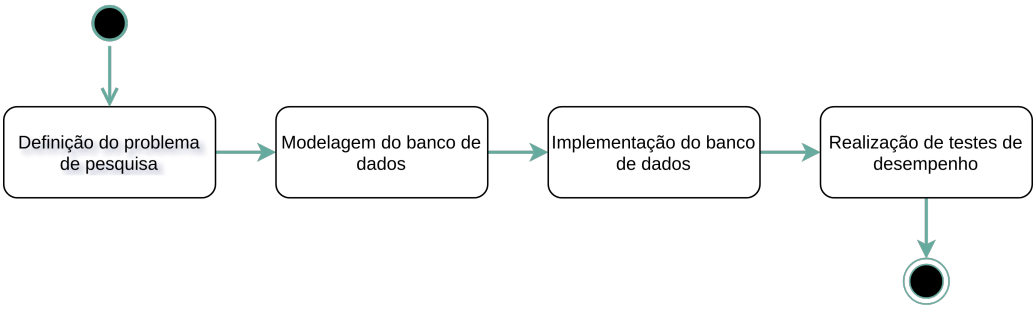
O presente artigo apresenta questões referentes à modelagem, à implementação e à avaliação de desempenho de um banco de dados espaço-temporal, voltado ao armazenamento de dados coletados automaticamente por redes de sensores, em um experimento de manejo de pastagem.

METODOLOGIA

A metodologia de pesquisa do presente trabalho foi definida em quatro etapas principais. A Figura 1 ilustra, em um diagrama de sequência, as etapas definidas nesta metodologia.

Durante a primeira etapa da presente metodologia, foi definido o problema de pesquisa. Tal problema foi caracterizado pela modelagem, implementação e avaliação de desempenho de

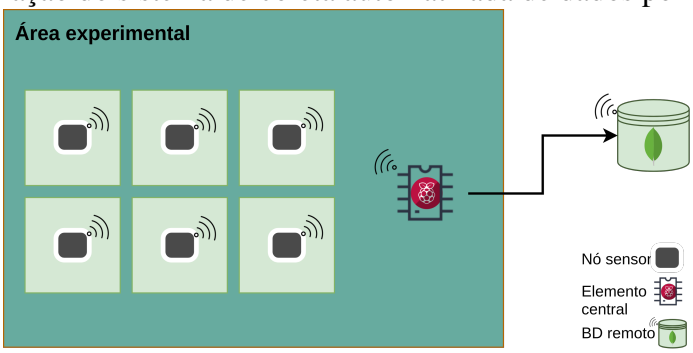
Figura 1: Diagrama de sequências das etapas da metodologia



Fonte: Próprio Autor (2019)

um banco de dados espaço-temporal voltado ao armazenamento de dados de um experimento de coleta automatizada de dados. Nessa etapa, foi caracterizado o estudo de caso a ser aplicado no experimento. Dessa forma, pôde-se delimitar os tipos de dados coletados, a partir dos elementos de coleta vislumbrados. O estudo será realizado nas imediações da Embrapa Pecuária Sul, situada em Bagé, na Região da Campanha do Rio Grande do Sul, no período de abril de 2019 a maio de 2020. A Figura 2 ilustra a organização do estudo de caso idealizado para a presente pesquisa.

Figura 2: Ilustração do sistema de coleta automatizada de dados por redes de sensores



Fonte: Próprio Autor (2019)

Com a finalidade de avaliar o crescimento de pastagens com base nas condições meteorológicas do ambiente, será realizado um experimento contendo redes de sensores para a coleta destes dados. No interior da área experimental, serão cultivados seis canteiros com a cultivar capim-sudão BRS Estribo. Cinco destes canteiros serão utilizados para avaliar o crescimento da forrageira em cenários de corte ideais. O sexto canteiro será utilizado no experimento de identificação da taxa de crescimento do capim-sudão quando submetido a diferentes alturas de corte. No centro de cada um desses canteiros será instalado um nó sensor. Esses nós sensores irão mensurar os valores de temperatura e umidade em determinados intervalos de tempo e se comunicarão com um elemento central através rede sem fio. O elemento central, por sua vez, possui a responsabilidade de transmitir os dados coletados para um banco de dados remoto. Dessa forma, a rede de sensores é composta pelos nós sensores e pelo elemento de comunicação central. Além de coleta automatizada de dados na área experimental, futuramente o sistema de

coleta automatizada de dados deverá armazenar dados meteorológicos registrados pelo Instituto Nacional de Meteorologia (INMET) e do *National Oceanic and Atmospheric Administration* (NOAA).

Na segunda etapa de desenvolvimento do trabalho foi realizada à modelagem do banco de dados espaço-temporal. Em um primeiro momento, analisou-se os dados que serão gerados durante o experimento. Essa análise levou em consideração o tipo de dado, a frequência da coleta, além da existência das componentes espacial e temporal nestes dados. Em um segundo momento, empregou-se a ferramenta OMTG-Designer (CS+X, 2019) na confecção de um diagrama utilizando o modelo OMT-G que descreve em detalhes o banco de dados projetado.

Por fim, na última etapa da presente pesquisa o banco de dados foi implementado e realizou-se testes de desempenho. Para isso, foram implementados programas na linguagem de programação Python com o objetivo de simular o comportamento dos sensores, isto é, gerando um grande volume de dados continuamente. Esses dados foram inseridos de forma automatizada no banco de dados MongoDB e as medidas de tempo empregadas nestas operações foram computadas.

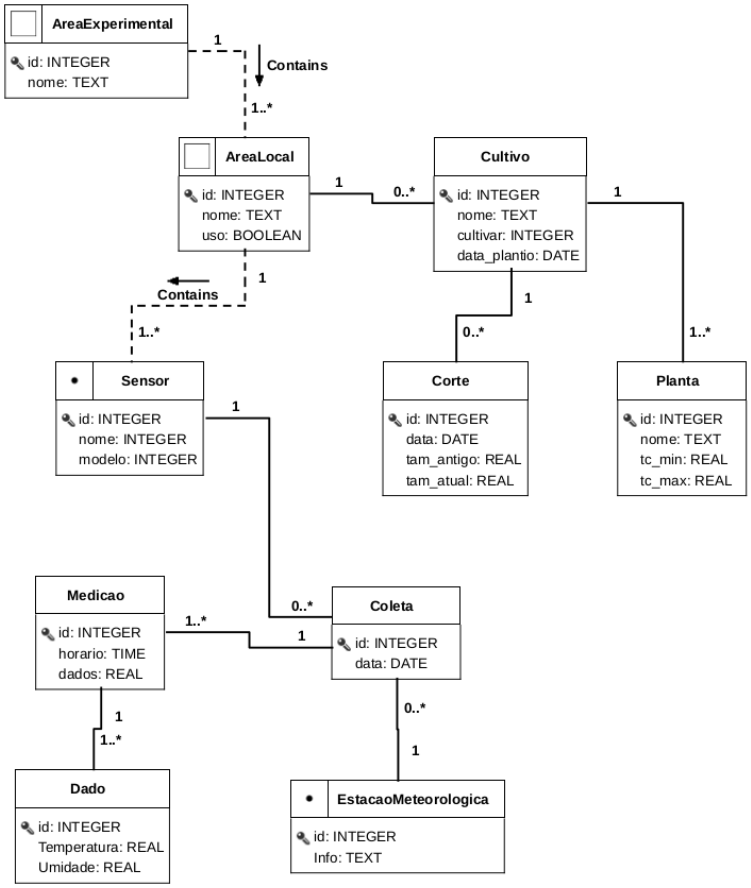
RESULTADOS E DISCUSSÃO

MODELAGEM

A partir da contextualização do problema foi possível identificar as entidades referentes ao estudo proposto, bem como os relacionamentos entre elas. Utilizou-se a ferramenta OMT-G Designer para criar um diagrama OMT-G que ilustra a organização do dados na base de dados. A Figura 3 apresenta o modelo OMT-G desenvolvido.

Os dados com componentes espaciais identificados foram: a área experimental, a área local, o sensor e a estação meteorológica. As áreas (*área experimental* e *área local*) foram modeladas como polígonos enquanto o *sensor* e a *estação meteorológica* foram modelados como pontos, conforme definido na especificação de dados GeoJSON, utilizada pelo MongoDB ((IETF), 2016). Cada área experimental contém uma ou mais áreas locais, nas quais será cultivada a pastagem. Na área local, será instalado um nó sensor. Tanto os sensores quanto as estações meteorológicas geram dados em determinados intervalos de tempo. A componente temporal foi inserida nas entidades *coleta* e *medição*. As coletas agrupam as medições por data de criação, enquanto as medições agrupam os dados por horário de criação. As entidades *cultivo*, *planta* e *corte* armazenam informações referentes à planta cultivada. A entidade *dado* contém todas as informações captadas por um nó sensor. Como o banco de dados foi modelado em um banco de dados padrão NoSQL, não existe um *schema* restrito. Dessa forma, o modelo pode ser adequado em tempo de execução à eventual ausência ou excesso de dados, visto que os dados são armazenados na forma de documentos, e estes podem conter outros documentos, como o especificado no formato JSON (ECMA, 2017).

Figura 3: Diagrama OMT-G do banco de dados projetado



Fonte: Próprio Autor (2019)

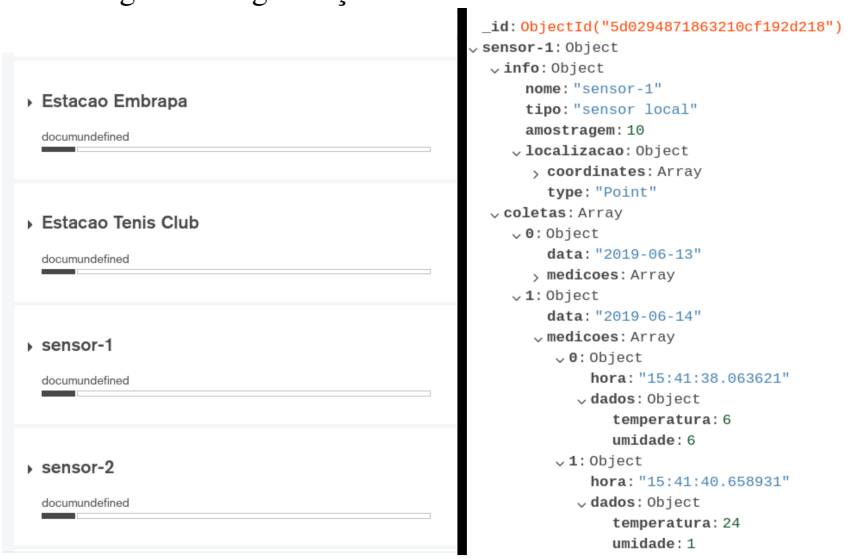
IMPLEMENTAÇÃO

Após a modelagem descrita anteriormente, o banco de dados foi implementado utilizando a tecnologia de banco de dados MongoDB. Foram desenvolvidos programas na linguagem de programação Python para a realizar a inserção de dados na estrutura pré-definida. A Figura 4 (A) ilustra a organização dos documentos inseridos no banco de dados. Utilizou-se a ferramenta MongoDB Compass (MONGODB, 2019a) para a apresentação destes dados. A Figura 4 (B) apresenta, no formato JSON, a organização hierárquica dos dados inseridos em um documento do banco de dados.

AVALIAÇÃO DE DESEMPENHO

Para avaliar o desempenho do banco de dados projetado, foram realizados dois experimentos. Para isso, foi implementado um programa na linguagem de programação Python, cuja funcionalidade é a geração e a inserção automática de dados no banco de dados. Os experimentos foram executados em um banco de dados local e sem a utilização de rede. Cada medição de desempenho foi repetida três vezes, e os resultados apresentados foram obtidos a partir da média destas medições. Tais testes foram projetados para avaliar o desempenho do banco de dados

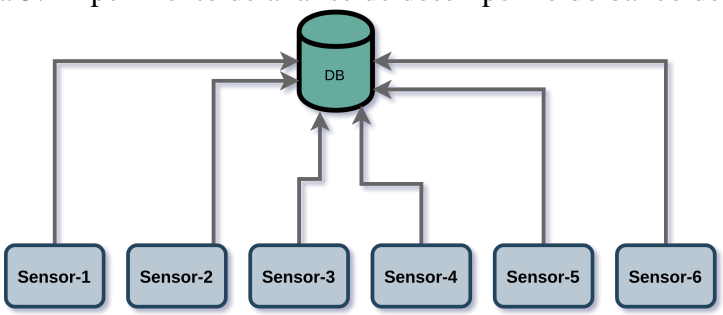
Figura 4: Organização dos dados no banco de dados



Fonte: Próprio Autor (2019)

em questão quando condicionado ao cenário exposto no estudo de caso. Ou seja, seis sensores inserindo dados com frequência constante. A Figura 5 ilustra a abordagem de experimento utilizada.

Figura 5: Experimento de análise de desempenho do banco de dados

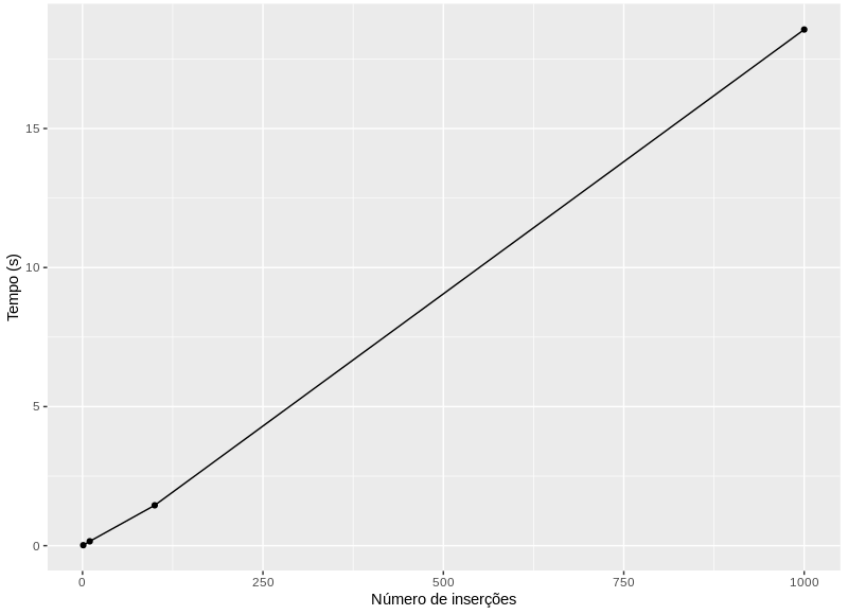


Fonte: Próprio Autor (2019)

O primeiro experimento objetivou avaliar o comportamento do banco de dados em um cenário de apenas um sensor inserindo dados continuamente. Foi medido o tempo de inserção de dados para o seguinte número de inserções: 1, 10, 100, 1000. A Figura 6 apresenta um gráfico dos resultados obtidos. Como pode ser observado nesse gráfico, para valores de até cem inserções, o tempo é inferior a dois segundos, já para mil inserções o tempo foi de aproximadamente 22 segundos.

O mesmo experimento foi repetido, entretanto, executando seis instâncias do programa em paralelo. Essa abordagem foi utilizada para simular o funcionamento de sensores em paralelo em um cenário real. A Figura 7 apresenta um gráfico dos resultados obtidos no segundo experimento. Para valores de até cem inserções, o tempo é inferior a dois segundos, já para mil

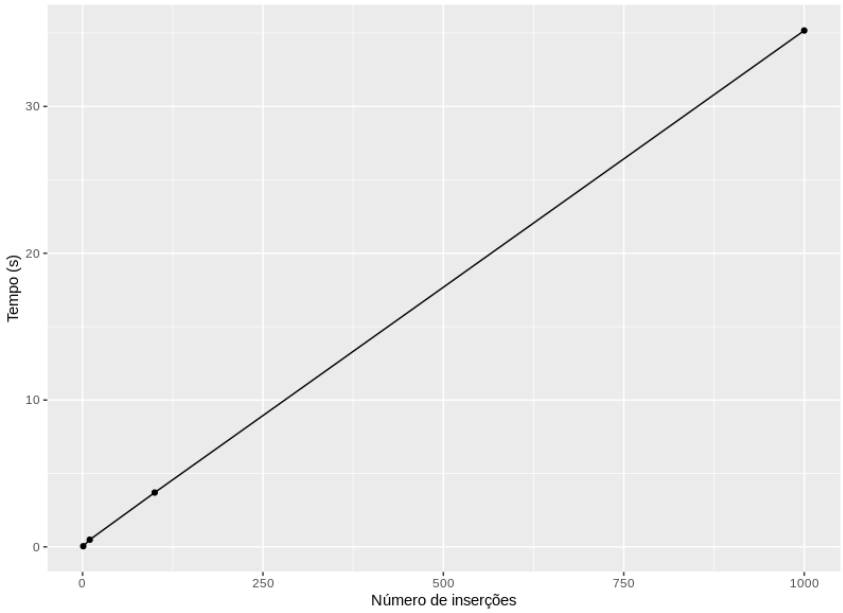
Figura 6: Gráfico desempenho de inserções geradas por um sensor



Fonte: Próprio Autor (2019)

inserções o tempo foi de aproximadamente 35 segundos.

Figura 7: Gráfico desempenho de inserções geradas por seis sensores simultaneamente



Fonte: Próprio Autor (2019)

Observa-se um crescimento linear do tempo em relação ao número de inserções, nos resultados obtidos no presente estudo. Além disso, o incremento do número de instâncias de elementos não alterou o comportamento linear observado. Esses resultados confirmam a escalabilidade do banco de dados MongoDB.

CONSIDERAÇÕES FINAIS

A partir do estudo realizado, conclui-se que o banco de dados MongoDB possui suporte a dados georreferenciados, contendo métodos de consulta e de inserção pré-definidos e documentação completa. Além disso, foi possível modelar e implementar um banco de dados espaço-temporal utilizando essa ferramenta. Os resultados do teste de desempenho demonstram que o banco de dados em questão possui um desempenho adequado para a utilização no experimento de coleta automatizada de dados. Observou-se um comportamento linear entre o tempo da operação e número de operações, dessa forma, os resultados obtidos nesse estudo confirmam a escalabilidade do MongoDB, como indicada por (FATIMA; WASNIK, 2016).

Como trabalhos futuros, pretende-se avaliar o desempenho da ferramenta em operações de consultas e contrastar com a literatura; repetir o experimento utilizando os bancos de dados MongoDB e PostgreSQL a fim de contrastar o desempenhos das duas tecnologias; avaliar o desempenho do banco de dados MongoDB em um cenário de coleta de dados real e investigar formas de visualização de dados georreferenciados.

AGRADECIMENTOS

Os autores agradecem à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAP ES) pela concessão da bolsa de mestrado.

REFERÊNCIAS

AGROCLIMATE. 2019. Disponível em: <http://agroclimate.org/>.

BAKER, K. et al. Point-trained models in a grid environment: Transforming a potato late blight risk forecast for use with the national digital forecast database. *Computers and electronics in agriculture*, Elsevier, v. 105, p. 1–8, 2014.

CHEN, C. X. Spatiotemporal databases. *Encyclopedia of GIS*, Springer, p. 2160–2160, 2017.

COSTA, J. A. A. da; QUEIROZ, H. P. de. *Régua de Manejo de Pastagens*. [S.l.], 2013.

CS+X, L. *OMT-G Designer*. 2019. Disponível em: <http://aqui.io/omtg-designer/>.

ECMA. *Standard ECMA-404: The JSON Data Interchange Syntax*. 2017. Disponível em: <https://www.json.org/>.

FATIMA, H.; WASNIK, K. Comparison of sql, nosql and newsql databases for internet of things. In: IEEE. *2016 IEEE Bombay Section Symposium (IBSS)*. [S.l.], 2016. p. 1–6.

FILHO, G. M. et al. Gdcana – programa cálculo de graus-dia para a cultura da cana-de-aÇúcar. In: . [S.l.: s.n.], 2008.

HENRY, E. et al. Precision apiculture: Development of a wireless sensor network for honeybee hives. *Computers and Electronics in Agriculture*, Elsevier, v. 156, p. 138–144, 2019.

IBGE. *Censo Agro 2017*. 2017. https://censos.ibge.gov.br/agro/2017. Acessado em 01/04/2019. Disponível em: <https://censos.ibge.gov.br/agro/2017>.

(IETF), I. E. T. F. *The GeoJSON Format*. 2016. Disponível em: <https://tools.ietf.org/html/rfc7946>.

KAMILARIS, A.; KARTAKOULLIS, A.; PRENAFETA-BOLDÚ, F. X. A review on the practice of big data analysis in agriculture. *Computers and Electronics in Agriculture*, Elsevier, v. 143, p. 23–37, 2017.

KARMAS, A.; TZOTSOS, A.; KARANTZALOS, K. Geospatial big data for environmental and agricultural applications. In: *Big Data Concepts, Theories, and Applications*. [S.l.]: Springer, 2016. p. 353–390.

MAPA. *Valor da Produção Agropecuária de 2019 deverá superar o deste ano*. 2018. http://www.agricultura.gov.br/noticias/valor-da-producao-agropecuaria-de-2019-devera-superar-o-deste-ano. Acessado em 14/06/2019. Disponível em: <http://www.agricultura.gov.br/noticias/valor-da-producao-agropecuaria-de-2019-devera-superar-o-deste-ano>.

MONGODB. *MongoDB Compass*. 2019. Disponível em: <https://www.mongodb.com/products/compass>.

MONGODB. *MongoDB Mobile*. 2019. Disponível em: <https://www.mongodb.com/products/mobile>.

PEREIRA, A. R.; ANGELOCCI, L. R.; SENTELHAS, P. C. *Meteorologia Agrícola*. [S.l.]: Universidade de São Paulo, 2007.

PERONDI, D. et al. Crop season planning tool: Adjusting sowing decisions to reduce the risk of extreme weather events. *Computers and Electronics in Agriculture*, Elsevier, v. 156, p. 62–70, 2019.

ROBAINA, R. P.; CAMARGO, S. da S. Georeferenciamento e bancos de dados nosql: Uma análise do panorama atual na agropecuária de precisão. *Congreso Argentino de AgroInformática*, v. 11, p. Aceito para publicação, 2019.

SAATH, K. C. d. O.; FACHINELLO, A. L. Crescimento da demanda mundial de alimentos e restrições do fator terra no brasil. *Revista de Economia e Sociologia Rural*, SciELO Brasil, v. 56, n. 2, p. 195–212, 2018.

WOLFERT, S. et al. Big data in smart farming–a review. *Agricultural Systems*, Elsevier, v. 153, p. 69–80, 2017.

POMODORO: ARQUITETURA PARA O MONITORAMENTO E CONTROLE DE UMIDADE DO SOLO EM UM CULTIVO PROTEGIDO DE TOMATE

Leonardo Rodrigues de Lima Teixeira[1], *Joffrey Peyrac de Albuquerque Galvão Junior*[1], *Anderson Patrício Fernandes dos Santos*[1]

[1] Escola Agrícola de Jundiaí, Universidade Federal do Rio Grande do Norte, Macaíba, Rio Grande do Norte, Brasil, leonardo@eaj.ufrn.br, joffreypag@gmail.com, andersonp17@bol.com.br

RESUMO

Este trabalho tem como objetivo o desenvolvimento de uma arquitetura, denominada Pomodoro, a qual permite a realização de monitoramento e controle da umidade do solo de um cultivo protegido de tomate. A arquitetura permite que haja a possibilidade de análise dos dados coletados mesmo em um ambiente sem conexão com a Internet (maioria das aplicações na agricultura), devido ao desenvolvimento de uma estratégia de transmissão do conjunto de dados de umidade salvo no sistema local em estufa para um banco de dados em nuvem usando um aplicativo Android. Os dados salvos em nuvem são utilizados na geração de gráficos para estudos dos profissionais da área agrícola.

PALAVRAS-CHAVE: Monitoramento e controle de umidade do solo, Aplicativo móvel, Cultivo de tomate.

ABSTRACT

The aim of this work is the development of a architecture named Pomodoro, which allows monitoring and control of soil moisture in a protected tomato crop. The architecture allows the collected data analysis of a environment without connection to the Internet (most applications in agriculture), due the development of a strategy of saved data transmission locally for a cloud database using an Android application. The saved data in the cloud are used in the generation of graphs for the studies of the professionals of the agricultural area.

KEYWORDS: Monitoring and control of soil moisture, Mobile application, Tomato crop.

INTRODUÇÃO

Há vários anos o setor agrícola brasileiro vem se consolidando, diversificando seus produtos e aumentando sua produção de maneira significativa e, como consequência, sua importância no Produto Interno Bruto (PIB) nacional, constituindo-se também como um dos maiores produtores de alimentos do mundo.

No ano de 2016, de acordo com a Embrapa (2019), o agronegócio como um todo gerou 23% do PIB e 46% do valor das exportações. Em 2017, o setor foi responsável por 19 milhões de trabalhadores ocupados. Agroindústria e serviços empregaram, respectivamente, 4,12 milhões e 5,67 milhões de pessoas, enquanto 227,9 mil pessoas estavam ocupadas no segmento de insumos do agronegócio.

Na diversidade agrícola presente no Brasil, encontra-se o tomate, tanto para processamento como de mesa. O cultivo deste, de acordo com Souza (2016), é destinado basicamente ao mercado local, e que trata-se de uma cultura de tratos culturais intensos e de alto custo de produção, além de ser uma hortaliça muito susceptível ao ataque de pragas e doenças.

O clima é um fator que influencia a produção de hortaliças. Dias seguidos de chuva, comumente no verão, danificam as hortaliças e criam condições favoráveis ao surgimento de doenças, e em épocas frias como no inverno acabam prolongando o ciclo das culturas (PURQUERIO; TIVELLI, 2006). Sobre essas circunstâncias, Purquerio e Tivelli (2006) defendem que como forma de proteger o cultivo dos ataques naturais pode-se adotar o cultivo protegido para controle das condições edafoclimáticas como: temperatura, umidade do ar, radiação, solo, vento e composição atmosférica. O ambiente protegido pode ser um túnel (baixo ou alto), uma estufa agrícola com ou sem pé direito ou até mesmo uma casa de vegetação.

A automação de estufas agrícolas controla com precisão os fatores que interferem no crescimento da planta, assim como é possível ler dados vindo de sensores, utilizando-os para possíveis tomadas de decisão (LOUREIRO *et al.*, 2018).

No Brasil, trabalhos de automação de estufa vem sendo desenvolvidos em diversos polos com seus respectivos fins. Como exemplo, temos os trabalhos de Loureiro *et al.* (2018) e o de Fernandes, Preuss e Silva (2017), que serão detalhados a seguir.

Loureiro *et al.* (2018) apresentam o desenvolvimento de um protótipo de automação para estufa usando coleta de dados de temperatura e umidade por sensores para controle de irrigação e ventilação da estufa e visão computacional para acompanhamento da taxa de crescimento. A partir de um sistema móvel, o usuário pode controlar remotamente

funcionalidades da estufa. Dentre as funcionalidades, temos: irrigação da estufa, acesso a informações coletadas por meio de sensores (temperatura e umidade) e análise de imagens para identificar a taxa de crescimento do cultivo.

Fernandes, Preuss e Silva (2017) também apresentam o desenvolvimento de um protótipo automatizado, controlando temperatura, umidade e luminosidade, utilizando um microcontrolador Arduino e um aplicativo para *smartphones* com sistema operacional Android.

Ambos os trabalhos obtiveram resultados em nível de protótipo, e em ambos, tínhamos a presença de Internet para transmissão dos dados. Este trabalho propõe o desenvolvimento de uma arquitetura, denominada Pomodoro, para o monitoramento e controle de umidade do solo em um cultivo protegido de tomate, localizado em uma estufa real onde não há sinal de Internet para transmissão dos dados.

A ideia é que, além de monitorar a umidade do solo para cultivo de tomate, a partir da leitura de sensores instalados no solo, e acionar uma bomba de irrigação, quando necessário, seja proposta uma estratégia de armazenamento dos dados adquiridos pelos sensores para um banco de dados na estufa em um primeiro momento (sem sinal de Internet), e que eles possam ser enviados, em um segundo momento (na presença de sinal de Internet) para a nuvem a partir de um aplicativo Android, tornando possível gerar um histórico de umidade do solo do cultivo, para futuros estudos.

MATERIAL E MÉTODOS

O trabalho foi implantado em uma estufa na Escola Agrícola de Jundiaí, Unidade Acadêmica Especializada em Ciências Agrárias, da Universidade Federal do Rio Grande do Norte.

Para o desenvolvimento desse projeto, foram utilizados um NodeMCU com microcontrolador do tipo ESP8266, um Arduino Pro Mini, e três sensores de umidade do solo, modelo FC-28, constituindo o módulo sensor, além de um microprocessador Raspberry Pi 3 model B com módulo RTC DS3231 (para atualização de data e hora do sistema), e um módulo relé para ativação da bomba, constituindo o módulo central. Esses equipamentos trabalham em conjunto: os microcontroladores e sensores de umidade do solo atuando com a coleta de dados da umidade do solo e enviando para o microprocessador, o qual salvará o dado recebido em um banco de dados local e tomará a decisão de acionamento ou não da bomba de irrigação da estufa através de um relé.

O NodeMCU é um *firmware* baseado em Lua para o microcontrolador ESP8266 embutido em sua composição. Como principal vantagem com relação ao Arduino é o fato de

possuir uma interface de conexão sem fio com a Internet, e como desvantagem o fato de possuir apenas uma entrada analógica.

Os sensores de umidade do solo FC-28 são sensores resistivos, os quais possuem duas hastes com polaridade da tensão de alimentação. Quando em contato com o solo, a corrente é transmitida de uma haste a outra a depender da presença de água no local. Quanto mais úmido, mais corrente é transmitida entre as hastes, tendo assim baixa resistência. Quanto mais seco, menos corrente elétrica haverá entre as hastes, e consequentemente a resistência é maior. Essa variação de resistência é interpretada pelo *driver* que acompanha o sensor, fornecendo dois tipos de saída: uma saída digital que retorna um sinal de nível lógico baixo (0) para quando o solo estiver seco e um sinal de nível lógico alto (1) para o solo molhado, e vem acompanhado de um trimpot para ajustar o limiar que define seco e úmido na saída digital; e uma saída analógica onde ele retorna valores na resolução de 10 bits (leitura entre os valores 0 e 1023).

Primeiramente, foi realizado o levantamento de requisitos funcionais do sistema, os quais são listados a seguir: coletar o valor de umidade do solo, observando que o tomate apresenta três fases de crescimento (equivalendo a três profundidades diferentes para coleta); ativar ou desativar irrigação a partir do valor da umidade do solo; registrar o tempo de ativação da bomba durante o dia completo; armazenar dados coletados pelos sensores em um banco de dados local na estufa; coletar dados salvos no banco de dados da estufa a partir de um aplicativo móvel; enviar dados salvos na estufa para um banco de dados na nuvem a partir de um aplicativo móvel; exibir gráfico de umidade do solo no aplicativo móvel; e configurar parâmetros de irrigação a partir de um aplicativo móvel.

Em seguida, o Raspberry Pi foi configurado para se tornar o servidor *Message Queue Telemetry Transport* (MQTT), o qual é um protocolo de transporte de mensagens no padrão *publish/subscribe*, sendo ideal para comunicação máquina para máquina (M2M) e Internet das Coisas (IoT) (BANKS; GUPTA, 2014). A ideia é que este servidor atue como um *Broker*, ou seja, um software servidor que recebe solicitações e as repassa, quando solicitado, atuando nos sentidos de receber e enviar os dados (OLIVEIRA, 2017).

O servidor MQTT escolhido foi o *Mosquitto* da *Eclipse Foundation* por ser um software de código aberto. A instalação e configuração dele no Raspberry Pi só foi possível após a prévia instalação do sistema operacional *Raspbian* no cartão miniSD de 8 GB.

Concluída a instalação e testes de envio de mensagens, foi realizada a implementação de comunicação do NodeMCU à rede WiFi do Raspberry Pi, possibilitando ao microcontrolador o envio de dados ao servidor MQTT Mosquitto.

Os sensores foram inseridos nas profundidades de 7,5 cm, 22,5 cm e 35 cm, que equivalem as camadas de solo exploradas pelas raízes da cultura do tomate em suas fases de crescimento, frutificação e maturação. No sistema, foram configurados 20 dias para a fase 1 (crescimento), 15 dias para a fase 2 (frutificação) e 15 dias para a fase 3 (maturação). E o inicio da fase 1 foi considerada a partir do plantio.

Os dados dos sensores são coletados pelo Arduino Pro Mini (devido a sua quantidade de entradas analógicas) e enviadas através do protocolo serial *Inter-intergrated Circuit* (I2C) para o NodeMCU, o qual envia os dados via WiFi para o módulo central (Raspberry Pi).

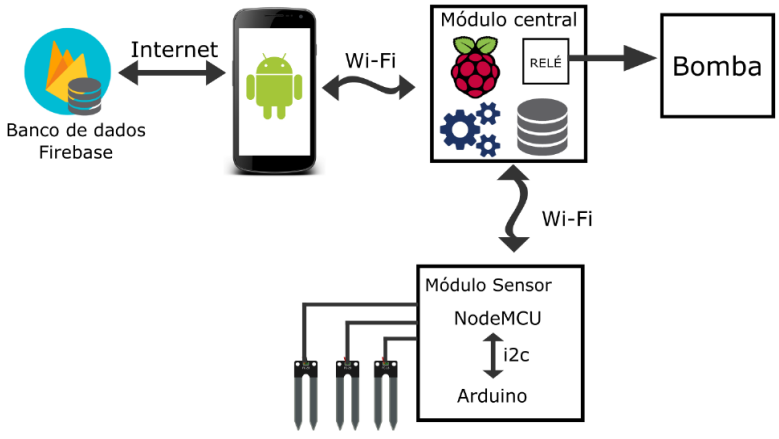
No módulo central, foi implementado um algoritmo em Python no Raspberry Pi, o qual é executado logo após a inicialização do sistema operacional. A função principal desse algoritmo é de receber os dados do tópico usado pelo módulo sensor, armazenar esses dados em um banco de dados MySQL e tomar a ação de ativar ou desativar a bomba de irrigação.

Também foi desenvolvido no módulo central um Web Service em Java, responsável por receber solicitações de um aplicativo móvel e enviar os dados do banco de dados MySQL local para o *smartphone*, deletando os dados do primeiro, devido às limitações de espaço.

O aplicativo móvel permite que, na presença de sinal de Internet, os dados sejam retirados do *smartphone* e salvos em um banco de dados na nuvem usando a tecnologia Firebase da Google. A partir desse momento, os dados poderão ser visualizados na forma de gráficos para análise dos dados coletados.

A Figura 1 apresenta em detalhes a arquitetura Pomodoro desenvolvida, com seus módulos central e sensor, protocolos e demais componentes de sistema.

Figura 1 - Arquitetura Pomodoro



Fonte: Autoria própria

O sistema de irrigação utilizado foi o gotejamento, com gotejadores espaçados a cada 30cm ao longo da mangueira, emitindo uma vazão de 1,5L/h. E a bomba é monofásica de ½ CV de potência.

O tomate utilizado é do tipo tomate para mesa, cultivar IPA-07 de crescimento determinado, e a semeadura foi realizada direto no campo, porém os sensores foram ligados após o estabelecimento da cultura, ou seja, após o pegamento das mudas, cerca de 20 dias após a semeadura.

Calibração dos Sensores

Após a realização de testes da arquitetura de comunicação, foi realizada a calibração dos sensores, em laboratório, com três amostras do solo retirado do local de cultivo, sendo uma amostra de cada profundidade de interesse, de acordo com as fases de crescimento do tomateiro. O objetivo foi coletar os valores dos sensores para as condições de capacidade de campo, isto é, a quantidade de água considerada ideal no solo que a planta absorva sem que haja excesso ou escassez.

Para se ter o solo em sua capacidade de campo, Bernardo, Soares e Mantovani (2008) sugerem encharcar o campo que será trabalhado o cultivo e cobri-lo para evitar a evaporação. Neste trabalho, o método sugerido foi adaptado para ser realizado em laboratório, ambiente onde não recebe influência da radiação solar.

O solo foi separado em três recipientes: a primeira amostra de 0 a 15 cm, a segunda de 15 a 30 cm e a terceira de 30 a 40 cm, que equivalem às profundidades que compreende a maior parte do sistema radicular do tomateiro. Em seguida os recipientes com o solo foram encharcados com água e dava-se um tempo para que o excesso drenasse pelo fundo do recipiente no qual foi retirado.

Os recipientes com o solo encharcado foram colocados dentro de um balde com solo seco, para que a água drenada do recipiente encontrasse o solo seco, tal qual acontece na natureza. A Figura 2 apresenta como foi montado o experimento para calibração dos sensores.

Com um tempo de 30 horas (estimado pelo Engenheiro Agrônomo Me. Anderson), após as amostras terem sido encharcadas, o solo deveria encontrar-se em sua capacidade de campo, momento esse que o sensor coleta o dado para encontrar o valor considerado ideal em umidade do solo para a planta.

A Tabela 1 mostra o comportamento das três camadas do solo após o encharcamento, em função do tempo. Os valores expressos são leituras analógicas na resolução de 10 bits (0 - 1023) dos sensores que serão convertidos em porcentagem de umidade na capacidade de campo.

Figura 2 -Teste em laboratório da capacidade de campo



Fonte: Autoria Própria

Os sensores são identificados como: A0 para o sensor da camada de 0 a 15 cm; A1 para o sensor da camada de 15 a 30 cm e A2 para o sensor da camada de 30 a 40 cm.

Tabela 1 - Análise dos sensores em laboratório

	29/08/2018						30/08/2018
	30 horas	+10 min	+10 min	+10 min	+10 min	+10 min	
Hora	15:55	16:05	16:15	16:25	16:35	16:45	08:35
A0	350	342	339	339	339	340	335
A1	481	482	481	478	474	471	447
A2	578	580	577	577	576	573	559

Como os valores variavam, foi realizada uma média dos valores dos dias 29 de agosto de 2018 às 16:45hs e 30 de agosto de 2018 às 08:35, resultando nos valores apresentados a seguir: A0 = 337,5; A1 = 459; e A2 = 566.

Estes são os valores considerados para o caso em que o solo está em 100% da sua capacidade de campo. Quanto mais seco, mais os sensores apresentam valores tendendo a 1023. Assim, a leitura do valor 1023 representa 0% da capacidade de campo. A partir dessas relações, foram determinadas funções lineares que relacionassem os valores lidos de cada sensor com as porcentagens de umidade, como a seguir: $y = -0{,}1458x + 149{,}1534$ (para A0), $y = -0{,}1773x + 181{,}3779$ (para A1) e $y = -0{,}2188x + 223{,}8324$ (para A2).

RESULTADOS E DISCUSSÃO

Após o desenvolvimento da arquitetura (tanto a parte de hardware como de software), os módulos central (Figura 3a) e sensor (Figura 3b) foram instalados na estufa. O módulo central foi instalado mais ao centro da estufa para garantir segurança contra fenômenos meteorológicos externos.

Foi realizado um orçamento para estimativa de custo para produzir horticultura com um sistema automático seguindo a arquitetura Pomodoro. O sistema custaria em torno de R$ 300,00, considerando sempre o menor valor encontrado para cada item e com valor total sem o frete incluso.

O aplicativo móvel é constituído de duas telas: a inicial (de configuração) e a de dados. Na tela inicial, o usuário pode visualizar o estado atual da bomba (ligado ou desligado), a fase e o dia que o sistema está operando; definir operação manual ou automática; ativar ou desativar manualmente a bomba (ação independente do sistema estar no manual ou automático); definir valores de umidade máxima e mínima para uma fase especifica; alterar a fase do sistema; alterar o dia do sistema; e definir tempo de irrigação e intervalo (funcionamento por tempo, independente dos sensores).

As Figuras 4a e 4b apresentam a tela inicial do aplicativo e a Figura 4c a tela de dados, a qual está definida para coletar os dados salvos no módulo central e armazená-los na memória interna do dispositivo móvel. Nessa mesma tela é possível, assim que o aparelho estiver em uma rede com conexão com a internet, enviar os dados que foram armazenados e visualizar no gráfico todos os dados que foram enviados ao banco de dados remoto.

Devido à grande quantidade de pontos presentes no gráfico e para melhor visualização das informações, elas não são exibidas no eixo x do gráfico. Para visualizar os detalhes (data, hora e porcentagem de umidade do solo) basta clicar na amostra desejada.

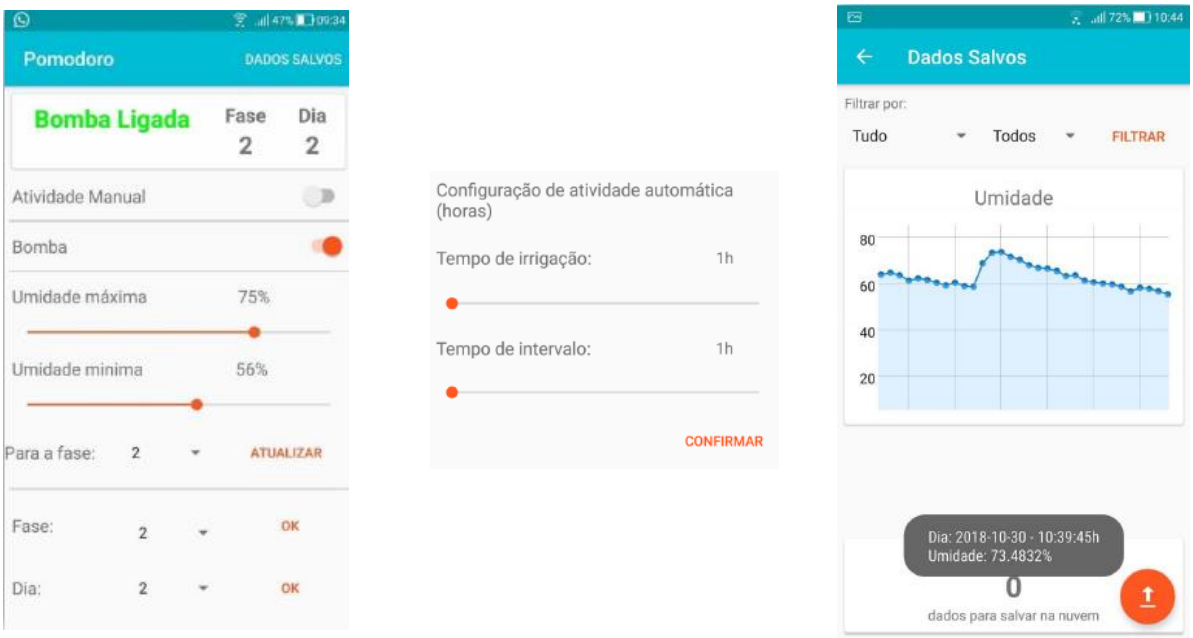
Figura 3 - (a) Módulo central. (b) Módulo sensor



(a) (b)

Fonte: Autoria própria

Figura 4 - (a) Tela inicial do aplicativo. (b) Continuação da tela inicial. (c) Tela de dados



(a) (b) (c)

Fonte: Autoria própria

CONCLUSÕES

Com o desenvolvimento da arquitetura Pomodoro, temos a possibilidade de implantar um sistema de irrigação de baixo custo para cultivo protegido, proporcionando economia de água e energia. Além disso, como principal diferencial, a arquitetura desenvolvida permite que o sistema seja instalado em um local sem acesso a Internet, não abrindo mão da possibilidade de análise dos dados coletados caso o usuário do sistema transfira os dados do módulo central para o seu smartphone, em um primeiro momento, e para a nuvem posteriormente, quando estiver conectado à Internet.

A arquitetura desenvolvida é escalável, atendendo do pequeno agricultor com uma estufa, até um grande produtor, com várias estufas de produção, adicionando apenas novos módulos sensores.

Apesar da arquitetura ter sido utilizada em um cultivo protegido de tomate, nada impede que outra cultura seja utilizada, bastando para tal realizar um estudo das fases de crescimento da cultura para definir a quantidade de sensores e calibrá-los de acordo com a capacidade de campo do solo a ser utilizado para a cultura.

Com relação a trabalhos futuros, temos vários desafios a superar, tais como: realizar a troca dos sensores resistivos FC-28, que oxidaram em aproximadamente um mês, por sensores que não oxidem com tanta facilidade; estudar a viabilidade no uso de sensores

capacitivos para obter uma maior precisão na coleta de dados; além de fazer um comparativo de produção, analisando número de frutos, consumo de água e de energia, dentre outras variáveis, entre um cultivo que utiliza o sistema e um que não utiliza.

REFERÊNCIAS

BANKS, A.; GUPTA, R. OASIS Standard: MQTT Version 3.1.1**.** 2014. Disponível em: <http://docs.oasis-open.org/mqtt/mqtt/v3.1.1/os/mqtt-v3.1.1-os.html>. Acesso em: 15 jun. 2019.

BERNARDO, S; SOARES, A. A.; MANTOVANI, E. C. Manual de Irrigação, 8. ed. Viçosa: Editora UFV, 2008. p. 26.

EMBRAPA. Trajetória da Agricultura Brasileira. Disponível em: <https://www.embrapa.br/visao/trajetoria-da-agricultura-brasileira>. Acesso em: 20 jun. 2019.

FERNANDES, D. G.; PREUSS, E.; SILVA, T. L.; Sistema Automatizado de Controle de Estufas para Cultivo de Hortaliças. Universidade Federal de Santa Maria. Rio Grande do Sul. 2017. Disponível em: <http://repositorio.ufsm.br/handle/1/12958>. Acesso em: 18 jun. 2019.

LOUREIRO, J. F.; BRITO R. C.; FAVARIM, F.; TODT, E. Automação de estufa agrícola integrando hardware e controle remoto pela internet. Revista de Computação Aplicada ao Agronegócio. Medianeira, v. 1. n. 1, p. 38-35. Disponível em:
< https://revistas.utfpr.edu.br/rcaa/article/view/7437/5290 >. Acesso em: 19 jun. 2019.

OLIVEIRA, S. Internet das coisas com ESP8266, ARDUINO e RASPBERRY PI, 1. ed. São Paulo: Novatec, 2017.

PURQUERIO, L. F. V.; TIVELLI, S. W. Manejo do ambiente em cultivo protegido. Manual técnico de orientação: projeto hortalimento. São Paulo: Codeagro, 2006. p. 15-29. Disponível em: <http://www.iac.sp.gov.br/imagem_informacoestecnologicas/58.pdf>. Acesso em: 19 jun. 2019.

SOUZA, M. A. Análise do desempenho de sistemas produtivos de tomate de mesa no DF e sua relação com a cadeia produtiva. 2016. 70 f. Dissertação (Mestrado em Agronegócios). Faculdade de Agronomia e Medicina Veterinária, Universidade de Brasília, 2016.

Abordagem neural para previsão de peso de contusões no processo de transporte bovino

Flávia Morales Antunes[1], *Ana Paula Lüdtke Ferreira*[1], *Tiago Alan Cunha Nardino*[2], *Naylor Bastiani Perez*[3], *Fabiano Nunes Vaz*[2], *Ricardo Zambarda Vaz*[2]

[1]Universidade Federal do Pampa
Bagé, Rio Grande do Sul, Brasil
flaviamoralesantunes@gmail.com, anaferreira@unipampa.edu.br

[2]Universidade Federal de Santa Maria
Santa Maria, Rio Grande do Sul, Brasil
tiagonardino@hotmail.com, fabianonunesvaz@gmail.com, rzvaz@terra.com.br

[3]EMBRAPA Pecuária Sul, Bagé, Rio Grande do Sul
naylor.perez@embrapa.br

RESUMO

O contínuo aumento das exigências do consumidor final quanto à procedência e qualidade da carne comprada no varejo vem fazendo com que os participantes da cadeia produtiva tenham maior atenção a ações potencialmente danosas, tanto do ponto de vista financeiro quanto da imagem final do produto. O processo de transporte do gado, das propriedades até o local de abate, produz estresse nos animais, potencializando o aparecimento de carnes escuras, rígidas, com pH alto e contusões, que tornam a carne imprópria para consumo humano. Esse trabalho visa mapear as condições do processo de transporte, do embarque ao descanso pós-desembarque, das propriedades rurais ao frigorífico, na probabilidade da carga apresentar, ao final do processo, menor ou maior peso total de contusões. A abordagem escolhida foi a de uma rede neural multicamada, com dados de treinamento/classificação coletados em um frigorífico do Rio Grande do Sul durante o período de um ano. A acurácia obtida foi de aproximadamente 94,5%, com alta correlação entre o dado previsto e o real.

PALAVRAS-CHAVE: Redes neurais, Inteligência artificial, Bem-estar animal.

ABSTRACT

The final consumer demands for meat quality and origin are continuously increasing, making all productive chain participants pay more attention to potentially harmful actions, both from the

financial and the final product image point of view. The livestock transportation process, from the rural properties to the place of slaughter, produces stress in the animals, producing a dark appearance, rigid meat with high pH and bruises, which makes it unfit for human consumption. This work aims to map the livestock transportation process conditions, from the truck boarding to the post-arriving rest, into the probable animal total bruise weight. The approach uses a multilayer perceptron neural network trained from data collected during one year. The accuracy obtained was approximately 94.5%, with a high correlation between the predicted and the actual data.
KEYWORDS: Neural Networks, Artificial Intelligence, Animal welfare.

INTRODUÇÃO

Segundo o Instituto Brasileiro de Geografia e Estatística (IBGE), no primeiro trimestre de 2017 foram abatidas 7,4 milhões de cabeças de bovinos. Já no mesmo período de 2018, esse número foi de 7,72 milhões, perfazendo um aumento de 4,4%. A exportação da carne bovina *in natura*, sem processamento, aumentou 31,1% do primeiro trimestre de 2017 para o mesmo período de 2018, aumentando o ingresso de divisas no país (IBGE, 2018). Apesar do título de maior produtor mundial e da extensão territorial brasileira, a eficiência do setor de produção pecuária ainda tem espaço para aumentar, especialmente por meio do uso sistematizado da tecnologia, aumentando a participação do setor no Produto Interno Bruto (PIB) e melhorando outros indicadores de produção (MAGGI et al., 2006).

As exigências a respeito da qualidade da carne vendida no comércio aumentam na medida em que os consumidores procuram produtos de qualidade, com garantias de segurança alimentar. Além de características como textura, maciez e sabor, os consumidores passaram a interessar-se também pelas condições de vida dos animais criados para abate. O bem-estar animal passou a ser um fator diferencial, tanto em termos éticos como em termos econômicos, visto que um percentual significativo do dos consumidores está disposto a pagar mais pela garantia de um produto com procedência de animais que não sofreram desnecessariamente em seu processo de criação (LAMA et al., 2013).

O processo de transporte dos animais, das propriedades até o local de abate, tem características que atingem diretamente questões de bem-estar: condições inadequadas de manejo no embarque, deslocamento, desembarque e período de espera, geram estresse e movimentações inadequadas que podem provocar ferimentos e contusões; quedas durante o transporte podem fazer com que o animal caído seja pisoteado, causando a perda parcial ou total da carcaça (STRAPPINI et al., 2009). A quantidade de estresse vivenciado relaciona-se, ainda, com a qualidade da carne no final do ciclo de vida do animal: carnes escuras, rígidas e com pH alto, que não são adequadas para o consumo humano. Carnes com contusões ou problemas relacionados ao estresse não podem ser vendidas no mercado e, mesmo que parcialmente aproveitadas, são destinadas a produtos com menor valor agregado, acarretando em prejuízo tanto para os

criadores como para os frigoríficos (FILHO, 2006).

A questão das contusões no processo de transporte tem sido estudada em diferentes contextos, com muitos trabalhos focados na descoberta de conhecimento sobre as causas de contusões no processo de transporte. Romero et al. (2013) usa regressão logística multivariada no estudo de 86 viagens e 1179 animais, avaliando diâmetro de hematomas e pH muscular. O teste Hosmer-Lemeshow foi usado e as variáveis de relevância foram a densidade carga, número e condições de parada e tempo de espera; os terneiros apresentavam menor probabilidade de alto pH muscular, em comparação com animais mais velhos. Frasão et al. (2014) estudou o efeito da distância percorrida na quantidade de contusões de animais, concluindo que a variável não é significativa, apesar de que 89,1% dos 320 animais selecionados (todas fêmeas) apresentavam pelo menos uma lesão. Novamente, as partes traseiras e lateral foram as mais atingidas 71,4%. Por outro lado,Vimiso e Muchenje (2013) encontra correlação da distância percorrida em relação tanto à quantidade e severidade de contusões quanto ao pH da carne dos animais abatidos. Mendonça et al. (2018) utilizam o critério Akaike (SAKAMOTO; ISHIGURO; KITAGAWA, 1986) para inferir causas de contusões, encontrando que o sexo do animal era o fator mais importante, tendo as fêmeas 91% mais chance de apresentar hematomas carcaça. A densidade inadequada de carga do caminhão, longos períodos de transporte e descarga também são relevantes. Camargo et al. (2018) faz uso de um algoritmo de agrupamento K-*means* (MACQUEEN, 1967), redes Bayesianas e árvores de decisão para inferir as causas de contusões, fazendo uso dos mesmos dados de (MENDONÇA et al., 2018). Com essa abordagem, as variáveis determinantes do número de contusões são a densidade da carga do caminhão e o sexo dos animais, mostrando que técnicas diferentes produzem o mesmo resultado. Huertas et al. (2010) avaliaram os efeitos das condições de transporte de gado bovino no bem-estar e na quantidade de contusões das carcaças; a coleta de dados foi realizada por um período de dois anos em 448 cargas distintas. A análise dos dados foi feita pelo cálculo da significância estatística entre as variáveis, usando o método χ^2. Os achados deste trabalho indicam que portas traseiras do caminhão do tipo guilhotina, aparatos para provocar a movimentação de animais, distâncias e tempo de deslocamento longos afetam significativamente o aparecimento de contusões na carcaça. Kline (2018) conduz um estudo sobre localização de contusões, com um número menos de indivíduos foi conduzido nos EUA e, apesar do número de animais contundidos no transporte ser menor, com 28,1% de carcaças com contusões visíveis, é relatado que diversas contusões somente são notadas durante o processo de corte. As regiões com maiores índices de hematomas foram as regiões traseira, costela e lombo. Huertas et al. (2015) analisou o impacto econômico da ocorrência de contusões sobre os mesmos dados obtidos em (HUERTAS et al., 2010). Os achados mostram que, das 15.157 carcaças observadas, 60% tinham, pelo menos uma contusão e que 42% apresentavam hematomas em ambos os lados e que as perdas financeiras são majoradas, pois a maior parte da carne contundida está localizada justamente nas áreas com maior valor de mercado: traseira, laterais, costela, ombro e lombo. O conhecimento sobre os aspectos de causa e efeito é relevante, visto que Huertas, Kempener e Eerdenburg (2018) mostram que a alteração no manejo, por

meio do treinamento do pessoal encarregado do processo de embarque, é efetivo na diminuição e seriedade das contusões.

O objetivo deste trabalho é a construção de um sistema de inferência em que, dadas as condições de transporte dos animais, possa ser inferido o peso esperado das contusões produzidas após o período de abate. Diferentemente dos trabalhos citados, em que o objetivo é encontrar as variáveis relevantes do processo, este trabalho visa prever as perda esperada para o produtor. A abordagem utilizada será a de redes neurais artificiais, capazes de aprender com dados de realidade e produzir um mapeamento entre os dados de entrada e a saída esperada (HAYKIN, 2001).

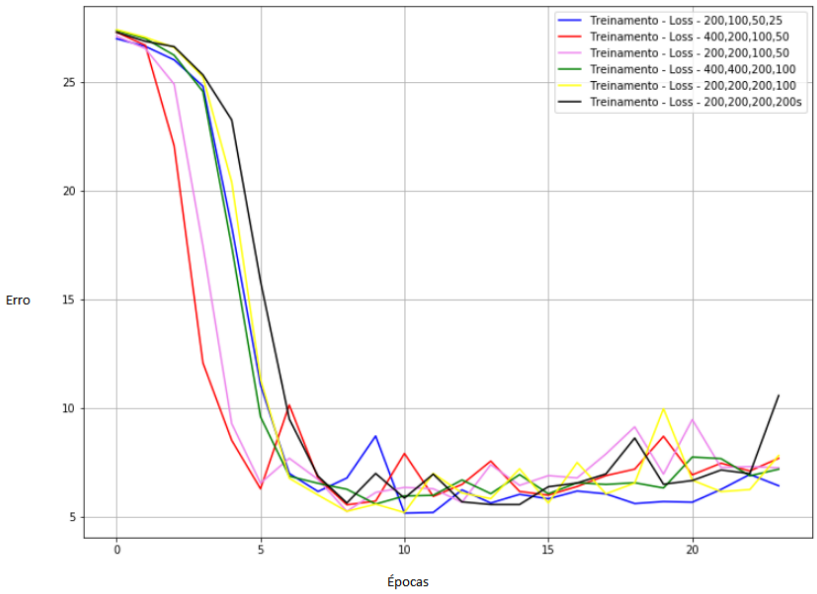
MATERIAL E MÉTODOS

O trabalho realizado caracteriza-se, metodologicamente, como um trabalho de pesquisa aplicada e experimental. Os dados utilizados foram coletados a partir de cargas recebidas pelo Frigorífico Silva, pelo período de 1 ano, e armazenados em planilhas Microsoft Excel, totalizando 3.850 carregamentos e cerca de 260.000 animais abatidos. As planilhas, contendo 63 variáveis ao todo, foram analisadas para verificação de quais variáveis faria sentido usar no processo. As linhas da planilha com erros de digitação e com dados faltantes foram eliminados do conjunto. Alguns dados foram agrupados em valores que faziam mais sentido do ponto de vista de treinamento; por exemplo, os horários de conferência foram reduzidos por meio da contagem de todos em cada viagem, criando uma nova coluna com o somatório de conferências e outra coluna com o número de conferências por quilômetro rodado. As 63 variáveis iniciais foram reduzidas para 28, envolvendo informações sobre o produtor, município, número de bois, vacas, touros, búfalos machos e búfalos fêmeas embarcados, tipo e condição da mangueira, tipo de caminhão, quilômetros percorridos, quilometragem percorrida em estrada de chão, número de pessoas no carregamento, condições da mangueira, condições do tempo no embarque, condições da estrada, tipo de manejo na propriedade, se o caminhão atolou no percurso, se houve um único ou mais do que um embarque, tempo de viagem, motorista, peso líquido, quantidade de animais, número de conferências ao longo da viagem, número de conferências por quilômetro rodado, peso da carcaça fria. O peso das contusões de cada animal foi usado como saída da rede e usado no processo de treinamento.

Os dados que não numéricos foram escalonados para utilizá-los na entrada da rede. A biblioteca usada para conversão foi a *feature columns* do pacote *Tensorflow*, usando para o desenvolvimento das arquiteturas de rede e particionamento dos conjuntos de dados. A linguagem Python foi usada para as implementações necessárias além do uso do TensorFlow, escolhida por ser uma linguagem de código livre, em desenvolvimento ativo, e que possui pacotes específicos para a construção de redes(GOOGLE BRAIN, 2019).

A arquitetura escolhida foi a de uma rede neural *feedforward multilayer perceptron* (MLP) com algoritmo de treinamento backpropagation (HAYKIN, 2001). O processo de treinamento foi

Figura 1: Gráfico do erro para 4 camadas ocultas



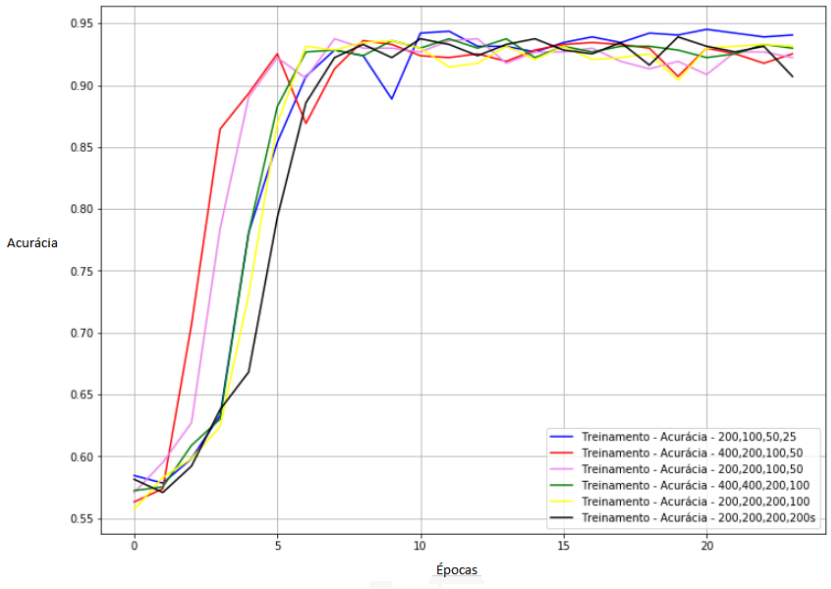
feito com 80% dos dados, escolhidos aleatoriamente do conjunto completo. A cada novo treinamento um novo conjunto de dados era gerado, para evitar *bias* no processo. Os 20% de dados deixados de fora de cada treinamento foram usados para teste do aprendizado da rede. A estrutura da rede foi contruída em um processo interativo de alteração da quantidade de camadas ocultas (de uma a quatro camadas) e da quantidade de neurônios em cada camada (de 50 a 400 neurônios). As funções de ativação, nas diferentes camadas serão escolhidas entre as funções sigmóide, tangente hiperbólica, Elu, Relu e SoftMax. A alteração das funções será feita alternando as funções entre as diferentes camadas. A taxa de aprendizado também será alterada nas diferentes simulações. Depois de definida a arquitetura, serão refeitos os treinamentos usando otimizadores *GradientDescendentOptimizer*, *AdamOptimizer*, *AdagradOptimizer* e *RMSPropOptimizer*.

Os gráficos apresentados foram gerados com a biblioteca `matplotlib` da linguagem Python.

RESULTADOS E DISCUSSÃO

O melhores resultados obtidos foram com uma arquitetura de rede de quatro camadas ocultas. O treinamento deve assegurar que o valor do erro médio diminua sistematicamente, mas sem que a rede passe a "decorar" os dados, perdendo sua capacidade de generalização. A Figura 1 apresenta o gráfico dos erro obtidos em diferentes arquiteturas de quatro camadas ocultas. O modelo que alcançou o menor valor de erro está representado pela linha azul do gráfico (200 neurônios na primeira camada oculta, 100 na segunda, 50 na terceira e 25 neurônios na quarta).A rede alcançou o menor valor do erro na época 10, o que ocasionou a terminação do treino. A Figura 4camacur apresenta, para as mesmas arquiteturas, os gráficos da acurácia obtida.

Figura 2: Gráfico da acurácia para 4 camadas ocultas



Os otimizadores *AdamOptimizer*, *AdagradOptimizer* e *RMSPropOptimizer* foram usados, mas não se mostraram superiores ao *GradientDescendentOptimizer*, algoritmo de descida do gradiente tradicional do método. A combinação de funções de ativação usadas foi a função Elu para ativação das camadas intermediárias e a função SoftMax para a camada de sáida. A taxa de aprendizagem com melhordesempenho foi 0,001.

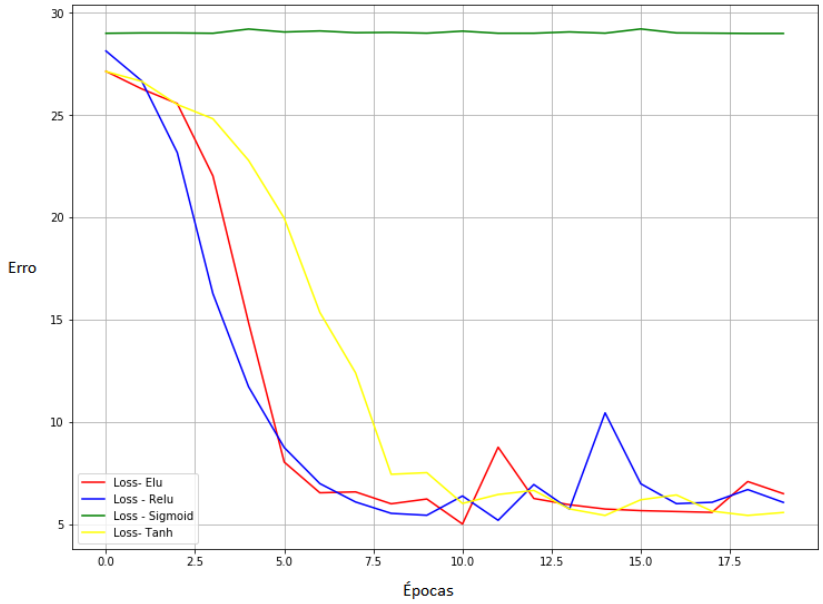
As funções de ativação das camadas ocultas utilizadas, além da função Elu, foram: *Relu*, *Sigmoid* e *Tanh* No gráfico representado na Figura 3 está os valores de erros obtidos.

A função que alcançou o menor valor do erro foi a função Elu (já utilizada inicialmente) em suas exatas 10 épocas de treinamento.

Após os testes com a substituição de otimizadores e funções de ativação a configuração do melhor resultado obtido foi de 4 camadas ocultas, a primeira com 200 neurônios, a segunda com 100, a terceira com 50 e a quarta com 25. Utilizando a função de ativação Elu, otimizador *GradientDescendentOptimizer*, taxa de aprendizagem de 0.001 e 2100 steps.

A acurácia foi de 94,5%. Para melhor visualização da taxa de acerto foi feita a matriz de confusão dos resultados representada na Tabela 1. Cada interseção na tabela representa o número de vezes que a rede, devendo dizer como valor de saída o elemento da linha, apresentou o elemento da coluna. A matriz de confusão apresenta, em sua diagonal principal, as vezes que a rede acertou o peso esperado das contusões para o dado de entrada recebido e nos demais espaços, o número de vezes que a rede errou. Como já era esperado, o maior número de erros na previsão foi em relação ao peso médio, visto que a diferenciação de leve para médio e de médio para pesado é arbitrária e abrupta. No caso, a rede previu um peso médio de contusões em 8 casos em que a saída deveria ser leve e em 17 que a saída deveria ser pesado, mas em 262 saídas ocorreu acerto. É importante notar que não houve nenhum caso em que a rede tenha errado uma classificação com 2 níveis de diferença: não ouve previsão de peso leve de contusões quando

Figura 3: Gráfico do erro com diferentes funções de ativação



deveria ser pesado ou vice-versa. Esse fato caracteriza as redes neurais como um bom método de previsão de peso de contusões no processo de transporte.

Tabela 1: Matriz de confusão

		Previsto		
		Leve	Médio	Pesado
Valor Real	Leve	48	8	0
	Médio	8	262	17
	Pesado	0	3	311

CONCLUSÕES

O transporte do gado de corte, das propriedades rurais até o frigoríficos, é causa de estresse para os animais transportados. Além das questões referentes ao bem-estar animal, o estresse gerado causa perdas financeiras, na medida que o gado estressado é mais propenso a atitudes agressivas ou de fuga, fazendo com que a qualidade da carne diminua e as perdas decorrentes de contusões impeçam seu aproveitamento. Estudos citados neste trabalho mostram que modificações no processo de embarque, transporte e desembarque, especialmente no que tange ao manejo dos animais, pode contribuir para a minimização deste problema.

Este trabalho fez uso de uma rede neural MLP para mapeamento das variáveis do processo de transporte, com dados coletados durante o período de um ano, em peso esperado de contusões decorrentes do processo nos animais. Os resultados mostram uma acurácia de 94,5% na predição do peso (referente à quantidade e gravidade) de contusões a partir da entrada das variáveis. A análise de sensibilidade da rede não foi realizada neste trabalho, mas pretende-

se que ela seja usada para o que o produtor possa inferir como alterações no seu processo de embarque e deslocamento ao frigorífico podem diminuir ou agravar as perdas decorrentes do não aproveitamento da carne. Da mesma forma, os frigoríficos podem fazer uso da rede para calcular o quanto precisam pagar para o produtor, em decorrência da perda esperada de seus animais, dadas as condições de embarque e as condições de deslocamento.

O motivo pelo qual a rede não atingiu uma acurácia mais próxima a 100% pode ser tanto pela limitação dos dados utilizados quanto por fatores ausentes dos dados. As condições meteorológicas do dia do transporte podem ter influência no fator estresse, especialmente se estiver muito quente e as paradas não forem feitas em local adequado. Como trabalhos futuros estão a inserção de dados climáticos como variáveis do modelo e a construção de métodos de coleta de dados mais confiáveis e automatizados, que permitam a escala da aquisição de dados para construção de melhores modelos de inferência.

AGRADECIMENTOS

Agradecemos ao Frigorífico Silva, de Santa Maria, pela cedência das instalações para a coleta dos dados usados neste trabalho.

REFERÊNCIAS

CAMARGO, M. da S. et al. Identificação de variáveis de relevância no Índice de contusões associadas ao transporte de gado de corte. In: *Congreso Argentino de AgroInformática*. [S.l.: s.n.], 2018.

FILHO, A. L. Produção De Carne Bovina No Brasil. *II Simpósio sobre desafios e novas tecnologias na bovinocultura de corte*, n. JANUARY 2006, 2006.

FRASãO, B. da S. et al. Quantity, location, and description of bruises in beef cattle slaughtered under sanitary inspection. *Acta Scientiae Veterinariae*, v. 42, p. 6, 2014. Publ. 1192.

GOOGLE BRAIN. *Tensorflow*. 2019. [Online; accessed 29-maio-2019]. Disponível em: <https://www.tensorflow.org/>.

HAYKIN, S. *Redes Neurais*. 2. ed. BOOKMAN, 2001. ISBN 9788573077186. Disponível em: <https://books.google.com.br/books?id=lBp0X5qfyjUC>.

HUERTAS, S. M. et al. Prevalence of carcass bruises as an indicator of welfare in beef cattle and the relation to the economic impact. *Veterinary Medicine and Science*, John Wiley & Sons, v. 1, p. 9–15, 2015.

HUERTAS, S. M. et al. Transportation of beef cattle to slaughterhouses and how this relates to animal welfare and carcase bruising in an extensive production system. v. 19, p. 281–285, 2010. ISSN 0962-7286.

HUERTAS, S. M.; KEMPENER, R. E. A. M.; EERDENBURG, F. J. C. M. van. Relationship between methods of loading and unloading, carcass bruising, and animal welfare in the transportation of extensively reared beef cattle. *Animals (Basel) 8, 119*, v. 17, n. 7, p. E119, 2018.

IBGE. Indicadores ibge - estatística da produção pecuária. *Instituto Brasileiro de Geografia e estatística*, 2018.

KLINE, H. C. *Carcass bruising location and bruise trim loss in finished steers, cows, and bulls at five commercial slaughter facilities*. 112p. p. Tese (Doutorado) — Colorado State University, 2018.

LAMA, M. de la et al. Atitudes dos retalhistas de carne ao bem-estar animal em espanha. *Meat Science*, v. 95, n. 3, p. 569–575, 2013.

MACQUEEN, J. Some methods for classification and analysis of multivariate observations. In: *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics*. Berkeley, Calif.: University of California Press, 1967. p. 281–297. Disponível em: <https://projecteuclid.org/euclid.bsmsp/1200512992>.

MAGGI, B. et al. . *Revista de Politica Agrícola*, 2006.

MENDONÇA, F. S. et al. Pre-slaughtering factors related to bruises on cattle carcasses. *Animal Production Science*, CSIRO, v. 58, n. 2, p. 385–392, 2018.

ROMERO, M. et al. Risk factors influencing bruising and high muscle ph in colombian cattle carcasses due to transport and pre-slaughter operations. *Meat science*, Elsevier, v. 95, n. 2, p. 256–263, 2013.

SAKAMOTO, Y.; ISHIGURO, M.; KITAGAWA, G. Akaike information criterion statistics. *Dordrecht, The Netherlands: D. Reidel*, Taylor & Francis, v. 81, 1986.

STRAPPINI, A. C. et al. Origin and assessment of bruises in beef cattle at slaughter. *Animal*, v. 3, n. 5, p. 728–736, 2009.

VIMISO, P.; MUCHENJE, V. A survey on the effect of transport method on bruises, ph and colour of meat from cattle slaughtered at a south african commercial abattoir. *South African Journal of Animal Science*, v. 43, n. 1, p. 105–111, 2013.

Use of artificial intelligence on the agroindustrial domain

Roberto Fray da Silva[1], Fernando Xavier[1], Gustavo Marques Mostaço[1], Antonio Mauro Saraiva[1], Carlos Eduardo Cugnasca[1]

[1] Departamento de Engenharia de Computação e Sistemas Digitais, Escola Politécnica da Universidade de São Paulo, São Paulo, São Paulo, Brasil, roberto.fray.silva@gmail.com, fxavier@usp.br, gmostaco@usp.br, saraiva@usp.br, carlos.cugnasca@usp.br

RESUMO

Atualmente, com o cenário de Big Data e IoT, uma quantidade crescente de dados é gerada por diversos dispositivos e aplicações. Isso representa uma ótima oportunidade para o desenvolvimento de pesquisas que busquem gerar informações úteis a partir desse volume de dados de maneira eficiente. As áreas de inteligência artificial (IA) e aprendizado de máquina englobam uma série de métodos e técnicas que possuem grande potencial para atingir este objetivo. No domínio agroindustrial, diversas pesquisas relacionadas ao desenvolvimento e implementação de IA têm sido realizadas, buscando principalmente tornar o processo produtivo mais eficiente e sustentável, dentre outros objetivos. Nesse sentido, esse trabalho realiza uma análise da produção científica envolvendo IA na agricultura, considerando o estado da arte de diversas aplicações bem como as iniciativas mais recentes. Devido a complexidade deste domínio, essa análise foi realizada através da categorização em três subdomínios: produção vegetal, produção animal e sustentabilidade e meio ambiente. O trabalho é finalizado com uma discussão dos resultados e a identificação das principais lacunas, que podem ser utilizadas para guiar as pesquisas futuras de IA no domínio agroindustrial.

PALAVRAS-CHAVE: Agricultura, Pecuária, Aprendizado de máquina.

ABSTRACT

Currently, with the Big Data and IoT scenarios, an increasing amount of data is generated by various devices and applications. This represents a great opportunity for the development of research that efficiently generates useful information from this volume of data. The areas of artificial intelligence (AI) and machine learning encompass several methods and techniques that have great potential to achieve this goal. In the agroindustrial domain, several pieces of researches related to the development and implementation of AI have been carried out. One of their main objectives is to increase the efficiency and sustainability of productive processes. In this sense, this work performs an analysis of the scientific production involving AI in agriculture, considering the state of the art of several applications, as well as the most recent initiatives. Due to the complexity of this domain, this analysis was performed through its categorization into three subdomains: plant production, animal production, and sustainability and environment. The paper also contributes with a discussion of the results and identification of the main gaps that can be used to guide future AI researches in the agroindustrial domain.
KEYWORDS: Agriculture, Livestock, Machine Learning.

INTRODUCTION

The agroindustrial domain is comprised of several companies, involved in all types of activities related to food production, from farm production to processing and transportation, warehousing and retailing. These agents, as well as their relations, can be analyzed through the supply chain perspective, which considers all companies directly or indirectly involved in fulfilling the customers' demands (CHOPRA; MEINDL, 2013).

Supply chains can generate a huge amount of data, and usually involve a variety of processes, such as environmental monitoring using sensors, product and information flow throughout the chain, process and product control inside each agent, among others. A considerable part of the current processes, especially at the farm level, could be improved by using artificial intelligence (AI) and machine learning (ML) methods and techniques. These can help in delivering recommendations and insights throughout the whole data lifecycle to support the agents during decision making (LIAKOS et al., 2018).

The term AI can be assigned to multiple definitions. Among those, Russel and Norvig (2016), define AI as the use of systems and techniques to model cognitive functions that are associated with human intelligence on machines. This encompasses learning, decision making, problem-solving, pattern recognition, among others.

ML is considered a component of AI and can be defined as the development of models that are used to analyze, explain and recognize patterns on observed data. Additionally, these models can be used to predict and evaluate the behavior of different variables or features (GHAHRAMANI, 2015).

Both AI and ML have many applications on the agroindustrial domain, ranging from the farm to the whole supply chain, including traceability, an essential topic for the agroindustrial supply chains (DABBENE; GAY; TORTIA, 2014). These encompass applications on crop production, animal production, sustainability, biodiversity, traceability, among others, which will be further explored in the Results section of this paper.

A broader evaluation of the use of AI and ML on this domain, clearly identifying its subdomains, their research problems, and the main relevant papers associated, could provide valuable information for current and future researchers on the fields of AI and ML. Therefore, this paper aims to identify the main agroindustrial subdomains, in which AI and ML are used, their open research questions and main papers, and identify important gaps that could be explored on future research, through an in-depth literature review.

METHODOLOGY

The methodology for this research can be divided into three main steps:

1. **Identification of the main agroindustrial subdomains** in which AI and ML are used and their characterization, based on a literature review. In this step, we also identified the main categories that comprised each subdomain;
2. **Search for papers in reference databases**, regarding the application of AI and ML models and techniques on the agroindustrial domain and its subdomains.
3. **Selection and analysis of the three most representative papers for each subdomain,** using as criteria the number of citations and the most relevant recent paper for them.

RESULTS

This section identifies the principal agroindustrial subdomains and present, for each, the selected papers related to the application of AI and ML.

The agroindustrial domain

Under the agroindustrial domain, three principal subdomains were identified and described, regarding its own problems and context. It is important to note that this classification was

related to the problems existing in these subdomains, and exclusively for those that could be solved by using ML and AI.

Liakos et al. (2018) characterized the agricultural subdomains as: (i) crop management; (ii) livestock management; (iii) water management; and (iv) soil management. Nevertheless, we believe that the water and soil papers could belong to a broader subdomain, that also encompasses environmental, social, and biodiversity papers.

For this reason, we adopt the following subdomains: (i) crop production; (ii) animal production; and (iii) social, environmental, and biodiversity. These will be further described in each of the following subsections. Even though these were considered separately in our analysis, in real life scenarios they are all mixed, and influenced by each other.

Crop production

Crop production is related to all processes involved in planting, growing and harvesting a specific crop. It is essential to be considered, since each crop has a growing pattern, as well as is influenced by soil, climate, pests, and diseases. Among the main activities related to crop production, there are: choose which crop to plant; manage soil fertility; plant the seeds; irrigate; manage pests and diseases; and finally harvest the crops. The main research questions in this subdomain were divided into the four categories below, and Table 1 relates the selected papers for each.

1. **Economics:** How to better predict product prices? Better predictions of future prices would allow farmers to improve their decision making in topics such as crop selection, product storage after harvest (aiming at selling off-season);
2. **Precision agriculture:** (i) how to develop and implement autonomous vehicles for processes such as planting, product application, and harvesting?; and (ii) how to perform in-field separation of products during harvest operations? Addressing the first question would allow for more remote and autonomous operation of the farm, while the second would lead to better product quality, as well as improvement on efficiency and loss reduction. Both would result in better profits for the farmers;
3. **Remote sensing:** how to monitor the lands and plant growth automatically using satellites and/or drone images? Solving these questions would allow for real-time problem detection, improving decision making, and becoming closer to the remote management of the farm.
4. **Pests and diseases:** how to automatically identify pests and diseases affecting a specific plantation field? This would allow for real-time problem identification,

improving decision making, resource use, and damage control, also leading to higher profits for the farmers.

Table 1 - Papers selected for crop production subdomain

Problem category	Papers	Models and Techniques*
1. Economics	Papageorgiou, Markinos, Gemtos (2011)	1-2-3-6
	Xiong, Li, Bao (2018)	2-3-7
	Abrougui et al. (2019)	1-3
2. Precision agriculture	Steen et al. (2016)	3
	Pérez-Ortiz et al. (2016)	2-4
	Bauer et al. (2019)	3
3. Remote sensing	Duro, Franklin, Dubé (2012)	2
	Kussul et al. (2017)	2-3
	Persello et al. (2019)	3-7
4. Pests and diseases	Mohanty, Hughes, Salathé (2016)	3
	Sladojevic et al. (2016)	3
	Gu, Wang, Wang (2019)	1-2-3-5-7

***** 1. Traditional statistics; 2. Traditional machine learning; 3. Deep learning; 4. Unsupervised learning; 5. Dimensionality reduction; 6. Fuzzy logic; 7. Others.

Animal production

This subdomain includes all activities related to animal breeding, rearing and finishing, such as cattle monitoring, health management, production management, disease identification, simulation, etc. For this subdomain, we divide the papers into the four categories below, and the selected papers are listed in Table 2.

1. **Economics:** how to better predict product prices? Just as in the case of crop production, better predictions of future prices would allow the farmers to improve their decision-making process;
2. **Precision livestock farming:** (i) how to automate different animal production processes?; and (ii) how to identify and monitor individual animal behavior, welfare, and production indexes among all stages of animal production? Solving these problems would result in animal welfare increase, better resource use and product quality improvement. This could also lead to better profits for the farmer;
3. **Computational simulations:** (i) how to operate the farm autonomously and optimally?; and (ii) how to better simulate and predict animal growth and welfare? Addressing these problems would allow for better resource use, animal welfare increase, and profit gains, while also contributing towards the development of a smart farm;

4. **Diseases and behavior identification:** (i) how to automatically identify and monitor diseases affecting the animals?; and (ii) how to automatically identify individual animal behavior? Solving these questions would allow for real-time problem identification, also improving decision making, resource use, and animal welfare.

Table 2 - Papers selected for animal production subdomain

Problem category	Papers	Models and Techniques*
1. Economics	Njubi et al. (2009)	3
	Ramírez-Morales et al. (2017)	3
	Shine et al. (2018)	2-3
2. Precision livestock farming	Gaber et al. (2016)	2-6
	Dunne et al. (2017)	2-4
	Rodríguez Alvarez et al. (2019)	3-7
3. Computational simulations	Shahinfar et al. (2012)	3-6
	Borchers et al. (2017)	2-3
	Shine et al. (2019)	2
4. Diseases and behavior identification	Carroll et al. (2014)	2-4
	Dutta et al. (2015)	1-2-4-5-6
	Machado et al. (2019)	2

***** 1. Traditional statistics; 2. Traditional machine learning; 3. Deep learning; 4. Unsupervised learning; 5. Dimensionality reduction; 6. Fuzzy logic; 7. Others.

Social, environmental and biodiversity

As mentioned before, this subdomain is related to the other two, since it involves elements directly linked to animal and vegetal production. However, the categories of this subdomain are focused on the interactions between production and the environment, such as water consumption, contamination risks, as well as impacts on climate change. For this subdomain, the papers were divided into three categories, described below, and are presented in Table 3.

1. **Contamination:** how to determine and quantify the impacts of contamination on soil, air, water, and people? Addressing this issue is essential to improve life on the planet, increasing social and environmental welfare;
2. **Resource use:** (i) how to estimate the footprint of different products?; (ii) how to better use the different resources on the farm? Just as the previous item, addressing these issues would improve both the health of the planet and resource use, also leading to better profits for the farmer;
3. **Climate change:** (i) how do agroindustrial operations and natural resource use affect the environment?; (ii) how to quantify the impacts of agriculture on climate change?;

among others. These issues are essential to improve decision making, of both companies and governmental agencies.

Table 3 - Papers selected for environment subdomain

Problem category	Papers	Models and Techniques*
1. Contamination	Rodriguez-Galiano et al. (2014)	2
	Bouzembrak, Marvin (2019)	7
2. Resource use	Khoshnevisan et al. (2015)	4-7
	Navarro-Hellín et al. (2016)	1-6
	Ma et al. (2019)	2-3
3. Climate change	Moriondo et al. (2013)	2
	Anaya-Romero et al. (2015)	7
	Feng et al. (2019)	1-2-7

***** 1. Traditional statistics; 2. Traditional machine learning; 3. Deep learning; 4. Unsupervised learning; 5. Dimensionality reduction; 6. Fuzzy logic; 7. Others.

DISCUSSIONS

In relation to the subdomain of plant production, it was possible to notice that several works could be classified in more than one category. With increasing data from IoT devices and Unmanned Aerial Vehicles (UAVs), perhaps more research may be useful to compare which techniques would be most appropriate for the volume of data that can be obtained by IoTs, UAVs, and satellites.

For the animal production subdomain, most of the work is related to the conditions of the animal, either for disease monitoring or to improve production. Although some works use devices such as collars and image acquisition, it is noted that there is more to be done in AI research in animal production that makes use of device data, such as those related to IoT, which may represent an opportunity for researchers in the field.

In environmental subdomain, several pieces of researches are related to the use of resources, such as water, soil, and energy, as well as to the contamination of groundwater and final products. In addition, work related to climate change was identified. In this sense, new research could be done to evaluate the resilience of agriculture in face of future climate change scenarios, such as the genetic improvement of crops and animals in order to adapt to temperature increases.

CONCLUSIONS

The agroindustrial domain is benefiting from the use of AI techniques, with increasing numbers of relevant research. Allied to the use of UAVs and IoT, AI has been increasingly

used in several activities of the agricultural chain, including crop and animal production, and also aspects related to the environment. With the use of forecasting models, producers can become more efficient and sustainable, by acquiring and treating information that can be fundamental for the decision-making process. Not least, works that seek to evaluate the effects of climate change on agriculture, this way evidencing possible impacts of production on the various scenarios of climate change, should be strongly encouraged. Finally, with the increased use of IoT in the field, more data will be generated, that needs treatment and analysis, resulting in an opportunity for further research on the use of AI in agriculture.

ACKNOWLEDGMENTS

This work was supported by Itaú Unibanco S.A. through the Itaú Scholarship Program, at the Centro de Ciência de Dados (C2D), Universidade de São Paulo, Brazil, by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brazil (CAPES) - Finance Code 001, and also by the National Council for Scientific and Technological Development (CNPq).

REFERENCES

ABROUGUI, K. et al. Prediction of organic potato yield using tillage systems and soil properties by artificial neural network (ANN) and multiple linear regressions (MLR). Soil and Tillage Research, v. 190, p. 202-208, 2019.

ANAYA-ROMERO, M. et al. Evaluating soil threats under climate change scenarios in the Andalusia Region, Southern Spain. Land Degradation & Development, v. 26, n. 5, p. 441-449, 2015.

BAUER, A. et al. Combining computer vision and deep learning to enable ultra-scale aerial phenotyping and precision agriculture: A case study of lettuce production. Horticulture Research, v. 6, n. 1, p. 70, 2019.

BORCHERS, M. R. et al. Machine-learning-based calving prediction from activity, lying, and ruminating behaviors in dairy cattle. Journal of dairy science, v. 100, n. 7, p. 5664-5674, 2017.

BOUZEMBRAK, Y.; MARVIN, H. JP. Impact of drivers of change, including climatic factors, on the occurrence of chemical food safety hazards in fruits and vegetables: A Bayesian Network approach. Food Control, v. 97, p. 67-76, 2019.

CARROLL, B. T. et al. Detecting symptoms of diseases in poultry through audio signal processing. In: 2014 IEEE Global Conference on Signal and Information Processing (GlobalSIP). IEEE, 2014. p. 1132-1135.

DABBENE, F.; GAY, P.; TORTIA, C. Traceability issues in food supply chain management: A review. Biosystems engineering, v. 120, p. 65-80, 2014.

DUNNE, R. et al. Behavior Classification of Dairy Cows Fitted with GPS Collars. In: Pacific-Asia Conference on Knowledge Discovery and Data Mining. Springer, Cham, 2017. p. 15-25.

DURO, D. C.; FRANKLIN, S. E.; DUBÉ, M. G. A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery. Remote sensing of environment, v. 118, p. 259-272, 2012.

DUTTA, R. et al. Dynamic cattle behavioural classification using supervised ensemble classifiers. Computers and Electronics in Agriculture, v. 111, p. 18-28, 2015.

FENG, P. et al. Incorporating machine learning with biophysical model can improve the evaluation of climate extremes impacts on wheat yield in south-eastern Australia. Agricultural and Forest Meteorology, v. 275, p. 100-113, 2019.

GABER, T. et al. Biometric cattle identification approach based on weber's local descriptor and adaboost classifier. Computers and Electronics in Agriculture, v. 122, p. 55-66, 2016.

GHAHRAMANI, Z. Probabilistic machine learning and artificial intelligence. Nature, 521(7553), p.452, 2015.

GU, S.; WANG, J.; WANG, Y.. Early discrimination and growth tracking of Aspergillus spp. contamination in rice kernels using electronic nose. Food chemistry, v. 292, p. 325-335, 2019.

KHOSHNEVISAN, B. et al. A clustering model based on an evolutionary algorithm for better energy use in crop production. Stochastic environmental research and risk assessment, v. 29, n. 8, p. 1921-1935, 2015.

KUSSUL, N. et al. Deep learning classification of land cover and crop types using remote sensing data. IEEE Geoscience and Remote Sensing Letters, v. 14, n. 5, p. 778-782, 2017.

LIAKOS, K. et al. Machine learning in agriculture: A review. Sensors, v. 18, n. 8, p. 2674-2703, 2018.

MA, Y. et al. Integration Agricultural Knowledge and Internet of Things for Multi-Agent Deficit Irrigation Control. In: 2019 21st International Conference on Advanced Communication Technology (ICACT). IEEE, 2019. p. 299-304.

MACHADO, G. et al. Identifying outbreaks of Porcine Epidemic Diarrhea virus through animal movements and spatial neighborhoods. Scientific reports, v. 9, n. 1, p. 457, 2019.

MOHANTY, S. P., HUGHES, D. P., & SALATHÉ, M. Using deep learning for image-based plant disease detection. Frontiers in plant science, v. 7, p. 1-10, 2016.

MORIONDO, M. et al. Projected shifts of wine regions in response to climate change. Climatic change, v. 119, n. 3-4, p. 825-839, 2013.

NAVARRO-HELLÍN, H. et al. A decision support system for managing irrigation in agriculture. Computers and Electronics in Agriculture, v. 124, p. 121-131, 2016.

NJUBI, D. M. et al. Milk yield prediction in Kenyan Holstein-Friesian cattle using computer neural networks system. Livestock Research for Rural Development. Volume 21, Article, v. 46, 2009.

PAPAGEORGIOU, E. I.; MARKINOS, A. T.; GEMTOS, T. A. Fuzzy cognitive map based approach for predicting yield in cotton crop production as a basis for decision support system in precision agriculture application. Applied Soft Computing, v. 11, n. 4, p. 3643-3657, 2011.

PÉREZ-ORTIZ, M. et al. Selecting patterns and features for between-and within-crop-row weed mapping using UAV-imagery. Expert Systems with Applications, v. 47, p. 85-94, 2016.

PERSELLO, C. et al. Delineation of agricultural fields in smallholder farms from satellite images using fully convolutional networks and combinatorial grouping. Remote Sensing of Environment, v. 231, p. 111253, 2019.

RAMÍREZ-MORALES, I. et al. Automated early detection of drops in commercial egg production using neural networks. British poultry science, v. 58, n. 6, p. 739-747, 2017.

RODRÍGUEZ ALVAREZ, J. et al. Estimating body condition score in dairy cows from depth images using convolutional neural networks, transfer learning and model ensembling techniques. Agronomy, v. 9, n. 2, p. 90, 2019.

RODRIGUEZ-GALIANO, V. et al. Predictive modeling of groundwater nitrate pollution using Random Forest and multisource variables related to intrinsic and specific vulnerability: A case study in an agricultural setting (Southern Spain). Science of the Total Environment, v. 476, p. 189-206, 2014.

RUSSELL, S. J.; NORVIG, P. Artificial intelligence: a modern approach. Malaysia; Pearson Education Limited,, 2016.

SHAHINFAR, S. et al. Prediction of breeding values for dairy cattle using artificial neural networks and neuro-fuzzy systems. Computational and mathematical methods in medicine, v. 2012, 2012.

SHINE, P. et al. Machine-learning algorithms for predicting on-farm direct water and electricity consumption on pasture based dairy farms. Computers and electronics in agriculture, v. 150, p. 74-87, 2018.

SHINE, P. et al. Annual electricity consumption prediction and future expansion analysis on dairy farms using a support vector machine. Applied Energy, v. 250, p. 1110-1119, 2019.

SLADOJEVIC, S. et al. Deep neural networks based recognition of plant diseases by leaf image classification. Computational intelligence and neuroscience, 2016.

STEEN, K. et al. Using deep learning to challenge safety standard for highly autonomous machines in agriculture. Journal of Imaging, v. 2, n. 1, p. 6, 2016.

XIONG, T.; LI, C.; BAO, Y. Seasonal forecasting of agricultural commodity price using a hybrid STL and ELM method: Evidence from the vegetable market in China. Neurocomputing, v. 275, p. 2831-2844, 2018.



MODELAGEM COMPUTACIONAL DE FLUXO EM APLICADOR DE AGROQUÍMICO AMPARADA POR SIMULAÇÃO COM EXPLORAÇÃO DE PARALELISMO

Mateus Specht[1], *Naylor Bastiani Perez*[2], *Leonardo Bidese de Pinho*[1]

[1]Universidade Federal do Pampa
Bagé, Rio Grande do Sul, Brasil
mateusspecht@unipampa.edu.br, leonardopinho@unipampa.edu.br

[2]Embrapa Pecuária Sul, Empresa Brasileira de Pesquisa Agropecuária
Bagé, Rio Grande do Sul, Brasil
naylor.perez@embrapa.br

RESUMO

A presente pesquisa analisa o funcionamento do aplicador seletivo de agroquımicos Campo Limpo, que promove o controle das ervas daninhas em pastagem por contato direto com aplicadores embebidos com a calda herbicida, visando entender melhor o escoamento e aprimorar sua precisão. Inicialmente a pesquisa é exploratória experimental e aponta para o uso de simulador fluidodinâmico, com especial atenção às técnicas de paralelismo, buscando a resolução de estudos de caso com restrições de tempo. A concepção de um ambiente de simulação foi conduzida a partir de dados obtidos em experimento laboratorial. A construção de um modelo da geometria em três dimensões para uso na construção da malha e, posteriormente, a execução do algoritmo de resolução das rotinas de cálculo demandam soluções de computação de alto desempenho. A aplicação de técnicas de paralelismo em diferentes ambientes computacionais propiciam a avaliação do tempo e gargalos da arquitetura. No caso estudado, o resultado da modelagem representada pelo simulador se mostrou satisfatória em relação aos fundamentos fluidodinâmicos e o resultado da análise do processamento demonstra um comportamento esperado para técnicas de paralelismo, quando avaliado o tempo de execução e a escalabilidade do ambiente computacional.

PALAVRAS-CHAVE: Computação Aplicada à Agropecuária, Computação de Alto Desempenho, Roçadeira Química.

ABSTRACT

The present research analyzes the operation of the selective agrochemical applicator Campo Limpo, which promotes the control of weeds in pasture by direct contact with applicators soaked with the herbicide syrup, in order to better understand the flow and improve its accuracy. Initially the research is experimental exploratory and points to the use of fluid dynamics simulator, with special attention to parallelism techniques, seeking the resolution of case studies with time constraints. The design of a simulation environment was conducted from the data obtained in a laboratory experiment. The construction of a three-dimensional geometry model for use in the construction of the mesh and, afterwards, the execution of the algorithm of resolution of the calculation routines demand high-performance computing solutions. The application of parallelism techniques in different computational environments allows the evaluation of the time and bottlenecks of the architecture. In the case studied, the modeling result represented by the simulator was satisfactory in relation to the fluidodynamic fundamentals and the result of the analysis of the process demonstrates an expected behavior for parallelism techniques, when the execution time and the scalability of the computational environment were evaluated.
KEYWORDS: Computation Applied to Agriculture, High Performance Computing, Chemical Trimmer.

INTRODUÇÃO

A indústria de alimentos de origem animal depende da capacidade da pecuária se sustentar no ambiente de competição contemporâneo. As mudanças climáticas e o uso intensivo das áreas modificam as condições de crescimento e a reprodução das plantas usadas como fonte de alimento de animais, tornando a atividade cada vez mais desafiadora.

Nas pastagens, as plantas mais nutritivas e preferencialmente escolhidas pelos animais são consumidas em uma velocidade maior que as demais e assim cria-se diferenças entre a população dos dois grupos em áreas de pastoreio. Além desta diferença do consumo, a competição entre as espécies fica mais acirrada com a presença de plantas com facilidades de adaptação ao ambiente.

As espécies indesejáveis acabam ganhando volume com a falta de consumo pelos animais, aumentando a competição com as plantas desejáveis por sol, água e nutrientes nas áreas destinadas a pecuária. Atualmente, existem diferentes técnicas para buscar o controle das plantas indesejáveis como o uso de agroquímicos (PEREZ, 2008). Ao longo dos anos implementos foram criados e aprimorados com o intuito de melhorar a qualidade da aplicação de agroquímicos com a contribuição da comunidade de usuários e dos fabricantes dos mesmos. A aplicação é comumente realizada por pulverização, contudo a técnica conta hoje somente com agroquímicos de ação ampla que controlam, indistintamente, diferentes gramíneas por meio de gotículas aspergidas no ar. O efeito de mortandade causado por pulverizadores convencionais acaba atingindo as plantas desejáveis, o que gera uma redução na oferta de alimento para os animais, por fim, uma redução na oportunidade de ganho de peso animal (PEREZ, 2010).

Implementos com diferentes formas de aplicar os agroquímicos são criados para não afetar as plantas importantes para a dieta animal. A Campo Limpo, mostrado na Figura 1, é um aplicador seletivo de agroquímico que permite atingir as plantas alvo com aplicadores umedecidos em um conjunto de tubulações e mangueiras acopladas, mostradas na Figura 1, ao longo de quatro metros de comprimento. Boa parte da concepção estrutural deste implemento se caracteriza pelo empirismo e pelo uso de peças fabricadas para fins diversos. O processo de desenvolvimento e melhoria da capacidade do equipamento ser mais preciso requer conhecimento envolvendo os fenômenos fluidodinâmicos ocorridos no escoamento de agroquímico no aplicador (PEREZ, 2010).

Figura 1: Campo Limpo e Hastes de Distribuição



Fonte: (PEREZ, 2008)

As modelagens fluidodinâmicas desenvolvidas em ambiente computacional (*Computacional Fluid Dynamics* - CFD) são vastamente utilizadas para diversos problemas de engenharia. Dentre os mais abordados estão os de aerodinâmica de veículos sendo aplicáveis nas mais diversas áreas da ciência. Existem abordagens cientificas na área de CFD que buscam possibilitar a modelagem de sistemas com equações de comportamento, cuja resolução é complexa e, muitas vezes, obtida somente a partir de técnicas computacionais (FERZIGER; PERIC, 2012).

A construção de modelos CFD podem exigir alto desempenho do ambiente computacional e técnicas para melhorar esse processamento são fundamentais para viabilizar a resolução das equações de estado do fluxo. O paralelismo é uma forma de execução simultânea para a resolução de problemas com o uso de diversos recursos computacionais, na qual os problemas são divididos em partes menores para serem resolvidos concorrentemente (KUMAR, 2002).

A capacidade de executar processos em unidades de processamento diferentes com técnicas de paralelismo viabilizaram o aumento da utilização de ferramentas de fluidodinâmica computacional. O uso de simuladores para obter um maior conhecimento sobre as características do escoamento podem tornar o processo de aplicação mais eficiente. O aperfeiçoamento desta máquina pode gerar um menor impacto ambiental por manter a aplicação precisa e focada nas plantas alvo, além de aumentar a eficiência das pastagens e o ganho de peso animal.

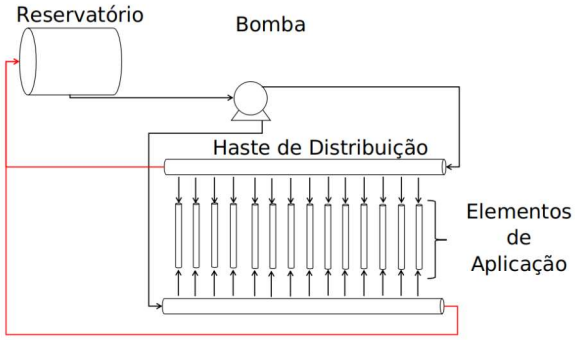
O trabalho possui o objetivo de desenvolver, por meio de simulação com exploração do potencial de paralelismo de ambientes computacionais, um modelo fluidodinâmico da distribuição

de calda de herbicida em uma mangueira do aplicador de agroquímico seletivo Campo Limpo destinado ao controle de ervas daninhas em pastagens, visando aprimorar o conhecimento sobre o equipamento.

MATERIAL E MÉTODOS

O escoamento do implemento está representado na Figura 2, o reservatório estoca a calda herbicida, a bomba transfere a calda para as hastes e estas distribuem a carga nos elementos de aplicação. A bomba é do tipo peristáltica e possui atuação para controlar a velocidade do rotor e possibilita a alteração da vazão para a haste.

Figura 2: Esquemático da Campo Limpo



Fonte: (Autor, 2019)

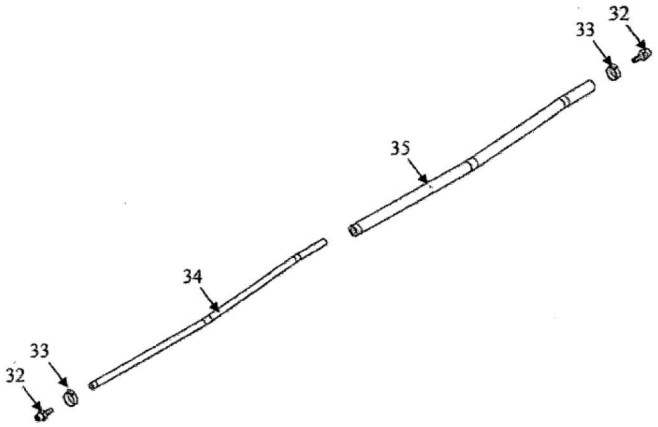
O elemento de aplicação é composto internamente por uma mangueira plástica com furações, representada pelo número 34 na Figura 3, e externamente por uma corda, número 35 Figura 3. Existem dois furos na mangueira simetricamente afastados a 5 cm das extremidades com diâmetros de 1 mm. O comprimento da mangueira é de 60 cm e o diâmetro interno de 5 mm e 10mm externo, dessa forma os furos possuíam comprimento de 2,5 mm ao longo da parede.

A calda entra pela extremidade das mangueiras por ambos os lados e sai da mangueira exclusivamente pelos furos descritos. Após sair pelos furos entra em contato com a corda, que umedece e, por fim, ao tocar a planta transfere um quantidade de calda para as folhas. Nos lados opostos das hastes é conectado um retorno para o reservatório com o objetivo de destinar a calda excedente, representados na Figura 3 pelas linhas vermelhas.

O OpenFOAM é um software de simulação CFD e é escolhido considerando as potencialidades de uso de paralelismo e por ser um software com código fonte aberto. Mohiuddin e Mathkour (2015) ranqueou um conjunto de programas a partir de critérios pré definidos e o OpenFOAM se destacou como uma opção deste conjunto por sua flexibilidade pode ser adaptado a diferentes problemas, em contrapartida não possui uma interface gráfica com usuário.

O simulador conta com diferentes etapas, representadas na Figura 4: pré-processamento, construção da geometria da malha de simulação, definição de condições iniciais e condições de contorno; solver, execução das rotinas de cálculo com o sistema de equação selecionado através

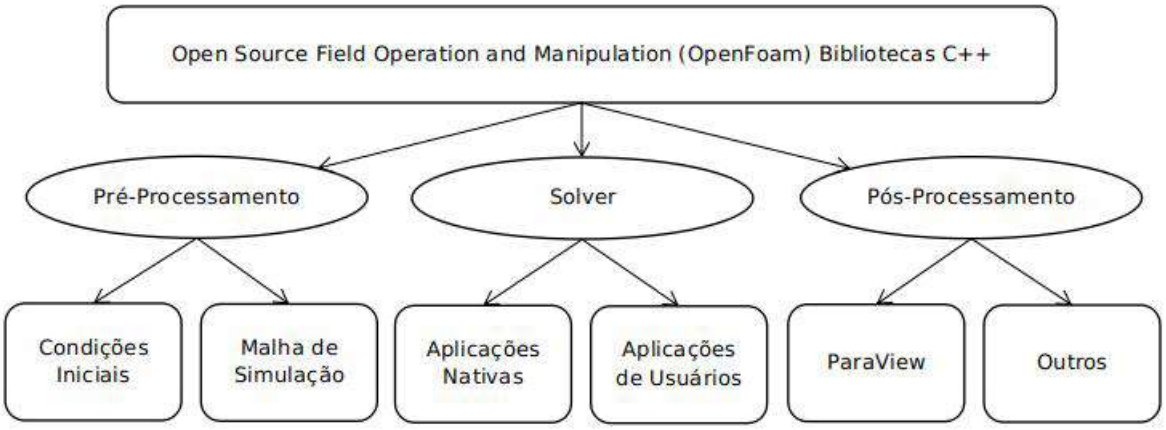
Figura 3: Esquemático do Elemento de Aplicação



Fonte: (PEREZ, 2011)

de rotina criada por usuário ou nativa do OpenFOAM; pós-processamento, visualização dos resultados obtidos com uso de ferramentas de manipulação nativas ou externas (compatibilidade com outros simuladores).

Figura 4: Esquema de Etapas da Simulação no OpenFOAM



Fonte: adaptado de (OPENCFD, 2019)

Neste trabalho a mangueira, representada na Figura 3 pelo número 34 localizada no interior do elemento de aplicação, é definida como volume de simulação considerando somente a parte interna do tubo com duas entradas e duas saídas. O intuito do foco em um elemento é explorar com um problema de menor complexidade as soluções disponibilizadas pelos simuladores CFD e demais ferramentas para a construção do ambiente computacional adequado.

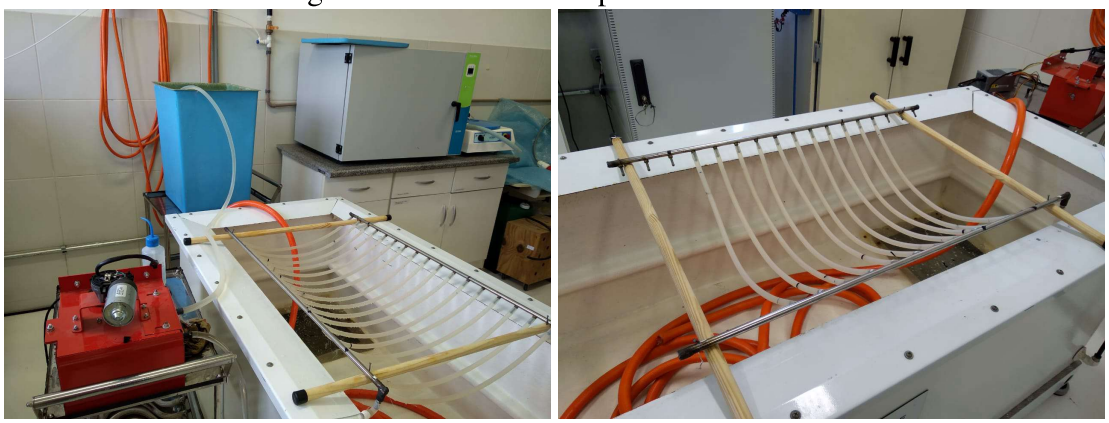
Nestas condições a partir da vazão de saída do tubo é possível estimar, a vazão nas entradas da mangueira e a partir da caracterização do escoamento determinar as condições iniciais e de contorno do sistema. A velocidade, as características geométricas e as físico-químicas definem o comportamento predominante do fluxo ou mais ou menos viscoso quantificado pelo número de Reynolds.

Experimento Fluidodinâmico

O experimento é construído para simular condições reais de operação da Campo Limpo e prover dados para a construção do ambiente de simulação e definição de condições iniciais das equações de estado do fluido. As cordas foram retiradas dos elementos de aplicação para facilitar a medida de vazão em cada furo da mangueira, através de uma proveta volumétrica e um cronômetro com medidas tomadas em triplicata para posteriormente a média dos dados ser utilizada como entrada para o simulador.

O reservatório foi alimentado com água potável e durante o experimento o volume de 40 litros foi mantido constante para evitar efeitos de pressão indesejáveis. A bomba peristáltica com um de seus êmbolos envia o fluido para as hastes de distribuição em direções opostas, assim como é executado na Campo Limpo e representado pela Figura 5. Após entrar nas hastes, a água sai pelos furos em todas as mangueiras e as medidas de vazão só ocorreram no momento em que todo o ar aprisionado nas tubulações havia sido retirado do sistema experimental.

Figura 5: Bancada de Experimento Fluidodinâmico



Fonte: (Autor, 2019)

Ambiente da Simulação

O *OpenFoam* possui algumas ferramentas nativas para a construção da geometria de simulação. O uso das ferramentas *BlockMesh* para construção da malha de fundo e *SnappyHexMesh* para o refinamento da malha de fundo na geometria final. A média dos dados do experimento fluidodinâmico é utilizada para definir o campo de velocidade das entradas das mangueiras.

A Equação 1 representa a expressão matemática para obtenção do número adimensional de *Reynolds*, que prevê uma forma de classificar o comportamento de um escoamento em laminar ou turbulento. A diferença dos diâmetros dos furos e da velocidade garante números de *Reynolds* distintos para as entradas e saídas. Na prática dentro do contexto dos simuladores essa classificação define qual linha de solver pode ser usada para a resolução do problema.

$$Re = \frac{D.V.\rho}{\mu} \tag{1}$$

Outra classificação do escoamento para definição do solver é em relação a variabilidade de suas propriedades ao longo do tempo. O tempo zero da simulação corresponde a uma situação de velocidade igual a zero em todo o sistema e no instante de tempo posterior a velocidade das entradas é a máxima. Essa condição de escoamento gera a variabilidade dos campos de velocidade no mesmo ponto ao longo do tempo, dessa forma sendo classificado como um escoamento em regime transiente.

O solver utilizado na simulação foi o *IcoFoam* que resolve as rotinas de cálculo, armazena em bibliotecas e todos os resultados possuem referência tridimensional na geometria de simulação. O software *ParaView* é indicado para a visualização dos resultados do *OpenFOAM* e possui ferramentas para análise e manipulação de dados.

Experimento Computacional

O experimento computacional é conduzido utilizando a aplicação *OpenFoam* na sua versão 6.0 com uso do *OpenMPI* versão 2.1.1, do *ParaView* 5.6.0 e é executada em um servidor *HP ProLiant ML350 G6* com um processador *Intel Xeon CPU E5620 @ 2.40 GHz* com 4 núcleos e 8 *Threads* , 8 GB de Memória *ECC DDR3*, Placa de Vídeo *NVIDIA Tesla C2075* com 6 GB de memória *GDDR5* e Disco Rígido de 500 GB configurados em *RAID 1* com o sistema operacional Linux Ubuntu versão 18.04.2 *LTS*.

Um ambiente de processamento é criado com um, dois e quatro processadores para analisar o tempo necessário para executar os casos em diferentes ambientes computacionais. O estudo será separado para cada etapa, uso da ferramenta *SnappyHexMesh* e do solver, com o intuito de constatar nuances entre as diferentes etapas e o desempenho do processamento considerando somente a diferença na distribuição do trabalho computacional.

A métrica utilizada para a comparação de desempenho entre dois processos realizando o mesmo trabalho é o *Speedup*, que, em outras palavras, é o o ganho de velocidade por utilizar os recursos computacionais semelhantes de formas diferentes. A Equação 2 representa a equação matemática para o *Speedup*, na qual o termo do numerador é o tempo de execução em um, dois ou quatro processadores e o do denominador é o tempo de execução do processo somente em com um processador (LIU, 2011).

$$Speedup = \frac{t_n}{t_1} \tag{2}$$

RESULTADOS E DISCUSSÃO

O sistema laboratorial teve seus componentes testados separadamente e a bomba apresentou falha em um dos rotores o qual estava com uma das mangueiras com sua estrutura colapsada. A estrutura possui uma posição fixa para as mangueiras sem batente móvel, o que acaba forçando as mangueiras a se manterem deformadas mesmo com a bomba fora de funcionamento

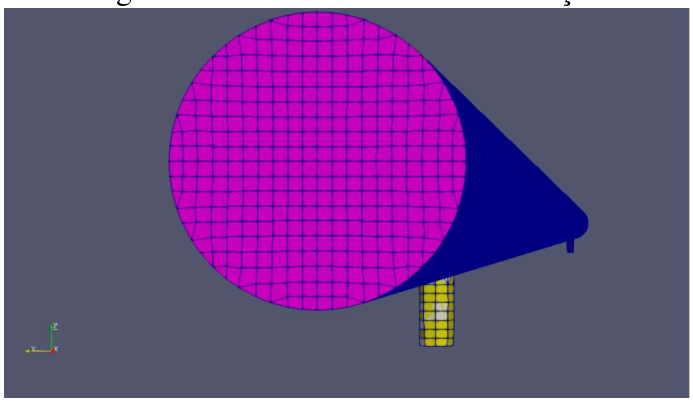
Os resultados das medições de vazão obtiveram uma média de 426 microlitros de agroquímico por segundo por furo com desvio padrão de 92 microlitros, considerando os trinta furos das man-

gueiras com medidas realizadas em triplicatas. Essa informação é usada para definir o campo de velocidade e por fim o cálculo de número de *Reynolds* com os resultados acima de 500, que é dentro dos parâmetros para escoamento laminar.

As fronteiras do sistema demarcam a interface entre o escoamento e a parede do tubo, de forma a representar somente o fluido e não a tubulação. A geometria de fundo criada com o *BlockMesh* e possui 70 centímetros de comprimento e um paralelepípedo de secção transversal quadrada com aresta de 6 milímetros. Estas medidas são estipuladas em função das dimensões da geometria descrita no arquivo STL, o qual será usado como forma para refinar a geometria de fundo na geometria final.

Os furos para transferência do agroquímico para a corda saem da geometria da tubulação principal de forma a representar trecho referente a parede da tubulação. Os furos estão definidos a uma distância de cinco centímetros da entrada do agroquímico em direção ao centro da mangueira, conforme é possível visualizar na Figura 6.

Figura 6: Geometria Final de Simulação

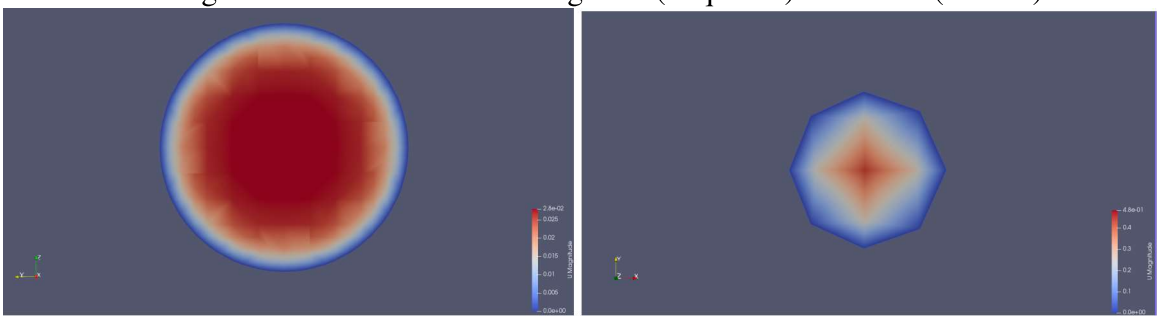


Fonte: (Autor, 2019)

Resultados expostos mostram uma redução de velocidade significativa no fluido, conforme é mostrado na Figura 7, nas proximidades da parede do tubo, no centro a velocidade máxima de 0,028 m/s aproximadamente. Já na seção do furo na parede da mangueira a velocidade máxima no centro é de 0,48 m/s, que é uma velocidade 17 vezes maior do que a velocidade no centro da mangueira. O aumento de velocidade é característico para escoamentos com diferentes diâmetros (MITCHELL et al., 2015).

As escalas do *ParaView* são adaptadas para a magnitude do conjunto de resultados mostrados na tela do software e a cada alteração da visão da geometria o sistema de escala precisa ser atualizado para o novo intervalo de dados. O sistema de coordenadas utilizado para a visualização conta com três eixos cartesianos e para cada unidade de tempo simulada há um conjunto de resultados já referenciados na geometria. Na idealidade o valor do *Speedup* na situação deste trabalho são de 2 e 4 para a comparação entre o tempo necessário para execução de um trabalho computacional com o uso de 2 e de 4 processadores,respectivamente, em relação a execução em série. Os valores obtidos na alteração de arquitetura foram dispostos no gráfico

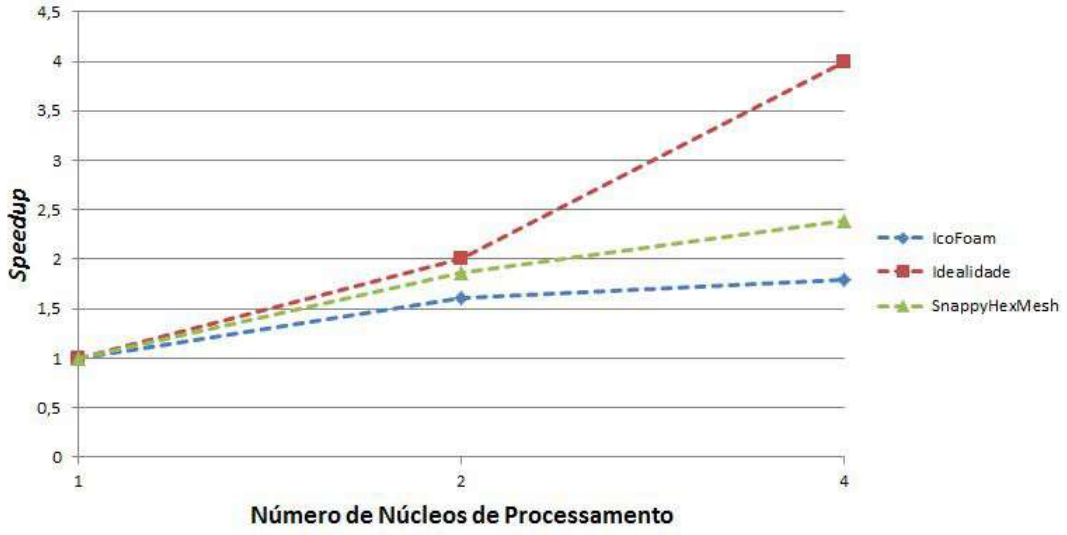
Figura 7: Corte Radial da Mangueira (Esquerda) e do Furo (Direita)



Fonte: (Autor, 2019)

representado pela Figura 8. Os resultados para as duas ferramentas estão abaixo da idealidade nos dois casos, embora no processamento com 2 núcleos estejam mais próximos do ideal. Em ambos os testes o *Solver* obteve o pior resultado em relação à ferramenta *SnappyHexMesh*.

Figura 8: *Speedup* da Ferramenta Geradora da Malha (*SnappyHexMesh*) e do *Solver* (*IcoFoam*)



Fonte: (Autor, 2019)

CONCLUSÕES

O experimento laboratorial foi realizado e auxiliou a determinação das condições de entrada do simulador e subsidiou a escolha do solver e das condições iniciais do processo. Nas condições testadas o fluxo não é capaz retornar ao reservatório e escoa completamente pelas mangueiras.

A malha de simulação criada para o sistema satisfez a necessidade para uma primeira avaliação dos fenômenos fluidodinâmicos. O comportamento modelado demonstra consistência com sistemas teóricos, nos quais a velocidade cresce de zero próximo da parede até o valor máximo no centro da tubulação. Além da diferença significativa entre os dados de velocidade simulados entre a entrada e a saída, que também é um comportamento compatível, quando considerada a diferença entre os diâmetros entre das mesmas.

As etapas de refinamento da malha e de execução do solver tiveram seu processamento comparado em diferentes ambientes computacionais e demostraram um comportamento distante do ideal principalmente para o composição de quatro núcleos.

REFERÊNCIAS

FERZIGER, J. H.; PERIC, M. *Computational methods for fluid dynamics*. Berlim: Springer Science & Business Media, 2012.

KUMAR, V. *Introduction to parallel computing*. Boston: Addison-Wesley Longman Publishing Co., Inc., 2002.

LIU, Y. *Hybrid Parallel Computation of OpenFOAM Solver on Multi-Core Cluster Systems*. Tese (Doutorado), 2011. Disponível em: <http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-31175>.

MITCHELL, J. W. et al. *Fox and McDonald's Introduction to Fluid Mechanics*. Estados Unidos: Wiley, 2015.

MOHIUDDIN, I.; MATHKOUR, H. Computational fluid dynamics application tools. In: IEEE. *2015 World Congress on Information Technology and Computer Applications (WCITCA)*. Hammamet, 2015. p. 1–5.

OPENCFD. *OpenFOAM®*. 2019. Disponível em: <https://openfoam.org>.

PEREZ, N. B. Aplicador manual de herbicida por contato: enxada química. *Embrapa Pecuária Sul-Comunicado Técnico (INFOTECA-E)*, Bagé: Embrapa Pecuária Sul., 2008.

PEREZ, N. B. Controle de plantas indesejáveis em pastagens: uso da tecnologia campo limpo. *Embrapa Pecuária Sul. Comunicado técnico*, Bagé: Embrapa Pecuária Sul., 2010.

PEREZ, N. B. *APERFEIÇOAMENTO EM DISPOSITIVO APLICADOR DE HERBICIDA*. Dez 2011.

Labeling images with the LabelIt! app

Vitor Trevisan[1], *Thiago Luís Lopes Siqueira*[1]

[1] Instituto Federal de Educação, Ciência e Tecnologia de São Paulo, Piracicaba, SP, Brasil,

t.vitor@aluno.ifsp.edu.br , prof.thiago@ifsp.edu.br

RESUMO

A inclusão digital e a utilização de sistemas baseados em internet-das-coisas (IoT), visão computacional, e aprendizado de máquina têm sido ampliadas na agropecuária brasileira. A combinação desses métodos computacionais robustos junto a tecnologias emergentes possui um grande potencial de impactar a agropecuária. Dentre os dados obtidos e transferidos por dispositivos IoT estão as imagens (fotografias). Já os algoritmos de aprendizado de máquina comumente requerem imagens rotuladas (anotadas) como entrada para posterior treinamento. Deste modo, uma ferramenta de rotulação de imagens possui especial relevância entre as imagens dos dispositivos IoT e os algoritmos de aprendizado de máquina, pois recebe imagens como entrada e fornece como saída imagens rotuladas que servem de entrada para os algoritmos. Neste artigo, nós descrevemos a ferramenta de rotulação LabelIt! que desenvolvemos para esta finalidade.

PALAVRAS-CHAVE: imagens, rotulação, aplicativo.

ABSTRACT

Digital inclusion and the adoption of solutions based on internet-of-things (IoT), computer vision and machine learning are increasing in Brazilian agriculture and livestock. The combination of these robust computational methods with emerging technologies have a great potential to impact agriculture and livestock. Images (pictures) are among the data obtained from and transferred by IoT devices. Machine learning algorithms often demand labeled (annotated) images to learn how to imitate the training dataset. Therefore, an image labeling tool has a particular relevance between the images of IoT devices and machine learning algorithms, because it has as input images and produces as output labeled images that can be

used by the algorithms. In this paper, present a tool for labeling images called LabelIt! that we developed to comply with this purpose.
KEYWORDS: images, labeling, app.

INTRODUCTION

According to a survey carried on 2017, more and more farmers in Brazil have a smartphone connected to the Internet (ABMRA, 2017). In the state of Rio Grande do Sul, for instance, seven in ten producers interviewed used a smartphone for management (GAÚCHAZH, 2018). Furthermore, 67% of these devices were connected to the Internet. The usage of smartphones not only connect farmers to the Internet, but also allow them to produce and collect data with the device, such as pictures, videos, audios, positions, routes, texts, among others. Clearly, the digital inclusion is increasing in Brazilian rural areas. As a result, there are plenty of opportunities to build solutions that combine both robust computational methods with emerging technologies to achieve outstanding impacts. Furthermore, specialists emphasize the necessity of building solutions for real problems, simplifying the presentation of solutions for the producers (despite the inherent complex internal design), and training farmers to use the solutions (SANTOS et al., 2019).

Internet-of-things (IoT), computer vision and machine learning have been increasingly adopted and combined to monitor the livestock in order to both maximize farm owners' yield and decrease production costs (LIAKOS et al., 2018). Sensors periodically collect data across interconnected (mobile) devices, which communicate and may feed a voluminous database. The latter is read by machine learning algorithms aiming at decision support, recommendation and event prediction. These technologies are applied, for instance, on behavior analysis (LOMBA et al., 2015; TSCHARKE; BANHAZI, 2016; GARCIA et al., 2017), welfare (LUKAS; TANUMIHARDJA; GUNAWAN, 2015), nutrition (LUIZ et al., 2018), race classification (SANTONI et al., 2015), among others.

The combination of computer vision and supervised machine learning systems requires tasks such as collecting data, training data, and training the model (ALVES, et al. 2018). A reasonable amount of representative data must be collected. Accurate labeling (annotation) is essential to transform collected data into training data, because machine learning algorithms learn to act like the input training datasets. A dataset containing thousands of images may be necessary to train an initial model.

Labeling often consist of drawing boxes around objects of an image and associating a label. Alternatively, a label can be assigned to the whole image. Therefore, labeling play a key

role, since voluminous image datasets collected using IoT cannot be processed by machine learning algorithms if labeling was not executed to transform collected data into training data. For instance, deep learning (a prominent subset of machine learning) became feasible after millions of images were collected from the web and labeled by human labelers (KRIZHEVSKY; SUTSKEVER; HINTON, 2012). Although labeling images is crucial, it is commonly considered a time-consuming and tedious task (VON AHN; DABBISH, 2004). Several tools for labeling images exist due to the relevance of this task. We briefly summarize some of them in the following.

The Image Labeler by Mathworks[1] allows the user to label images interactively into the MATLAB desktop environment. A label has a name (e.g. vehicle), a shape (rectangle, line or point) and a description. Sublabels are similar and keep a part-whole relationship with a label (e.g. taillight of a vehicle). Attributes and their values describe a given (sub)label (e.g. vehicle type and car or truck). Labels can be drawn by the user on the image (or video frame).

LabelImg[2] is a desktop application written in Python designed for image annotation that provides a graphical user interface based on Qt. The user is able to draw rectangles over an image and assign a label to each one. Thereafter, the user can choose a proper file format to save the annotation (e.g. XML).

The Labelbox[3] is a web application that provides an image labeling interface where the user can annotate images and videos in various ways. Additionally, the user is able to customize the interface to suit the needs of a particular labeling task. It requires sign up and sign in and allow collaborators to be included in a given labeling project. Images are uploaded from a local computer or even from cloud computing services. Machine learning models created by different contributors can be run and have their performance compared.

The Image Labeler by Google[4] displays a previously labeled image and ask users whether it was correctly categorized. The user answers “yes” or “no”, or skip if not sure. Different categories of images can be selected in a menu. The aim of the system is to help improve the quality of image classification and searching. In the past, it had a game to entertain users during the task.

The VGG Image Annotator is a standalone software for manual annotation of image, audio and video. It runs as an offline application in most modern web browsers (DUTTA;

[1] https://www.mathworks.com/help/vision/ug/get-started-with-the-image-labeler.html

[2] https://github.com/tzutalin/labelImg

[3] https://support.labelbox.com/docs

[4] http://g.co/imagelabeler

ZISSERMAN, 2019). Several shapes of labels are available and the user can manage label types and attributes (e.g. describing as both “blurred” and “with good illumination”).

We verified a gap between IoT systems, which collect and transfer agriculture and livestock image data, and machine learning algorithms. Thus, our objective became to provide an information system with a labeling tool to fill the gap. In specific, we the tool should run in mobile devices, omit the complexity of the labeling task, allow multiple users to label images with their perceptual abilities, maintain users’ labeling work safely stored to be reloaded whenever necessary. This paper describes the LabelIt! app that we developed to satisfy these requirements.

MATERIAL AND METHODS

A project hosted by our institute enabled interactions with agtech companies. We studied and executed existing labeling software. A literature review focused on IoT, computer vision and machine learning applied to agriculture and livestock. The software has been developed through incremental prototyping and its source code was written according to object-oriented programming.

To design our information system, we separated the tasks to be run on the server from those to be run on the clients. On the one hand, users catalog, authentication permissions, and labels database remains on the server. On the other hand, a client displays images and allows for labeling them. To enable these tasks to be accomplished on both sides of the system, we utilized cloud computing and mobile app development. The server was deployed on a cloud computing platform and inherits its advantages, such as availability, flexibility and scalability. The client was implemented as an app that runs on smartphones and tablets. The communication between the clients and the server is either through a mobile data network or a Wi-Fi network.

The Google Cloud Platform was chosen as it supplies a server-side infrastructure for storage and networking as well as authentication and role-based access controls based on accounts (that can be linked to users’ e-mail accounts, for example). As for the client-side, the mobile app development comprised both Android and Firebase platforms. The source code was written in Java programming language. We used the Android Studio integrated development environment.

We used Cloud Firestore to store the database at server-side. It is a semi-structured NoSQL document database organized as collections and documents (REDMOND, 2012), rather than tables. A collection has several documents. A document is described by several

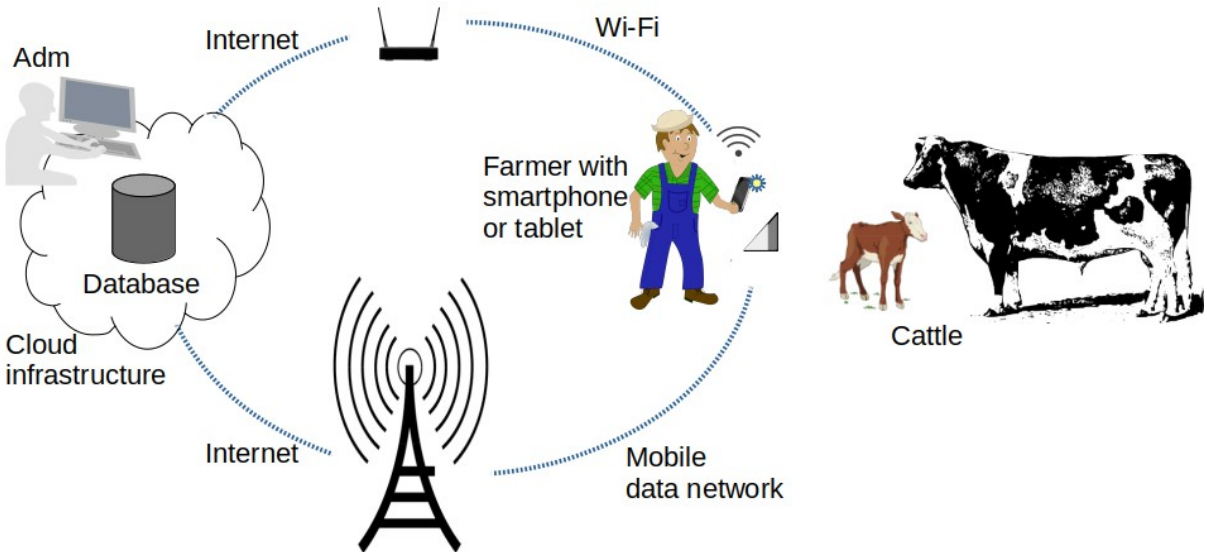
attributes. Documents in the same collection can share the same attributes or have different ones. An attribute may have assigned either a value, a document or a collection.

We also went to a farm and photographed around one hundred different scenes of beef cattle (Canchim and Nelore) and dairy cattle, which were either grazing, confined, or milking. These photos composed the real dataset. We also downloaded a few pictures of bovine, caprine and ovine cattles, containing animals of the same or different species, genders, ages or races. These pictures composed the downloaded dataset.

RESULTS AND DISCUSSIONS

We developed an information system whose architecture is depicted in Figure 1. On the client side, a farmer is the end user of the system. It uses a mobile device with camera such as a smartphone or a tablet to take pictures, e.g. cattle. The user executes the LabelIt! app to open the pictures taken (or even other images) and label them. The labeled data are sent to the server through the Internet and inserted (or updated) in the database. Although only one farmer was depicted, our system supports multiple clients. On the server side, a database administrator creates the database and manages access rules for users. Unauthorized users are not allowed to connect to the database and therefore cannot execute the app. The following subsections detail both the app and the remote database. Thereafter, we discuss the results.

Figure 1 – System architecture.



The LabelIt! app

On the client side, the LabelIt! app provides a straightforward interface for the user to allow the creation and removal of labels on images. Figure 2 illustrates screenshots of the app. First, a welcome screen customized for livestock asks whether the user can start labeling or needs to login, as shown in Figure 2a. Second, the login screen, asks the user to provide the credentials to access the remote database, as shown in Figure 2b. If the user is authorized or has already logged in before, then it is possible to continue. Third, the available pictures are listed and the user can select one of them to label, as shown in Figure 2c. Fourth, the labeling screen is displayed and the user can: (i) add as many labels are necessary by clicking on the "+" button; (ii) submit the labeling work by clicking on the "➢" button; or (iii) remove all labels by clicking on the bin button. After submitting, the user can choose another image to label and is redirected to the screen shown in Figure 2b.

It is not necessary to draw every edge of the rectangle. Instead, only one of its diagonals is sufficient. To add a label, the user first draws a rectangle on the image by sliding the finger on the screen. Then, a dialog box appears and the user chooses from a predefined list of labels or types a customized label, as shown in Figure 2d. Thereafter, the rectangle is colored and the label appears at its bottom. All rectangles are colored differently, as shown in Figure 2e.

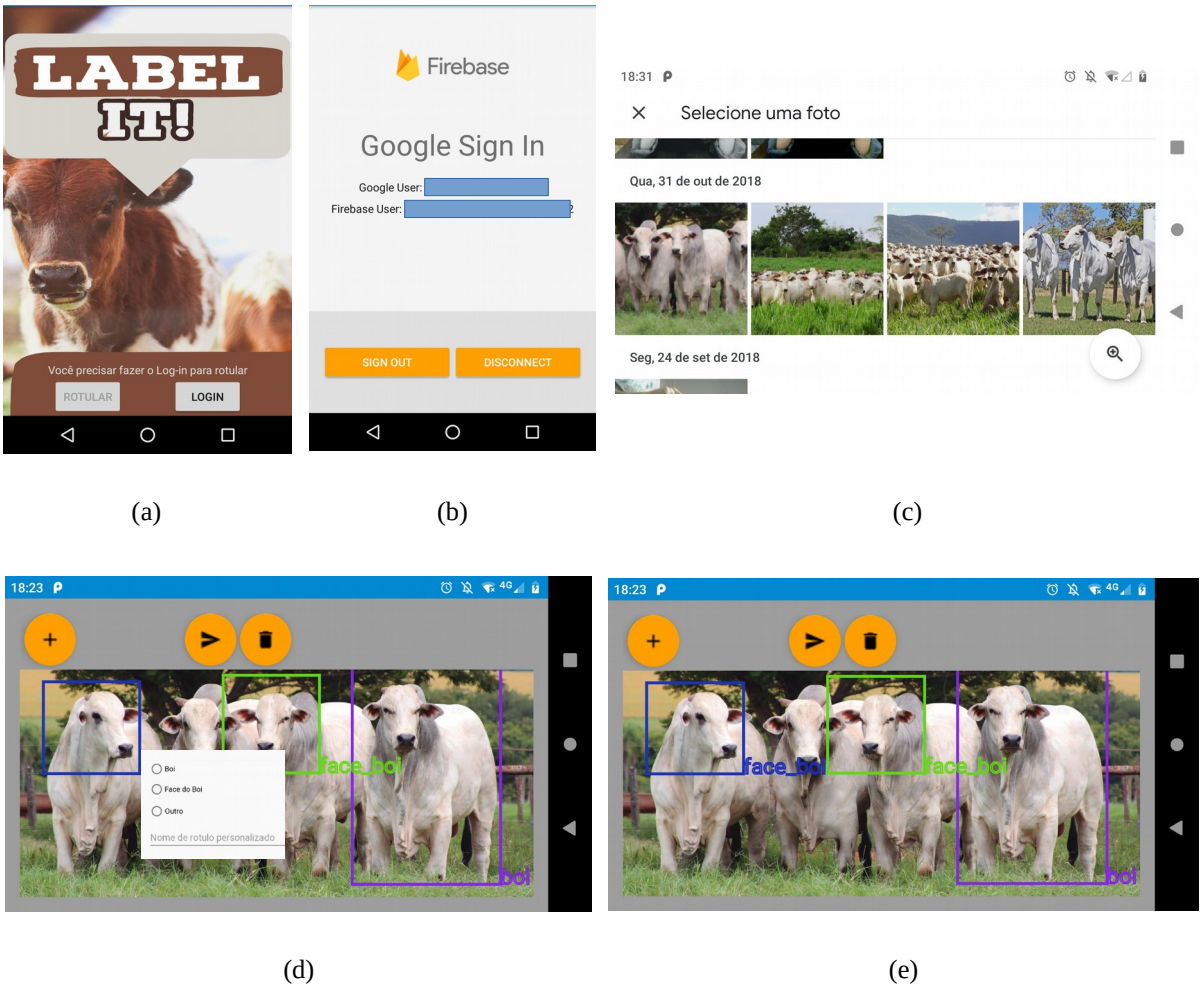
The LabelIt! App does not require advanced technical skills from the end user (farmer). To sum up, the user needs to know its username and password and be careful enough to draw the rectangles and do not damage the screen. If the user wants to label the pictures taken, then it is necessary to know how to operate the camera of the device.

Our app takes advantages of mobility. If the user decides to label images where there is no Internet connection, the app will keep data stored locally. As soon as the device is connected again and the app is executed, the user's labeling work is sent to the server. This resource is equally useful if the connection is unstable.

After the implementation, we executed and tested the app in devices. The first ran Android 8 (Oreo) and had 8 cores, 4GB of RAM, 64GB of internal storage and a camera with 12 megapixels and 4290x2800 resolution. The second ran Android 8 and had 8 cores, 3GB of RAM, 32GB of internal storage and a camera with 12 megapixels and 4000x3000 resolution. The manufacturers of the two devices were different.

All tests and executions succeeded in both devices. To assess portability, we upgraded the operating system of the second device to Android 9 (Pie) and noted that all functionalities of the app worked as well. The app consumed only around 18MB of internal storage.

Figure 2 – Screenshots of the LabelIt! app.



(a) (b) (c)

(d) (e)

The database

A NoSQL document database was created and stored at server-side using Cloud Firestore to organize data as collections and documents, facilitate the creation of part-whole relationships (hierachies) and allow writing queries. Each label is a document described by an unique ID, a descriptive name, and a rectangle whose coordinates are pixels of an image. Each image is a document that belongs to a database, has a filename, has a size in terms of height and width in pixels, and is assigned several labels by users. Thus, a user can draw several rectangles on the same image to label it, as well as different users can assign one or more labels to the same image. Figure 3 describes the structure of the documents Label and Image and the documents referring to the three labels assigned to the image shown in Figure 2e.

One might ask how documents are created, read and deleted. An image document is created whenever a user labels an image and submits the work. One label document is created for each label assigned by a user. Whenever a user opens an image that was previously

labeled, the corresponding label documents are read. If the user clears the labels of such an image and submit, then the corresponding label documents are deleted.

The database administration does not require advanced technical skills from the user that will play the administrator role. The essential is to have a basic background on the creation of NoSQL document-oriented databases inside Firebase and on the configuration of access for end users.

Figure 3 – Label and Image document structures and four sample documents.

<table>
<tr>
<td colspan="2">Label
{
"id": "String",
"name": "String",
"xmax": "Number",
"xmin": "Number",
"ymax": "Number",
"ymin": "Number"
}</td>
<td colspan="2">Image
{
"database": "String",
"filename": "String",
"height": "Number",
"width": "Number",
"labels": ["Label"]
}</td>
</tr>
<tr>
<td>{
"id": "L1roxo",
"name": "Boi",
"xmax": 880,
"xmin": 600,
"ymax": 300,
"ymin": 10
}</td>
<td>{
"id": "L2verde",
"name": "Face do boi",
"xmax": 500,
"xmin": 400,
"ymax": 290,
"ymin": 170
}</td>
<td>{
"id": "L3azul",
"name": "String",
"xmax": 200,
"xmin": 40,
"ymax": 290,
"ymin": 170
}</td>
<td>{
"database": "pastejo",
"filename": "bovinos.jpg",
"height": "300",
"width": "900",
"labels": ["L1roxo, L2verde, L3azul"]
}</td>
</tr>
</table>

Discussion

We selected 60 pictures from the real dataset and produced around 150 labels considering the predefined values “ox”, “ox’ head”, “cow”, “cow’s head”, or the customized value “calf”. Regarding the downloaded dataset, the variety of pictures illustrated different scenarios and livestock applications. Furthermore, we emphasize that agricultural and environmental pictures can also be labeled by using our app.

Moreover, our system differs from existing solutions (cf. Introduction) in several aspects. Some of them are highlighted in Table 1. Our LabelIt! App was the only tool designed for mobile devices among the surveyed tools. Its remote storage (in a database) prevent the end user to worry about local files. Its authentication ensures that only authorized users can execute the software. It allows multiple end users to label images and save their work. Predefined, customized and sublabels are enabled, as shown in Figure 2d, where a part-whole is exemplified, i.e. face and animal. The end user cannot export its labeling work, but only the administrator role can do it in the database. On the client side, a straightforward interface omits the complexity of the labeling task. On the server side, permission to label images are set up by an administrator and the labeling work is stored for the authorized users.

Like the Limelight system by Tscharke and Banhazi (2016), a LabelIt! user can manually label or score a given behavior (offline or online). Unlike Lomba et al. (2015), Lukas Tanumihardja and Gunawan (2015), Garcia et al. (2017) and Luiz et al. (2018), we focus on images rather than strictly on sensors to collect input data. Like Santoni et al. (2015), users capable of distinguishing races can annotate them correctly with LabelIt!.

Table 1 – A comparison of existing labeling tools.

Feature	Labeling tool					
	Image Labeler M.	LabelImg	Labelbox	Image Labeler G.	VGG	**LabelIt!**
Platform	Desktop	Desktop	Web	Web	Web	**Mobile**
Storage	Local	Local	Remote	Remote	Local	**Remote**
Storage type	File	File	File	Database	File	**Database**
Authentication	No	No	Yes	Yes	No	**Yes**
Users	Single	Single	Multiple	Multiple	Single	**Multiple**
Predefined label	Yes	Yes	Yes	Yes	Yes	**Yes**
Customized label	Yes	Yes	Yes	No	Yes	**Yes**
Sublabel	Yes	No	Yes	No	No	**Yes**

CONCLUSION

Digital inclusion and the adoption of solutions based on IoT, computer vision and machine learning are increasing in Brazilian agriculture and livestock. In this work, we developed a system that enables labeling images obtained from mobile (IoT) devices. We focused on labeling to ensure that machine learning algorithms can benefit from labeled input.

We developed the LabelIt! app for farmers to label images using a smartphone or a tablet. Labels are assigned to rectangles drawn by the user on the image according to the user's perceptual abilities. The straightforward interface of the app does not require advanced technical skills. Not only a single, but multiple users can label images in their own devices. Users are benefited from mobility, as they can label images anywhere, but need an internet connection to submit the work. A remote database stored on a cloud computing platform maintains the labeling work performed by the users. Data are kept safe, as access is granted only to authorized users by an administrator.

With our system, we intend to: (i) provide a simple solution for a real problem in livestock; (ii) facilitate the labeling task for end users; and (iii) provide proper output to be used as input for machine learning algorithms. As future work, we intend to run usability tests and use labeled data generated by LabelIt! as input for machine learning algorithms.

ACKNOWLEDGMENTS

The authors thank to: PRE-IFSP for funding the 1st author's student scholarship - Call 618/18, project 307969.1720.249199.08062018; agtech professionals from @Tech for giving talks to students; and Embrapa Southeast Livestock for allowing photographs of cattle.

REFERENCES

ABMRA. Pesquisa ABMRA. 2017. Disponível em: <http://www.abmra.org.br/2016/index.php/pesquisa-abmra/>. Acesso em: out 2018.

ALVES, T. et al. From knowledge based vision systems to cognitive vision systems: a review. In: INTERNATIONAL CONFERENCE ON KNOWLEDGE BASED AND INTELLIGENT INFORMATION AND ENGINEERING SYSTEMS, 2018, Belgrade. Procedia Computer Science: Elsevier, 2018, n. 126, p. 1855-1864.

DUTTA, A.; ZISSERMAN, A. The VIA Annotation Software for Images, Audio and Video. In: 27th ACM International Conference on Multimedia, 2019, Nice. Proceedings… 4p.

GARCIA, A. et al. Comportamento de fêmeas bovinas de corte em pastagens sem arborização, avaliado por sistema wireless de monitoramento. In: III WORKSHOP INTERNACIONAL DE AMBIÊNCIA DE PRECISÃO, 2017, Campinas. ISSN 1981-2086. Circular Técnica: Embrapa, 2018, n. 82, 9 p.

GAÚCHAZH. Smartphone é ferramenta de gestão para sete em cada dez produtores rurais gaúchos. 2018. Disponível em: <https://gauchazh.clicrbs.com.br/economia/campo-e-lavoura/noticia/2018/10/smartphone-e-ferramenta-de-gestao-para-sete-em-cada-dez-produtores-rurais-gauchos-cjng5nfhi06gu01rx73frt8ct.html>. Acesso em: nov 2018.

KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. Imagenet classification with deep convolutional neural networks. In: INTERNATIONAL CONFERENCE ON NEURAL INFORMATION PROCESSING SYSTEMS, 25., 2012, Lake Tahoe. Proceedings… p. 1097-1105.

LIAKOS, K. et al. Machine learning in agriculture: a review. Sensors. v. 18, n. 2674, 29 p., 2018.

LOMBA, L. et al. O uso de inteligência artificial na identificação do comportamento bovino. In: CONGRESSO BRASILEIRO DE AGROINFORMÁTICA, 10, 2015, Ponta Grossa. Anais… 12 p.

LUIZ, O. et al. Computational System to Support Bovine Nutritional Behavior. International Journal of Agricultural and Environmental Information Systems. v.9, n.3, p.22-37.

LUKAS; TANUMIHARDJA, W.; GUNAWAN, E. On the application of IoT: monitoring of troughs water level using WSN. In: IEEE CONFERENCE ON WIRELESS SENSORS, 2015, Melaka. Proceedings... p. 58-62.

SANTONI, M. et al. Cattle Race Classification Using Gray Level Co-occurrence Matrix Convolutional Neural Networks. In: INTERNATIONAL CONFERENCE ON COMPUTER SCIENCE AND COMPUTATIONAL INTELLIGENCE, 2015, Jakarta. Procedia Computer Science: Elsevier, 2015, n. 59, p. 493-502.

SANTOS, P. et al. Encontro de inovação em pastagens: relatório final. ISSN 1980-6841. Embrapa Pecuária Sudeste, 2019. São Carlos, SP. 30 p.

TSCHARKE, M.; BANHAZI, T. A brief review of the application of machine vision in livestock behaviour analysis. Journal of Agricultural Informatics. v. 7, n. 1, p. 23-42, 2016.

VON AHN, L.; DABBISH, L. Labeling images with a computer game. In: CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS, 2004, Vienna. Proceedings… p. 319-326 .



Automação Inteligente para Coletores e Canais de Irrigação com Prototipação Arduíno e Sistema Web

André M. Silva[1], *Ezequiel M. Muxito*[1], *Miguel M. Awad*[1], *Cezar L. Duarte*[1], *Matheus V. Silva*[1]

[1]Centro Universitário Adventista de São Paulo (UNASP) - Campus São Paulo
Departamento de Ciência e Engenharia de Computação, São Paulo, SP – Brasil
andre.marcos@unasp.edu.br, ezequielmuxito@gmail.com, contatoawad@outlook.com
cezar_cld_unasp@hotmail.com, matheusvitor26@hotmail.com

RESUMO

Este artigo apresenta um trabalho de pesquisa em andamento em Internet das Coisas (IoT) e suas aplicações na área da agricultura. Após o levantamento teórico e validação de trabalhos correlatos, a atividade prática foca na construção de uma proposta de automação integrada a componentes de controle através da Internet promovendo melhoria para um processo de produção agrícola. Como estudo de caso, a pesquisa propõe a explorar as necessidades e características de um ambiente real de produção hortifrutigranjeiros na região da grande São Paulo. Esta parceria permite um cenário de colaboração, principalmente para as validações do projeto construído.
PALAVRAS-CHAVE: Internet das Coisas; Agricultura Inteligente; Arduíno.

ABSTRACT

This paper presents an ongoing research work in IoT (Internet of Things) and its applications in the agriculture. After surveys and validation of related works, the practical activity focuses on the construction of a proposal of automation integrated to control components through the Internet promoting improvement for some agricultural production process. As a case study, the research aims to explore the needs of a real and small farm. This environment allows a scenario mainly for the validations of the project.
KEYWORDS: Internet of Things; Smart Agriculture; Arduino Prototyping.

INTRODUÇÃO

A Internet das Coisas, ou simplesmente chamada de IoT vem mudando a forma na qual a computação e, principalmente, a automação vem mudando a rotinas das atividades diárias (EVANS, 2011). Porém, ainda que timidamente, estas tecnologias já vêm impactando fortemente em áreas como a educação, negócios, medicina, comunicação, governos, e na agricultura

(INAMASU Y. E BERNARDI, 2014). Entre os principais motivos de preocupação ou empecilho para adoção de tecnologias IoT como facilitadores nos diversos contextos, estão o custo para aquisição da infraestrutura e a complexidade de implantação nos mais diversos meios.

Uma das principais grandes áreas consumidoras de projetos baseados em IoT, é a agricultura (FAKHDDIN, 2018). Ao mesmo tempo que a agricultura é um setor produtivo bastante evoluído, também apresenta um deficit (GOMES, 2018) quanto a aplicação de tecnologia que promova a melhoria, precisão e a automação de processos produtivos (SILVA A. M., 2015). Um dos principais fatores para o não aproveitamento em larga escala da tecnologia no campo destaca-se o alto custo e a possibilidade de vantagens ainda prematuras. Por exemplo, em vários setores, a possibilidade e facilidade da exploração de artefatos computacionais integrados com sensores de coleta e mecanismos de tomada de decisão (KAPLAN, 2014), só têm sido facilmente acessíveis a partir dos últimos anos.

OBJETIVO

Considerando o cenário apresentado, o principal objetivo deste trabalho é o estudo do impacto da IoT no setor da agricultura, em especial das suas aplicações para produções de pequena e média escala, com o intuito principal de experimentar uma proposta de melhoria e controle de um sistema de irrigação automatizada para agricultura de precisão que seja de baixo custo e de fácil implantação. Em outras palavras, propõe-se a construção de um protótipo para simular e testar o comportamento dos pivôs junto à uma nova solução tecnológica de monitoramento e controle. Este protótipo deve ser basicamente composto de três ambientes bem definidos: o ambiente mecânico automatizado a partir da plataforma Arduíno; o sistema central baseado em *webservice* para integração; e por último, o ambiente de controle funcional, baseado em um sistema web.

METODOLOGIA

Como metodologia de trabalho este projeto envolve inicialmente pesquisas em bibliografias especializadas na área de IoT. Também serão considerados trabalhos correlatos na área de automação por dispositivos Arduíno.

Como atividade prática propõe o desenvolvimento de um protótipo para automatização de um pivô agrícola que, basicamente, conta com o desenvolvimento de dois ambientes: o ambiente mecânico automatizado a partir da plataforma de prototipação Arduíno; e o ambiente de controle funcional, que é baseado em um sistema por um dispositivo web através de uma comunicação por servidores de aplicação e a rede Internet.

Esta proposta é voltada a atender problemas, desafios e necessidades reais de produtores agrícolas. Tal característica será possível através de parcerias externas de pesquisas junto a uma empresa de tecnologia atuante no mercado angolano e uma cooperativa paulistana de médio porte atuante nas atividades de produção e distribuição de hortifrutigranjeiros. Estas

parcerias permitirão, além de outros benefícios, a possibilidade de um ambiente real de estudo de caso para levantamento de requisitos, orientações especializadas e laboratório para testes de validação do protótipo construído.

ATIVIDADE PRÁTICA

Estudo de Caso

O estudo de caso permitiu a consideração de um problema real existente. Desta forma, como especificação prática, este trabalho envolve atividades reais para identificação de problemas, levantamento de requisitos, prototipação e sessões de *brainstorming* realizados junto a especialistas e potenciais usuários do sistema em campo.

Este laboratório, que também servirá como um ambiente de prova e observação para os resultados obtidos, é uma cooperativa que, além de agenciar escoamentos, revendas e representações de hortifrutigranjeiros, também possui sua própria produção de hortaliças, operando, desde 1980, em uma área de aproximadamente cinco mil metros quadrados. Nesta área, localizada na região de Itapecerica da Serra, ao sul da cidade de São Paulo, aproximadamente 56 funcionários são responsáveis pela produção própria dos mais variados tipos de hortaliças: tais como: acelga, agrião, alface, couve, rúcula, repolho e salsa.

Com esta característica, a empresa consegue atender um grande leque de clientes fixos, tais como hotéis, padarias, escolas, feiras e mercados em toda cidade, com atendimento tanto no atacado, como no varejo, através de pontos de venda no entorno urbano. Ainda que apresente um grande porte de acordo com a área de atuação e ramo de negócio, esta cooperativa não explora tecnologia computacional. A maior parte da produção em campo é apoiada por esforço manual equipada com tratores e motores elétricos para irrigação. Por este motivo, possui algumas deficiências ou problemas, na qual a implantação de recursos computacionais poderia evitar ou melhorar. Dentre os principais problemas estão os relacionados ao processo de irrigação da produção, tais como:

- Dificuldade de ativação e controle do processo;
- Problema de manutenção de históricos e estatísticas do processo;
- Diferentes frequências e intervalos para cada tipo de produto;
- Distintos tipos de volume de água para específicos tipos de hortaliças e
- Tratamentos especiais, ou mudança de regras conforme temperatura ou pluviosidade;

Principalmente, por este motivo, esta cooperativa tem sido um campo de oportunidade para melhoria do processo através da implantação e uso da tecnologia da informação. Isto com o objetivo de melhorar processos, garantir qualidade e melhorar a agilidade diante da concorrência.

Desenvolvimento da Proposta

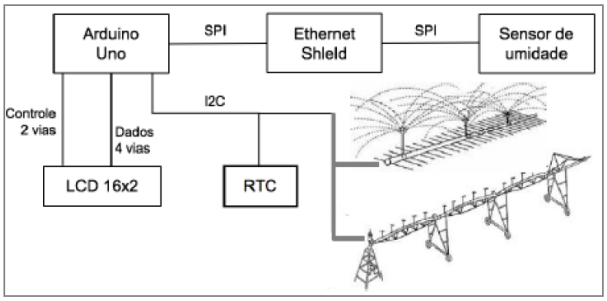
Para o desenvolvimento do trabalho é proposto o desenvolvimento de um protótipo para parte embarcada, através de Arduíno e uma aplicação web para controle do processo de irrigação real

á definido. As principais funcionalidades do protótipo são:

1. Leitura da umidade do solo,
2. automatismo da decisão do disparo ou interrupção mecânica do processo de irrigação,
3. histórico de leituras realizadas,
4. apresentação dos registros,
5. painel para ativação manual do processo de irrigação e
6. painel para configuração e agendamento de ciclos de irrigação.

Basicamente, está proposta é formada por um conjunto de elementos eletrônico integrados a placa Arduíno Uno R3 (ARDUINO, 2015) para formar o circuito que permita as funcionalidades desejadas (Figura 1).

Figura 1: Esquemático da estrutura para proposta



Fonte (Elaborada pelo autor)

Estas funcionalidades, de acordo com cada especificação de infraestrutura, web, automação, sensores ou *webservice*, estão apoiadas através das tecnologias apresentadas na Tabela 1.

Tabela 1: Apoio Tecnológico por Módulo

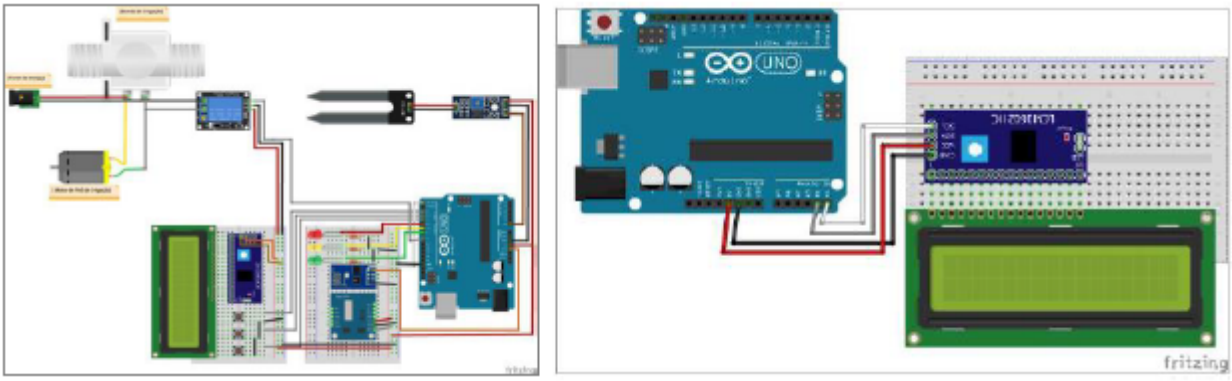
Módulo	**Apoio tecnológico**
Automação	Hardware Arduíno e IDE para codificação Shields e breakout para Arduino Sketch das bibliotecas para os shields e breakout
Web	Microsoft Visual Studio Code para edição de código PHP como linguagem de programação Web MySQL como gerenciador da base de dados

A primeira etapa do desenvolvimento do protótipo foi a construção de um simulador para o circuito e suas implementações físicas. O circuito foi projetado e desenvolvido com auxílio do software Fritzing (SCHMIDT, 2011), o qual possibilitou realizar simulações com os componentes disponível do Arduíno. Além de permitir uma antecipação de como o sistema funcionaria, as simulações possibilitaram testes para compreender e analisar o comportamento eletrônico do protótipo, ainda em laboratório ((FRITZING, 2018)).

A montagem da prototipagem foi realizada pelo "Protoboard" (Figura 2) e outros recursos para manipulação de projetos Arduíno, tais como: placa Arduíno Uno, Protoboard, três *leds* de cores (vermelha, amarela e verde), três resistores de 10 ohms, um relé, uma bomba de água,

um motor representando pivô de irrigação, uma fonte de alimentação para a bomba de água e o motor, um sensor de umidade do solo hidrométrico, um módulo I2C para *display* LCD, um *display* LCD, três botões para funções de operações, um módulo RTC (Real Time Clock), e um módulo ESP8266 ESP-01 de wifi. A figura 2 (a) apresenta a disposição e integração destes componentes conforme projetados.

Figura 2: Protótipo dos componentes integrados para módulo de automação (a) e LC2 (b)



Fonte (Elaborada pelo autor)

RESULTADOS

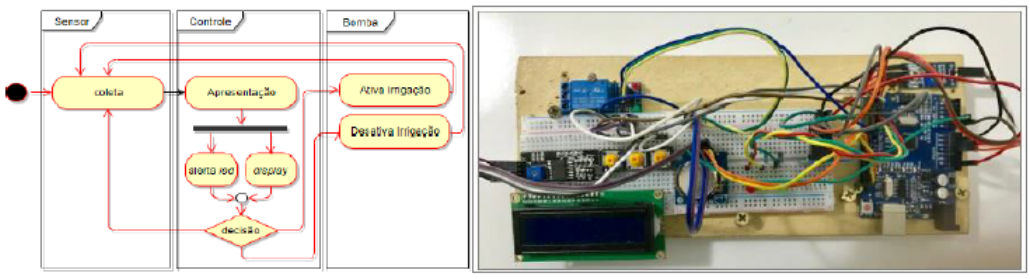
Atualmente o projeto se encontra na fase de validação do protótipo que representa o modelo proposto. Este ambiente é composto por uma modelagem customizada, componentes de automação integrado e operacional, e de sensores que simulam canais de alimentação de informações para decisões de execução. Esta fase de avaliação conta com a participação de *stakeholders* multifuncionais reais, entre eles: usuários produtores, agrônomos e analistas de sistemas. Entre outros objetivos, esta fase objetiva validar as configurações dos sensores como base de tomada de decisão e, a partir destes resultados, indicar melhorias novas configurações de acordo com um comportamento mais real. Em paralelo à fase de validação, temos, também em andamento, a finalização do módulo web de controle. Estes módulos permitirão tanto a configuração para agendamentos para execução do processo quanto permitindo informações estatísticas de leitura e execução do sistema (SILVA A. M., 2018).

A Figura 3 (a) representa graficamente o processo como um todo ao longo de um fluxo de execução do protótipo obtido e em fase de teste. Basicamente o processo é executado em *loop*. Na primeira etapa do processo o dispositivo de coleta sensorial é responsável por identificar a taxa de umidade no solo e enviar pra central de automação. Esta, por sua vez, dispara sinais para os dispositivos de apresentação: tanto para os *leds* de alerta, quanto para o *display* de apresentação de resultado. Para apresentação os dados coletados são trabalhados conforme o algoritmo de tratamento, que enquadra os valores obtidos dentro das taxas de porcentagem 0 a 100%referente a sensação do senso comum de umidade.

Após apresentação dos dados, a central de automação, segundo configurações de tratamento, toma a decisão de hibernar até a próxima leitura ou disparar comandos para a bomba de irrigação. Dependendo dos valores recebidos e do status atual da bomba, este comando pode ser de ativação ou desativação. A Figura 3 (b) apresenta o módulo de automação como resultado obtido e seus componentes de circuitos integrados responsáveis pelo controle do que vem a ser a central de ativação automática do processo de irrigação.

Este módulo é o principal componente do projeto. Ele centraliza o recebimento das leituras de umidade, toma a decisão de ativar ou interromper a bomba de irrigação de acordo com os valores de configuração pré-estabelecidos. Este módulo também possui um dispositivo de apresentação por *display* para facilitar a interação humana com o componente.

Figura 3: Representação esquemática (a) da proposta para a estrutura do pivô de irrigação e Módulo (b) central do processo de controle



Fonte (Elaborada pelo autor)

A Figura 4 (a) demonstra o projeto do protótipo executável de uma versão inicial de teste para ativação das funcionalidades do módulo de automatização por Arduíno. Através da interface web do SAPI o usuário pode consultar leituras coletadas, configurar um algoritmo de atuação do pivô de irrigação, especificar ciclos para tornar esse processo repetitivo, ou até ativar ou desativar o processo de uma forma interativa.

A partir deste módulo web o operador pode interferir no processo automatizado de execução. O módulo web tem o papel de apresentar dados coletados, apresentar de uma forma mais amigável e de permitir o agendamento de acionamentos do processo; entre eles:

- **leitura dos índices de umidade coletados**: além dos dispositivos de apresentação de dados coletados apresentados no nível de circuito (*display* e *leds*), o sistema web permite que o usuário visualize as informações de coleta diretamente na interface web, provendo ao técnico operador um ambiente mais amigável para consultar os dados.
- **agendamentos de ativação**: a partir do sistema web, o operador do sistema tem a permissão de configurar os pontos de ativação do sistema, por exemplo, indicando horas e ciclos de repetição para o processo (Figura 4 (b)).
- **ativação e interrupção do processo**: uma vez necessário o operador pode, através do módulo web, iniciar imediatamente a execução da irrigação. Também é possível interromper um processo já iniciado ou interromper um agendamento qualquer.

Figura 4: Protótipo do (a) módulo web de configuração e ativação do processo (b)



Fonte (Elaborada pelo autor)

O usuário também poderá, em uma próxima versão, realizar análises de informações coletadas de acordo com o comportamento do pivô de acordo com cada leitura e execução do dispositivo apresentadas tabela de históricos e em gráficos.

Sensibilidade do sensor e tipos de solo

Com o propósito de validar o comportamento do protótipo em diferentes tipos de ambientes em campo, foram considerados três tipos de diferentes de solos: terra vermelha, terreno arenoso terra escura. Cada situação foi submetida a um mesmo volume de água para que fosse possível entender os diferentes resultados do sistema quanto ao nível de absorção de água e a taxa de conservação de umidade pra cada tipo de terra.

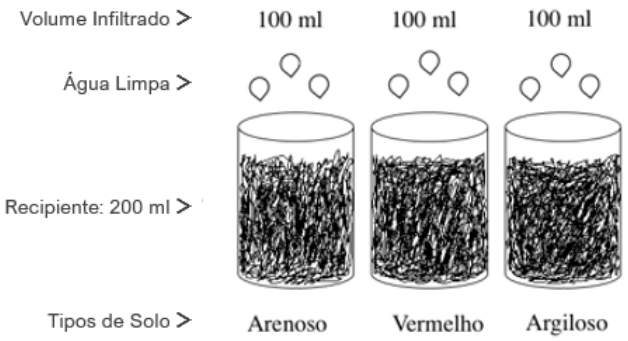
Ao receber 100 ml de água, o solo mais arenoso, apresentou uma leitura de umidade aproximada de 64%. No total, após 4 minutos e 32 segundos, o solo absorveu 100 ml de água a ponto de permitir a captação da leitura pelo sensor de umidade hidrométrico. Após 32 minutos de tempo cronometrado, o sensor detectou uma varação na umidade de 64% para 48% do nível de umidade, uma queda de 16% de umidade nesta janela de observação.

Na amostragem de terra vermelha o sensor de umidade detectou um nível de umidade de 74% que se manteve estável dentro da faixa de captação durante o derramamento de água. No total foram precisos 7 minutos e 44 segundos para que o solo identificasse os 100 ml de águas usados nesta configuração. Após 32 minutos o sensor apresentou uma variação no nível de umidade de 74% para 61% de umidade, uma queda de 13% do nível de umidade.

Já o sensor no solo de terra escura (terra preta) detectou um nível de umidade de 78% após um período de 10 minutos e 27 segundos, que foram necessários para absorção detectada da água. Apôs 32 minutos de tempo cronometrado houve uma variação na umidade de 78% para 67%, uma queda no nível de umidade de 11%. Os testes de variação dos tipos de solos foram feitos em recipientes pequenos (copos plásticos de 200 ml), tal apresentado na Figura 5.

Estas simulações permitiram concluir que o sistema apresenta comportamentos distintos para os diferentes modelos de terra. O solo arenoso, se apresentou ter, maior nível de absorção de água, e nível de conservação de umidade menor, ou seja, é um tipo de solo que rapidamente absorve a água, porém muito rapidamente perde umidade. Esta característica exige uma maior

Figura 5: Tipos de Terreno



Fonte (Elaborada pelo autor)

necessidade, em pequenos intervalos de tempo que o processo de irrigação seja ativado. O solo de terra vermelha mostrou ter menor nível de absorção de água, ou seja, absorve a água de forma muito lentamente. Em contrapartida apresenta um maior poder de conservação de umidade, afrouxando a dependência de água, e de um processo menos rigorosamente definido. Já o solo de terra preta apresentou ser mas eficiente nos testes. Esse tipo de solo tem um nível de absorção de água mais lento que outros tipos, porém apresenta um alto poder de manutenção da umidade. A Tabela apresenta os dados destes testes e as análises construídas.

Tabela 2: Análise dos tipos de solo

Itens	**Arenoso**	**Vermelho**	**Preto**
Recipiente	200 ml	200 ml	200 ml
Umidade inicial	0	0	0
Quantidade de água	100 ml	100 ml	100 ml
Umidade do sensor	64%	74%	78%
Tempo absorção de água	4 m 32 min.	7 m 44 min.	10 m 27 min.
Janela de observação	32 min	32 min	32 min
Queda de umidade (%)	16	13	11
Variação de umidade	48 a 64%	61 a 74%	67 a 78%
Nível de conservação	baixo	médio	alto

Validação dos Resultados Parciais

Além da exploração do estudo de caso como uma base para consideração de problemas e requisitos, esta convivência com o mundo real permitiu uma validação mais assertiva dos resultados obtidos. Desta forma, o protótipo construído foi submetido a uma avaliação de especialistas, que além de apontarem grandes benefícios, também levantaram possíveis problemas com a implantação do sistema.

Entre as vantagens possíveis com a implantação da proposta, o sistema realmente atende os principais objetivos quanto a redução da carga de trabalho manual, eliminação de atividade-

des repetitivas e dos processos de medição. Muitas das responsabilidades humanas podem ser integralmente substituídas ou melhoradas com a implantação do sistema. Entre os problemas apontados, a maioria deles já eram considerados e seguem sendo trabalhados em desenvolvimento para uma próxima versão do protótipo, tais como, a possibilidade de configuração por tipo específico de produção e a interação por interface móvel, do tipo *smartphone*.

Quanto melhorias, esta avaliação apontou que uma possível integração com sistemas públicos de previsões do tempo poderia trazer muitos ganhos ao sistema. Por exemplo, na eminência de chuva em um curto espaço de tempo, o sistema inteligente poderia adiar ou até cancelar um ciclo de irrigação. Esta funcionalidade poderia, de imediato contribuir com redução de esforço, energia, água, permitindo um projeto mais sustentável.

Contudo, um dos problemas identificados ainda não era considerado pelo projeto, e deve, necessariamente, ser incluído como funcionalidade tratada em versões futuras do projeto. Este problema é relacionado a profundidade das leituras realizadas, ou seja, dependente do tipo de cultura, a amplitude de leitura da umidade deve ser diferente. Para determinados produtos a taxa de umidade deve ser relativa a uma profundidade mais rasa, para outras mais profunda. E esta diferenciação não estava sendo considerada pelo algoritmo de tratamento.

A Tabela 3 resume os principais apontamentos identificados ao longo da etapa de validação da proposta por especialistas, produtores e potenciais operadores do sistema. Nesta tabela, **Apontamentos** referem-se aos principais problemas identificados durante as sessões de validação dos resultados parciais obtidos; **Avaliação** é o tratamento ou classificação atribuído, podendo ser uma melhoria ou um problema para o projeto; E por fim, **Negociação** refere-se à negociação definida entre os envolvidos no projeto, podendo ser: uma característica que já está em produção na versão corrente (desenvolvimento), uma funcionalidade já definida mas que ainda não está em desenvolvimento (projetada), uma melhoria que se julgou importante mas não obrigatória para o projeto em andamento (análise) ou uma correção obrigatoriamente necessária para uma futura implantação do projeto (correção programada).

Tabela 3: Validação de implantação de protótipo

Apontamentos	**Avaliação**	**Negociação**
Configuração e agendamento por tipo de produção	Melhoria. Flexibilidade para agendamentos	Desenvolvimento
Ativação e configuração por interface móvel	Melhoria. Facilita uso em campo	Projetado
Umidade e profundidade por produto ou solo	Problema para uma versão produção	Programada
Integração com sistema de previsão de tempo	Melhoria. Otimiza processo e reduz gastos	Análise

CONCLUSÃO

A realização deste trabalho vem permitindo uma real e rica experiência de integração entre as áreas de computação e agronomia. Apesar das várias dificuldades enfrentadas ao considerar um modelo real como estudo de caso, os resultados parciais encontrados têm sido bastante satisfatórios. O modelo encontrado e representação das informações através de um ambiente

mais amigável constituem em uma apresentação mais estimulante para usuários que até então desconsiderava o uso deste tipo de ferramenta ao manipular mecanismos motores na lida com o processo de produção.

A conclusão do trabalho é uma contribuição direta para área de computação, mas também se preocupa com resultados voltados as problemas, desafios e necessidades reais de produtores agrícolas. Tal envolvimento e colaboração é possível a partir de parcerias de pesquisas junto empresas de tecnologia atuante no mercado brasileiro e angolano, que permitem, além de outros benefícios, a possibilidade de um ambiente de estudo de caso para levantamento de requisitos, orientações especializadas e, principalmente, laboratórios para testes do protótipo construído

REFERÊNCIAS

ARDUINO. *Site oficial do Arduino com Tutorial com exemplos das funções do Core e Libraries do Arduino*. Glandrive Artigo Online, 2015. Disponível em: <http://www.arduino.cc/en/Tutorial/HomePage/>.

EVANS, D. *A Internet das coisas como a próxima evolução da internet está mudando tudo: auxílio para a sua estruturação*. [S.l.]: Cisco Internet Business Solutions Group (IBSG), 2011. (White Paper).

FAKHDDIN, H. *Smart Agriculture: 13 Trends To Watch Out For. Teknowledge Software*. 2018. Disponível em: <http://teks.co.in/site/blog/smart-agriculture-13-trends-to-watch-out-for>.

FRITZING. *Eletronics Made Easy*. 2018. Disponível em: <http://fritzing.org/home/>.

GOMES, J. *Agricultura Inteligente, ou Agricultura de Precisão*. Glandrive Artigo Online, 2018. Disponível em: <https://www.glandrive.pt/agricultura-inteligente-ou-agricultura-de-precisao-a-tecnologia-ao-servico-da-agricultura-agricultura-high-tech/>.

INAMASU Y. E BERNARDI, A. C. *Agricultura de precisão: resultados de um novo olhar*. [S.l.]: Embrapa Pub., 2014. (Capítulo de livro Agricultura de Precisão).

KAPLAN, J. *Artificial Intelligence: Think Again*. [S.l.]: Communications of the ACM, 2014. (Vol.60). ISBN 10.1145/2950039.

SCHMIDT, M. *A Quick-Start Guide; The Pragmatic Programmers*. [S.l.]: Pragmatic Bookshelf, 2011. ISBN 9781934356661.

SILVA A. M., e. A. A. A. E. *Ambiente facilitador de decisões agrícolas para produtores Angolanos baseado em informações geográficas e técnicas de Business Intelligence*. [S.l.]: XSBIAgro, Congresso Brasileiro de Agro informática e III SBIAgro IC, 2015.

SILVA A. M., e. M. E. M. *Agricultura Inteligente - Proposta de automação de pivôs e canais de irrigação com prototipação por Arduino e Webservice*. [S.l.]: Cimatech V Congress of Industrial Management and Aeronautical Technology, 2018. ISSN 2447-5378.

HORUS GREEN: SISTEMA PARA MONITORAMENTO DO CRESCIMENTO DA ALFACE (*Lactuca sativa L.*) EM CULTIVO HIDROPÔNICO UTILIZANDO PROCESSAMENTO DIGITAL DE IMAGENS

Leonardo Rodrigues de Lima Teixeira[1], *Suéliton dos Santos Silva*[1], *Anderson Patrício Fernandes dos Santos*[1]

[1] Escola Agrícola de Jundiaí, Universidade Federal do Rio Grande do Norte, Macaíba, Rio Grande do Norte, Brasil, leonardo@eaj.ufrn.br, tonmelodicmetal@gmail.com, andersonp17@bol.com.br

RESUMO

A alface (*Lactuca sativa L.*) é uma das hortaliças mais consumidas pela população mundial, tendo em vista sua composição nutricional e benefícios à saúde. Dentre os métodos de cultivo desta hortaliça, temos o cultivo hidropônico. Estudos mostram que modificações na solução nutritiva tem efeito no crescimento da área foliar da alface. Este trabalho propõe um sistema de medição de área foliar não destrutivo, utilizando um aplicativo móvel, processamento digital de imagens (PDI) e um serviço de armazenamento e processamento em nuvem para acompanhamento do crescimento durante o cultivo de forma simplificada.

PALAVRAS-CHAVE: Aplicativo móvel, Processamento digital de imagens, Taxa de crescimento de alface.

ABSTRACT

Lettuce (*Lactuca sativa L.*) is one of the vegetables most consumed by the world population, considering its nutritional composition and health benefits. Among the methods of cultivation of this vegetable, one has the hydroponic culture. Studies show that the modification of some components of the nutrient solution has an effect on lettuce leaf area growth. This work proposes a non destructive foliar area measurement system using a mobile application, digital image processing and a cloud storage and processing service to monitor growth during cultivation in a simplified way.

KEYWORDS: Mobile application, Digital image processing, Lettuce growth rate.

INTRODUÇÃO

No Brasil existe a possibilidade do cultivo de hortaliças folhosas durante todo o ano devido às condições climáticas, assim como pela adoção de novas técnicas de cultivo que utilizam com mais eficácia a água, fertilizantes e área de cultivo.

O cultivo da grande maioria das hortaliças folhosas utilizam o solo como principal fonte de nutrientes. Em menor quantidade são encontrados cultivos que utilizam técnicas de hidroponia. Este tipo de cultivo emprega uma solução nutritiva composta por água e sais minerais, sendo a mesma responsável por prover todos os nutrientes para o crescimento da planta (BEZERRA NETO; BARRETO, 2012).

Neste cenário, a alface apresenta-se como uma das culturas mais importantes e rentáveis financeiramente em sistemas hidropônicos (FAVARATO;GUARÇONI;SIQUEIRA, 2017). Sua facilidade de manejo, tolerância à pragas e ciclo de produção rápido, aliados aos benefícios característicos dos sistemas hidropônicos como a oferta e controle de nutrientes que a planta necessita através da solução nutritiva, tornam o cultivo desta hortaliça bastante promissor (MACÊDO, 2004; BEZERRA NETO, BARRETO, 2012).

As deficiências na nutrição podem ser percebidas no formato e tamanho da alface, sendo este último o principal fator levado em consideração para a colheita da hortaliça e para o consumidor final. Com isso, evidencia-se a importância de métodos de medição do crescimento baseados em área foliar, os quais possibilitam a análise das plantas verificando se as mesmas estão se desenvolvendo como o esperado para a cultura.

Na literatura, encontram-se diversos métodos de medição de área foliar e consequentemente do crescimento da alface e de outras espécies vegetais. Estes métodos podem ser destrutivos ou não destrutivos, diretos ou indiretos, diferindo principalmente na forma de obtenção da área foliar como descrito por Santos *et al.* (2014) e Reis *et al.* (2000).

O presente trabalho foi inspirado em métodos de medição foliar não destrutivos. Esses utilizam-se de ferramentas matemáticas, programas de manipulação de imagens e modelos matemáticos na mensuração da massa seca (MS), massa fresca (MF) e área foliar. Ribeiro (2016), Pereira *et al.* (2011), Easlon e Bloom (2014) e Jung *et al.* (2015) são exemplos de trabalhos que desenvolveram métodos não destrutivos de medição.

Diante dos trabalhos desenvolvidos, observou-se a importância dos métodos de análise não destrutivos que aliados a modelos matemáticos podem fornecer algumas informações sobre a espécie analisada. Massa fresca, altura e área foliar são algumas das variáveis que podem ser medidas, permitindo ao pesquisador ou produtor tomar as medidas necessárias para manter o controle do plantio. Entretanto, ainda nota-se uma carência de aplicações que

facilitem o acompanhamento do crescimento durante todo o período do cultivo de forma simplificada.

Esse trabalho tem como objetivo a criação de um sistema, denominado Horus Green, que facilite as aquisições e armazenamento das informações relacionadas ao crescimento da cultura da alface em sistema hidropônico, mais precisamente a área foliar, o que facilitará a consulta posterior dessas informações pelo produtor da alface ou pesquisador da área.

MATERIAL E MÉTODOS

Primeiramente, foi realizado o levantamento de requisitos do sistema, cujos requisitos funcionais são listados a seguir: permitir o cadastro e manutenção de usuários de acesso ao aplicativo; cadastrar experimentos; capturar e guardar imagens para cada experimento separadamente; enviar imagens capturadas para o serviço de armazenamento em nuvem; exibir uma lista com datas e dados processados referentes a cada um dos experimentos; exibir gráficos de área foliar e percentual de crescimento referente aos experimentos que estão sendo realizados; e exibir todas as informações sobre o experimento que está em execução, incluindo informações como a quantidade de imagens processadas e a idade atual da planta.

Os requisitos não funcionais se referem às regras e/ou restrições das funções previamente descritas pelos requisitos funcionais. Seguem os requisitos não funcionais do sistema: o sistema operacional do *smartphone* utilizado deve ser Android e possuir uma versão igual ou superior à 5.1 Lollypop; o aplicativo deve ser sincronizado pelo menos uma vez antes da próxima captura de fotos, porque a aplicação só consegue armazenar uma única imagem por vez para cada experimento, sem que esteja sincronizado com a nuvem; a câmera do *smartphone* deve preferencialmente ter uma resolução igual ou superior a 12 Megapixels.

O sistema proposto neste trabalho é composto por três componentes principais: aplicativo móvel, firebase e PDI. Os componentes aplicativo móvel e PDI se comunicam um com o outro utilizando o módulo firebase. Na Figura 1, é apresentado o diagrama de componentes do sistema.

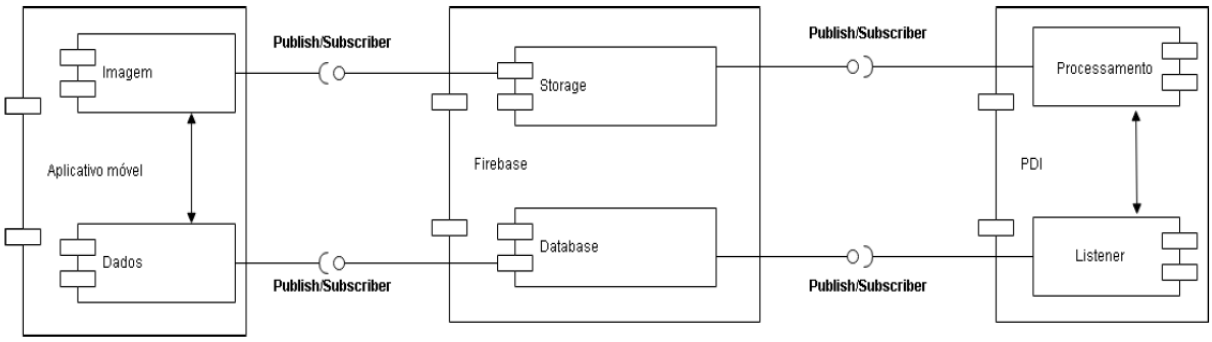
O aplicativo móvel tem como principal finalidade prover a interface entre o sistema e o usuário. Seu escopo incorpora dois componentes abstratos menores que se relacionam entre si, sendo eles: o módulo de dados e o módulo de imagem.

O módulo de dados é responsável por se relacionar com o banco de dados (*Real Time Database*), permitindo que toda a aplicação móvel seja sincronizada e tenha seus dados publicados para a posterior consulta.

O módulo de imagem por sua vez, se comunica com o serviço de armazenamento de arquivos em nuvem (*Storage*), criando pastas e referências para as imagens que são enviadas

para processamento, e posteriormente realizando a consulta das imagens já processadas, sendo estas últimas gravadas a partir de outro componente do sistema, o PDI.

Figura 1 - Diagrama de componentes



Fonte: Autoria própria

O componente firebase é implementado e mantido pela Google. Sua infraestrutura provê uma ótima integração com os outros dois módulos do sistema, garantindo a consistência dos dados e escalabilidade do sistema de acordo com a necessidade. Foram utilizados apenas dois serviços do Firebase: o *Real Time Database* e o *Storage*, os quais funcionam de forma muito parecida utilizando referências, onde as demais partes do sistema poderão assinar ou criar novas referências com novos dados a qualquer momento.

O componente PDI é o responsável por buscar e processar as imagens enviadas pela aplicação móvel para o *Storage*, sendo dividido em dois módulos: *listener* e processamento.

O *listener* é o responsável pela leitura dos dados contidos no *Real Time Database* e a verificação se existe alguma imagem nova para processamento. Após a verificação de que existe uma nova imagem para processamento, a foto é carregada e processada pelos métodos do módulo de processamento para medição de área foliar e percentual.

Por se tratar de uma aplicação que utiliza banco de dados não relacional (NoSQL), a sua modelagem teve como principal característica um arranjo hierárquico entre os objetos das classes de entidades que compõem o modelo do sistema: Usuário, Experimento, Solução Nutritiva, Captura e Crescimento.

Para o processamento digital de imagens deste trabalho, foram utilizadas apenas 3 etapas do processamento digital proposto por Gonzalez e Woods (2010): aquisição de imagens, filtragem e realce de imagens, e segmentação.

As capturas foram realizadas sempre no mesmo período da manhã entre 7:15 e 7:45 horas, tendo apresentado melhores resultados de iluminação a partir das 7:30, à medida em que o sol incide cada vez mais sobre o filme agrícola de polietileno que encobre a estufa.

Os materiais utilizados para a aquisição das imagens foram: um suporte com quatro apoios (quadrúpede), um *smartphone* Samsung modelo SM-J500M com resolução de 12.8 Megapixels, responsável por gerar uma imagem de dimensões 4128 x 3096, e um monopod, o qual se conectava com o *smartphone* através de uma conexão *bluetooth* para aquisição das imagens (sem contato do usuário).

O quadrúpede, apresentado na Figura 2, foi construído utilizando canos de PVC e outros materiais improvisados, com o objetivo de sustentar o *smartphone* que servirá para captura das fotos. O tamanho dos canos de PVC utilizados na construção são de 1,45m. Graças ao formato de pirâmide, a altura do quadrúpede montado sobre as canaletas, ficou de aproximadamente 1,5m.

Figura 2 - Suporte quadrúpede utilizado na captura das imagens



Fonte: Autoria própria

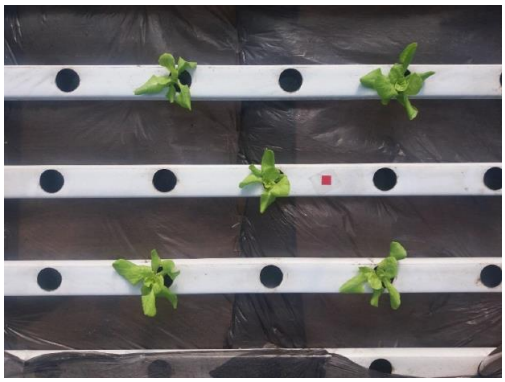
Para evitar que ervas daninhas aparecessem nas imagens capturadas, foram colados sacos de lixo previamente abertos abaixo das canaletas e colados uns aos outros por meio de fita adesiva. Na Figura 3, temos um exemplo de foto tirada através do sistema de aquisição (smartphone, monopod e quadrúpede).

Para a etapa de filtragem e realce, foram utilizadas duas técnicas: o *resize* e a aplicação do filtro *blur*. O *resize* teve a finalidade de diminuir as dimensões da imagem para diminuir o processamento requerido para processar uma foto. Após a realização de testes, chegou-se à conclusão que as dimensões com a melhor relação custo/benefício foram: 1032 x 874. Com isso o tempo necessário para o processamento ficou em torno de 20 segundos. Já com a utilização do filtro *blur* foi possível homogeneizar as curvas do histograma da imagem, retirando os ruídos de alta frequência e delineando melhor as partes de interesse da imagem.

Na etapa de segmentação utilizou-se a divisão do processamento em duas rotinas: segmentação do quadrado vermelho e segmentação da área foliar verde. Para a limiarização do quadrado vermelho, foi utilizada a seguinte fórmula: ExcVermelho = {R+(455-G)}/4; e para a

área foliar: ExcVerde = {G + (455-B)}/4. Essas fórmulas e valores foram obtidos a partir de tentativa e erro, pois as soluções encontradas na literatura não tiveram um bom resultado.

Figura 3 - Padrão de aquisição de imagens



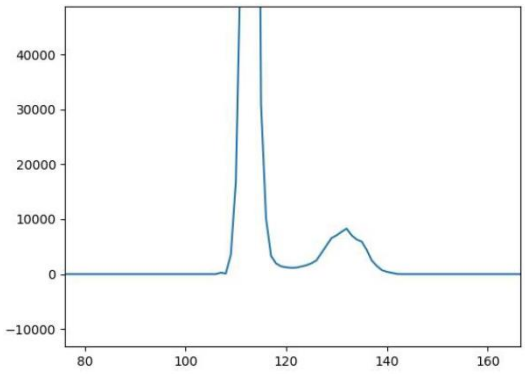
Fonte: Autoria própria

O algoritmo proposto neste trabalho possui estratégia e funcionamento similar ao encontrado em Easlon e Bloom (2014), com o aplicativo Easy Leaf, principalmente no que diz respeito à estratégia de inferência de área verde. Esta consistiu da utilização de um quadrado vermelho de 4 cm² como referência para encontrar a área verde da imagem.

Para identificação da área de interesse foi utilizado o histograma da imagem segmentada. A etapa de segmentação com operações aritméticas nas camadas da imagem, fornece uma imagem cujo o histograma apresenta dois picos: um grande e outro bem menor, sendo este o de interesse que pode ser visto na Figura 4.

Sabendo da posição do pico de interesse, o algoritmo calcula o limiar (*thresholding*) de corte. Limiar esse que isola a parte de interesse, que neste caso é do início do 2° pico até o fim do histograma. De posse do limiar de corte, é possível realizar a contagem dos pixels identificados como objeto pelo algoritmo.

Figura 4 - Picos do histograma



Fonte: Autoria própria

Conhecendo a quantidade de pixels e a área do quadrado vermelho (aproximadamente 4 cm²), e conhecendo a quantidade de pixels verdes identificados, é possível calcular a área

verde da imagem através de A = (4*PG)/PR, onde A é a área verde, PG a quantidade de pixels verdes identificados, PR a quantidade de pixels vermelhos identificados, e o 4 representa a área do quadrado vermelho em cm².

A última etapa que o algoritmo executa é a de geração de uma imagem colorida para que o usuário possa verificar os objetos identificados pelo processamento através do aplicativo.

O experimento foi executado em uma estufa de aproximadamente 168 m² para cultivo de alface hidropônico situado na Escola Agrícola de Jundiaí/UFRN no Estado do Rio Grande do Norte, porém apenas uma pequena porcentagem desta área foi utilizada para o experimento. No local, há a presença de vento durante todo o dia.

Com a estrutura de captura montada, posicionada e devidamente calibrada, pôde-se dar início as capturas das imagens e validação do aplicativo. Para esta, foram feitas capturas também com o aplicativo Easy Leaf, para posterior comparação com os valores obtidos através do aplicativo Horus Green.

O aplicativo Horus Green tem como principal diferencial em relação ao Easy Leaf a possibilidade de persistência dos dados processados. Isso garante que as imagens possam ser capturadas e enviadas, e os resultados das medições de área foliar e as fotos originais ou processadas, ficarão disponíveis para posterior consulta e análise dos dados. Além disso, o sistema conta com um algoritmo de segmentação automático, que garante que o valor de área foliar não sofrerá interferência humana, garantindo assim um valor mais aproximado ao real.

Foram utilizadas 5 mudas da cultivar baba verão do tipo lisa repolhuda, com idade de 15 dias, dispostas como na Figura 3. Este formato foi escolhido para evitar a sobreposição de folhas, pois o sistema leva em consideração apenas as folhas visíveis.

Para o experimento, foram feitas 6 capturas, com intervalo de aproximadamente 72 horas entre a primeira e a segunda captura, e a partir desta o intervalo foi de aproximadamente 24 horas para as demais.

RESULTADOS E DISCUSSÃO

A tela principal do aplicativo Horus Green desenvolvido é apresentada na Figura 5. Na barra superior representado pela letra a) tem-se o botão responsável pela sincronização do aplicativo com a base de dados. Este botão possui 4 estados que indicam a presença ou não de Internet e se existem dados ou não a serem sincronizados. Se o usuário tentar enviar uma imagem sem Internet, o aplicativo salvará a imagem localmente evitando que ela se perca.

Ainda na barra superior, representado pela letra b), tem-se o botão de adicionar novo experimento, indicando por exemplo qual a variedade da alface e as dosagens dos nutrientes.

Na parte de conteúdo da tela, vê-se a listagem de experimentos cadastrados. Para cada experimento existem duas opções: clicar no corpo do elemento ou clicar no ícone da câmera. A primeira opção exibe os gráficos e informações do experimento e a segunda chama a tela responsável por adicionar uma nova foto ao experimento.

A tela de confirmação de captura aparece após o botão “tirar foto” da tela principal ser clicado e a imagem ser capturada. Esta tela tem como finalidade a confirmação da captura feita, onde é verificado se a imagem está dentro dos padrões e se não há nenhuma inconformidade.

A última tela do aplicativo é a que apresenta os gráficos e as informações do experimento. É formada por três abas, uma para o gráfico de área, outra para o gráfico de percentual de crescimento em relação à primeira captura feita, e outra para as informações do experimento que foram cadastradas. Esta tela e suas abas podem ser visualizadas na Figura 6.

Figura 5 - Tela principal do aplicativo móvel



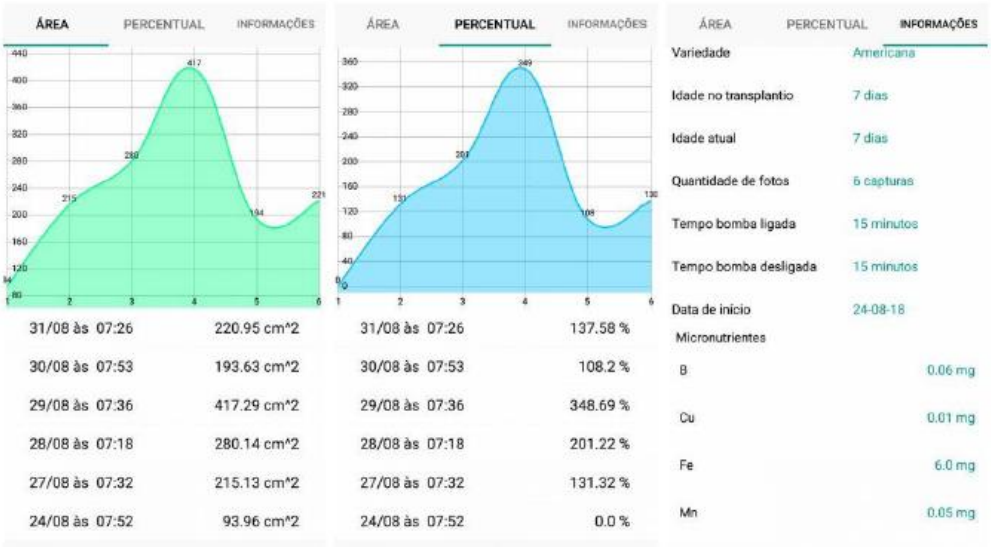
Fonte: Autoria própria

Durante o período do experimento houve uma pane temporária no bombeamento da solução nutritiva, fazendo com que as mudas de alface viessem a murchar. Este ocorrido não atrapalhou o experimento, mas gerou uma curva de crescimento diferente da que seria se as plantas estivessem sempre crescendo.

Para validar o sistema proposto neste trabalho, foi utilizado o aplicativo Easy Leaf. Este software tem como sua principal funcionalidade encontrar a área verde de uma planta da espécie *Arabidopsis*. Porém, como não há grandes diferenças entre o verde desta espécie citada e o verde da alface, o aplicativo foi utilizado para medir a área foliar das alfaces e possibilitar uma análise comparativa.

Na Tabela 1, temos a comparação dos valores obtidos com os dois aplicativos.

Figura 6 - Abas de gráficos e informações do experimento



Fonte: Autoria própria

Tabela 1 - Comparação entre os valores gerados pelo Horus Green e pelo Easy Leaf

Data	Área foliar – Horus Green	Área foliar – Easy Leaf
24/08/2018	93.96 cm^2	87.09 cm^2
27/08/2018	215.13 cm^2	216.99 cm^2
28/08/2018	280.14 cm^2	275.08 cm^2
29/08/2018	417.29 cm^2	386.74 cm^2
30/08/2018	193.63 cm^2	191.10 cm^2
31/08/2018	220.95 cm^2	198.92 cm^2

Constatou-se que o maior erro entre os valores de área foliar medidos foi de 30.55 cm² no dia 29/08/18, e o menor erro foi de 1.86 cm² no dia 27/08/18. O valor de erro médio entre as capturas foi de 11.48 cm².

No ambiente utilizado neste trabalho, a grande problemática foi a mudança constante nas posições das folhas de alface ocasionadas pelas intensas massas de ar que transpassavam às telas que delimitam a estufa. Acredita-se que esse seja o fator principal para as discrepâncias entre os valores medidos, no entanto, é importante destacar que essas diferenças são muito pequenas em relação aos valores de área foliar medidos.

CONCLUSÕES

Conclui-se que o sistema Horus Green pode ajudar aos produtores e principalmente aos pesquisadores do seguimento de cultivo hidropônico no acompanhamento do crescimento da alface, sendo o software de grande serventia para a realização de medições mais fiéis à realidade e mais precisas, substituindo métodos empíricos (olho) e destrutivos (colheita para medições diversas de tamanho, peso, massa fresca, etc).

Como trabalhos futuros, temos a captura de vídeo para analisar a taxa de crescimento, possibilitando a captura de vários momentos consecutivos extraindo *frames*, e gerando uma média dos resultados de área foliar obtidos com os mesmos; transferir a tarefa de *resize* do módulo PDI para o *smartphone*, diminuindo a quantidade de dados enviados; possibilitar o enfileiramento de várias imagens no *smartphone* sem necessidade de sincronização diária; além de submeter o algoritmo ao processamento de uma grande quantidade de imagens, com o intuito de encontrar falhas e adaptar o mesmo para vários cenários.

REFERÊNCIAS

BEZERRA NETO, E.; BARRETO, L. P. As Técnicas de hidroponia, Anais da Academia Pernambucana de Ciência Agronômica, Recife, v. 8 - 9, p.107-137, 2011/2012.

EASLON, H. M.; BLOOM, A. J. Easy Leaf Area: Automated digital image analysis for rapid and accurate measurement of leaf area. Applications in Plant Sciences. 2014.

FAVARATO, L. F., GUARÇONI, R. C., SIQUEIRA, A. P. Produção de alface de primavera/verão sob diferentes sistemas de cultivo. Revista Científica Intelletto. v. 2, n. 1. 2017.

GONZALEZ, R. C., WOODS, R. E. Processamento Digital de Imagens. 3ª ed. São Paulo: Pearson, 2010.

JUNG, D. H.; PARK, S. H.; HAN, X. Z.; KIM, H. J. Image Processing Methods for Measurement of Lettuce Fresh Weight. Journal of Biosystems Engineering. v. 40, n. 1, p. 89-93. 2015.

MACÊDO, C. Ajuste de modelos de crescimento da alface (*Lactuca sativa L.*) cultivada em sistema hidropônico e estimativa da variação da condutividade elétrica da solução nutritiva. Tese - Pós-graduação - Curso de meteorologia agrícola, Universidade Federal de Viçosa, Viçosa, MG. 2004.

PEREIRA, F. M.; LACERDA, F. T. L.; STINGHEL, M.; PASSOS, G.; HELL, L. R.; CAZAROTI, E. P. F.; WESTPHAL, R.; BREMENKAMP, D. M.; GALON, K.; COMETTI, N. N. Método não destrutivo para análise de crescimento de alface(*Lactuca sativa L.*) hidropônica e predição de colheita com o uso de imagens digitais. Anais do XV Encontro Latino Americano de Iniciação Científica e XI Encontro Latino Americano de Pós-Graduação – Universidade do Vale do Paraíba, São José dos Campos. 2011.

REIS, T. E. S., REIS, L. C., BARROS, O. N. F. Comparação de métodos de determinação de área: superfície foliar do feijoeiro. Geografia, Londrina.v. 9, n. 2, p. 151-157.2000.

RIBEIRO, H. F. Uso de imagens digitais na diagnose do teor de nitrogênio foliar em plantas de alface. Dissetação (Dissertação em produção vegetal) – Universidade Federal do Tocantins, Gurupi, TO. p. 62, 2016.

SANTOS, S. N., DIGAN, R. C., AGUILAR, M. A. G., SOUZA, C. A. S., PINTO, D. G., MARINATO, C. S., ARPINI, T. S. Análise comparativa de métodos de determinação de área foliar em genótipos de cacau. Bioscience Journal, Uberlandia, v. 30, p. 411-419. 2014.

Agriflix plataforma de conhecimento sobre o agronegócio

Valter Castelhano de Oliveira[2]*,Luiz Antonio Daniel*[2]

[2]Agriflix, São Paulo, Brasil, valter@agriflix.com.br, daniel@agriflix.com.br

RESUMO

A gestão e o fomento do conhecimento tecnológico tende a ser um gargalo nas projeções de protagonismo do Brasil, calcadas no Agronegócio. As inovações tecnológicas viabilizam a propagação de conteúdos nas mais diversas mídias, com destaque para a exacerbada produção de vídeos, quando associados ao agronegócio carecem de tutoria e validação nas mais diversas áreas técnicas e científicas do agronegócio. Este trabalho visa apresentar o desenvolvimento, sustentado por um processo de engenharia, de um sistema de serviço que disponibiliza uma plataforma colaborativa denominada Agriflix, com multiacesso de conteúdo audiovisual de qualidade sobre o agronegócio, selecionado e tutorado por especialistas, disponibilizado gratuitamente aos usuários e sustentado por um modelo de negócio inovador baseado em apoiadores financeiros e de conteúdo. No trabalho, são apresentados os fundamentos teóricos associados ao processo de engenharia de serviço e a modelagem associada ao sistema de serviço e desenvolvimento da plataforma.

PALAVRAS-CHAVE: sistema de serviço, multimídia, agronegócio, conhecimento

ABSTRACT

The management and promotion of technological knowledge tends to be a bottleneck in the projections of prominence in Brazil, based on Agribusiness. Technological innovations make possible the spread of content in all kind of media, especially the exacerbation of video production, but when associated with agribusiness, these videos don't present enough tutoring and validation in technical and scientific areas. This work aims to present the development, supported by an engineering process, of a service system that provides a collaborative platform called Agrifix, with multi-access of high quality audiovisual content about agribusiness, that were selected and tutored by specialists. The content is made available for

free to users and sustained by innovative business model based on financial and content supporters. This work is going to present the theoretical foundations associated with the service engineering and modeling processes associated to the system service and platform's development.
KEYWORDS: *service system, multimedia, agribusiness, knowledge.*

INTRODUÇÃO

O Brasil como um dos principais protagonistas da produção de alimentos no mundo, tente-se mostrado eficaz na aplicação de tecnologias de gestão e de produção agrosilvopastoril, haja vista que nos últimos 30 anos evoluíram sua produtividade de grãos de 1,5 ton/ha para aproximadamente 4,0 ton/ha (CONAB, 2017).

No que se refere a produção integrada sustentável para aumento de renda e produtividade, o Sistema de Integração Lavoura, Pecuária e Floresta "iLPF", tem-se efetivado como o paradigma mais produtivo da agricultura tropical, permitindo a utilização das terras na totalidade do ano, tornando-se o meio produtivo sustentado mais eficiente meio produtivo do agronegócio brasileiro (EMBRAPA, 2019).

Com base em uma agricultura com alto incremento no desenvolvimento tecnológico, de 1990 a 2017, o total das exportações agrícolas superou US$ 1,3 trilhão e ajudou a garantir saldos comerciais positivos para o Brasil (NEVES, 2018).

O setor de desenvolvimento tecnológico que mais tem contribuído para esse avanço técnico e cientifico da agricultura brasileira é sem dúvida a Tecnologia de Informação e Comunicação "TIC", que alavanca todos os aspectos de automação em operações da produção agrícola, nos aspectos da inteligência artificial em gestão e administração operacional, bem como a formação de bancos de dados e conhecimentos em "*big data in cloud*".

Operações e gestão de máquinas e equipamentos agrícolas, irrigação, condicionamento ambiental fechado, tecnologia de aplicação, armazenamento, logística, entre outros, são exemplos de eficácia face ao desenvolvimento e aplicação da informática e suas derivações em automação, mecatrônica, inteligência artificial.

No contexto do agronegócio, a gestão, disseminação e fomento do conhecimento tecnológico tende a ser um gargalo nas projeções de protagonismo do Brasil, implicando na busca por caminhos inovadores. Quanto à disseminação de conhecimento, as inovações tecnológicas viabilizam a propagação de conteúdos nas mais diversas mídias, com destaque para a exacerbada produção de vídeos associados ao agronegócio, com variado grau de

qualidade, pertinência e aplicação, carecem de curadoria e validação nas mais diversas áreas técnicas e científicas do agronegócio.

Este trabalho visa apresentar o desenvolvimento, sustentado por um processo de engenharia, de um sistema de serviço que disponibiliza uma plataforma colaborativa denominada Agriflix, que oferece multiacesso de conteúdo audiovisual de qualidade sobre o agronegócio, selecionado e tutorado por especialistas, disponibilizado gratuitamente aos usuários e sustentado por um modelo de negócio inovador baseado em apoiadores financeiros e de conteúdo.

No trabalho são apresentados fundamentos de engenharia de serviço, incluindo ciência de serviço, paradigma orientado a serviço, co-criação de valor, capacidade cognitiva e intenções.

Como resultado, apresenta-se um nível superior da modelagem do sistema de serviço Agriflix, direcionada pela engenharia de serviço, visando o processo de desenvolvimento, implantação e produção da plataforma e, também, dos processos de sustentação e operação do negócio.

ENGENHARIA DE SERVIÇO

Serviço é a aplicação de competências, incluindo conhecimento e proficiência, visando o benefício de terceiros (LUSCH; VARGO; WESSELS, 2008; SPOHRER et al., 2008; VARGO; AKAKA, 2009) e a configuração adequada entre esses elementos direcionada para a produção de valor ao cliente caracteriza-se como sistema de serviço (MAGLIO et al., 2009; SPOHRER et al., 2007). A engenharia de serviço enfoca estas definições, ao abordar o projeto de desenvolvimento de sistemas de serviço por meio da aplicação de técnicas e métodos que buscam a viabilização e manutenção de serviços consolidando a interação e co-criação entre os envolvidos com o sistema (OLIVEIRA, 2013; OLIVEIRA; SILVA, 2015).

A principal relação levantada pelos estudiosos da ciência de serviço (LUSCH; VARGO; WESSELS, 2008; SPOHRER et al., 2007; STANICEK; WINKLER, 2010) é a relação denominada de “co-criação de valor”, que ocorre quando consumidores e fornecedores se encontram acoplados para a realização do serviço.

A co-criação de valor representada por uma relação de colaboração entre consumidores e sistemas de serviço, resulta em um processo de satisfação do consumidor, e constitui uma relação extremamente subjetiva e singular para cada consumidor. Isso resulta em um problema importante na tentativa de automação de sistemas de serviços, pois exige que os sistemas de serviço tenham maior habilidade de percepção (das características e respostas dos

consumidores) e de adaptabilidade (para se alterar e se adequar às necessidades do consumidor) (OLIVEIRA; SILVA, 2015).

Assim, identifica-se a necessidade de se obter sistemas com maior capacidade cognitiva atendendo plenamente a ideia de co-criação de valor, co-design e co-produção (DUTRA; OLIVEIRA; SILVA, 2014), através da noção de sistemas de serviço, aplicados em cadeias de serviços e rede de agentes (artificiais e pessoas), tornando o desafio ainda maior, pois o design e a modelagem destes sistemas de serviço levam em conta fatores que antes não eram considerados (OLIVEIRA; SILVA, 2011).

A transformação para uma economia baseada em serviços também influencia outros setores da economia: na indústria de manufaturados, onde a automação está intimamente ligada à integração e digitalização, tornando mais fácil o acoplamento entre produto e serviço; na agricultura, com a introdução de processos mais precisos (agricultura de precisão) e mais integrados (integração no plantio de vários produtos, combinado com o tratamento da terra, como na Integração Lavoura-Pecuária-Floresta – ILPF (CORDEIRO et al., 2015).

Estas mudanças viabilizam a substituição dos modelos de produção clássicos, orientados a bens tangíveis, por modelos orientados a serviço, apoiados por sistemas de informação cognitivos (DUTRA et al., 2013; SILVA, 2014).

Na agricultura o principal problema, uma vez modelado o processo principal de produção agrícola, é a regulação da distribuição de insumos e recursos, oferecendo insumos para a integração com sistemas de serviço, resultando nos Agroserviços como estratégia de negócios (OLIVEIRA; SILVA; DANIEL, 2013).

Viabilizar a automação envolvendo máquinas, pessoas e processos como recursos, onde o produto final agora é o serviço (intangível, inseparável, variável e perecível) (GRÖNROOS, 2004; LOVELOCK; WRIGHT, 2002), potencializa a importância de tratar o design dos sistemas de serviço de forma sistemática (KIM, 2009; MAGLIO et al., 2009; OSTROM et al., 2010, 2015; SPOHRER et al., 2007; STANICEK; WINKLER, 2010).

O uso de modelos no processo de design de sistemas de serviço segue a tendência atual proposta pela OMG[1] e pelo INCOSE[2]. A aderência do design por modelos (*model driven design*) (KENT, 2002) a sistemas de serviço se justifica pela premissa de que o processo de automação aplicado aos sistemas de serviço seja implementado por meio dos sistemas de informação, como uma forma interativa e que admite: co-criação, flexibilidade, aderência a

[1] OMG - *Object Management Group* (http://www.omg.org/).

[2] INCOSE - The International Council on Systems Engineering (http://www.incose.org).

requisitos voláteis, aderência a sistemas com um alto número de *viewpoints* e facilidade de reutilização.

O desafio é aliar boas práticas à integração com métodos e processos consolidados para tratar do desenvolvimento de sistemas de serviço, com o uso de métodos formais de forma harmônica, técnica e economicamente viável, a rastreabilidade na análise dos pontos de vista durante a fase de eliciação e análise de requisitos (OLIVEIRA, 2013). Além disso, a interação entre os agentes para validar requisitos devem receber tratamento especial e esta é uma questão-chave na engenharia de serviço, uma vez que os sistemas de serviço devem resultar da convergência entre diversas e diferentes expectativas de várias classes de agentes, para gerar valor a partir do processo de automação (DUTRA, 2016).

Outro aspecto prático importante é que o processo de análise de requisitos deve contemplar a parte cognitiva, e consequentemente os sistemas de serviço devem ser inteligentes, a necessidade dos envolvidos quanto a cognição precisa ser eliciada e uma possibilidade é a inclusão do tratamento de intenções (BRESCIANI et al., 2004) durante a fase de análise de requisitos, o que aponta para uma abordagem baseada no método i* (YU, 1997) criado em 1997 na tese de doutorado de Eric Yu (University of Toronto) e desenvolvido mais recentemente até atingir estágios que a habilitam a implementações de grande porte.

MODELAGEM SISTEMA DE SERVIÇO

O *Service-Oriented Framework to the Design of Information System Service*) SoftDiss - é uma disciplina de projeto orientada a serviço, destinada à especificação de sistemas de serviço, suportando as fases iniciais do desenvolvimento, incluindo a eliciação, modelagem e análise de requisitos, através da configuração dinâmica e cognitiva de tecnologias, pessoas, sistemas externos e das informações compartilhadas entre estes elementos (OLIVEIRA, 2013). Os SoftDiss é implementado no Enterprise Architect (SPARX, 2019) e utiliza o *Service-Oriented Modeling Framework* – SOMF (BELL, 2008) como base para a realização das etapas iniciais de modelagem e especificação do sistema de serviço.

A primeira etapa da modelagem de um sistema de serviço é compreender o domínio onde está inserido e os agentes envolvidos no seu desenvolvimento visando a eliciação e análise de seus requisitos (OLIVEIRA, 2013).

O Documento de Visão[3] tem a finalidade de coletar, analisar e definir as necessidades e características de nível superior do sistema de serviço, enfocando os recursos esperados pelos envolvidos e usuários, além de esclarecer as razões destas necessidades. Um importante elemento do documento de visão é a Descrição do Problema, que forneça uma descrição resumindo o problema que está sendo resolvido pelo sistema de serviço (Figura 1).

Figura 1 – Documento de Visão Agriflix: Descrição do Problema

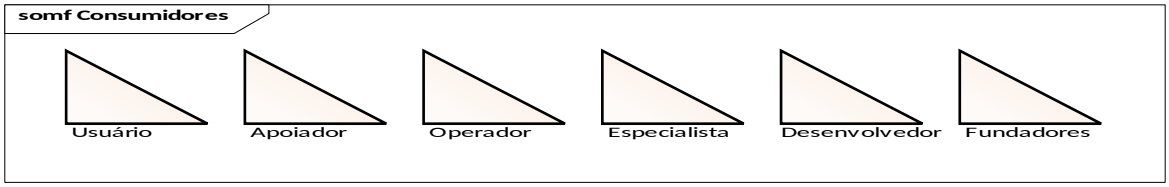
O problema	de disponibilizar conhecimento sobre o Agronegócio na forma de vídeos, séries e documentários
afeta	consumidores de conteúdo associado ao Agronegócio no Brasil e América Latina, incluindo empresas, profissionais, estudantes, agricultores e público urbano em geral.
cujo impacto é	a inexistência de uma plataforma colaborativa, onde especialistas possam indicar episódios de séries de vídeos organizados e contendo informação útil e confiável
uma boa solução seria	a estruturação de uma plataforma multi acesso (apps, web, etc) que reúna usuários e especialistas do agronegócio em um ecosistema de cooperação, calcado no modelo de rede social, compartilhando conhecimento na forma de séries e episódios organizados, recomendados e validados por especialistas do agronegócio brasileiro e da américa latina.

Fonte: autoria própria

Para o atendimento eficiente das reais necessidades dos usuários e envolvidos, é preciso identificar e considerar todos eles como parte do processo de Modelagem de Requisitos. O Documento de Visão fornece um perfil dos envolvidos e dos usuários que integram o projeto, e dos principais problemas que, de acordo com o ponto de vista deles, poderão ser abordados pela solução proposta (não apresentado neste trabalho).

Considerando que um sistema de serviço é formado por uma cadeia de serviços, onde consumidores internos e externos atuam como elementos de interligação em processo de co-criação de valor, identificar estes potenciais consumidores é importante para a eliciação e documentação dos requisitos. A Figura 2 apresenta o Diagrama de Consumidores SOMF inspirado no levantamento de envolvidos e usuários registrados no Documento de Visão.

Figura 2 – Consumidores na cadeia de serviços



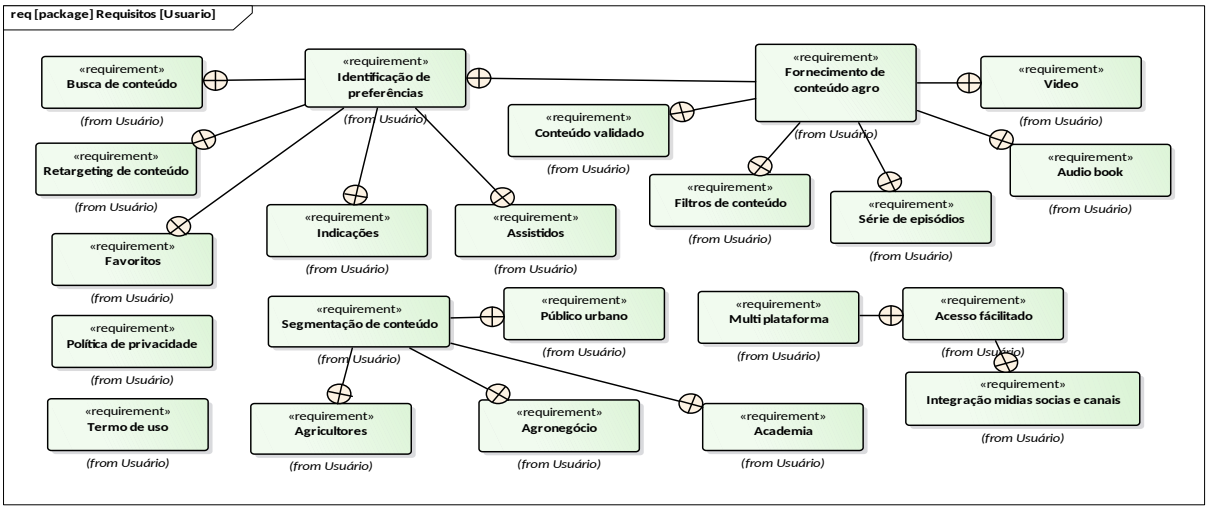
Fonte: autoria própria

A Figura 3 apresenta um dos diversos diagrama de requisitos SysML (*System Modeling Language*) (FRIEDENTHAL; MOORE; STEINER, 2009), resultado da eliciação de requisitos (necessidades levantadas e documentadas), neste caso do consumidor denominado

[3] Documento de Visão (https://www.ibm.com/support/knowledgecenter/pt-br/SSWMEQ_4.0.6/com.ibm.rational.rrm.help.doc/topics/r_vision_doc.html)

Usuários do Agriflix. Cabe observar que, por exemplo, requisitos como “Retarging de Conteúdo” e “Conteúdo Validado” implicam na inclusão de recursos cognitivos no sistema de serviço.

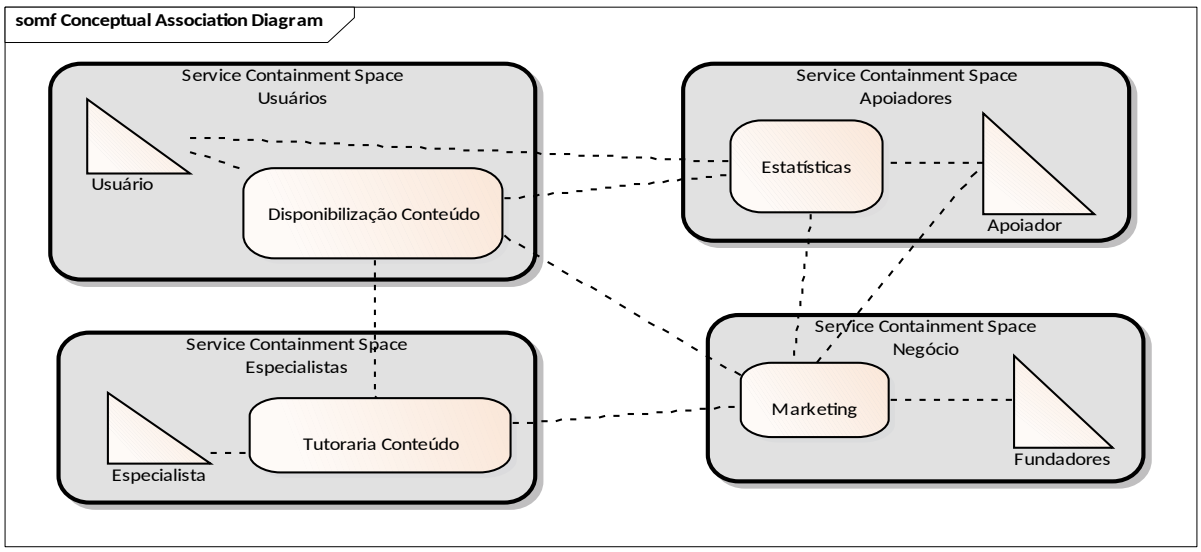
Figura 3 - Requisitos dos Usuários Agriflix



Fonte: autoria própria

A definição dos atributos auxilia nas decisões e priorizações no decorrer da modelagem SOMF, principalmente, como base para a identificação dos serviços. Considerando-se os requisitos e a documentação obtida para o Agriflix foram selecionados os atributos Colaboração, Credibilidade, Neutralidade e Educação. O desenvolvimento dos diagramas de atributos e árvore de decisão do SOMF (BELL, 2008; OLIVEIRA, 2013) são a base para a definição dos Serviços de alto nível que compõem a cadeia ou rede serviços que implementam o Agriflix.

Figura 4 – Diagrama de Associação Conceitual do Agriflix



Fonte: autoria própria

No ciclo de vida de desenvolvimento do Agriflix estes diagramas (e todos os demais modelos já desenvolvidos e suprimidos devido a restrições no tamanho da publicação) foram revistos e atualizados diversas vezes e, certamente, até a etapa de descontinuidade do sistema ainda deverão sofrer diversas alterações, pois o *feedback* é inerente a todo sistema de serviço e deve ser exercitado para capturar as mudanças nas necessidades dos envolvidos, realimentando a modelagem desde os níveis mais altos (propósitos, intenções e requisitos) até os elementos mais detalhados da implementação dos serviços básicos.

CONSIDERAÇÕES FINAIS

No trabalho foram apresentados os fundamentos de engenharia de serviço, incluindo ciência de serviço, paradigma orientado a serviço, co-criação de valor, capacidade cognitiva e intenções, que sustentam e justificam a necessidade de um processo de modelagem de sistemas de serviço, visando o sucesso e longevidade de sua implementação, sempre alinhado com dinâmica das necessidades dos envolvidos com o sistema de serviço.

A modelagem de nível superior do sistema de serviço Agriflix, mesmo sendo parcialmente apresentada devido a restrições impostas para a publicação, permite compreender a importância da modelagem para os sistemas de serviço e apresentar a dinâmica utilizada no processo de modelagem e a sustentação dos fundamentos da engenharia de serviços.

A engenharia de serviços, sustentando o processo de modelagem do sistema de serviço, utilizando ferramentas (Enterprise Architect) e frameworks (SOMF, SoftDiss) são a base para a manutenção e alcance do propósito do sistema de serviço Agriflix, que é ser a maior e mais completa plataforma de conhecimento multimídia sobre o Agronegócio do Brasil e da América Latina.

REFERÊNCIAS

BELL, M. **Service-Oriented Modeling: Service Analysis, Design, and Architecture**. New Jersey: Wiley, 2008.

BRESCIANI, P.; PERINI, A.; GIORGINI, P.; GIUNCHIGLIA, F.; MYLOPOULOS, J. Tropos: An Agent-Oriented Software Development Methodology. **Autonomous Agents and Multi-Agent Systems**, [s. l.], v. 8, n. 3, p. 203–236, 2004. Disponível em: <http://link.springer.com/10.1023/B:AGNT.0000018806.20944.ef>

CONAB. **Acompanhamento da safra Brasileira, Grãos**. 2017. Disponível em: <https://www.conab.gov.br/info-agro/safras/graos/boletim-da-safra-de-graos?start=20>. Acesso em: 28 jun. 2019.

CORDEIRO, L. A. M.; VILELA, L.; KLUTHCOUSKI, J.; MARCHÃO, R. L. **Integração lavoura-pecuária-floresta: o produtor pergunta, a Embrapa responde**. DF: Embrapa, 2015. Disponível em: <http://mais500p500r.sct.embrapa.br/view/pdfs/90000033-ebook-pdf.pdf>

DUTRA, D. de S. **An Intention-Based Service Design Discipline for the Product-Service Architecture**. 2016. USP - Universidade de São Paulo, [s. l.], 2016.

DUTRA, D.; OLIVEIRA, V. C. De; SILVA, J. R. J. R.; DE OLIVEIRA, V. C.; SILVA, J. R. J. R. Manufacturing as Service : the challenge of Intelligent Manufacturing. In: 11TH IFAC WORKSHOP ON INTELLIGENT MANUFACTURING SYSTEMS (IMS13) 2013, São Paulo, SP, Brasil. **Anais**... São Paulo, SP, Brasil

DUTRA, D. S.; OLIVEIRA, V. C. De; SILVA, J. R. A Service-oriented Approach to Technology Based Industry. In: INDIN IEEE CONGRESS 2014 2014, Porto Alegre, Brazil. **Anais**... Porto Alegre, Brazil

EMBRAPA. **Integração Lavoura-Pecuária-Floresta**. 2019. Disponível em: <https://www.embrapa.br/web/rede-ilpf>. Acesso em: 28 jun. 2019.

FRIEDENTHAL, S.; MOORE, A.; STEINER, R. **A Practical Guide to SysML: The Systems Modeling Language**. First edit ed. San Francisco, CA: Morgan Kaufmann, 2009.

GRÖNROOS, C. **Marketing : gerenciamento e serviços**. Rio de Janeiro: Elsevier/Campus, 2004.

KENT, S. Model driven engineering. In: INTEGRATED FORMAL METHODS 2002, **Anais**... : Springer, 2002. Disponível em: <http://www.springerlink.com/index/9VUQB4HP8FYG2ADV.pdf>. Acesso em: 1 ago. 2011.

KIM, H. Service science for service innovation. **Journal of Service Science**, [s. l.], v. 1, n. 1, p. 1–7, 2009. Disponível em: <http://link.springer.com/10.1007/s12927-009-0001-3>. Acesso em: 15 maio. 2013.

LOVELOCK, C.; WRIGHT, L. **Principles of Service Marketing and Management**. Upper Saddle River, NJ: Prentice Hall, 2002.

LUSCH, R.; VARGO, S.; WESSELS, G. Towards a conceptual foundation for service science: Contributions from service-dominant logic. **IBM Systems Journal**, [s. l.], p. 1–9, 2008.

MAGLIO, P. P.; VARGO, S. L.; CASWELL, N.; SPOHRER, J. The service system is the basic abstraction of service science. **Information Systems and e-Business Management**, [s. l.], v. 7, n. 4, p. 395–406, 2009. Disponível em: <http://www.springerlink.com/index/10.1007/s10257-008-0105-1>

NEVES, M. F. **Agronegócio: Uma Máquina de Geração de Renda ao Brasil**. 2018. Disponível em: <https://noticiasagricolas.com.br/noticias/agronegocio/206417>. Acesso em: 28 jun. 2019.

OLIVEIRA, V. C. De. **Modelagem e design de sistemas de serviço para automação**. 2013. Escola Politécnica da Universidade de São Paulo, [s. l.], 2013.

OLIVEIRA, V. C. De; SILVA, J. R. Process Oriented Design Service: Case Study for Automated Information System. In: 21ST BRAZILIAN CONGRESS OF MECHANICAL ENGINEERING 2011, Natal, RN, Brasil. **Anais**... Natal, RN, Brasil

OLIVEIRA, V. C. De; SILVA, J. R. A service-oriented framework to the design of information system service. **Journal of Service Science Research**, [s. l.], v. 7, n. 2, p. 55–96, 2015. Disponível em: <http://link.springer.com/10.1007/s12927-015-0003-2>

OLIVEIRA, V. C. De; SILVA, J. R.; DANIEL, L. A. Engenharia de serviço aplicada ao agronegócio. **Reverte - Revista de Estudos e Reflexões Tecnológicas da Faculdade de Indaiatuba**, [s. l.], n. 13, p. 14, 2013. Disponível em: <http://fatecindaiatuba.edu.br/reverte/index.php/revista/article/view/77>

OSTROM, A. L.; BITNER, M. J.; BROWN, S. W.; BURKHARD, K. a.; GOUL, M.; SMITH-DANIELS, V.; DEMIRKAN, H.; RABINOVICH, E. Moving Forward and Making a Difference: Research Priorities for the Science of Service. **Journal of Service Research**, [s. l.], v. 13, n. 1, p. 4–36, 2010.

OSTROM, A. L.; PARASURAMAN, A.; BOWEN, D. E.; PATRICIO, L.; VOSS, C. a. Service Research Priorities in a Rapidly Changing Context. **Journal of Service Research**, [s. l.], v. 18, n. 2, p. 127–159, 2015. Disponível em: <http://jsr.sagepub.com/cgi/doi/10.1177/1094670515576315>

SILVA, J. R. New Trends in Manufacturing : Converging to Service and Intelligent Systems. **JRSa IFAC2014**, [s. l.], 2014.

SPARX. **Sparx Systems Pty Ltd.** 2019. Disponível em: <http://www.sparxsystems.com.au/>. Acesso em: 28 jun. 2019.

SPOHRER, J.; MAGLIO, P. P.; BAILEY, J.; GRUHL, D. Steps Toward a Service Systems. **IEEE Computer Society**, [s. l.], n. January, p. 71–77, 2007.

SPOHRER, J.; VARGO, S. L.; CASWELL, N.; MAGLIO, P. P. The Service System Is the Basic Abstraction of Service Science. In: PROCEEDINGS OF THE 41ST ANNUAL HAWAII INTERNATIONAL CONFERENCE ON SYSTEM SCIENCES (HICSS 2008) 2008, **Anais**... : Ieee, 2008.

STANICEK, Z.; WINKLER, M. Service Systems Through The Prism of Conceptual Modeling. **Service Science Journal**, [s. l.], v. 2, p. 112–125, 2010.

VARGO, S. L.; AKAKA, M. A. Service-Dominant Logic as a Foundation for Service Science: Clarifications. **Service Science**, [s. l.], v. 1, n. 1, p. 32–41, 2009. Disponível em: <http://sersci.com/magazine/V1N1.pdf#page=37>

YU, E. S. Towards modelling and reasoning support for early-phase requirements engineering. In: PROCEEDINGS OF ISRE '97: 3RD IEEE INTERNATIONAL SYMPOSIUM ON REQUIREMENTS ENGINEERING 1997, **Anais**... : IEEE Comput. Soc. Press, 1997. Disponível em: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=566873>

Uso de redes neurais convolucionais para detecção de laranjas no campo

João Camargo Neto[1], Sônia Ternes[1], Kleber Xavier Sampaio de Souza[1], Inácio Henrique Yano[1] , Leonardo Ribeiro Queiros[1]

[1] Embrapa Informática Agropecuária, Campinas, São Paulo, Brasil,

joao.camargo@embrapa.br, sonia.ternes@embrapa.br, kleber.sampaio@embrapa.br,

inacio.yano@embrapa.br, leonardo.queiros@embrapa.br

RESUMO

A laranja e seus derivados são um dos principais produtos do agronegócio brasileiro, além de uma das cadeias produtivas que mais emprega mão de obra por hectare, o que mostra o alto grau de impacto econômico e social desta cultura para o país. Uma estimativa de produção eficiente pode auxiliar os produtores tanto no manejo de sua lavoura quanto na adoção de estratégias de vendas com a indústria. Este trabalho descreve o processo de treinamento e teste de uma rede neural de aprendizado profundo para a detecção e contagem de frutos verdes a partir de imagens digitais de pés de laranja obtidas no campo. Os resultados para as imagens de teste apresentaram índice de mais de 90% de precisão, com cerca de 90% de revocação para a rede neural. Isso indica que a metodologia utilizada é bastante promissora.

PALAVRAS-CHAVE: Visão computacional, Aprendizado profundo, Yolo-v3, Citros.

ABSTRACT

Oranges and derivatives are one of the main products of Brazilian agribusiness and one of the sectors that employs more labor force per hectare, which shows the high degree of economic and social impact of this crop for the country. An efficient yield estimation can assist producers in managing their crop and in adopting sales strategies with industry. This work describes the process of training and testing a convolutional neural network for detecting and counting green fruits from digital images of orange trees obtained in the field. The results for the test images showed an index of more than 90% precision, with about 90% recall to the neural network. This indicates that the methodology used is very promising.

KEYWORDS: Computer vision, Deep learning, Yolo-v3, Citrus.

INTRODUÇÃO

Segundo o Levantamento Sistemático da Produção Agrícola (IBGE, 2019) a área plantada de laranja em 2018 foi de 605.752 ha e a estimativa para 2019 é de 594.541 ha, a produção em 2018 foi de 16.677.091 t e estimativa de 16.730.652 t para 2019.

Esses números demonstram a importância econômica e social que a citricultura tem para o país, como também relatado por Neves e Trombim (2017), que verificaram que em 2017 a citricultura foi responsável por gerar cerca de 200 mil empregos diretos e indiretos, além do setor gerar um PIB de US$ 6,5 bilhões de dólares em todos elos de sua cadeia produtiva. Por ser uma atividade que exige grande quantidade de mão de obra, manejos de insumos e superação de desafios agronômicos, a citricultura tem um terreno fértil para desenvolvimento e uso de tecnologias, dentro da filosofia da Agricultura de Precisão, que permitam automatizar e otimizar tarefas dentro do cultivo de plantas de laranjas, buscando melhorar a eficiência de produção e retorno econômico.

Diferente de outras *commodities* como milho e soja, a produção da laranja não pode ficar estocada na propriedade aguardando melhores condições de comercialização. Por esse motivo, uma estimativa de produção eficiente pode auxiliar os produtores em estratégias de vendas com a indústria. Um grande problema é o alto custo dessa estimativa que atualmente é intensiva em mão de obra e tempo.

Buscando contribuir para solucionar essa problemática enfrentada pelos citricultores, o objetivo deste trabalho é usar técnicas de visão computacional baseadas em redes neurais de aprendizado profundo para detecção e contagem de laranjas verdes.

A utilização de redes neurais de aprendizado profundo para identificação e contagem de frutos visíveis na copa da árvore pode ser o primeiro passo para a construção de um sistema robusto de estimativa do número de frutos no pomar. As redes neurais de aprendizado profundo vem sendo utilizado com sucesso em diversos tipos de aplicações. Na área agrícola para detecção de frutas utilizando a rede neural convolucional - Faster R-CNN (SA *et al*., 2016), para a contagem do número da panícula por unidade de área em experimentos de melhoramento genético de sorgo (GHOSAL *et al*., 2019), na detecção e classificação (gramíneas e folhas largas) de plantas daninhas na cultura de soja (FERREIRA *et al*., 2017). Em Tian *et al*. (2019) relata-se o uso de rede YOLO-V3 para identificação de maçãs em vários estádios de desenvolvimento da fruta para estimativa de produção. Já em Koirala *et al*. (2019b) são feitas identificações de mangas utilizando-se diversas redes neurais convolucionais, entre as quais uma versão modificada da YOLOv3, denominada MangoYOLO, todas as redes testadas obtiveram resultados superiores a 90% de taxa de

acerto, a partir de imagens noturnas com iluminação artificial, este recurso foi utilizado para minimizar os efeitos das variações na luminosidade nas fotos, causadas pela iluminação natural.

Apesar de existirem alguns artigos na literatura relatando a eficiência do uso de redes neurais convolucionais de aprendizado profundo em aplicações agrícolas, ainda existem vários desafios a serem superados para disponibilizar esta tecnologia em aplicações reais no campo. A metodologia aqui relatada busca avaliar o potencial do uso de rede neurais para detecção de frutos verdes a partir de fotos de árvores tomadas em condições reais de campo, sem preocupação com iluminação específica, tipo de equipamento ou resolução das imagens.

MATERIAL E MÉTODOS

Sistema de detecção usando as redes convolucionais YOLO

O sistema usado neste trabalho para a detecção de frutos diretamente a partir de imagens coletadas em campo foi um tipo de rede neural convolutiva batizada de YOLO (*You Only Look Once*) pelo seu criador Joseph Redmon (REDMON, 2016). Em particular, usou-se uma versão modificada da YOLO-v3 (REDMON; FARHADI, 2018), que ao invés das 53 camadas convolucionais originais, possui 75. A rede possui ainda outras 31 camadas dos tipos *shortcut*, *yolo*, *route* e *upsample*, perfazendo 106 camadas.

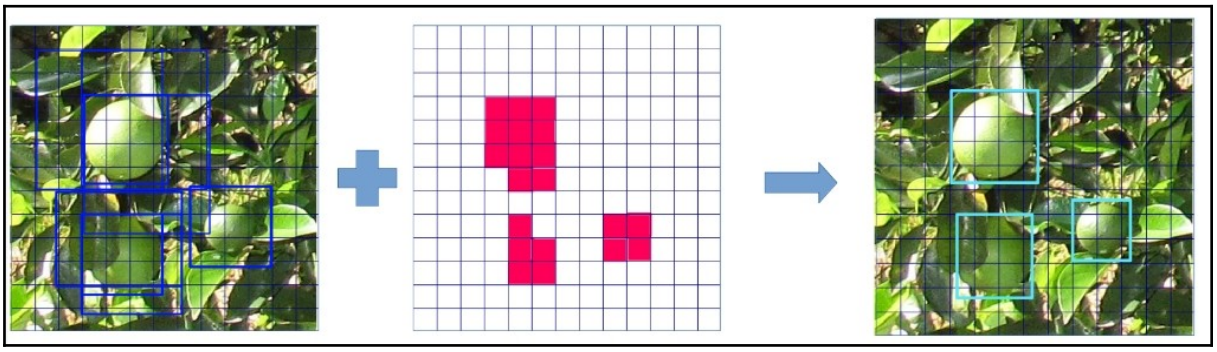
Existem dois repositórios de onde podem ser baixados os fontes da YOLO-v3 para serem compilados, um deles desenvolvido por Joseph Redmon (REDMON, 2019) e outro, que é um *fork* deste primeiro, desenvolvido por Alexey A. B. (ALEXEY, 2019). Preferimos usar esta última implementação, pois ela apresenta várias melhorias em relação à original, tais como a melhora no desempenho da detecção, redução no tempo de treinamento, otimização da alocação de memória e cálculo de medidas de desempenho dos modelos, dentre outras. Estes cálculos foram essenciais para produzirmos os gráficos da próxima seção.

O tamanho do *kernel* de detecção é determinado pela fórmula N x N x (B x (5+C)), onde N é a dimensão da imagem na última camada de convolução, B é o número de *bounding boxes*, e C é o número de classes. O número 5 vem do fato de que para cada ponto onde ocorre a detecção, tem-se cinco medidas: as quatro coordenadas de cada *bounding box* e um valor de probabilidade de haver um objeto naquela área. Como temos uma única classe (laranja) no projeto e usamos 3 *bounding boxes*, o dimensionamento do *kernel* de detecção é N x N x 18. O número N é variável conforme a escala de detecção. Para melhorar a precisão, o YOLO v3 implementa 3 escalas. A menor escala ocorre quando a rede realizou sua subamostragem máxima e a imagem foi reduzida em 32 vezes, de 416 x 416 para 13 x 13.

Este detector fica na 82ª camada e tem dimensão 3.042 (13 x 13 x 18). Após este ponto, a rede volta a aumentar (*up-sampling*), até atingir 26 x 26 e, para melhorar o desempenho, é concatenada com a rede de mesmo tamanho lá da camada 61 e a nova detecção ocorre na camada 94, agora em uma grade de 26 x 26. Nesta detecção, o *kernel* é um tensor de dimensão 12.168. O processo se repete para a última escala de detecção, com novo aumento de 26 x 26 para 52 x 52. Novamente há uma concatenação com as camadas de mesma dimensão lá do início, na camada 36, e novas convoluções, levando a uma detecção final na 106ª camada. Para esta última detecção, o *kernel* é um tensor de dimensão 48.672.

Como exemplo de funcionamento do YOLO, na primeira detecção, com grade 13 x 13, na Figura 1 à esquerda tem-se o sistema testando todos os *bounding boxes* para cada uma das células da grade. Para cada camada de detecção existem 3 possíveis *bounding boxes*. Os *bounding boxes* com maior probabilidade de haver um objeto de interesse naquela região recebem valores maiores. O sistema também estabelece a probabilidade de haver um fruto naquela região, ilustrado na parte central da Figura 1. Por fim, o sistema aplica a regra Bayesiana P(fruto) = P(objeto) x P(fruto/objeto). Toda convergência é guiada pela otimização de uma função de perda, cujos detalhes podem ser encontrados na literatura.

Figura 1 – Funcionamento da rede neural convolucional YOLO



Fonte: gerada pelo Autores

Aquisição e tratamento das imagens

O sistema foi treinado com 2.035 imagens, utilizando placa gráfica GPU, divididas em 1.832 para treino e 203 para validação (10% do conjunto de treinamento) durante a etapa de treinamento. Para teste das redes treinadas usou-se um outro conjunto de dados com 1.030 imagens. Cada uma dessas imagens tem dimensão de 416x416 pixels e foram recortadas das imagens originais com maior tamanho. As imagens originais foram capturadas usando diferentes dispositivos e resoluções, tais como câmeras fotográficas e celulares, sendo sua grande maioria fornecida pelo Programa de Estimativa de Safra (PES) do Fundo de Defesa da Citricultura (Fundecitrus). A diversidade de meios de captura se justifica por querermos testar

o sistema em uma situação real de campo, sem impor condições específicas de equipamentos, de resolução ou de condições de iluminação. Adicionalmente, os frutos capturados nas imagens são de diferentes variedades de laranjas e, embora predominantemente verdes, também representam vários graus de maturação.

Anotação das imagens

Após recortadas, cada uma das imagens foi manualmente anotada e verificada por usuários humanos em um sistema desenvolvido para este fim. Neste sistema desenha-se um retângulo ao redor da área que se identifica como um fruto ou parte visível deste. A Figura 2 ilustra duas imagens reais com 5 e 4 frutos anotados, respectivamente. Como se pode perceber, a primeira imagem está muito mais iluminada que a segunda, situação em que o reflexo da luz do sol tanto nas laranjas quanto nas folhas pode confundir o sistema de identificação neural. Outro complicador é que existem frutos em tons de verde mais escuro localizados em região de sombra (mais escura) e frutos parcialmente oclusos pela folhagem.

Figura 2 – Exemplo de duas imagens da base de treino com os frutos anotados

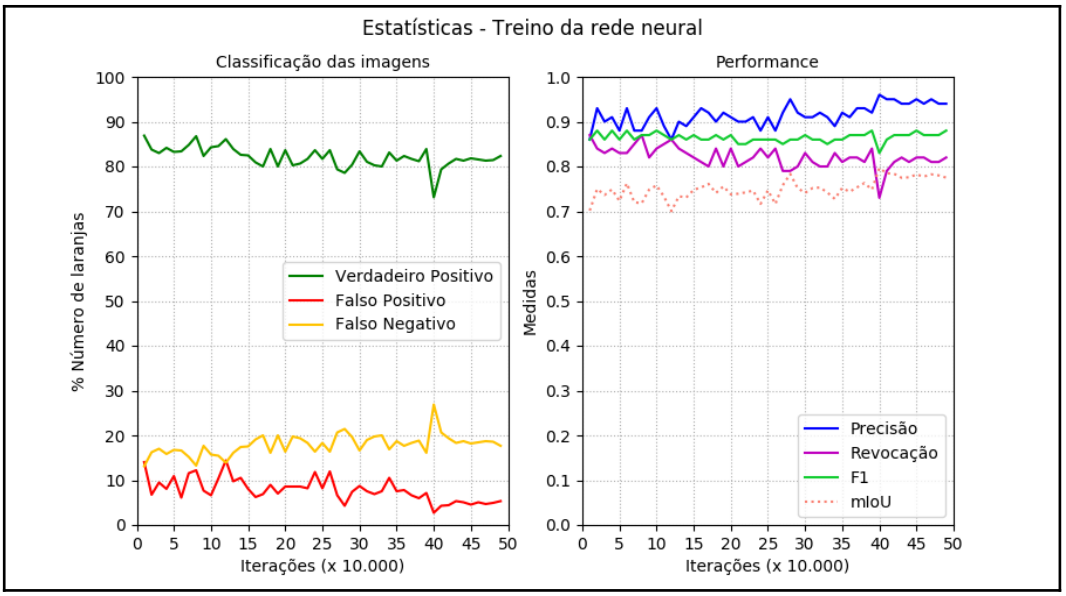


Fonte: gerada a partir da versão original cedida pelo PES/Fundecitrus

RESULTADOS E DISCUSSÃO

Durante o treinamento (base contendo 2.035 imagens) foram executadas 490.000 iterações da rede neural, permitindo observar por meio de gráficos diversos aspectos do comportamento dos modelos. A cada 1.000 iterações o modelo foi salvo, ou seja, foram geradas 490 redes neurais para serem testadas. Na Figura 3 são apresentadas as estatísticas de desempenho dessas redes frente à base de validação (203 imagens contendo 770 frutos anotados) durante a etapa de treinamento.

Figura 3 – Estatísticas obtidas durante a fase de treino da rede neural



Fonte: gerada pelo Autores

No gráfico à esquerda são apresentadas as porcentagens de frutos identificados corretamente (verdadeiro positivo ou VP), frutos não identificados pela rede (falso negativo ou FN) e partes de folhas/galhos identificados erroneamente como frutos (falso positivo ou FP). Ao final do treinamento obteve-se a identificação correta de mais de 80% dos frutos existentes na base.

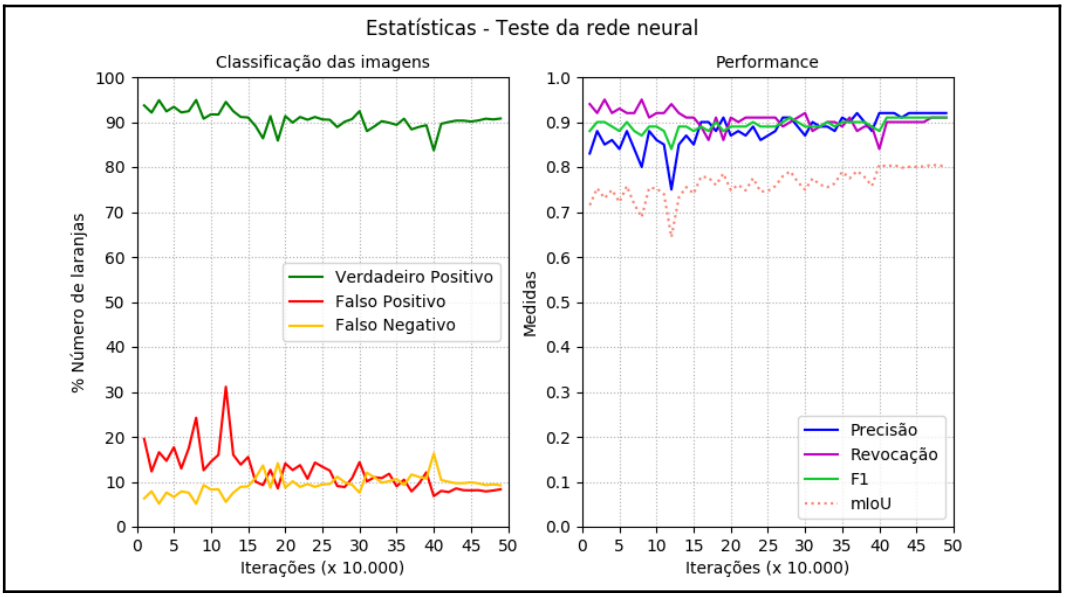
No gráfico à direita são apresentadas as curvas relativas à performance da rede. A curva azul indica a precisão do modelo, dada pela fração VP/(VP+FP), ou seja, quantas detecções corretas o modelo produziu. A curva roxa relaciona-se à métrica de revocação que é calculada pela fração VP/(VP+FN), indicando a frequência com que o modelo identifica corretamente uma laranja. Observando o gráfico percebe-se uma tendência de melhora na precisão, com pequena queda na revocação. Em geral, com o aumento da precisão ocorre uma perda na revocação, pois quanto mais preciso o sistema se torna em identificar a classe alvo (laranja), menos ele detecta alguns objetos que tenham menor probabilidade de serem verdadeiros positivos. Quem estabelece o compromisso entre a precisão e a revocação é a medida F1 (curva verde), dada por: F1 = 2 * (Precisão * Revocação)/(Precisão + Revocação). Percebe-se que essa medida, após a estabilização, teve uma melhoria por volta de 450.000 iterações. Por fim, percebe-se uma estabilização a partir de 400.000 iterações do valor da métrica *mean intersect over union* ou mIoU (curva pontilhada rosa), que fornece a medida de que há um objeto de interesse nas várias regiões identificadas pelos *bounding boxes*.

Ressalta-se que as quedas em todas as curvas na iteração 400.000 se dá pela mudança automática no valor do passo do gradiente de otimização da função de perda citada na seção

anterior. Após um ajuste inicial, a rede retoma seu equilíbrio.

Na Figura 4 estão apresentados os resultados das estatísticas obtidas do processo de detecção para a base de teste (1.030 imagens não usadas no treinamento).

Figura 4 – Estatísticas obtidas para a fase de teste da rede neural



Fonte: gerada pelo Autores

Observando as curvas do gráfico à esquerda, embora os modelos obtidos com menor número de iterações tenha uma alta porcentagem de laranjas identificadas (verdadeiro positivo), há ainda um número expressivo de falsos positivos. A situação melhora quando se testa os modelos obtidos com maior número de iterações, pois há um aumento da precisão sem implicar numa queda substancial na revocação, conforme mostrado no gráfico à direita. Por exemplo, a partir de 200.000 iterações, quando os modelos tendem a ser mais estáveis, e em especial, após 400.000 iterações, há uma queda considerável da porcentagem de falsos positivos e falsos negativos, com valores de precisão e revocação acima de 0,9.

O modelo correspondente à iteração 450.000, escolhido com base nos resultados mostrados na Figura 4, foi utilizado para teste da detecção dos frutos em imagens contendo uma árvore inteira. Para este modelo, os valores alcançados nas estatísticas da Figura 4 são: precisão = 0,95; revocação = 0,82; F1 = 0,88; mIoU = 78,2%. A Figura 5 mostra a imagem usada no teste, que possui tamanho original de 2.068x3.333 pixels.

Um experimento foi realizado utilizando uma janela deslizante, que percorre toda a imagem. A Figura 6 mostra o resultado para diversos tamanhos de janela deslizante. No eixo das abcissas é apresentado o tamanho da janela deslizante e no eixo das ordenadas a porcentagem de frutos detectados com relação ao total real de 206 frutos contidos na imagem.

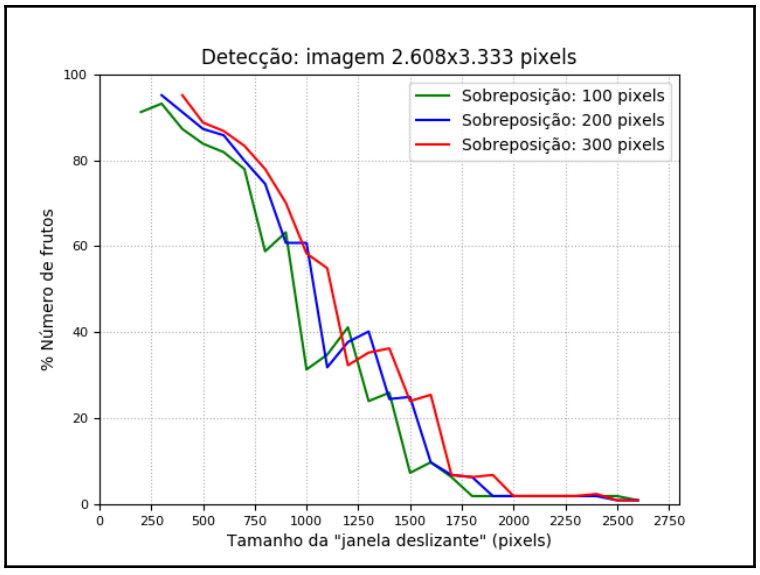
Figura 5 – Imagem original usada para testes de detecção de frutos por planta



Fonte: PES/Fundecitrus

Como pode ser observado, conforme o tamanho dessa janela tende às dimensões da entrada da rede (416x416 pixels), mais frutos são detectados. Isto se deve ao fato que imagens com dimensões muito maiores que a configuração de entrada da rede são reduzidas, e neste processo de redução o objeto a ser detectado pela rede é eliminado ou reduzido a tamanhos que impossibilitam a sua identificação. Quando é apresentada à rede uma imagem maior que 1.700x1.700 pixels praticamente não é detectado nenhum fruto. Para imagens menores que 1.500x1.500 pixels a rede já começa a detectar frutos, sendo que a maioria dos frutos são detectados com dimensões de imagens próximas da configuração da rede.

Figura 6 – Porcentagem de detecção de frutos em função do tamanho da janela deslizante



Fonte: gerada pelos Autores

A Figura 7a mostra o resultado da detecção dos frutos utilizando uma janela deslizante de 416x416 pixels com uma sobreposição entres janelas de 300 pixels. Nesta árvore os frutos estão totalmente verdes e há uma maior quantidade de frutos na parte interna da árvore, como mostrado em detalhes na Figura 7b, dificultando a detecção devido à oclusão. Neste exemplo foram detectados 208 frutos, nenhum falso positivo e 21 falsos negativos.

Figura 7 – Resultado da detecção de frutos verdes (a); e frutos na parte interna da árvore (b)



(a)



(b)

CONCLUSÕES

Os resultados descritos mostram o alto potencial de uso da rede neural de aprendizado profundo aqui descrita para a identificação de frutos verdes em pés de laranja. Sabe-se entretanto, que a variedade do pé de laranja também tem influência no resultado geral da detecção, pois algumas espécies têm a folhagem mais intensa, aumentando a oclusão de frutos. Assim, pretende-se buscar uma melhora nos resultados aumentando o número de imagens nas bases de treino e teste da rede neural, contemplando diferentes variedades de citros. Será avaliado também o impacto de diferentes tamanhos de *bounding boxes* em novos testes de detecção em árvores inteiras.

AGRADECIMENTOS

Agradecemos aos coordenadores do PES/Fundecitrus pela disponibilização de fotos de árvores e à NVIDIA Corporation pela doação da placa GeForce usada no processamento e teste da rede neural.

REFERÊNCIAS

Instituto Brasileiro de Geografia e Estatística - IBGE. **Levantamento Sistemático da Produção Agrícola.** https://www.ibge.gov.br/estatisticas/economicas/agricultura-e-pecuaria/9201-levantamento-sistematico-da-producao-agricola.html?=&t=resultados/ Acesso em 04 de julho de 2018.

ALEXEY A. B. Yolo-v3 and Yolo-v2 for Windows and Linux. https://github.com/AlexeyAB/darknet#yolo-v3-in-other-frameworks. Acessado em 05/07/2019.

FERREIRA, A. S.; FREITAS, D.M.; SILVA, G. G.; PISTORI, H.; FOLHES; M.T. Weed detection in soybean crops using ConvNets. Comput. Electron. Agr, v. 143, p. 314-324, 2017.

KOIRALA, A., WALSH, K.B., WANG, Z. *et al*. Deep learning for real-time fruit detection and orchard fruit load estimation: benchmarking of 'MangoYOLO. Precision Agric (2019). https://doi.org/10.1007/s11119-019-09642-0

SA, I.; GE, Z.; DAYOUB, F.; UPCROFT, B.; PEREZ, T.; MCCOOL, C. Deepfruits: A fruit detection system using deep neural network. Sensors, v. 16, n. 8, 2016.

GHOSAL, S.; ZHENG, B.; CHAPMAN, S.; POTGIETERo, A.; JORDAN, D.; WANG, X.; SINGH, A.; SINGH, A.; HIRAFUJI, M.; NINOMIYA, S.; GANAPATHYSUBRAMANIAN, B.; SARKAR, S.; GUO, W. A Weakly Supervised Deep Learning Framework for Sorghum Head Detection and Counting. Plant Phenomics, v. 2019, 14p., 2019. https://doi.org/10.34133/2019/1525874.

NEVES, M. F.; TROMBIM, V. G. Anuário da citricultura 2017. 1. ed. SÃO PAULO CITRUSBR, 2017. 60 p.

REDMON, J. *et al*.You only look once: Unified, real-time object detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, Las Vegas. IEEE: p. 779-788, 2016.

REDMON, J.; FARHADI, A. YOLOv3: An incremental improvement. arXiv. Technical Report preprint arXiv:1804.02767, 2018.

REDMON, J. Darknet: Open Source Neural Networks in C. https://pjreddie.com/darknet/ Acessado em 05/07/2019.

TIAN, Y.; YANG, G.; WANG, Z.; WANG, H.; LI, E.; LIANG, Z. "Apple detection during different growth stages in orchards using the improved YOLO-V3 model," Computers and Electronics in Agriculture, vol. 157, pp. 417–426, 2019.

SOX: UMA PLATAFORMA PARA COLETA MASSIVA DE DADOS E O MONITORAMENTO CONTÍNUO DE VARIÁVEIS AMBIENTAIS

Leandro da Silva Camargo[1], *Érico Marcelo Hoff do Amaral*[1], *Leandro Bochi da Silva Volk*[1], *Marcel da Silva Camargo*[1]

[1]Programa de Pós-graduação em Computação Aplicada, Universidade Federal do Pampa e Embrapa Pecuária Sul, Bagé, Rio Grande do Sul, Brasil, leandrocamargo@ifsul.edu.br, erico.amaral@unipampa.edu.br, leandro.volk@embrapa.br, marcel@sigasistemas.com.br

RESUMO

A interação entre solo e planta é fundamental para produzir insumos que fazem parte da cadeia alimentar dos seres vivos. O conhecimento das variáveis ambientais envolvidas nessa interação e que causam impacto na produtividade sem que ocorra a extração, destruição e contaminação por químicos das amostras do solo coletado é desafiador. A adoção de sensores para prover a avaliação do solo pode ser considerada solução eficaz e conveniente frente às técnicas laboratoriais padrão e métodos químicos. Nesse sentido, o presente trabalho apresenta uma arquitetura composta por nó sensor que viabiliza a coleta dos dados de forma contínua, bem como uma estrutura computacional em nuvem que permite o suporte e a persistência de volumes massivos de dados, característicos desse tipo de sistema. A metodologia de construção ocorreu em quatro etapas, a primeira que identifica as variáveis relacionadas ao solo e a produtividade da área. A segunda referente a definição dos equipamentos disponíveis, passando para a coleta e validações. A terceira etapa consiste na definição da arquitetura e integração das camadas. A última trata da aplicação de uma bateria de testes de estresse, carga, funcionalidades e estabilidade do sistema. Esta solução demonstrou suportar acessos simultâneos com até 32 usuários, onde foram executadas 100.000 operações de gravação de documentos. A plataforma batizada de SOX pode ser acessível a muitos usuários, aplicável às condições da área produtiva, sensível às tensões ambientais e de manejo ao longo do tempo e, permite o armazenamento das séries históricas dos valores em estrutura de Banco de Dados digital.

PALAVRAS-CHAVE: Rede de sensores, Variáveis ambientais, Persistência dos dados.

ABSTRACT

The interaction between soil and plant are fundamental inputs for production, these resources make up the living things food chain. Environmental variables knowledge involved in process have a productivity impact, without extraction, destruction and chemical contamination, soil sample collected is challenging. Sensors use for soil evaluation is effective replacing way laboratory techniques chemicals use, because they present a low ambient impact and cost. The present work presents an electronic sensors network that allow data collection in a continuous way, as well a cloud computing which structure supports data massive volumes. To support the platform construction, an exploratory study was carried out bringing to light some functional and operational challenges imposed by agricultural scenario on technological solutions. The SOX platform construction methodology took place in four stages, the first one dealing with variables related identification to the soil and the productivity of the area. The second stage refers available equipment definition, passing for values collection and validations. The third step is architecture definition and integration of the layers. The last step deals with a battery applications of stress tests, load, functionalities and platform stability. The SOX has been shown to support simultaneous access with up to 32 users, where 100,000 document recording operations were performed. The SOX platform can be accessible to many users, applicable to productive area conditions, sensitive to environmental stresses and management over time, and allows values storage historical series in a digital database structure.

KEYWORDS: Network of sensors, Environmental data, Data persistence.

INTRODUÇÃO

O momento tecnológico atual é de uma grande proliferação de recursos de hardware e software, muitos dos quais sem possibilidades de integração de modo a satisfazer as demandas dos usuários. Outros, porém, possuem características adequadas, mas com limitações que dificultam seu uso. Numa aplicação de software, mais especificamente, o aspecto de real relevância é a capacidade de gerenciamento de dados, tanto na estruturação, armazenamento e pesquisa, como na visualização e distribuição.

Nesse sentido, devido aos avanços tecnológicos, bem como pelo aumento na demanda por alimentos, há uma forte pressão por parte da Organização das Nações Unidas para Agricultura e Alimentação (FAO), para que os agricultores utilizem ferramentas e tecnologias

inovadoras (FAO, 2014). “Os principais componentes da agricultura inteligente: o sensoriamento de dados, a comunicação, o armazenamento e o processamento, são essenciais para obter recursos preditivos robustos” (IVANOV; BHARGAVA; DONNELLY, 2015, p. 77). Tais atributos presentes na agricultura inteligente são potencializados com o uso de sensores e a estrutura de persistência propostos nesse trabalho.

Contudo, o uso desses sensores para a coleta das informações impõe desafios à Internet das coisas (*Internet of Things* - IoT), tais como a eficiência energética, a tolerância à intempérie e a disponibilidade dos dados. Conforme Wang (2011), outros tantos desafios relacionados à área de cobertura das redes e limitações na perspectiva do acesso aos dados. Nesse sentido, a construção de uma arquitetura baseada em sensores com a capacidade de oferecer o monitoramento permanente das condições ambientais e armazenamento dos valores capturados em uma estrutura de Banco de Dados se torna fundamental para apoiar a produção de alimentos.

Variáveis ambientais relacionadas ao solo

Atualmente para um aumento na produtividade o controle do solo e o ajuste às práticas convencionais de seu preparo têm se mostrado insuficientes para enfrentar os desafios da produtividade de forma adequada e sustentável evitando a degradação dos recursos naturais. Com relação às condições biológicas, segundo Kabiri *et al.* (2016), as atividades microbianas têm sido utilizadas para medir a influência das práticas de manejo do solo na sua qualidade. Essa atividade pode ser avaliada com o uso de sensores sensíveis à gases provenientes desses organismos.

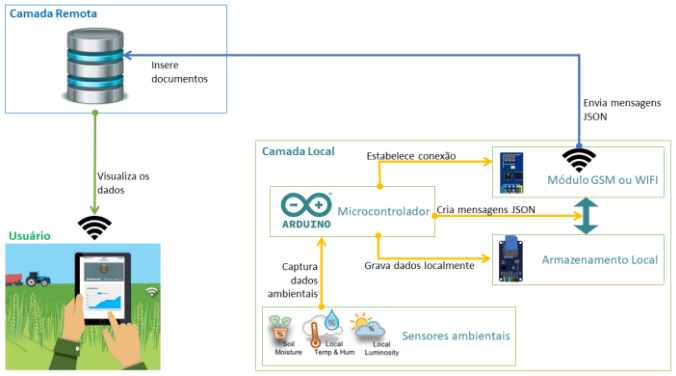
A qualidade do solo é afetada por sua estrutura, temperatura, umidade e aeração, essas condições impactam direta e indiretamente os serviços ecossistêmicos e consequentemente a produtividade (BARRIOS, 2007). Assim, a leitura contínua dessas variáveis pode permitir a análise da quantidade dos insumos, bem como a percepção dos efeitos da adubação e da rotação de culturas, avaliando o impacto sobre a densidade populacional dos microrganismos e as alterações percebidas nesse ambiente.

Esboço da plataforma SOX

A solução construída recebeu o nome de SOX (acrônimo para *soil observation and exploration*), está baseada em sensores que coletam valores ambientais, transmitem para uma estrutura de armazenamento dos dados e permite a visualização através de um navegador Web. A Figura 1 apresenta um esboço do funcionamento da plataforma SOX, onde é

estabelecida a conexão da camada local com a remota, o que ocorre através de uma conexão com a Internet, a qual pode acontecer através de uma rede WIFI ou GPRS, dependendo do sistema de comunicação implementado na solução.

Figura 1 - Esquema de funcionamento da plataforma SOX.



Fonte: Autor (2019)

Uma vez estabelecida conexão entre as camadas da plataforma, ocorre o envio dos documentos no formato de mensagens JSON. O documento é criado e contém os valores ambientais, esses estarão automaticamente disponíveis ao usuário da aplicação, que pode ser tanto o agricultor quanto o consultor técnico.

Trabalhos correlatos de soluções IoT voltadas à agricultura

As soluções de IoT devem apresentar como características o baixo custo, baixo consumo de energia dos dispositivos e, auxiliar no processo de tomada de decisão (MEKALA; VISWANATHAN, 2017, p.7). Além de oferecer um fácil entendimento do agricultor, mesmo que este detenha pouca experiência no uso de recursos tecnológicos.

Considerando tais características foram realizadas análise das arquiteturais de algumas soluções de IoT destinadas à agricultura. Na solução proposta por Salvi *et al.* (2017), embora ofereça uma interface para a visualização dos dados, tal solução utiliza uma aplicação *open source*, genérica, desenvolvida por terceiros, destinada à visualização de qualquer dado gerado a partir de sensores. Onde esses aplicativos demandam de instalação e, assim, exigem a utilização de um dispositivo computacional com sistema operacional Android, consumindo recursos e necessitando de *download* na loja de aplicativos. Outro limitante do projeto está a cargo da nuvem onde os dados são registrados, pois também é um serviço de terceiros, o qual exige a configuração e compatibilização com o aplicativo para que os valores sejam visualizados.

A proposta de Navarro-Hellín (2015) abrange a maior parte das características desejáveis em uma solução IoT aplicada ao agronegócio. Oferecendo uma interface com o

usuário da aplicação, auxilia na tomada de decisão, possui uma persistência local dos dados, de forma temporária e, a persistência remota, utilizando a transmissão através da rede GSM. Esse sistema possibilita a manipulação de sensores para a coleta de dados, bem como atuadores. As demais soluções utilizam rede de dados WiFi, ou seja, demandam de uma estrutura de comunicação local, contudo essa estrutura garante uma comunicação contínua dos dados, sem a preocupação com o volume de dados transmitidos e, também, com o consumo energético demandado nessa transmissão. Ainda, é possível observar que estas soluções implementadas com a rede de dados WiFi não possuem uma camada de persistência local dos dados coletados.

Segundo Adetunji e Joseph (2018) foi implementada uma abordagem compacta, usando Arduino e sensores relevantes para obter transferência de dados para a plataforma em nuvem. O modelo proposto é capaz de coletar os dados referentes a umidade do solo (porcentagem de umidade volumétrica), a umidade, a temperatura ambiente, o ponto de orvalho e a temperatura do solo.

O modelo de monitoramento ambiental proposto por Jha *et al*. (2017), apresenta uma ferramenta para o monitoramento da umidade do solo e temperatura usando dispositivos IoT. Para tal, fora utilizada uma placa Arduino com sensores de solo, temperatura e umidade, os quais suportam a coleta remotamente dos dados no campo. Uma vez recebidos, os dados são analisados e discutidos. Este trabalho viabiliza as ações preventivas para reduzir as perdas de culturas e consequentemente aumentar a produtividade.

A Tabela 1 agrupa as características arquiteturais dos trabalhos correlatos elencados, apontando a presença ou não dos atributos. As características abordadas são referentes: à visualização dos dados em uma interface com o usuário; ao apoio na tomada de decisão; ao tipo de banco de dados oferecido pela aplicação; ao padrão de armazenamento dos dados, estruturado, não estruturado ou orientado à documentos; à forma de comunicação dos dados; ao modelo dos sensores utilizados, atuadores ou de coleta dos parâmetros ambientais.

Tabela 1 – Comparativo das características presentes nos trabalhos selecionados.

Referências	Interface de visualização dos dados	Apoio à Tomada de Decisão	Padrão e Banco de Dados[1]	Comunicação de Dados[2]
JHA *et al.*, 2017	Não	Sim	E - Sim	C – WiFi
SALVI *et al.*, 2017	Não	Sim	N - Não	C – WiFi
ADETUNJI; JOSEPH, 2018	Sim	Não	E - Sim	C - WiFi
NAVARRO-HELLÍN, 2015	Sim	Sim	N - Não	I - GSM
Plataforma SOX	Sim	Sim	D - Sim	I - GSM

[1]Padrão do Banco: D – orientado a documentos (NoSQL); E – estruturado (SQL); N – não se aplica.
[2]Transmissão dos Dados: C – contínua; I – intermitente.

A plataforma SOX, proposta neste trabalho, oferece uma estrutura que abrange as características descritas no trabalho de Mekala e Viswanathan (2017), as quais são desejáveis em uma ferramenta de IoT, com recursos em *cloud computing*, destinada à agricultura

inteligente. Oferecendo uma interface acessível a qualquer dispositivo computacional e auxiliando à tomada de decisão, através dos dados armazenados de forma organizada, em uma base de dados.

Esses requisitos apresentados garantem a coleta e visualização dos dados ambientais, praticamente, em tempo de execução. Mas o grande diferencial frente aos demais trabalhos analisados está a cargo do sistema gerenciador de banco de dados orientado a documentos, tecnologia que permite a integração desta solução de coleta de parâmetros ambientais com imageamento de áreas produtivas georreferenciadas, bem como, com a tolerância a alta variabilidade de sensores, inclusive, suportando os desligamentos temporários sem causar distorções na fase de análise dos dados.

Por fim, os trabalhos correlatos que compõem esta revisão corroboram com a importância da temática apresentada nesse trabalho, bem como reforçam o potencial da plataforma, pois pode contribuir com os pesquisadores na aplicação do conceito da agricultura inteligente e, consequentemente, na construção de uma volumosa base de dados que futuramente sirva às pesquisas de mineração de dados, estatística e predição.

MATERIAL E MÉTODOS

Para a elaboração desse trabalho foi realizada uma pesquisa exploratória para identificar os principais sensores com potencial para a coleta dos dados ambientais. As atividades envolvidas no processo de pesquisa foram divididas em quatro etapas.

Na primeira etapa foi iniciado o trabalho de pesquisa bibliográfica avaliando algumas metodologias para análise dos indicadores da qualidade do solo, identificando suas potencialidades e limitações, bem como, o estudo dos sensores que suportem a coleta de parâmetros do solo *in situ*, sujeitos às intempéries na área produtiva.

Na segunda etapa da pesquisa, foram realizados testes com um conjunto de sensores para observar a relação entre as variáveis coletadas, tais como umidade, condutividade, compactação, nutrientes e gás carbônico presentes no solo. Com esse volume inicial de dados é possível identificar relações entre as variáveis coletadas.

A terceira etapa refere-se à construção da plataforma SOX, composta pelo conjunto de sensores e técnicas para leitura, transmissão, pré-processamento e armazenamento. O desafio dessa etapa consiste em integrar as tecnologias automatizando ao máximo todo o processo envolvido da coleta à visualização dos dados, dirimindo eventuais manutenções nos sensores ou obtenção das leituras realizadas no ambiente.

A quarta etapa avalia a capacidade da plataforma SOX em sustentar a carga e o estresse gerado pelo volume de acessos concorrentes, bem como a escalabilidade oferecida pela plataforma para suportar a cobertura da área produtiva. Para avaliação dos resultados será utilizada uma ferramenta de *benchmark*, confiável e que permita simulações dos valores, dos tipos de operações e da estrutura dos dados, as quais são equivalentes as geradas pelos sensores em um ambiente real.

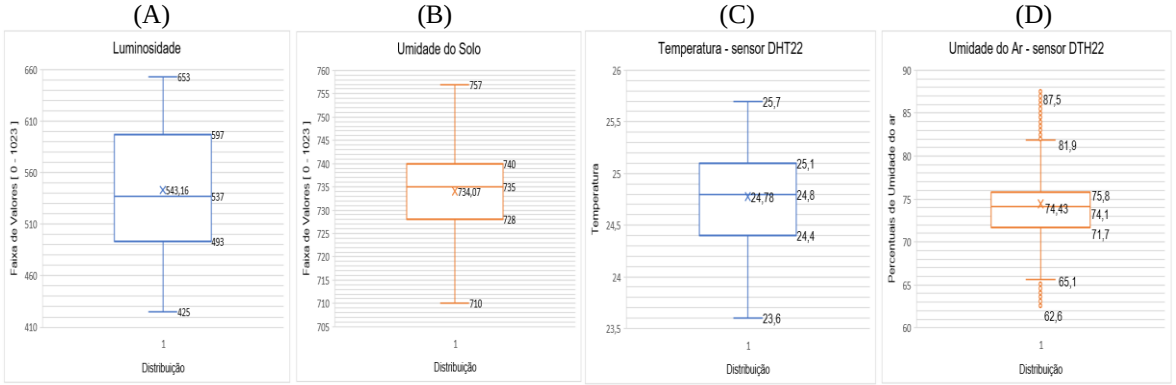
RESULTADOS E DISCUSSÃO

Para avaliar a estabilidade e desempenho da plataforma SOX, os seus componentes passaram por um conjunto de testes de conjuntos isolados. Onde se criou grupos conforme o nível de dependência dos dispositivos e, tais conjuntos passaram por cargas de trabalho específicas, com características similares ao ambiente real.

Teste de Estabilidade dos sensores

Para avaliar a estabilidade dos sensores, foi realizada uma leitura por minuto, durante vinte e quatro horas do dia, no período de um mês. A distribuição desses valores é apresentada na Figura 2, a qual demonstra uma estabilidade bastante importante dos sensores, mesmo quando expostos a um regime constante de leituras.

Figura 2 - Distribuição dos valores de luminosidade (A), umidade do solo (B), temperatura (C) e umidade do ar (D), resultantes de 43.000 amostras, obtidas durante um mês de leituras em um ambiente controlado.



Fonte: Autor (2019)

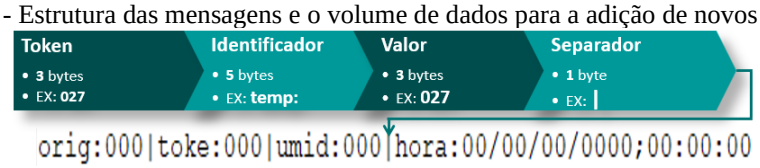
Os valores não apresentam uma flutuação muito significativa, tampouco a presença de *outliers*, demonstrando a estabilidade do módulo local da plataforma SOX. Pois no período de testes laboratoriais nenhuma falha de coleta ou desgaste do conjunto foi observada.

Teste de estrutura das mensagens e comunicação dos dados

O tamanho mínimo de uma mensagem gerada pelo sistema é de 54 Bytes, pois basicamente, um registro possui a identificação da origem, um token (uma cópia de dos valores coletados),

a data e hora da gravação do registro, adicionado ainda, os separadores de campos, nesse caso, um *pipeline* (|). A Figura 3 exibe o *layout* de mensagem.

Figura 3 - Estrutura das mensagens e o volume de dados para a adição de novos sensores.



Fonte: Autor (2019)

O valor de cada campo possui tamanho padronizado de 3 Bytes e outros 5 Bytes da respectiva identificação da chave. Para cada novo sensor adicionado ao sistema a mensagem é incrementada em 12 Bytes, sendo 8 bytes conforme já descrito anteriormente, acrescidos de 3 Bytes que integram o campo de token e 1 Byte do separador desse novo campo.

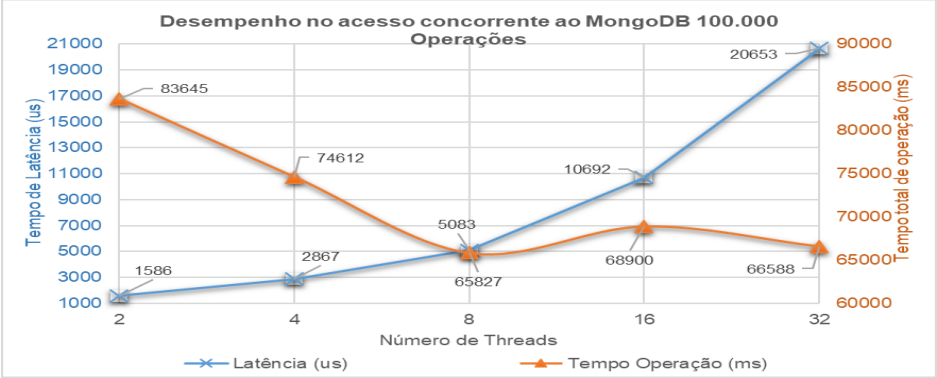
Considerando uma memória com capacidade de armazenamento de 8 GB (8.589 Bytes), a camada local pode oferecer uma autonomia para armazenar o volume de dados gerados por quatro sensores, com uma coleta por hora, pelo período de 36 meses. Ressaltando que o armazenamento local dos dados é provisório, pois a persistência ocorre na camada remota.

Teste de acesso concorrente

Para verificar a capacidade da plataforma SOX em suportar acessos concorrentes ao banco de dados MongoDB, o qual armazenará um conjunto de dados, foi utilizada a ferramenta de *benchmark* YCSB, que permite a análise experimental e padronizada de desempenho baseado em métricas e cargas de trabalho similar ao cenário real.

A Figura 5 apresenta o cruzamento dos valores de tempo de latência, ou seja, do tempo médio que corresponde a requisição e a completude na resposta da operação e, o gráfico apresenta ainda, o tempo total para a realização das operações. Contudo é necessário destacar que o tempo de latência está expresso em microssegundos (µs), já os tempos totais para a realização das operações são colocados em milissegundos (ms). Estes valores estão apresentados em função dos *threads* que estão na base do gráfico (eixo x).

Figura 5 - Avaliação do desempenho do MongoDB com acessos concorrentes à base de dados.



Fonte: Autor (2019)

Conforme evidenciado pela Figura 5, o tempo total para a execução das 100.000 operações é mais elevado quando executados em 2 ou 4 *threads*, por outro lado, entre 8, 16 e 32 *threads*, o tempo total para a conclusão desse mesmo volume de operações apresenta uma variação de apenas 3.000 milissegundos, ou seja, uma diferença de 3 segundos. Já a latência aumenta exponencialmente à medida que são inseridas novas *threads*, com uma amplitude de valor entre 1.500 e 20.000 microssegundos, ou seja, o valor máximo observado é de 0,02 segundos. Demonstrando que a camada remota mantém o desempenho, mesmo com o acesso concorrente e a manipulação de 100.000 documentos, cada qual equivalente a 100 sensores.

CONCLUSÕES

A plataforma SOX, possibilita a coleta, transmissão e o armazenamento de um volume massivo de parâmetros ambientais, de forma contínua no tempo e no espaço, mesmo em um cenário com baixa disponibilidade de recursos para a comunicação dos valores observados. Suportando a escalabilidade, pois durante os testes em laboratório foram simulados os acessos com até 32 usuários de forma simultânea, onde executaram 100.000 operações, cada qual produziu um volume de dados equivalentes a gerada por 100 sensores. Ratificando possibilidade de distribuição de uma rede de sensores espacializados na área produtiva de uma cultura, conforme a recomendação do consultor técnico.

A arquitetura sob a qual a plataforma SOX foi concebida permite seu uso em diferentes cenários relacionados a infraestrutura de rede de dados, inclusive na ausência total de recursos garantindo uma persistência local dos dados com autonomia compatível com o ciclo de desenvolvimento dos cultivares.

REFERÊNCIAS

ADETUNJI, K. E.; JOSEPH, M. K. Development of a cloud-based monitoring system using 4duino: applications in agriculture. In: INTERNATIONAL CONFERENCE ON ADVANCES IN BIG DATA, COMPUTING AND DATA COMMUNICATION SYSTEMS (ICABCD), 2018, Durban, South Africa. Anais, IEEE, [S.l]. ago. 2018. p. 4849-4854.

BARRIOS, Edmundo. Soil biota, ecosystem services and land productivity. Ecological Economics, v. 64, n. 2, p. 269-285. 2007.

FAO - Food and Agriculture Organization. Walking the Nexus Talk: Assessing the Water-Energy-Food Nexus in the Context of the Sustainable Energy for All Initiative. Food and Agriculture Organization of the United Nations. Rome, Italy. p.150. 2014.

IVANOV, S.; BHARGAVA, K.; DONNELLY, W. Precision Farming: Sensor Analytics. Ieee Intelligent Systems, Institute of Electrical and Electronics Engineers (IEEE), USA. v. 30, n. 4, p. 76-80. 2015.

JHA, R. K.; KUMAR, S.; JOSHI, K.; PANDEY, R. Field monitoring using IoT in agriculture. In: INTERNATIONAL CONFERENCE ON INTELLIGENT COMPUTING, INSTRUMENTATION AND CONTROL TECHNOLOGIES (ICICICT), 2017, Thalassery, India, Anais, IEEE, USA., jul. 2017. p. 1417-1420.

KABIRI, V.; RAIESI, F.; GHAZAVI, M. A. Tillage effects on soil microbial biomass, SOM mineralization and enzyme activity in a semi-arid Calcixerepts. Agriculture, Ecosystems & Environment, Elsevier BV. v. 232, p. 73-84, set. 2016.

MEKALA, M. S.; VISWANATHAN, P. A Survey: Smart agriculture IoT with cloud computing. In: INTERNATIONAL CONFERENCE ON MICROELECTRONIC DEVICES, CIRCUITS AND SYSTEMS (ICMDCS), 2017, Vellore, India, Anais. IEEE, [S.l.], ago. 2017. p. 1-8.

NAVARRO-HELLÍN, H.; TORRES-SÁNCHEZ, R.; SOTO-VALLES, F.; et al. A wireless sensors architecture for efficient irrigation water management. Agricultural Water Management, Elsevier, [S.l.], v. 151, p.64-74, mar. 2015.

SALVI, S.; FRAMED, J. S. A.; SANJAY, H. A.; HARSHITA, T. K.; FARHANA, M.; JAIN, N.; SUHAS, M. V. Cloud based data analysis and monitoring of smart multi-level irrigation system using IoT. In: INTERNATIONAL CONFERENCE ON I-SMAC (IOT IN SOCIAL, MOBILE, ANALYTICS AND CLOUD) (I-SMAC), 2017, Palladam, India, Anais, IEEE, [S.l.], fev. 2017. p. 752-757.

WANG, B. Coverage problems in sensor networks. Acm Computing Surveys, Association for Computing Machinery (ACM), v. 43, n. 4, p. 1-53, 2011.



Analysis of a ZigBee Wireless Sensor Network Based on IEEE 1451 Standard

Tércio A. S. Filho[1], *Juliano Mendes Ribeiro*[4], *Alexandre C. R. Silva*[4], *Sérgio F. Silva*[2], *Antonio Carlos de Oliveira Júnior*[2,3]

[1]Instituto de Biotecnologia
Universidade Federal de Catalão (UFCAT)
tercioas@ufg.br, sergio.f.silva@gmail.com

[2]Instituto de Informática
Universidade Federal de Goiás (UFG)
antonio@inf.ufg.br

[3]Fraunhofer AICOS Portugal
antonio.junior@fraunhofer.pt

[4]Universidade Estadual Paulista - "UNESP"
Programa de Pós-Graduação em Engenharia Elétrica - PPGEE
juliano.mendes.r@gmail.com,acrsilva@dee.feis.unesp.br

ABSTRACT

In this paper, a mesh type wireless sensor network based on the IEEE 1451 standard is presented to obtain meteorological data. The network consists of a single NCAP node and multiple WTIM nodes. The NCAP node is responsible for managing the network, performing plug and play recognition of WTIM nodes, sending request commands, and receiving data. A WTIM node consists of one or more transducers, a data transfer interface and a processing unit. Thus, with the developed network, it was possible to collect environmental data at different points within a city landscape, to perform analysis of the communication distance between the WTIM nodes and to monitor the number of bytes transferred according to each network node.

KEYWORDS: Wireless sensor networks, IEEE 1451, ZigBee.

INTRODUCTION

The use of technology for environmental monitoring has become an important tool in modern society such as for management of public health and inspection of flora. In addition, such a tool allows studies focusing on climate and micro-climate analyses, and weather forecasting. These studies can provide essential knowledge for a better management of resources and planning in a specific region. Possible areas that can benefit of such knowledge include agriculture, livestock, fish farming, apiculture, among others.

The environment consists of the interaction between physical, biotic and anthropic means. As this interaction is dynamic, environmental characteristics tend to change naturally over time. The dynamics of the environment is expressed by processes or factors of the environment that need to be monitored, aimed to take control measures when these factors can cause damage to life [Gore e Deshpande 2017]. For example, the introduction of gases, vapors and particulate matter from industrial activities in the atmospheric air can change significantly the quality of the air and may cause from discomforts to diseases in humans.

Sensor networks have been applied in applications from environmental [Dan et al. 2017] to buildings [Sánchez-Rosario et al. 2015], machines [Eren 2011], and even to health analysis [Sali e Parvathi 2017]. These analyses have became feasible by the emergence of sensor network technologies that allow periodic monitoring over a long period of time, enabling the understanding of complex phenomena, planning and diagnoses, which are very expensive to obtain without a complete automation of data acquisition. Sensor networks consist of interconnected nodes that communicate through a suitable interface and are essentially distributed systems for the purpose of automatically supervising and monitoring a given process [Erazo et al. 2015]. These networks can be developed using various types of technology, such as the field programmable gate array (FPGA), microcontroller, programmable logic controller (PLC) and microcomputer. A promising network technology with such features is the mesh sensor network.

A wireless mesh sensor network (WMSN) can be defined as a network of devices, capable of measuring aspects of the environment and disseminating the information collected through the network. The data are transmitted through the network nodes, possibly via multiple jumps until a master node, using a wireless interface. The data collected by a sensor network can be heterogeneous and distributed. In addition, the data should be collected over a long period of time and automatically sent to a computer storage infrastructure usually forming large repositories of data. As desirable features, a WMSN should be fault tolerant, scalable, low cost, have a low power consumption, be programmable, and be secure regarding aspects such as access control, message integrity, confidentiality, and replay protection.

In this paper, the design and implementation of a sensor network based on the Institute of Electrical and Electronics Engineers (IEEE) 1451 standard is presented that has capabilities of the aforementioned features. The network consists of two types of node, the network capable application processor (NCAP) and transducer interface module (TIM). The TIM is a module

that has a signal conditioning circuit, A / D (Analog / Digital) and D / A (Digital / Analog) converters, non-volatile storage for Transducer Eletronic Data Sheet (TEDS) and logic required to perform communication with the NCAP block. When developed based on the IEEE 1451.5 subcommittee it is defined as WTIM (Wireless Transducer Interface Module) where communication is done wirelessly between NCAP and WTIM. The NCAP is a local processing capability capable of receiving either high or low abstraction data through an external or local network, processing that data and sending it to the TIMs through a standardized interface based on one of the subcommittee of the IEEE 1451 standard [Adamo et al. 2015].

The network was implemented in a city to test in the field and to verify the quality of the link between the network nodes relating to the distance between nodes, packet loss and signal strength of each node. The results obtained show that the nodes can communicate in a range up to 3.2 km without obstacles between them, with the rate of packet loss between the nodes around 50%, which is a acceptable rate considering the features of a Zigbee network.

The developed network can be used to collect large amounts of meteorological data such as, temperature, wind direction, humidity, rainfall, and storing the collected data in a memory card connected to NCAP. The data collected over a long period of time using the sensor network can generate a theoretical basis necessary to perform data mining and data analysis, which can be used to extract hidden patterns embedded in the data. An example of such an application is to systematically monitor atmospheric conditions to infer the effect of air pollution spreading.

The remainder of this paper is organised as following: Section 1 presents the node design and implementation. Section 2 presents the tests performed with the network that were aimed to verify communication performance through mathematical analysis and the Section 3 presents the conclusion.

1 IEEE 1451: ARCHITECTURE OF NCAP AND WTIM

The IEEE 1451 standard defines two nodes within a wireless sensor network, the NCAP node and the WTIM node, and a set of common communication interfaces for connecting them. The IEEE 1451 standard is divided in seven different committees [Kumar et al. 2015]. In this paper, the wireless nodes were developed based on: IEEE Std 1451.0: describes common functions, protocols, and TEDS; IEEE Std 1451.1: which specifies the NCAP; IEEE Std 1451.5: specifies wireless communication protocols, being: wireless fidelity (WiFi), Zigbee or Bluetooth.

The NCAP network node was divided into different layers. Layer 0 presents the hardware and its interfaces, in which the Ethernet network was used to communicate with the external network and a Zigbee module connected to the Tinker Board. At Level 1, the Debian Linux operating system interfaces with hardware and applications, Level 2, applications for the development of the protocol manager, and Level 3, the protocol manager. Section 1.1 presents NCAP and Section 1.2 the WTIM, focusing on their features, connection and communication among the devices.

1.1 Network Capable Application Processor - NCAP

The NCAP was developed for the purpose of creating an expandable, autonomous network that can perform data collection for a long period of time. Its functionality includes plug-and-play discovery of the WTIM nodes, the management of sensor readings through the autonomous communication between the WTIMs at predefined time intervals, and the storage of the data or the data transmission to a external network. The NCAP can communicate with one or more modules using the identification device (ID) and the standard IEEE 1451.5 (Zigbee), and send the data to external network.

Thus, a packet can be routed among several nodes in the network until it reaches the target device. The WTIMs have a logical functionality in which the objective is to receive frames from the NCAP network node, check if the frame corresponds to WTIM, and if it is, the frame is decoded and the action required is performed such as reading a specific sensor. If it is not, the frame is forwarded to the next WTIM.

The NCAP can also be connected to a user network through of the Ethernet external interface. Therefore, the development of the NCAP was divided in two levels: hardware/interfaces and software. Section 1.1.1 presents the configuration of hardware and interfaces, and in Section 1.1.2 the software is described.

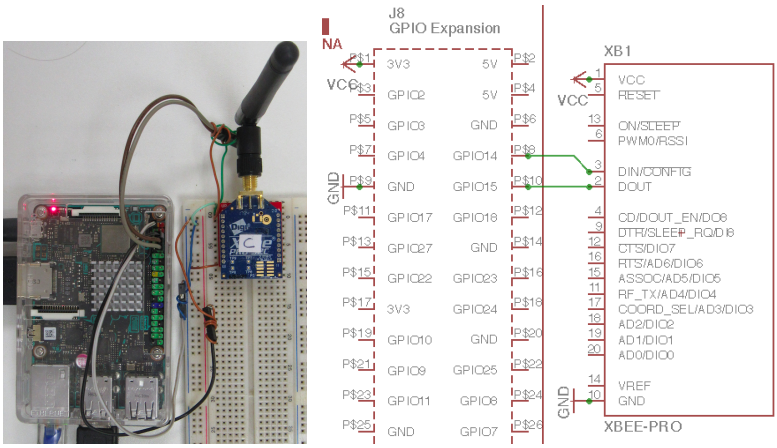
1.1.1 Hardware and Interfaces

The NCAP was developed using the *Tinker Board* manufactured by ASUS, that has a quad−core ARM-based processor, the Rockchip RK3288 operating at 1.8 GHz, 2 GB of random access memory (RAM), and the following interfaces: Ethernet, Bluetooth, high-definition multimedia interface (HDMI) output, 4 universal serial bus (USB) ports, a graphics processing unit (GPU) and 40 pins, configured as 28 general purpose input/output (GPIO) pins (shared with serial peripheral interface (SPI)/UART/ inter-integrated circuit I2C pins). The transceiver module used was the XBee Pro S2C, manufactured by Digi International, having 2.4 GHz of frequency band, transmission rate of 250 Kb/s and can reaches a range up to 3200 m of communication distance according to the proprietary specification. Figure 1(a) presents the NCAP node developed and Figure 1(b) the circuit.

The Tinker Board and the transceiver module communicate through UART interfaces and work with different voltage levels; the Tinker Board works with 5 V and the transceiver module at 3.3 V. Thus, the signals sent through the Tinker Board were conditioned using a resistive voltage divider circuit.

1.1.2 NCAP Software

At high abstraction level several technologies are used in developing transducer network nodes. In [Rasyid et al. 2016] was presented the development of a semantic sensor web (SSW), which was a combination of sensor networks, web services, database and semantic web technologies



(a) (b)

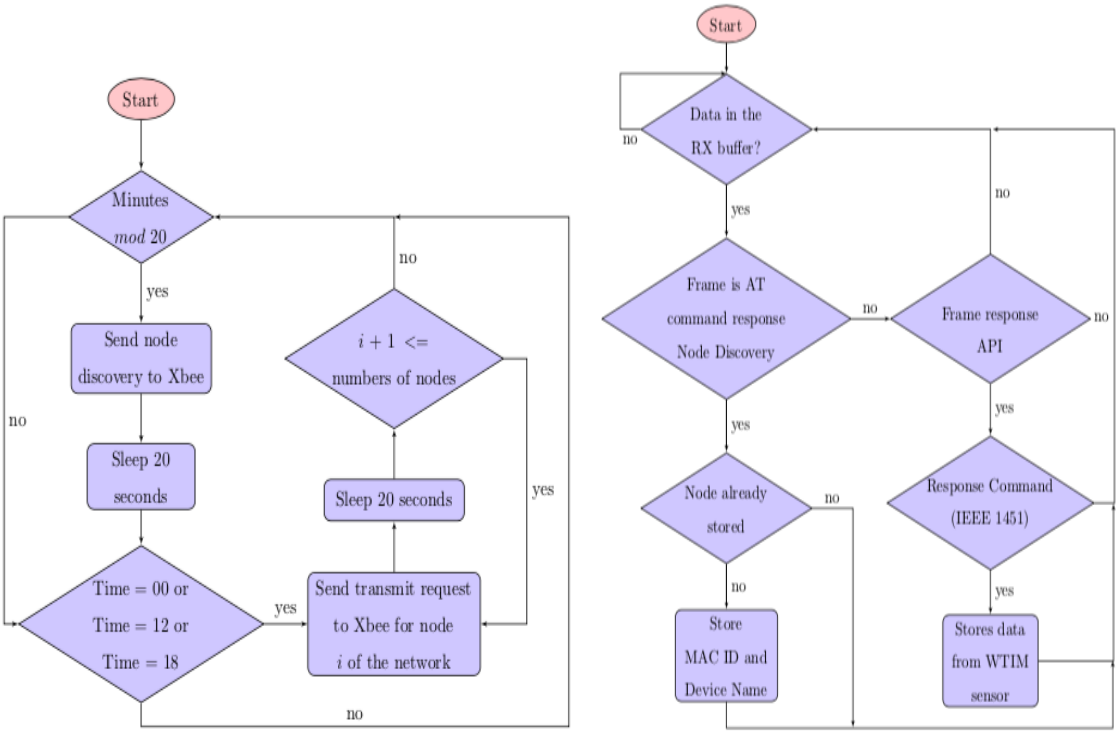
Figura 1: a: NCAP node developed in laboratory. b: NCAP circuit diagram.

for monitoring environmental conditions. The sensors were distributed on average by gas, temperature and humidity. The data were collected in the environment, stored in a database, and made available to the software through an application programming interface (API). In order to provide the data obtained by the API, a semantic web was used, which created the data interoperability in a system such as Resource Description Framework (RDF) which is a language used to process metadata and protocol and RDF query language (SPARQL) which is an integration class between RDF and hypertext preprocessor (PHP).

In this paper, for the development of the NCAP, the TinkerOS Debian operating system was installed, which is a Debian version for the device, available on the manufacturer's website. Next, a set of tools was installed: Apache server, PHP 7 and structured query language (MySQL) server database.

The management system (SiG-1451) was developed for the web environment, using the client-server model with three layers from the model-view-controller architecture. This system architecture together with the CodeIgniter PHP framework were adopted by providing an organization and modularization of the system through the separation of operations between model, view and controller.

When a new WTIM node is inserted in the mesh sensor network, the coordinator sends a discovery node command. In response to a command, the description of the WTIM, called TEDS, are transmitted to the NCAP and stored in file. At the end of the TEDS transmission, the NCAP will read the file and process the information, making it available via a web application. After a WTIM discovery command, the NCAP can send commands for data request to WTIMs based on MAC ID obtained through command discovery and stored in file. When NCAP receives the response, data relating to the sensor connected to the WTIM is stored in a file. As data acquisition is requested by the NCAP, the user can specify the time and periodicity of the data request for any WTIM. Figure 2(a) presents the flow chart of NCAP that represents transmission/request software and Figure 2(b) presents the software to obtained the MAC ID.



(a) (b)

Figura 2: a: Flow chart of NCAP that represents transmission/request software. b: software to obtained the MAC ID.

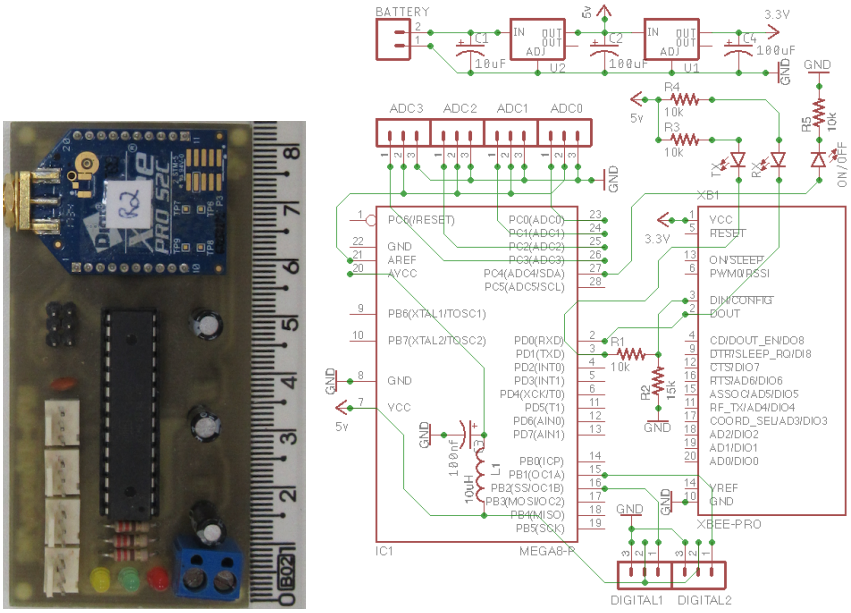
1.2 Wireless Transducer Interface Module - WTIM

The WTIM nodes were developed using sensors, signal conditioning circuits, a wireless communication interface and a microcontroller. The microcontroller used was Atmel Atmega8 based on the reduced instruction set computer (RISC) architecture, running 16 million instructions per second (MIPS) with clock freqeuncy of 16 MHz (external clock), with 8 kB of flash memory and 1 kB of RAM. The transceiver module used was the XBee PRO S2C radio based on the Zigbee standard. Figure 3(a) shows the WTIM node on its printed circuit board (PCB) and Figure 3(b) presents the schematic of the developed electronic circuit.

The WTIM firmware was developed with the use of embedded C language, which was programmed to support for 6 channels, being 4 sensors connected through A/D (analog-to-digital) converter inputs and 2 sensors connected through digital ports. The data acquisition and communication functions are controlled by the NCAP network node.

The WTIM firmware is illustrated by the flowchart in Figure 4. When the WTIM is started, the UART interface, the interrupts and the microcontroller channels are initialized. The communication and data acquisition functions wait for the NCAP request. The request and response commands are based on IEEE 1451 frames. The IEEE 1451 command are encapsulated in Zigbee frames in the data field and transmitted over the network.

The NCAP sends a request command to read a sensor channel from a WTIM and the



(a) (b)

Figura 3: a: WTIM node developed in laboratory. b: WTIM circuit diagram.

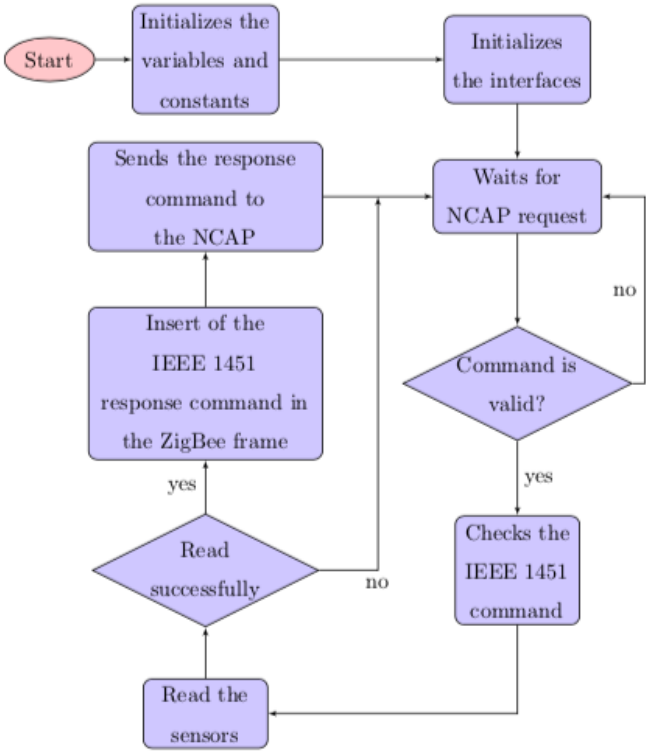


Figura 4: WTIM developed in laboratory.

command based on the IEEE 1451 standard is encapsulated in a Zigbee frame. The packet is decapsulated upon arriving at its destination WTIM and the sensor channel to be read is identified. The sensor data is placed into the reply command based on IEEE 1451 standard, encapsulated in a Zigbee frame, and transmitted to the NCAP.

Figure 3(a) presents the laboratory developed WTIM module for field data acquisition. The sensor used to data acquisition was BMP280 sensor that measures pressure, temperature and humidity. The characteristics of the BMP280 sensor are: supply voltage of 1.8 V and 5V DC, I2C interface (up to 3.4 MHz), SPI (up to 10 MHz).

2 ANALYSIS AND TESTS OF COMMUNICATION BETWEEN NCAP AND WTIM

The data collected by each WTIM are stored in a file according to each sensor in the module as shown in Table 1. The files are named according to the ID address of the transmitter module.

ID	Date	Hour	Temperature	Humidity	Pressure
1	2017-08-25	0	22.5	52	922.2
2	2017-08-25	12	21.5	53	924.2
3	2017-08-26	18	28.3	35	920.4
4	2017-08-26	0	22.6	48	922.4
5	2017-08-26	12	22.6	46	923.4

Tabela 1: Sample of the data captured.

Aiming to verify the communication range between nodes, a mesh network routing test was performed in which the NCAP and the WTIM nodes were positioned according to Figure 5. An important aspect is that each node allows direct visualization of one or more adjacent nodes, so there was no physical barrier to prevent communication.



Figura 5: Distribution of the WTIMs in the city.

In Figure 5, the edges indicate pairs of nodes with an established communication link, and the numbers on the edges corresponds to link quality of each node, being the number on the left the signal quality of the left (or upper) side node. The distance between the nodes is shown in Figure 5, where: NCAP for WTIM-3 (2.4 Km); NCAP to WTIM-2 (1.5 Km); NCAP for WTIM-1 (1.11 Km); WTIM-1 to WTIM-3 (3.15 Km) and WTIM-2 to WTIM-3 (2.4 Km). Based on each node, the NCAP and WTIMs had the follow coordinates: NCAP - Latitude: 18°9'27.89"S - Longitude: 47°55'13.06"W; WTIM-1 - Latitude: 18°10'0.73"S - Longitude: 47°54'56.97"W; WTIM-2 - Latitude: 18°10'10.93"S - Longitude: 47°55'37.49"W; WTIM-3

- Latitude: 18° 9’19.22"S - Longitude: 47°56’33.60"W. Link quality is given by link quality indication (LQI), which is a number between 0 and 255, where 0 is the lowest quality indicator and 255 is the highest quality indicator. As can be noted, WTIM-1 and WTIM-3 established a low quality link shown by the yellow edge in Figure 5. The rede edge indicates that there is not a two-away communication link between the NCAP and WTIM-3. The green line in Figure 5 represents success of communication between nodes. Table 2 presents measurements relating to the NCAP communication with each node in the network, “Sent” being the amount of packages sent by NCAP to each node, “Tx error” the amount of errors in the packages transmission, “Rx error” the amount of errors in the packages receiving, “Received” the amount of packages correctly received by the corresponding WTIM, and “Success” summarizes the percentage of packages correctly received.

	Tx error	Rx error	Sent	Received	Success	Elevation NCAP/WTIM
NCAP to WTIM-1	215	232	1000	553	55,3%	880 / 935 meters
NCAP to WTIM-2	260	185	1000	555	55,5%	880 / 868 meters
NCAP to WTIM-1	260	350	1000	390	39%	880 / 893 meters
Average	245	255,7	1000	499,3	49,9%	

Tabela 2: Network performance management.

As can be noted, the most effective communication was from NCAP to WTIM-2, followed closely by NCAP to WTIM-1. Communication from NCAP to WTIM-3 resulted in the lowest success rate that can be attributed to the multi-hop communication as the NCAP did not have direct communication with WTIM-3, as can bee seen in Figure 5. Another interesting aspect is the fact that, although the NCAP communication for both WTIM 1 and 2 have approximately the same success rate, the link NCAP to WTIM-1 had a smaller packets loss in the transmission whereas the link NCAP to WTIM2 had a smaller packets loss when receiving.

3 CONCLUSIONS

In this paper was presented a mesh sensor network based on the IEEE 1451 standard, where the NCAP node in a plug-and-play manner detects the WTIM nodes, send requests to read the WTIM’s sensors and routes the sensor data back in the network to the NCAP node that stores the data. All the nodes developed have low cost and can communicate over long distances, with an acceptable packets loss. Based on these characteristics, a mesh network can be easily expanded to perform a data acquisition in large areas without modifications in the network. The tests performed proved that the network is feasible for data collection purposes. In this paper the NCAP was implemented using a structured programming in a microcontroller to be

autonomous, extensible, and the mesh network was evaluated in a real word scenario, analyzing aspects of communication between the nodes.

ACKNOWLEDGMENTS

The present work was carried out with the support of the Coordination of Improvement of Higher Education Personnel - Brazil (CAPES) and National Council for Scientific and Technological Development (CNPq) - Processes: 431552 / 2016-9, 447984 / 2014-4 and 309193 / 2015-0, (PQ Grant).

REFERÊNCIAS

ADAMO, F. et al. A smart sensor network for sea water quality monitoring. *IEEE Sensors Journal*, v. 15, n. 5, p. 2514–2522, May 2015. ISSN 1530-437X.

DAN, L. et al. Intelligent agriculture greenhouse environment monitoring system based on the android platform. In: *2017 International Conference on Smart Grid and Electrical Automation (ICSGEA)*. Changsha, China: [s.n.], 2017. p. 358–361.

ERAZO, M. et al. Design and implementation of a wireless sensor network for rose greenhouses monitoring. In: *6th International Conference on Automation, Robotics and Applications (ICARA)*. Queenstown, New Zealand: [s.n.], 2015. p. 256–261.

EREN, H. Energy conscious application of zigbee wireless networks in machine health monitoring systems. In: *2011 IEEE Sensors Applications Symposium*. San Antonio, Texas, USA: [s.n.], 2011. p. 88–92.

GORE, R. W.; DESHPANDE, D. S. An approach for classification of health risks based on air quality levels. In: *2017 1st International Conference on Intelligent Systems and Information Management (ICISIM)*. [S.l.: s.n.], 2017. p. 58–61.

KUMAR, A. et al. Current status of the ieee 1451 standard-based sensor applications. *IEEE Sensors Journal*, v. 15, n. 5, p. 2505–2513, 2015.

RASYID, M. U. H. A. et al. Development of semantic sensor web for monitoring environment conditions. In: *2016 International Seminar on Intelligent Technology and Its Applications (ISITIA)*. [S.l.: s.n.], 2016. p. 607–612.

SALI, S.; PARVATHI, C. S. Integrated wireless instrument for heart rate and body temperature measurement. In: *2nd International Conference for Convergence in Technology (I2CT)*. Maharashtra, India: [s.n.], 2017. p. 457–463.

SÁNCHEZ-ROSARIO, F. et al. A low consumption real time environmental monitoring system for smart cities based on zigbee wireless sensor network. In: *2015 International Wireless Communications and Mobile Computing Conference (IWCMC)*. [S.l.: s.n.], 2015. p. 702–707.

Segmentação de Imagem Obtida por RPA Asa Fixa na Detecção de Erosão em Área Agrícola

Yasmin Volmer Tullio[1], Ricardo Kwiatkowski da Silva[1], Selma Regina Aranha Ribeiro[1], Maria Salete Marcon Gomes Vaz[1], Jessyca Kawana Franquito[1]

[1] Departamento de Informática, Universidade Estadual de Ponta Grossa, Ponta Grossa, Paraná, Brasil, yastullio@gmail.com, ricardoks@live.com, selmar.aranha@gmail.com, salete@uepg.br, jefranquitto@hotmail.com.

RESUMO

A erosão hídrica é um dos fatores que provocam a perda de solos. Por ser um problema que causa danos à produção, são necessárias técnicas para controle das áreas erodidas. O sensoriamento remoto permite a visualização dessas áreas a fim de monitorar e auxiliar na busca de soluções para esse processo. A análise de imagem baseada em objetos geográficos (GEOBIA), realizada a partir do sensoriamento remoto, abrange técnicas que permitem a melhor visualização de diferentes aspectos de uma área. Dessa forma, foi realizado o presente estudo, que apresenta a aplicação das técnicas de GEOBIA a fim de demonstrar a delimitação do processo erosivo sobre uma área. Os resultados confirmam que esse delineamento pode ser realizado a partir da visualização dos vetores gerados com a segmentação.

PALAVRAS-CHAVE: GEOBIA, Sensoriamento remoto, Segmentação.

ABSTRACT

Water erosion is one of the factors that cause soil loss. Because it is a problem that causes production damage, techniques are needed for the control of eroded areas. Remote sensing allows the visualization of these areas and helps in the search of solutions for this process. Geographic object-based image analysis (GEOBIA), based on remote sensing, addresses issues that allow the visualization of the different aspects of an area. Thus, the present study was carried out, which presents an application of GEOBIA techniques, to demonstrate a

delimitation of the erosive process under one area. The results confirm that this delineation can be performed from the visualization of the vectors generated with a segmentation.
KEYWORDS: GEOBIA, Remote sensing, Segmentation.

INTRODUÇÃO

Umas das principais condições que causam a degradação do solo é a erosão hídrica, que pode ser causada ou influenciada por diversos fatores, como por exemplo erodibilidade do solo, cobertura do solo, clima, topografia, uso da terra e manejo (DECHEN, S. C. F. et al., 2015). Essa degradação pode tornar o solo improdutivo, já que empobrece o local de origem da erosão e o ambiente ao seu redor (BAGIO, B. et al, 2016).

O monitoramento da área erodida é necessário para desenvolver estratégias e determinar as prioridades para o controle da erosão (UDDIN, K. et al, 2016). Nesse cenário, sensoriamento remoto é uma alternativa para monitorar e obter conhecimento a respeito de uma cultura e seu estado no momento em que os dados foram adquiridos.

O sensoriamento remoto é definido como a ciência que tem por objetivo a obtenção de informações sobre um objeto sem que haja contato físico com ele (JENSEN, 2000). Para se conhecer um área, é necessário realizar análises e interpretações através do processamento dos dados obtidos. (JUNIOR, E. R. H; DAUGHTRY, C. S. T., 2018).

A análise de imagens para mapeamento pode ser realizada através de técnicas de classificação das mesmas. FRANCISCO, C. N., ALMEIDA, C. M. de (2012) apresenta a análise de imagem baseada em objetos geográficos (geographic object-based image analysis - GEOBIA) como tendo melhor desempenho, em relação aos métodos tradicionais (pixel por pixel), para a realização dessa tarefa.

Em GEOBIA, a segmentação de imagens é realizada através de medidas baseadas em homogeneidade. A imagem é particionada em segmentos de espaço contínuos, separados e homogêneos, com o objetivo de encontrar objetos significativos dentro da imagem.

Uddin, et al (2018) realizaram um estudo utilizando GEOBIA para identificar as alterações na cobertura da terra e como essas mudanças influenciam a distribuição de áreas de risco de erosão de solo no Nepal. Os resultados apresentaram as discrepâncias entre os dados de erosão do solo nas décadas anteriores em relação aos dados obtidos no ano 2010. No trabalho supra citado foram realizadas análises nacionais, estaduais e distritais, onde os

autores conseguiram identificar quais os distritos com maior necessidade de conservação do solo.

O presente trabalho utilizará dados de sensoriamento remoto aéreo e técnicas de segmentação de imagem e GEOBIA para delimitar a formação e a existência de erosão em áreas agrícolas. O estudo tem por objetivo apresentar as características encontradas em áreas erodidas que podem ser observadas por meio da GEOBIA.

MATERIAL E MÉTODOS

O objeto de estudo deste trabalho é uma área da Fazenda Escola Capão da Onça, da Universidade Estadual de Ponta Grossa, localizada na cidade de Ponta Grossa, estado do Paraná, em uma área marcada por processos erosivos. Para realização das análises foram obtidas imagens por meio de sensoriamento remoto, através do RPA (Remoted Piloted Aircraft) asa fixa Ebee, com uma câmera Sony Cyber-Shot DSC-WX220 de 18.2MP, utilizando uma distância focal de 35mm a uma altura de voo de 350, na data de 13 de abril de 2019. A resolução espacial obtida é de 10 cm e corresponde às três bandas espectrais do visível: vermelho (R), verde (G) e azul (B).

As imagens foram processadas no software Agisoft PhotoScan Professional versão 1.3.3, para geração da ortoimagem. Para delimitar a erosão e criar o banco de dados relacional foi utilizado o software ENVI 5.0 onde estão implementadas as ferramentas de filtragem e segmentação, empregadas na análise GEOBIA.

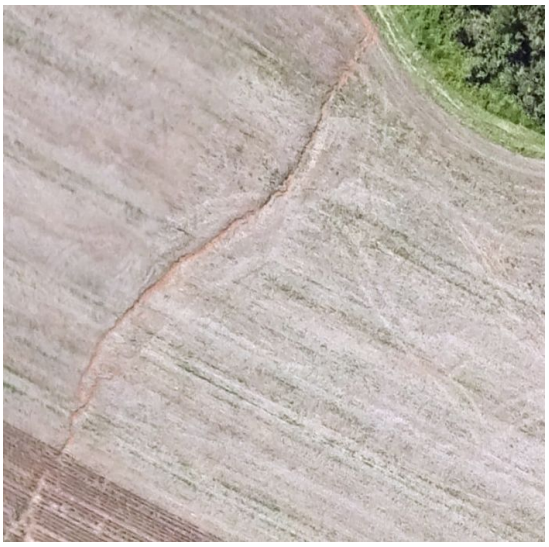
Visando a obtenção de uma segmentação com melhor resultado na delimitação e delineamento da feição erosiva, inicialmente aplicou-se a ferramenta de convolução e morfologia, utilizando filtro mediana de kernel 3x3, para suavização da imagem.

A ferramenta utilizada para realizar a segmentação é a Feature Extraction - Segment Only, que avalia os segmentos criados e mescla-os se suas propriedades espectrais forem semelhantes. Sua implementação foi realizada com os algoritmos Edge e Full Lambda Schedule. O primeiro identifica objetos com limites distintos e usa um método de detecção de bordas para calcular uma imagem gradiente, enquanto o segundo mescla iterativamente segmentos adjacentes com base em uma combinação de informações espectrais e espaciais.

Seguiu-se à geração automática do banco de dados relacional por meio da opção Export Segmentation Vectors, onde os vetores obtidos a partir da segmentação, constituem uma tabela, em que cada atributo descreve uma característica específica de um determinado vetor.

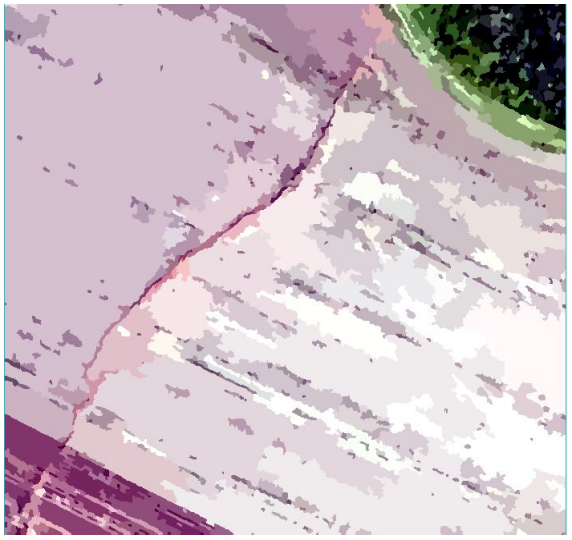
As figuras 1, 2 e 3 representam respectivamente as etapas de realização deste trabalho. A figura 1 refere-se a imagem original composição verdadeira cor (R banda 3, G banda 2 e B banda 1). A figura 2 apresenta o resultado do processo de aplicação do algoritmo Edge, com nível 42.0, e Full Lambda Schedule, com nível 94.0. E, por fim, a figura 3 apresenta os vetores gerados por meio da implementação Export Segmentation Vectors.

Figura 1



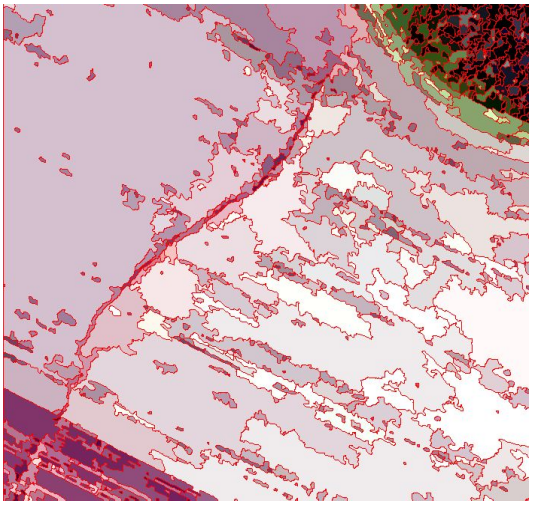
Fonte: Os autores

Figura 2



Fonte: Os autores

Figura 3



Fonte: Os autores

Como alternativa, para automatizar o processo de identificação da área erodida, foi utilizado o GIS (Sistema de Informações Geográficas) ArcGis 10.3. Nele foram realizados métodos de consulta a gráficos e medidas para construir uma fórmula que selecionasse apenas a área de interesse.

RESULTADOS E DISCUSSÃO

A exportação dos vetores gerados a partir da segmentação originou uma base de dados composta por 38 atributos que apresentam as características específicas de cada vetor. Desses atributos, foram selecionados os que melhor descrevem as feições da erosão a partir da análise da imagem, sendo eles (ITT Inc, 2008):

FX_ELONG: Este atributo demonstra a direção geral das feições por meio da razão entre os eixos maiores e menores, gerados a partir de um retângulo que envolve o objeto, demonstrando se um atributo é alongado ou não;

FX_AREA: Apresenta a área do polígono;

FX_LENGHT: Apresenta o perímetro da área;

FX_COMPACT: Esta medida se refere ao grau de compactação do polígono, com base na forma mais compacta que é o círculo;

FX_SOLID: Medida de forma que compara o polígono gerado com base em uma forma convexa que o envolve;

TX_RAN_Bx: Intervalo médio de dados dos pixels que compõem a região gerada com base em um kernel de tamanho previamente determinado;

TX_AVG_Bx: Valor médio dos pixels que compõem a região dentro do kernel;

FX_MAIN_DI: Indica a direção do maior eixo do polígono em relação ao eixo X;

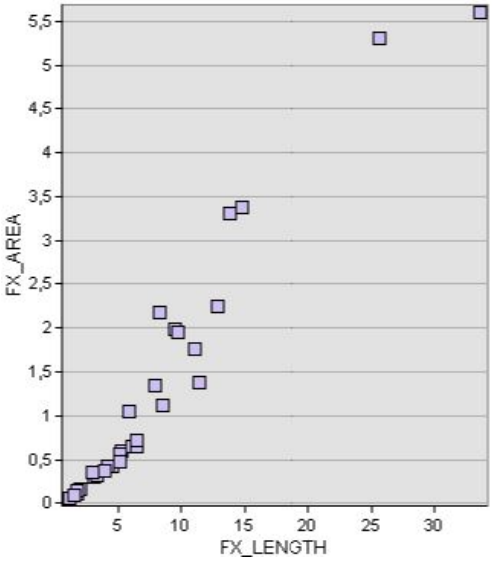
TX_VAR_Bx: Variação média dos pixels que compõem a região dentro do kernel;

MINBx: Valor mínimo dos pixels que compõem o vetor em cada uma das bandas.

Esses dados auxiliam no processo de reconhecimento da área e planejamento de medidas para reduzir a perda de solo causada pela erosão. Exemplos de informações geradas automaticamente pelo banco de dados relacional que fornecem subsídios para esse fim, são o fluxo de escoamento e o tamanho da erosão. O primeiro pode ser conhecido ao observar os valores dos dados referentes à direção predominante dos polígonos, e o segundo pode ser calculado a partir dos valores de área e perímetro de cada polígono.

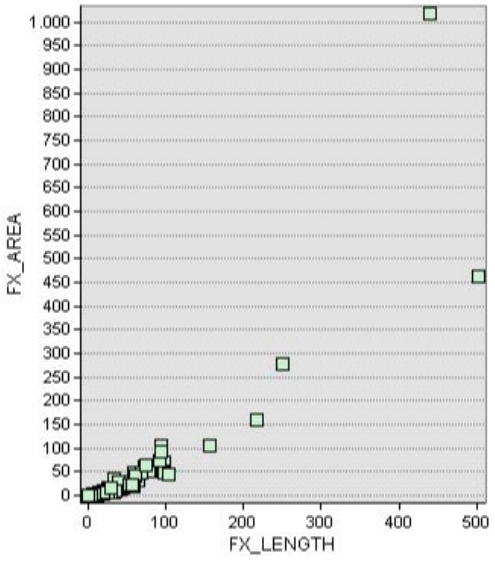
Além disso, os valores de área e perímetro dos polígonos podem servir de auxílio para verificar as diferenciações existentes no terreno, já que a geração dos mesmos depende de variação de informações. Dessa forma, polígonos com área e perímetro menores apresentam maior variabilidade. As figuras 4 e 5 demonstram os gráficos de dispersão dos polígonos em relação aos valores de área e perímetro. Com isso, é possível observar que a variação dos valores gerados na área da erosão, apresentados na figura 4, são menores do que os de polígonos gerados em toda a imagem, apresentados na figura 5.

Figura 4



Fonte: Os autores

Figura 5



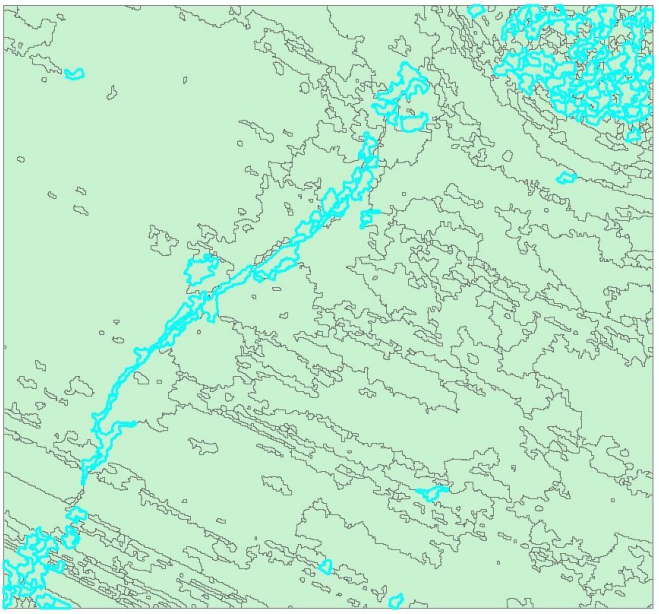
Fonte: Os autores

Após uma análise dos dados gráficos, gerou-se uma fórmula que identifica os polígonos existentes na parte erodida da imagem.

"FX_AREA" < 7 AND "FX_AREA" > 1.6 AND "FX_LENGHT" < 40 AND "FX_MAIN_DI" > 80

A fórmula utiliza os valores de área (FX_AREA), perímetro (FX_LENGHT) e direção (FX_MAIN_DI) dos polígonos, definindo o intervalo em que cada atributo se torna representativo para a delimitação da erosão. Com essa fórmula, o processo de reconhecimento da parte erodida se faz de forma automática, identificando apenas os elementos que são distintos de terrenos homogêneos. A figura 6 apresenta a delimitação da área por meio dessa fórmula, em que os polígonos selecionados a partir dela estão contornados pela cor azul. Foram selecionados os polígonos que representam a área erodida. Onde se encontra a mata ou vegetação apresentaram-se ruídos, o que não gera problemas para a delimitação da erosão.

Figura 6



Fonte: Os autores

Devido ao RPA ser uma alternativa de sensoriamento remoto menos custosa financeiramente do que, por exemplo, utilizar aeronaves tripuladas, os voos podem ser feitos com maior frequência. Assim, o monitoramento da área erodida pode ser realizado repetidas vezes, obtendo-se informações a respeito do estado em que se encontra a área a cada voo.

Dessa forma, unindo as atividades da GEOBIA e o sensoriamento remoto, é possível abstrair informações a respeito de mudanças ocorridas na área. Após um longo período de

chuva, é possível observar se houve evolução no processo erosivo, verificando-se numericamente os dados de área e perímetro informados no banco de dados, e se houve variações no fluxo de escoamento. Do mesmo modo, pode-se verificar, após medidas mitigadoras serem aplicadas, se a erosão foi amenizada. A figura 7 apresenta o delineamento da erosão a partir dos vetores gerados.

Figura 7



Fonte: Os autores

CONCLUSÕES

A partir da análise da representação da área pela segmentação, verificou-se que é possível identificar as feições da erosão ao visualizar os polígonos gerados, por meio de uma análise qualitativa. Isso mostra que a delimitação da erosão utilizando a ferramenta de segmentação Feature Extraction - Segment Only pode ser realizada e pode auxiliar no monitoramento dos processos erosivos que atingem a área.

Uma oportunidade de estudo para trabalhos futuros é verificar a possibilidade de, através de técnicas de mineração de dados, encontrar um modelo de classificação que possa identificar a erosão em uma área através dos dados obtidos através da GEOBIA.

REFERÊNCIAS

DECHEN, S. C. F., et al. Perdas e custos associados à erosão hídrica em função de taxas de cobertura do solo. Bragantia. Vol. 74, n. 2, pág. 224-233. 2015.

BAGIO B. et al. Water Erosion in Different Slope Lenghts on Bare Soil. Revista Brasileira de Ciência do solo. 2017.

UDDIN, K. et al. Estimation of Soil Erosion Dynamics in the Koshi Basin Using GIS and Remote Sensing to Assess Priority Areas for Conservation. PLoS ONE. Pág. 1-19. 2016.

JENSEN, J. R. Remote Sensing of the Enviroment an Earth Resource Perspective. Ed. Prentice Hall, Inc. 2000.

JUNIOR, E. R. H, DAUGHTRY. What good are unmanned aircraft systems for agricultural remote sensing and precision agriculture? Internacional Journal of Remote Sensing. Vol. 39, n 15-16, pag. 5345-5376. 2018.

FRANCISCO, C. N.; ALMEIDA, C. M. de. Interpretação de Imagens Orbitais por Meio de Sistema Especialista para o Mapeamento de Cobertura da Terra em Região Montanhosa. Sociedade & Natureza. Vol. 24. n. 2. 2012.

UDDIN, K. et al. Assessment of Land Cover Change and Its Impacto n Changes in Soil Erosion Risk in Nepal. Sustainability. Vol. 10. 2018.

ITT, Inc. ENVI Feature Extraction Module User's Guide, 2008.

Modelagem e persistência de dados sensoriais para a pecuária de precisão

Leandro Borges da Silva[1], *Hana Karina Salles Rubinsztejn*[1]

[1] Universidade Federal de Mato Grosso do Sul, Campo Grande, Mato Grosso do Sul, Brasil, leandroborges@outlook.com, hana@facom.ufms.br

RESUMO

Para manter-se competitivo na cadeia de produção bovina, é importante inovar na produção e gestão do agronegócio. Desse modo, surgem alguns desafios, como o aumento da demanda relacionada ao bem-estar animal, segurança alimentar, rastreabilidade bovina, sustentabilidade, entre outros. A pecuária de precisão pode contribuir para atingir esses objetivos. Diversos estudos para a pecuária de precisão são desenvolvidos em parceria entre a Embrapa Gado de Corte e a Faculdade de Computação (FACOM), da Universidade Federal do Mato Grosso do Sul (UFMS). Esses são projetos que possuem uma heterogeneidade de sensores e geram um grande volume de dados provenientes de diferentes fontes e padrões. Este artigo tem como finalidade definir um modelo de dados semântico e físico para persistência dos dados oriundos dos diversos sensores e sistemas utilizados na pecuária de precisão, visando a integração e compartilhamento padronizado dessas informações.

PALAVRAS-CHAVE: Modelagem de Dados, NoSQL, Pecuária de Precisão, Integração.

ABSTRACT

In order to remain competitive in cattle production chain, it is important to innovate in the production and management of agribusiness. Thus, there are some challenges such as increased demand related to animal welfare, food safety, bovine traceability, sustainability, among others. In this context, precision livestock can contribute to achieving these objectives. Several studies for precision livestock are developed between Embrapa Gado de Corte and Faculdade de Computação (FACOM) of the Universidade Federal do Mato Grosso do Sul (UFMS). These projects have a heterogeneity of sensors, and generate a large volume of data from different sources and with differents standards. This paper aims to define a semantic and physical data model for data persistence from the various sensors and system used in precision livestock, aiming the integration and sharing information.

KEYWORDS: Data modeling, Precision Livestock, NoSQL, Integration.

INTRODUÇÃO

No Brasil, o rebanho bovino vem crescendo a cada ano e, com esse desenvolvimento, alguns desafios surgem, como o aumento da demanda relacionada ao bem-estar animal, segurança alimentar, rastreabilidade dos animais, sustentabilidade, entre outros. Portanto, mudanças são exigidas na produção e na gestão do agronegócio, fazendo com que os participantes da cadeia de criação bovina busquem inovar para se manterem competitivos e, principalmente, melhorar o bem-estar animal e a produção em si.

Nesse contexto, a pecuária de precisão pode contribuir para esses objetivos. A pecuária de precisão pode ser definida como a gestão de bovinos a partir da Tecnologia da Informação e Comunicação, que garante boas práticas na produção de carne (CÁCERES et al., 2011). O desenvolvimento de ferramentas para o monitoramento dos animais durante sua vida, de forma automática, é uma de suas finalidades (BERCKMANS, 2006).

Há algumas oportunidades que tornam a pecuária de precisão viável. A obtenção de dados do rebanho, rastreando individualmente os animais, por exemplo, auxilia o produtor na tomada de decisões, diminuindo gastos desnecessários. Essas informações podem diminuir o tempo de manejo, que afeta diretamente em seu resultado econômico, o que assegura a redução de custos e até mesmo o ganho de produtividade.

Em uma parceria firmada entre a Embrapa Gado de Corte e a Faculdade de Computação (FACOM), há uma ampla pesquisa em pecuária de precisão. Alguns estudos relevantes se destacam: a predição de comportamento animal e técnicas de inteligência artificial, baseado em posições georreferenciadas; controle sanitário e manejo reprodutivo, registrando diariamente o peso dos animais, por meio de uma Balança de Passagem; integração lavoura-pecuária floresta (ILPF) para conforto-térmico animal, coletando informações ambientais, como por exemplo, temperatura e umidade relativa do ar, e condições fisiológica dos bois, como temperatura corporal e batimentos cardíacos; e rastreamento bovino.

Além da heterogeneidade de sensores, protocolos e padrões utilizados nos projetos mencionados, também há a geração de um grande volume de dados provenientes de diferentes fontes e diversos padrões de informações. Cada um dos trabalhos desenvolve tecnologias que funcionam isoladamente, capturando e processando os dados dos sensores.

Este trabalho pretende definir um modelo semântico de informações, bem como um modelo físico para persistência de dados oriundos dos diversos sensores utilizados na pecuária de precisão. Também tem o propósito de gerenciar dados heterogêneos e grandes volumes de dados providos por sensores e fornecer informações padronizadas.

MATERIAL E MÉTODOS

Uma revisão sistemática foi realizada para identificar os trabalhos na área de pecuária de precisão, além dos estudos da Embrapa Gado de Corte e, dessa forma, levantar variáveis para um dicionário de dados da pecuária de precisão.

Dividido em três fases, o mapeamento sistemático foi baseado no processo descrito por Kitchenham (2004), que contempla o planejamento, a condução e, por último, a extração dos dados e análise dos resultados. O planejamento é a etapa em que são definidas as necessidades da pesquisa, assim como o protocolo do mapeamento sistemático. No estágio de condução é feita a identificação dos estudos primários, em que eles são avaliados de acordo com critérios de inclusão e exclusão estabelecidos. Por último, é realizado o levantamento das informações após a seleção dos estudos primários, bem como as respostas das questões de pesquisa.

Planejamento

Após a pesquisa ser realizada e a base quantitativa ser considerada, foram estabelecidas a abrangência e as proximidades dos resultados no contexto mencionado anteriormente, e assim definida uma boa base qualitativa. Na fase do planejamento foram estipuladas as questões de pesquisa (QP), conforme apontado abaixo:

1. Quais trabalhos apresentam uma proposta para modelagem de dados no domínio da pecuária de precisão?
2. Quais fazem uso de metodologias e processos de modelagem de dados?
3. Quais possuem uma modelagem flexível e adaptável ao contexto?

A construção das *strings* de busca foi definida em três áreas: automação, modelagem de dados e pecuária. Em seguida, foram estabelecidas *strings* genéricas de buscas, as quais foram adaptadas de acordo com a base de busca pesquisada. A pesquisa foi feita a partir das principais bases utilizadas em revisões sistemáticas nas mais diversas áreas da computação, sendo elas: ACM *Digital Library* (ACM), IEEE *Xplore Digital Library* (IEEE), *Web of Science* e *Science Direct* (DYBA; DINGSOYR; HANSSEN, 2007). Do cruzamento dos conceitos-chave mencionados acima, o resultado gerado é a *string* apresentada no Quadro 1.

Quadro 1 – *String* padrão de busca em inglês do mapeamento sistemático.

(AUTOMATION or SENSOR) AND (DATAMODEL or DATA MODEL or DATA MODELING) AND (LIVESTOCK or AGRICULTURE or FARM or CATTLE)

Fonte: elaborado pelos autores.

Para conseguir uma amostragem adequada e padronizada, é importante refinar o resultado da busca dos estudos primários. Com isso, foram criadas duas regras: Critérios de Inclusão (CI) e Critérios de Exclusão (CE). São dois os critérios de inclusão: estudos que apresentem coleta de dados por sensores na agropecuária (CI1) e estudos que apresentem modelo de dados coletados por sensores (CI2).

Já os critérios de exclusão abrangem: estudo publicado fora do período de 2007 a 2017 (CE1); estudo que não contemple, satisfatoriamente, os critérios de inclusão (CE2); estudo que não esteja em inglês ou em português (CE3); estudo incompleto, indisponível ou similar a outros já elencados (CE4); e estudo não publicado em anais de eventos, em revistas, como relatórios técnicos, capítulos de livros, teses ou dissertações (CE5). Foram incluídos os estudos primários relevantes e excluídos os irrelevantes no contexto do mapeamento.

Condução da busca

O processo de execução das buscas considerou as configurações de pesquisa em algumas bases de dados, aplicando a *string* de busca sobre título, palavras-chave e resumo. Com a *string* definida e utilizando a sintaxe para cada base de pesquisa, foi realizado o processo de busca. Dos estudos resultantes, 296 no total, todos foram avaliados por título, resumo, data de publicação e idioma. Após aplicar os critérios de exclusão, foram excluídos 280 registros.

Dos 16 estudos resultantes após a aplicação dos critérios de exclusão, a maioria contempla todos os critérios de inclusão. Ou seja, descrevem soluções que apresentam coleta de dados por sensores, assim como um modelo de dados.

Extração dos dados e Análise dos resultados

De acordo com os resultados obtidos após a extração dos dados dos estudos selecionados, foi executada a análise de cada um dos trabalhos. Ao total, foram considerados 16, conforme ilustra o Quadro 2. Nenhum estudo atendeu plenamente as questões de pesquisa, porém todos apresentam dados coletados pelos sensores ou modelos de dados, mesmo que implicitamente.

Quadro 2 – Trabalhos selecionados para extração de dados.

Título	Autoria
Sensor data modeling and validating for wireless soil sensor network.	Wang; Damevski; Chen (2015).
Development of Knowledge Sharing System for Agriculture Application.	Uchinuno et al. (2013).
REnvDB, a RESTful Database for Pervasive Environmental Wireless Sensor Networks.	Stefani et al. (2010).
An Extensible Software Platform for Cloud-Based Decision Support and Automation in Precision Agriculture.	Tan; Hou; Zhang (2016).
Application of Intelligence Information Fusion Technology in Agriculture	Jiayu et al. (2015).

Monitoring and Early-warning Research.	
FieldTouch: an innovative agriculture decision support service based onmulti-scale sensor platform.	Kiyoshi et al. (2014).
Farming the Web of Things.	Taylor et al. (2013).
Temperature sensor network for prediction of possible start of brood rearing by indoor wintered honey bees.	Zacepins et al. (2011).
Microenvironment Information Acquisition and Processing in FarmlandWater Potential Soft-sensing.	Han et al. (2010).
Global Positioning of Off-Road Vehicles by Sensor Fusion for Precision Agriculture.	Lin et al. (2008).
Networked Biological System by Wireless Sensors.	Mariño et al. (2008).
Web-based cattle behavior service for researchers based on the smartphone inertial central.	Debauche et al. (2017).
Comparison of methods to determine methane emissions from dairy cows in farm conditions.	Huhtanen et al. (2015).
Greenhouse environmental control system based on SW-SVR Yukimasa.	Kaneda et al. (2015).
A Data-Driven approach to soil moisture collection and prediction.	Hong; Kalbarczyk; Iyer (2016).
Intelligent Livestock Farm Management for Context Data Model Design.	Kim; Cho; Yoe (2012).

Fonte: elaborado pelos autores.

Também foram investigados trabalhos e pesquisas realizados pela parceria da Embrapa Gado de Corte com a FACOM. Ao todo, foram analisados seis estudos relevantes em diversas áreas, como a predição de comportamento animal, rastreamento bovino, integração lavoura-pecuária-floresta, entre outras, conforme descritos no Quadro 3.

Quadro 3 – Trabalhos selecionados para extração de dados.

Título	Autoria
Identificação do comportamento bovino por meio do monitoramento animal.	Jesus (2014).
Identificação do comportamento bovino a partir dos dados de movimentacão e do posicionamento do animal.	Lomba (2015).
Termus uma plataforma para aferição remota de temperatura de bovinos da raça *brangus*.	Santos Neto (2010).
Análise comportamental de bovinos baseada em trajetórias semânticas aplicada à pecuária de precisão.	Oliveira (2013).
Microclima e conforto térmico em sistemas em integração no cerrado.	Ajala (2013).
Algoritmos para obtenção de frequência cardíaca e respiratória em bovinos.	Bressan (2016).

Fonte: elaborado pelos autores.

RESULTADOS E DISCUSSÃO

Após análise dos trabalhos, foram levantadas todas as variáveis e informações de cada um dos projetos, assim como seus valores, medidas e tipos dos dados. A partir dessas variáveis pré-selecionadas, foi criado um quadro com um nome comum de cada uma das variáveis, tal como a unidade de medida mais utilizada nas pesquisas, quando disponíveis.

As variáveis são categorizadas em quatro tipos: comportamentais, microclima, fisiológicos e contextuais. Os dados gerados por sensores relacionados ao comportamento de cada animal, como sua trajetória em uma área de pastejo, são classificados como

comportamentais. Informações como a temperatura do solo, radiação solar ou qualquer outro aspecto relacionado ao meio ambiente, são dados de microclima. Elementos fisiológicos são os dados de cada animal ao longo do tempo, como o peso e sua frequência cardíaca. Por último, os dados contextuais são os provenientes de sensores que aferem alguma alteração em insumos da propriedade, como a quantidade de suplemento em um cocho, total de água em um bebedouro, se uma porteira está aberta ou fechada, entre outros.

No Quadro 4 são descritas todas as variáveis já com a nomenclatura padronizada para o idioma Inglês.

Quadro 4 – Entidades do dicionário de dados.

Entidade	Tipo	Definição
Accelerometer	Comportamental	Aceleração própria de um sistema.
AirTemperature	Microclima	Temperatura do ambiente.
AnimalSpeed	Comportamental	Velocidade do animal.
AnimalWeight	Fisiológico	Peso do animal.
BlackGlobeTemperature	Microclima	Temperatura do globo negro.
BodyTemperature	Fisiológico	Temperatura corporal do animal.
CH4	Microclima	Gás metano expelido pelo animal.
CO2	Microclima	Dióxido de carbono do ambiente.
DewPointTemperature	Microclima	Temperatura do ponto de orvalho.
DryBulbTemperature	Microclima	Temperatura do bulbo molhado.
GateOpened	Contextual	*Status* aberto ou fechado de um portão.
GeographicCoordinate	Contextual	Coordenadas geográficas do GPS.
Gyroscope	Comportamental	Referência da direção.
HeartRate	Fisiológico	Batimento cardíaco do animal.
Magnetometer	Comportamental	Força, orientação e direção do campo magnético.
PH	Microclima	Nível de pH da água ou do solo.
Precipitation	Microclima	Quantidade pluviométrica.
RelativeHumidity	Microclima	Humidade relativa do ar.
RespiratoryFrequency	Fisiológico	Frequência respiratória do animal.
RetalTemperature	Fisiológico	Temperatura retal do animal.
SoilTemperature	Microclima	Temperatura do solo.
SoilMoisture	Microclima	Humidade do solo.
SoilWaterPotencial	Microclima	Potencial hídrico do solo.
SoilNitrogen	Microclima	Porcentagem de nitrogênio no solo.
SolarRadiation	Microclima	Nível de radiação solar.
WaterTemperature	Microclima	Temperatura da água.
WetBulbTemperature	Microclima	Temperatura do bulbo seco.
WindSpeed	Microclima	Velocidade do vento.

Fonte: elaborado pelos autores.

Estabelecidas e padronizadas as variáveis, cada uma delas simboliza agora uma entidade, com diferentes atributos entre elas. Os *metadados*, que contém as definições e representações dos dados, são comuns a todas entidades.

- **Entidade**: nome da entidade que foi definida após a revisão sistemática. Cada uma representa um dado coletado por sensores;
- **Atributo**: os atributos são as características de cada entidade. Uma chave primária, valor e unidades de medidas são os principais atributos das entidades;
- **Domínio**: um domínio é o tipo do valor que o atributo irá receber. O valor pode ser numérico, texto, *booleano* ou uma data;
- ***Default***: valor padrão do atributo. Esse *metadado* é flexível e opcional;
- **Mínimo**: o valor mínimo do atributo, quando obrigatório. Caso esse *metadado* esteja em branco no dicionário de dados, não há valor mínimo permitido;
- **Máximo**: o valor máximo do atributo, quando obrigatório. Caso esse *metadado* esteja em branco no dicionário de dados, não há valor máximo permitido;
- **Descrição**: descreve o que é o atributo ou informações adicionais que possam ser utilizadas pelo usuário do dicionário de dados.

Nesse contexto, o dicionário de dados permite que se obtenham informações sobre todos os objetos e entidades do modelo de forma textual, e seu objetivo é ser claro e consistente. Na Tabela 1 é descrito um exemplo de uma entidade chamada "AnimalWeight".

Tabela 1 – Exemplo de entidade para dicionário de dados.

Entidade: **AnimalWeight**					
Atributo	Domínio	Default	Mínimo	Máximo	Descrição
Value	Decimal (4,2)		0	1999	Peso do animal.
Unit	Varchar	Kg			Unidade de medida (Kg).
resource	String				Id, nome ou *tag* do animal.
Date	Timestamp				Data e hora do registro.

Fonte: elaborada pelos autores.

Como é possível observar, essa entidade possui os atributos ***value***, ***unit***, ***resource*** e ***date***. Nesse exemplo, o atributo ***value*** é do tipo *Decimal* de tamanho igual a 4, com duas casas decimais, que armazena o peso do animal. O atributo ***unit*** é apenas para referência da unidade de medida do dado a ser persistido. O atributo ***resource*** é descrito no dicionário do tipo *String*, que registra um "id", nome ou *tag* como o identificador do animal. Já o atributo ***date*** do tipo *Timestamp* refere-se à data de registro do sensor, no formato *Timestamp*.

Alguns tipos de entidades, como imagem, som e vídeo, ainda não são contemplados. Caso haja a necessidade de armazenar algum tipo de dado que o modelo ou dicionário de dados não abrange, é necessário criar manualmente a entidade e registrá-la previamente. Criar entidades automaticamente é um dos trabalhos futuros.

Para a gestão e armazenamento de dados, a solução foi implementada utilizando banco de dados NoSQL, que são os bancos não-relacionais. O objetivo do NoSQL é atender aos requisitos de gerenciamento no armazenamento e processamento de dados em larga escala, sejam eles semiestruturados ou não estruturados, que necessitam de alta disponibilidade e escalabilidade. Intuito esse que motivou a escolha desse tipo de banco de dados. Para o domínio da pecuária de precisão, baseado em estudos diversos e na necessidade de trabalhar com enormes quantidades de dados heterogêneos, o banco de dados NoSQL demonstrou ser a melhor solução para o projeto. Para fins de testes, foi utilizado o banco de dados MongoDB.

Os dados provenientes dos sensores são armazenados em uma estrutura individual. Cada *Schema* mapeia um tipo de dado no MongoDB, e determina a forma dos documentos. A estrutura do *Schema* é basicamente a mesma em todas entidades, com os atributos ***uid***, ***value***, ***date***, ***resource***, ***dataStorage*** e ***syncedAt***. Cada chave, representada pelo atributo, define uma propriedade no documento e possui um *SchemaType* associado. Os *SchemaTypes* permitidos são *string*, *number*, *date*, *buffer*, *boolean*, *mixed*, *objectId*, *array*, *decimal128* e *Map*.

CONCLUSÕES

Com o aumento das exigências na produção e gestão do agronegócio, principalmente em relação aos participantes da cadeia de produção bovina, o uso de tecnologias computacionais e a inovação no setor é primordial para manter-se competitivo. A pecuária de precisão pode auxiliar na automatização de processos, aumento de produção e bem-estar animal, bem como a produção bovina como um todo.

Um mapeamento sistemático foi executado, a fim de identificar trabalhos e estudos relacionados à modelagem e armazenamento de dados no contexto da pecuária. A partir dessas pesquisas, foi feito um levantamento de sensores e tipos de dados utilizados no domínio da pecuária de precisão e áreas correlatas.

Foi definido um modelo semântico e padronizado de dados inerentes ao domínio da pecuária de precisão, bem como a implementação de um modelo físico para persistência de dados em banco de dados não-relacionais. Um dicionário de dados foi elaborado com o intuito de guiar futuras implementações e integração com o modelo.

O banco de dados NoSQL foi a solução para o armazenamento dos dados, em que eficiência e escalabilidade são requisitos que uma arquitetura de dados para a pecuária de precisão necessita. O MongoDB, que é um banco NoSQL orientado a documentos, foi a opção escolhida na implementação do projeto, principalmente por ter uma documentação ampla e uma comunidade de desenvolvedores bastante ativa.

REFERÊNCIAS

AJALA, N. *et al.* Microclima e conforto térmico em sistemas em integração no cerrado. *In:* Jornada Científica EMBRAPA Gado de Corte, 9. 2013, Campo Grande, MS.[*Anais*...]. Campo Grande, MS: EMBRAPA Gado de Corte, 2013. Embrapa Gado de Corte – Resumo em anais de congresso (ALICE). [S.l.]. 2013.

BERCKMANS, D. Automatic on-line monitoring of animals by precision livestock farming. *Livestock Production and Society*. Wageningen Academic Publishers, Wageningen, p. 287-295, 2006.

BRESSAN, P. O. *Algoritmos para obtenção de frequência cardíaca e respiratória em bovinos.* Dissertação (Mestrado em Ciência da Computação) – Programa de Pós-graduação da Faculdade de Computação, Universidade Federal de Mato Grosso do Sul, Campo Grande, 2016.

CÁCERES, E. N. *et al.* Computational precision livestock–position paper. *In: II Workshop of the Brazilian Institute for Web Science Research*. [*S. l.: s. n.*], 2011, p. 9, 2011.

DEBAUCHE, O. *et al.* Web-based cattle behavior service for researchers based on the smartphone inertial central. *Procedia Computer Science*, Elsevier, v. 110, p. 110-116, 2017.

DYBA, T.; DINGSOYR, T.; HANSSEN, G. K. Applying systematic reviews to diverse study types: An experience report. *In: IEEE. Empirical Software Engineering and Measurement*. ESEM 2007. First International Symposium on. [*S. l.*], 2007, p. 225-234, 2007.

HAN, L. *et al.* Microenvironment information acquisition and processing in farmland water potential soft-sensing. *In: IEEE. Intelligent Control and Automation* (WCICA), 8th World Congress on. [*S. l.*], 2010, p. 6943-6948, 2010.

HONG, Z.; KALBARCZYK, Z.; IYER, R. K. A data-driven approach to soil moisture collection and prediction. *In: IEEE. Smart Computing* (SMARTCOMP), IEEE International Conference on. [*S. l.*], 2016, p. 1-6, 2016.

HUHTANEN, P. *et al.* Comparison of methods to determine methane emissions from dairy cows in farm conditions. *Journal of dairy science*, Elsevier, v. 98, n. 5, p. 3394-3409, 2015.

JESUS, L. de. *Identificação do comportamento bovino por meio do monitoramento animal.* Dissertação (Mestrado em Ciência da Computação) – Programa de Pós-graduação da Faculdade de Computação, Universidade Federal de Mato Grosso do Sul, Campo Grande, 2014.

JIAYU, Z. *et al*. Application of intelligence information fusion technology in agriculture monitoring and early-warning research. *In: IEEE. Control, Automation and Robotics* (ICCAR), International Conference on. [*S. l.*], 2015, p. 114-117, 2015.

KANEDA, Y. *et al.* Greenhouse environmental control system based on sw-svr. *Procedia Computer Science*, Elsevier, v. 60, p. 860-869, 2015.

KIM, H.; CHO, Y.; YOE, H. Intelligent livestock farm management for context data model design. *Green and Smart Technology with Sensor Applications*, Springer, p. 291-296, 2012.

KITCHENHAM, B. *Procedures for performing systematic reviews*. Keele, UK, Keele University, v. 33, n. 2004, p. 1-26, 2004.

KIYOSHI, H. *et al.* Fieldtouch: an innovative agriculture decision support service based on multi-scale sensor platform. *In: IEEE. Global Conference (SRII),* Annual SRII. [S.l.], 2014, p. 228-229, 2014.

LIN, H. *et al.* Global positioning of off-road vehicles by sensor fusion for precision agriculture. *In: IEEE. Intelligent Control and Automation,* 2008. WCICA. 7th World Congress on. [*S. l.*], 2008, p. 7005-7010, 2008.

LOMBA, L. F. D. *Identificação do comportamento bovino a partir dos dados de movimentacão e do posicionamento do animal.* Dissertação (Mestrado em Ciência da Computação) – Programa de Pós-Graduação da Faculdade de Computação, Universidade Federal de Mato Grosso do Sul, Campo Grande, 2015.

MARIÑO, P. *et al.* Networked biological system by wireless sensors. *In: IEEE. Control, Automation, Robotics and Vision*, 2008. ICARC, 10th International Conference on. [*S. l.*], 2008, p. 218-222, 2008.

OLIVEIRA, M. T. P. d. *Análise comportamental de bovinos baseada em trajetórias semânticas aplicada à pecuária de precisão*. Dissertação (Mestrado em Ciência da Computação) – Programa de Pós-graduação da Faculdade de Computação, Universidade Federal de Mato Grosso do Sul, Campo Grande, 2013.

SANTOS NETO, Q. I. dos. *Termus uma plataforma para aferição remota de temperatura de bovinos da raça brangus*. Dissertação (Mestrado em Ciência da Computação) – Programa de Pós-graduação da Faculdade de Computação, Universidade Federal de Mato Grosso do Sul, Campo Grande, 2010.

STEFANI, F. D. *et al.* Renvdb, a restful database for pervasive environmental wireless sensor networks. *In: IEEE. Distributed Computing Systems Workshops (ICDCSW*), IEEE 30th International Conference on. [*S. l.*], 2010, p. 206-212, 2010.

TAN, L.; HOU, H.; ZHANG, Q. An extensible software platform for cloud-based decision support and automation in precision agriculture. *In: IEEE. Information Reuse and Integration (IRI)*, IEEE 17th International Conference on. [*S. l.*], 2016, p. 218-225, 2016.

TAYLOR, K. *et al.* Farming the web of things. *IEEE Intelligent Systems,* IEEE, v. 28, n. 6, p. 12-19, 2013.

UCHINUNO, T. *et al.* Development of knowledge sharing system for agriculture application. *In: IEEE. Advanced Applied Informatics (IIAIAAI)*, IIAI International Conference on. [*S. l.*], 2013, p. 108-111, 2013.

WANG, J.; DAMEVSKI, K.; CHEN, H. Sensor data modeling and validating for wireless soil sensor network. Computers and Electronics in Agriculture, Elsevier, v. 112, p. 75–82, 2015.

ZACEPINS, A. *et al.* Temperature sensor network for prediction of possible start of brood rearing by indoor wintered honey bees. *In: IEEE. Carpathian Control Conference (ICCC)*, 12th International. [*S. l.*], 2011, p. 465-468, 2011.

Influência das diferentes faces de exposição ao sol nos índices vegetativos e relativo de clorofila em cafés especiais

Eduardo Antonio Speranza[1], *Célia Regina Grego*[1], *Gustavo Costa Rodrigues*[1], *Ariovaldo Luchiari Junior*[1]

[1] Embrapa Informática Agropecuária, Campinas, São Paulo, Brasil,

eduardo.speranza@embrapa.br, celia.grego@embrapa.br, gustavo.rodrigues@embrapa.br, ariovaldo.luchiari@embrapa.br

RESUMO

A cafeicultura tem se destacado no Brasil nos últimos anos, atingindo altos níveis de exportação. Desse modo, a busca por tecnologias que permitam melhorar a produtividade e a qualidade do produto final é uma realidade. Nesse contexto, ferramentas de Agricultura de Precisão (AP) possuem papel estratégico, pois possibilitam ao produtor obter informações georreferenciadas com alta precisão de atributos referentes ao solo e à cultura que podem ser analisadas e transformadas em informações úteis para o manejo da produção. Neste artigo, foi realizado um processo de mineração de dados obtidos em campos experimentais de cafeicultura por meio de sensores proximais que permitem o cálculo de índices de vegetação e de clorofila presente nas plantas. Foram utilizadas diferentes combinações de atributos para a geração de mapas potenciais de zonas de manejo que permitem caracterizar a variabilidade espacial da cultura com relação à diferentes faces de exposição ao sol. Como resultado, foi verificado que as faces de exposição ao sol das duas áreas analisadas influenciam muito no índice relativo de clorofila (IRC) presente nas plantas, potencializando a utilização desse atributo para o delineamento de zonas de manejo que permitam diferenciar a qualidade do produto final obtido.

PALAVRAS-CHAVE: Cafeicultura, Agricultura de precisão, Mineração de dados, Sensoriamento Proximal, Zonas de Manejo.

ABSTRACT

Coffee cultivation has been notable in Brazil in recent years, achieving high exportation levels. In this way, the search for technologies to improve yield and the quality of the final product is a reality. In this context, Precision Agriculture (PA) tools have a strategic role, since they enable the farmer to obtain high precision and georeferenced information of soil and crop attributes which we can analyze and transform into useful information for crop management. In this paper, we performed a data mining process using proximal sensors data from experimental fields of coffee cultivation, which allow the calculation of vegetation and chlorophyll indices present in the plants. We used different sets of attributes to build potential management zone maps describing the spatial variability of the crop regarding different faces of sun exposure. As a result, we verified that the sun exposure of the two analyzed experimental areas greatly influence the relative index of chlorophyll (IRC) present in the plants, boosting the use of this attribute for the delineation of management zones to distinguish the quality of the final product.

KEYWORDS: Coffee cultivation, Precision agriculture, Data mining, Proximal Sensing, Management Zones.

INTRODUÇÃO

O Brasil tem se destacado nos últimos anos como maior produtor e exportador mundial de café, possuindo cerca de 300 mil estabelecimentos produtores onde mais de 80% deles são considerados produção familiar (CONCAFÉ, 2019). Os tipos de café classificados como especiais possuem valores de venda até 40% maiores do que os convencionais, o que desperta maior interesse por parte dos produtores. Apesar dos aspectos sensoriais serem os mais utilizados para identificar um café especial, aspectos ambientais do local de produção, tais como microclima, tipo de solo, altitude, direção do vento, umidade, radiação solar e face de exposição ao sol, podem influenciar diretamente na qualidade do produto final (GASPARI-PEZZOPANE et al., 2005; FERREIRA et al., 2012).

Nesse sentido, ferramentas de Agricultura de Precisão (AP) podem ser utilizadas para auxiliar na caracterização ambiental dos sistemas de produção de café, permitindo caracterizar a variabilidade espacial da cultura agrícola com relação aos aspectos supracitados (SILVA; ALVES, 2013). No contexto da AP, sensores proximais que permitem verificar a

variabilidade dos índices de vegetação e relativo de clorofila presentes na cultura, em conjunto com técnicas de geoestatística e de mineração de dados, podem auxiliar na identificação de regiões dentro de uma mesma área de cultivo, com potencial para obtenção de um produto final de maior qualidade. Na literatura, existem diversos trabalhos realizados com esse tipo de sensor para identificação da variabilidade espacial em fruticultura. Trought e Bramley (2011) utilizaram sensoriamento proximal para mapeamento do vigor das plantas em viticultura para correlacionar essa variável com atributos importantes para detecção da qualidade dos frutos, tais como disponibilidade de sólidos solúveis, acidez e pH. Rodrigues Junior et al. (2011) utilizaram sensores proximais de índice de clorofila e análise foliar para definir zonas de manejo para uma cultura de café. Oldoni et al. (2017) utilizaram sensores de índice de clorofila e vegetação para definir zonas de manejo para irrigação em viticultura, com a expectativa de se obter regiões diferenciadas de produção relacionadas à qualidade do vinho que será produzido. Com relação à utilização desses dados, Souza et al. (2010) relatam que métodos geoestatísticos e de mineração de dados, quando combinados, podem transformá-los em informações mais robustas e assertivas.

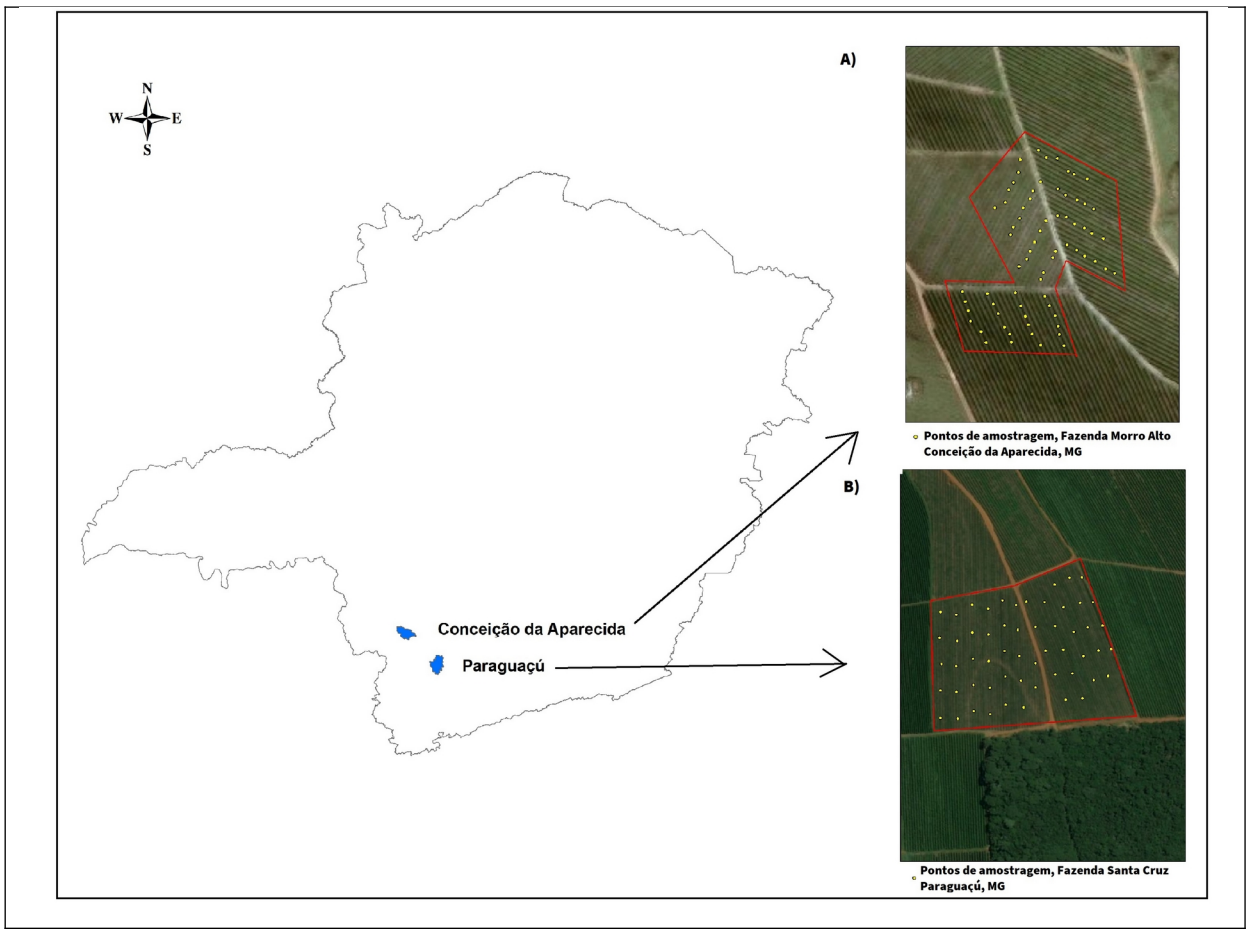
Diante desse contexto, este trabalho descreve um processo de mineração de dados realizado para investigar a correlação existente entre índices de vegetação e índice relativo de clorofila (IRC) com relação à face de exposição ao sol em culturas de café especiais, com o objetivo de definir variáveis capazes de auxiliar no delineamento de zonas de manejo que permitam diferenciar a qualidade do café produzido.

MATERIAL E MÉTODOS

Áreas de estudo

Os dados analisados no âmbito deste estudo foram coletados em trabalho de campo realizado em novembro de 2017, em duas áreas de produção de cafés especiais localizadas no sul de Minas Gerais. A Área 1 corresponde a parcelas de dois talhões, com variedade Catuaí, totalizando aproximadamente 5 ha da área de produção da Fazenda Santa Cruz, localizada no município de Paraguaçu (Figura 1a). Nessa área, foram identificadas duas faces distintas de exposição ao sol. A Área 2 corresponde a parcelas de 3 talhões, com as variedades Catucaí IAC-144, Obatã e Catuaí Amarelo, totalizando aproximadamente 1,5 ha da área de produção da Fazenda Morro Alto, localizada no município de Conceição da Aparecida (Figura 1b). Nessa área, foram identificadas 3 faces distintas de exposição ao sol.

Figura 1 – Localização e pontos de medidas das áreas de café especial em estudo: (a) Fazenda Santa Cruz, em Paraguaçu (MG); (b) Fazenda Morro Alto, em Conceição da Aparecida (MG).



Fonte: dos autores

Metodologia de coleta de dados

Considerando as diferentes faces de exposição ao sol para as duas áreas de estudo, foram definidos pontos para coleta de dados relacionados à biomassa e clorofila das plantas. Para a Área 1, foram georreferenciados 32 pontos localizados na face 1 e 24 na face 2. Para a área 2, foram georreferenciados 24 pontos localizados na face 1, 22 pontos na face 2 e 24 pontos na face 3.

Nos pontos amostrais, as medidas com o sensor Crop Circle[1] foram realizadas no terço médio das plantas (Área 1) e no terço superior (Área 2). Esse sensor mede seis diferentes bandas espectrais que cobrem as regiões do verde, vermelho, azul, infravermelho próximo e *red-edge*. Combinações dessas bandas permitiram a obtenção do Índice de Vegetação da Diferença Normalizada pelo Vermelho (NDVI) (ROUSE et al., 1973) e do Índice de Vegetação da

[1] Crop Circle, Holland Scientific, Lincoln, Nebraska, USA

Diferença Normalizada pelo *Red-Edge* (NDRE). Os dados de IRC foram coletados com o sensor SPAD-502[2], sendo as medidas realizadas em 10 folhas completamente expandidas, nos mesmos pontos e posições nos quais foram tomadas as medidas para o Crop Circle. Esses três atributos da cultura (NDVI, NDRE e IRC) foram selecionados e utilizados no trabalho de mineração de dados, que se iniciou com a análise estatística descrita a seguir.

Análise estatística

Os dados coletados em campo foram inicialmente analisados de maneira exploratória utilizando-se estatística descritiva. A seguir, foi realizada uma análise geoestatística composta pelo semivariograma experimental e ajustado para cada um dos atributos selecionados conforme indicado por Vieira (2000). Os semivariogramas forneceram os parâmetros de ajuste utilizados pelo método de interpolação espacial por krigagem ordinária (Oliver e Webster, 1990), para que fossem gerados os mapas de NDVI, NDRE e IRC para as três áreas com 1 metro de resolução espacial.

Análise de agrupamento e correlação entre mapas

A partir dos mapas gerados para os índices NDVI, NDRE e IRC, foram planejadas as análises de agrupamento considerando diferentes combinações de atributos de entrada, proporcionando a obtenção de potenciais mapas de zonas de manejo. Para a execução dos agrupamentos, foi utilizado o algoritmo Fuzzy c-means (FCM) (KITCHEN et al., 2005; CÓRDOBA et al. 2013; PERALTA et al., 2015; SCHENATTO et al., 2017). Por se tratar de um algoritmo de agrupamento particional, o FCM exige que seja definida a quantidade de grupos (neste contexto, zonas de manejo) em que o conjunto de dados de entrada deve ser dividido. Assim, especificamente para este trabalho, foram estabelecidas duas zonas para a Área 1 e três zonas para a Área 2, que correspondem à quantidade de faces de exposição ao sol pré-definidas para essas áreas.

Para efeito de comparação com os mapas de agrupamento gerados, as faces de exposição ao sol de cada uma das áreas foram rotuladas manualmente com a utilização de um Sistema de Informações Geográficas (SIG), considerando a mesma grade de pontos utilizada para a geração dos mapas de NDVI, NDRE e IRC. A correlação entre os mapas potenciais de zonas de manejo obtidos com diferentes combinações das variáveis NDVI, NDRE e SPAD com relação aos mapas rotulados de faces de exposição ao sol foi verificada a partir de duas

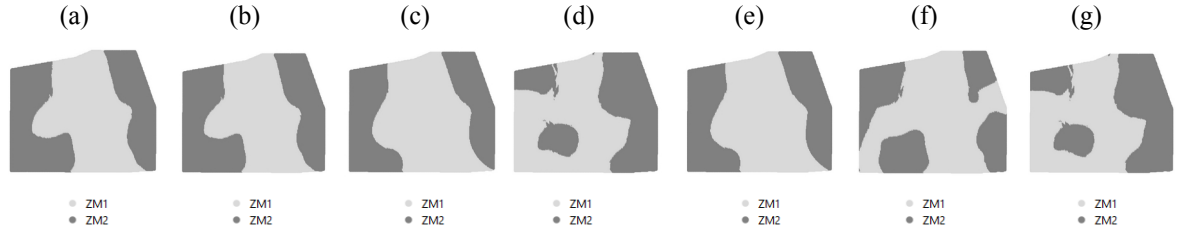
[2] SPAD-502, Konica Minolta, Tokyo, Japan

medidas distintas de concordância: função *agree*, disponível no pacote *irr* do software R (GAMER et al., 2012), que retorna a porcentagem de concordância entre os mapas; e a partir do índice *kappa* (COHEN, 1960), disponível no pacote *psych* do software R (REVELLE; REVELLE; 2015), cujo valor retornado permite classificar a concordância entre mapas em seis classes: pobre, leve, boa, moderada, substancial e quase perfeita.

RESULTADOS E DISCUSSÃO

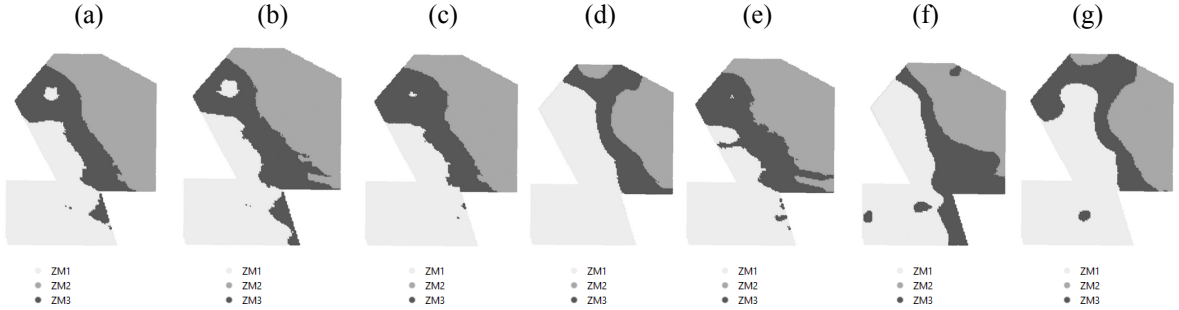
A partir da metodologia descrita na seção anterior, foram executadas análises de agrupamento considerando dados das duas áreas estudadas, utilizando sete combinações de atributos como dados de entrada para o algoritmo de agrupamento: I) NDVI, NDRE e IRC; II) NDVI e NDRE; III) NDVI e IRC; IV) NDRE e IRC; V) NDVI; VI) NDRE; e VII) IRC. Como resultado, foram obtidos mapas de potenciais zonas de manejo considerando as diferentes combinações de atributos para as Áreas 1 e 2 (Figura 2 e Figura 3).

Figura 2 – Potenciais zonas de manejo para a Área 1 (identificadas por diferentes tons de cinza), considerando diferentes combinações de índices de vegetação e clorofila: (a) I; (b) II; (c) III; (d) IV; (e) V; (f) VI; (g) VII.



Fonte: dos autores

Figura 3 – Potenciais zonas de manejo para a Área 2 (identificadas por diferentes tons de cinza), considerando diferentes combinações de índices de vegetação e clorofila: (a) I; (b) II; (c) III; (d) IV; (e) V; (f) VI; (g) VII.



Fonte: dos autores

A partir desses resultados, foram realizadas análises correlacionando os mapas de potenciais zonas de manejo com os mapas compostos pelas faces de exposição ao sol de cada

área. A Tabela 1 exibe os resultados obtidos, considerando os critérios de concordância descritos na seção anterior e as combinações de atributos de I a VII para a Área 1.

Tabela 1 – Resultados de concordância – Área 1

Índice	Combinações (Atributos de Entrada)						
	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)
agree (%)	45.1	54.8	51.2	71.9	50.8	54.7	76.1
kappa	-0.09	0.09	0.04	0.44	0.03	0.11	0.52
classe (*kappa*)	pobre	leve	leve	moderada	leve	leve	moderada

Por apresentar apenas duas faces distintas de exposição ao sol, a concordância dos potenciais mapas com duas zonas de manejo para a Área 1 com relação às faces utilizando o método *agree* apresentou valores maiores que 50% na maioria dos casos. Entretanto, como o cálculo do índice *kappa* é ponderado pela probabilidade de concordância entre rótulos das amostras, que nesse caso é de 50%, os resultados mostraram concordâncias no máximo moderadas, quando foram utilizados os atributos NDRE e IRC (IV) combinados e, principalmente, quando foi utilizado o atributo IRC de maneira isolada (VII). As combinações de atributos que utilizaram o NDVI (I, II, III e V) apresentaram concordância apenas leve na maioria dos casos. A mesma análise foi realizada para a Área 2, e os resultados obtidos são exibidos na Tabela 2.

Tabela 2 – Resultados de concordância – Área 2

Índice	Combinações (Atributos de Entrada)						
	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)
agree (%)	46.3	6.68	2.56	8.76	12.7	7.99	66.8
Kappa	0.52	-0.47	-0.51	-0.79	-0.41	-0.44	0.77
classe (*kappa*)	moderada	pobre	pobre	pobre	pobre	pobre	substancial

Para o caso da Área 2, a probabilidade de concordância entre os pontos diminui, pelo fato de estarem sendo utilizadas três faces de exposição ao sol. Desse modo, os índices obtidos pelo método *agree* se aproximaram mais dos índices fornecidos pelo método *kappa*. Para essa área, os melhores resultados foram obtidos utilizando-se todos os atributos (I), com concordância moderada entre os mapas; e utilizando apenas o atributo IRC (VII), com concordância substancial. Nesse caso, para as combinações restantes, que se utilizaram apenas dos índices NDVI e NDRE, a concordância foi pobre em todos os casos.

Em resumo, os resultados obtidos para as duas áreas mostraram alta correlação das faces de exposição ao sol com o índice IRC, medido pelo sensor SPAD-502. Os resultados mais próximos aos obtidos única e exclusivamente com os valores de IRC são por meio de

combinações que consideram também esse atributo como entrada, como nos casos da combinação IV para a Área 1 e da combinação I para a Área 2. Outra questão importante observada nos resultados é um melhor desempenho geral do índice NDRE com relação ao NDVI para esse contexto. Essa diferença pode ser explicada pelo fato dos dados terem sido coletados no início do período reprodutivo, logo após a floração, onde existia uma mistura de folhas completamente expandidas e outras com crescimento vegetativo inicial. Entretanto, por existirem poucos trabalhos na literatura com análises similares (RODRIGUES JUNIOR et al., 2011), as respostas ao NDVI e NDRE devem ser melhor exploradas em análises futuras, considerando cada estádio de desenvolvimento do café.

CONCLUSÕES

Os resultados deste artigo mostraram uma correlação substancial entre os dados obtidos com o sensor SPAD-502 e as faces de exposição ao sol estabelecidas nas duas áreas de cultivo. Desse modo, pode-se concluir que a variabilidade espacial da quantidade de clorofila presente nas plantas é importante para o delineamento de zonas de manejo individuais para as diferentes faces, com o objetivo de identificar cafés de maior e menor qualidade dentro de uma mesma área de cultivo. Os índices de biomassa NDVI e NDRE, amplamente utilizados para identificação de variabilidade espacial de diferentes culturas agrícolas, quando combinados com a quantidade de clorofila disponível, também apresentaram resultados satisfatórios. Entretanto, sua utilização para o delineamento de potenciais zonas de manejo capazes de determinar cafés de diferentes qualidades precisa ser melhor investigada.

O trabalho descrito neste artigo é apenas o início de atividades de pesquisa que ainda serão realizadas nas áreas piloto. Como trabalhos futuros, além da inclusão e análise de atributos de solo, tais como textura e condutividade elétrica, também está prevista a utilização de sensores capazes de monitorar o microclima das áreas. Esses dados servirão para potencializar as análises realizadas até o momento, proporcionando a tomada de decisão mais precisa com relação ao manejo da cultura e que priorize melhorias na qualidade do produto final obtido.

AGRADECIMENTOS

Agradecemos aos proprietários e funcionários das Fazendas Santa Cruz e Morro Alto, pela disponibilidade das áreas para realização desta pesquisa.

REFERÊNCIAS

COHEN, Jacob. A coefficient of agreement for nominal scales. Educational and psychological measurement, v. 20, n. 1, p. 37-46, 1960.

CONCAFÉ. Observatório do Café. Evolução da cafeicultura brasileira nas duas últimas décadas. Disponível em <http://www.sapc.embrapa.br/arquivos/consorcio/publicacoes_tecnicas/Consorcio-Embrapa-Cafe-Evolucao-12-3-19.pdf>. Acesso em 25 jun. 2019.

CÓRDOBA, M. et al. Subfield management class delineation using cluster analysis from spatial principal components of soil variables. Computers and electronics in agriculture, v. 97, p. 6-14, 2013.

FERREIRA, W.P.M. et al. As características térmicas das faces noruega e soalheira como fatores determinantes do clima para a cafeicultura de montanha. Série Documentos, v. 10. Brasília, DF: Embrapa Café, 2012, 34 p.

GAMER, Matthias et al. Package 'irr'. Various coefficients of interrater reliability and agreement, 2012.

GASPARI-PEZZOPANE, Cristina de et al. Influências ambientais no rendimento intrínseco do café. Bragantia, Campinas, v. 64, n. 1, p. 39-50, 2005.

KITCHEN, N. R. et al. Delineating productivity zones on claypan soil fields using apparent soil electrical conductivity. Computers and Electronics in Agriculture, v. 46, n. 1-3, p. 285-308, 2005.

OLDONI, H. et al. Vegetation indexes for management of irrigated wine vine orchard. In: II Latin-American Conference on Plant Phenotyping and Phenomics for Plant Breeding, 2, 2017, São Carlos, SP. Proceedings... São Carlos: Embrapa Instrumentação, 2017, 2017.

OLIVER, Margaret A.; WEBSTER, Richard. Kriging: a method of interpolation for geographical information systems. International Journal of Geographical Information System, v. 4, n. 3, p. 313-332, 1990.

PERALTA, N. R. et al. Delineation of management zones to improve nitrogen management of wheat. Computers and Electronics in Agriculture, v. 110, p. 103-113, 2015.

REVELLE, W.; REVELLE, M. W. Package 'psych'. The Comprehensive R Archive Network, 2015.

RODRIGUES JUNIOR, F. A. et al. Geração de zonas de manejo para cafeicultura empregando-se sensor SPAD e análise foliar. Revista Brasileira de Engenharia Agrícola e Ambiental, vol.15, n.8, pp. 778-787, 2011.

ROUSE JR, J. W. et al. Monitoring vegetation systems in the Great Plains with ERTS. 1974.

SCHENATTO, K. et al. Normalization of data for delineating management zones. Computers and Electronics in Agriculture, v. 143, p. 238-248, 2017.

SOUZA, Z. M. et al. Análise dos atributos do solo e da produtividade da cultura de cana-de-açúcar com o uso da geoestatística e árvore de decisão. Ciência Rural, v. 40, n. 4, p. 840-847, 2010.

SILVA, F. M.; ALVES, M. C. Cafeicultura de precisão. Lavras: Editora UFLA, 227 p., 2013.

TROUGHT, M. C. T; BRAMLEY, R. G. V. Vineyard variability in Marlborough, New Zealand: characterising spatial and temporal changes in fruit composition and juice quality in the vineyard. Australian Journal of Grape and Wine Research, v. 17, n. 1, p. 79-89, 2011.

VIEIRA, S. R. Uso de geoestatística em estudos de variabilidade espacial de propriedades do solo. In: NOVAIS, R. F.; ALVAREZ V, V. H.; SCHAEFER, C. E. G. R. (Org.). Tópicos em Ciência do Solo. Viçosa: Sociedade Brasileira de Ciência do Solo, 2000. v. 1, p. 1-54.

COMPARAÇÃO DO DESEMPENHO ENTRE MODELOS LINEARES PARA PREVISÃO DO PREÇO DO CAFÉ

Carolina Deina[1], Diego Solak Castanho[2], Matheus Henrique do Amaral Prates[2], Sergio Luiz Stevan Junior[2], Fernanda Cristina Corrêa[2], Hugo Siqueira[2], Flavio Trojan[1]

[1] Departamento de Engenharia de Produção, Universidade Tecnológica Federal do Paraná, Pato Branco, Paraná, Brasil, caroldeina@gmail.com, trojan@utfpr.edu.br

[2] Departamento de Engenharia Elétrica, Universidade Tecnológica Federal do Paraná, Ponta Grossa, Paraná, Brasil, matheushaprates22@gmail.com, diegocastanho@alunos.utfpr.edu.br, sstevanjr@gmail.com, fernandacorrea@utfpr.edu.br, hugosiqueira@utfpr.edu.br

RESUMO

O café há mais de um século é um importante ativo para a economia Brasileira. Por conta disso, este trabalho buscou investigar o comportamento de algumas ferramentas para a previsão de preço de café tipo arábica e robusta. A metodologia fundamentou-se na realização de uma comparação entre dois métodos lineares de previsão: modelo Auto Regressivo (AR) e Suavização Exponencial (SE). Os dados utilizados foram obtidos pelos Centro de Estudos Avançados em Economia Aplicada no Brasil (CEPEA). Após a normalização dos dados, foram identificados os atrasos mais relevantes para então realizar as previsões para 1, 3, 6 e 12 passos à frente sobre os últimos 24 meses de cada série de dados. Para as séries analisadas, o modelo AR obteve os melhores desempenhos gerais.
PALAVRAS-CHAVE: Café Arábica, Café Robusta, Previsão de preços, Modelos lineares.

ABSTRACT

Coffee is an important asset for the Brazilian economy. Because of this this work sought to investigate the behavior of some tools for forecasting coffee prices of the type Arabic and robust. The methodology was based on the realization of a comparison between two linear prediction methods: Auto Regressive (AR) model and Exponential Smoothing (SE). The data used in this research were obtained by the Center for Advanced Studies in Applied Economics in Brazil (CEPEA). After the normalization of the data, the most relevant delays were

identified to then carry out the forecasts for 1, 3, 6 and 12 steps ahead. For the analyzed series, the AR model achieved the best overall performances.
KEYWORDS: Arabica Coffee, Robusta Coffee, Price Forecast, Linear Models.

INTRODUÇÃO

O Brasil mantém a posição de maior produtor mundial de café desde a década de 1880 (ICO, 2019). Para o período de 2017-2018 foram produzidos cerca de 51 milhões de sacas de 60 kg, representando 31,94% da produção mundial. Já nos anos de 2018-2019, a produção atingiu seu recorde (60,1 milhões de sacas), representando um crescimento de 17,84% em relação ao período anterior (ICO, 2019). Em fevereiro de 2019, também alcançou recordes de exportação, enviando para o comércio exterior cerca de 3,4 milhões de sacas de 60 kg, número 36,3% superior ao mês de fevereiro de 2018 (ICO, 2019).

A representação expressiva do café para a economia brasileira, não está apenas na produção e exportação do grão, já que o Brasil também é um grande consumidor. Em 2018, foram vendidos internamente cerca de 20,5 milhões de sacas de 60 kg de café, o que fez do país o segundo maior consumidor da bebida no mundo, ficando atrás apenas dos Estados Unidos (ICO, 2019).

O consumo interno e a exportação do grão movimentam a balança do comercio nacional. Os agricultores são grandes responsáveis por contribuírem significativamente para essa movimentação mas, apesar de seus esforços, suas vidas financeiras não estão asseguradas. A causa desse problemática, segundo Hong (2016) e Cuaresma et al. (2018), está relacionada a volatilidade e tendências nos preços do mercado mundial. Fatores como esse deixam evidente a importância de se estudar estratégias que possam gerenciar riscos. Conforme Cuaresma et al. (2018), metodologias de previsão de preços de commodities cada vez mais são recomendadas para essa finalidade. Quando os resultados gerados pelas previsões são consistentes, não auxiliam apenas na tomada de decisão do produtor, mas também em toda a cadeia produtiva do grão (BACCI et al., 2019).

Estudos recentes têm sido desenvolvidos sobre esse tópico em diferentes partes do mundo. Milas et al. (2004) realizaram previsões de preços de quatro qualidades de café utilizando modelos de correlação de erro linear e não linear. Fousekis (2017) utilizou previsões para valor de risco (VaR) para prever preços de duas séries de café. Yashavanth et al. (2017) realizaram a previsão de preço da mercadoria na Índia utilizando modelos de previsão como o vetor auto regressivo (VAR) e o ARIMA tradicional e a versão fracionada

para prever preços de quatro diferentes tipos de grãos. Cuaresma et al. (2018) fizeram a previsão dos preços de café utilizando outras variáveis financeiras e macroeconômicas.

Visto que modelos lineares de previsão são de grande aplicação em diversos estudos correlatos, este trabalho tem como objetivo comparar o desempenho da Suavização Exponencial (SE) e Auto Regressivo (AR) para previsão de preços dos cafés Arábica e Robusta no Brasil.

Nas próximas seções serão apresentados os métodos SE e AR, bem como a metodologia utilizada para tratamento e previsão dos dados, e os resultados e conclusão encontrados nesse estudo.

SUAVIZAÇÃO EXPONENCIAL

Três variações das técnicas de Suavização Exponencial (SE) são amplamente utilizadas na literatura em decorrência da sua simplicidade de aplicação e resolução de problemas de previsão. São elas: suavização exponencial simples, suavização exponencial de Holt e suavização exponencial de Holt-Winters (TRATAR et al., 2016). A popularidade desses modelos é devida a sua flexibilidade para resolver problemas em que a série temporal apresente comportamento sazonal, tendencioso e/ou cíclico (TAYLOR, 2003).

O método SE simples, é utilizado quando uma série apresenta apenas um comportamento cíclico, o qual é suavizado por uma constante (α), com valor entre 0 e 1 (MENEGHINI et al., 2018). Já o método SE de Holt é adequado quando, além de ciclicidade, a série apresentar tendência. As constantes que suavizam esses comportamentos são: de suavização de ciclo (α) e de tendência (β), os quais possuem valores que variam entre 0 e 1 (GARDNER; MCKENZIE, 1989). Seu cálculo é expresso na Equação (1) e os parâmetros de nível e tendência são encontrados, respectivamente através das Equações (2) e (3) (HYNDMAN; ATHANASOPOULOS, 2018).

$$\hat{Z}_{t+n} = L_t + n\,T_t \tag{1}$$

$$L_t = \alpha\, Z_t + (1-\alpha).(L_{t-1} + T_{t-1}) \quad 0 \le \alpha \le 1 \tag{2}$$

$$T_t = \beta(L¿¿t - L_{t-1}) + (1-\beta).T_{t-1}¿ \quad 0 \le \beta \le 1 \tag{3}$$

sendo que $\hat{Z}_{t+n}$representa o valor previsto para o período, $t+n$, $(1-\alpha)$ é taxa exponencial dos pesos de ponderação dos dados,$\hat{Z}_t$é o valor real para o período de tempo *t, n* representa períodos de tempos e $(1 - \beta)$ é taxa exponencial dos pesos de ponderação dos dados.

Por fim, o método SE de Holt-Winters é utilizado quando, além de ciclo e tendência, a série apresentar padrões de sazonalidade, os quais são suavizados pela constante (γ) (HYNDMAN; ATHANASOPOULOS, 2018).

MODELO AUTORREGRESSIVO

O modelo Autoregressivo (AR) é comumente utilizado para séries que são invariáveis ao longo do tempo (HYNDMAN; ATHANASOPOULOS, 2018). O modelo AR de ordem (p), é representado pela Equação (4).

$$y_t = \varphi_1 y_{t-1} + \ldots + \varphi_p y_{t-p} + \varepsilon_t \quad (4)$$

a qual considera que uma série temporal de dados é representada por y_t, sendo $y_{t-1} + \ldots + y_{t-p}$ os termos passados da série, φ_p os coeficientes livres e ε_t o erro de previsão intrínseco.

Os valores dos coeficientes livres desse modelo são definidos de forma determinística por meio da aplicação direta das equações de Yule-Walker (BOX et al., 1994).

METODOLOGIA

Para realização das previsões, inicialmente coletaram-se dados de preços do café das qualidades Arábica e Robusta da base de dados disponível no site do Centro de Estudos Avançados em Economia Aplicada da Escola Superior de Agronomia Luiz de Queiroz (ESALQ) - Universidade de São Paulo (USP) (CEPEA, 2019). Os dados mensais do café Robusta correspondem ao período de novembro de 2001 até dezembro de 2018, enquanto os do Arábica estão entre setembro de 1996 e dezembro de 2018. Essa média de preços mensais (em reais) é referente a um saco de 60 kg, negociado no mercado *spot* do estado de São Paulo.

Tendo os dados coletados, o próximo passo foi fazer uma análise gráfica dos mesmos para identificar se havia presença de comportamento tendencioso, cíclico, sazonal e aleatório (flutuação). Com base nesse estudo definiu-se o modelo de suavização exponencial utilizado para a previsão.

Verificou-se que os dados possuíam tendência na variação dos dados. Para retirar tal comportamento, aplicou-se uma diferenciação (Equação 5) para torná-los estacionários e então ser possível definir o parâmetro (p) do AR.

$$y'_{t+1} = y_{t+1} - y_t \quad (5)$$

em que y'_t é a diferença entre o valor atual da série menos o valor anterior.

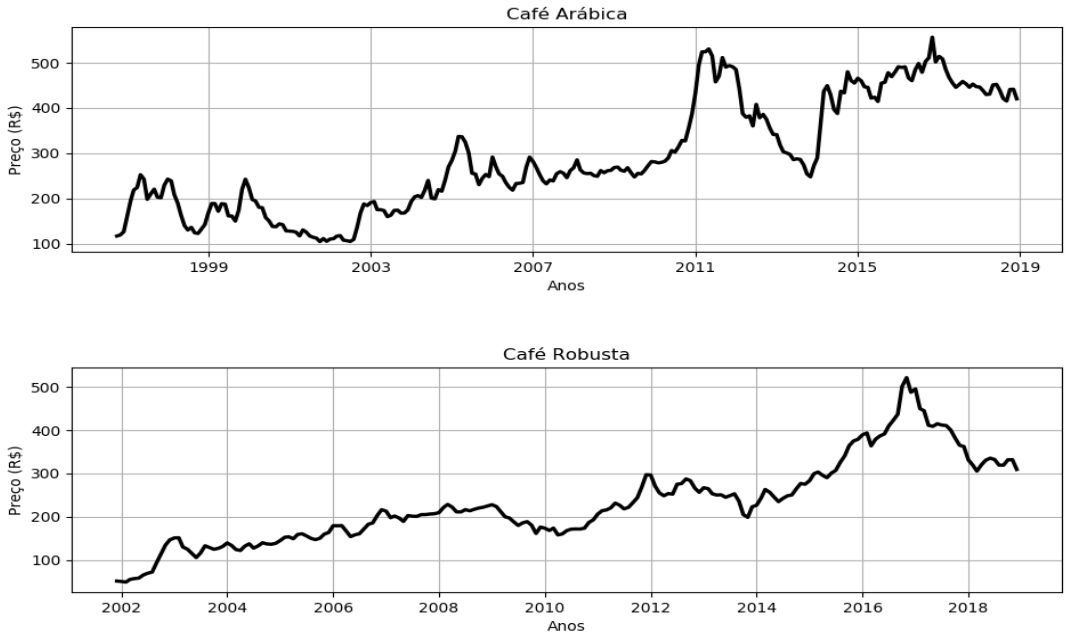
Definidos os pormenores da aplicação do SE e AR, as previsões foram realizadas com o auxílio do software Minitab®, para 1 (próximo mês), 3 (próxima estação), 6 (estação

inversa) e 12 (próximo ano) passos à frente. Neste caso, a previsão multi-passos é realizada de forma recursiva (BELOTTI et al., 2019). Ao final, os resultados das previsões pelos modelos propostos foram comparados pelo Erro Quadrático Médio (MSE), Erro Absoluto Médio (MAE) e Erro Percentual Absoluto Médio, descritos na próxima seção (MAPE) (SILVA et al., 2018; BACCI et al., 2019).

RESULTADOS E DISCUSSÃO

Seguindo a metodologia proposta, os 271 dados mensais de preços do café Arábica e os 206 dados mensais de preços praticados do café Robusta foram inicialmente coletados na base de dados do CEPEA os quais foram colocados em um gráfico para identificar seu padrão de comportamento (Figura 1). Isto contribui para verificar sua regularidade e consistência.

Figura 1 – Dados de preços praticados para o café Arábica e Robusta referente ao período de setembro de 1996 até dezembro de 2018 e novembro de 2001 até 2018, respectivamente.



Fonte: Autoria própria

Pela análise gráfica da Figura 1, verificou-se que as duas séries possuem tendência de crescimento dos preços ao longo do tempo, ciclos de variação e ausência de padrões sazonais.

Dessa forma, o modelo de SE escolhido para realizar as previsões foi o Holt (MENEGHINI et al., 2018). De acordo com Montgomery et al. (1990) tem os parâmetros que

suavizam as constantes de ciclo α e tendência β, encontrados através de testes paramétricos, pois não há um método absoluto ou determinístico para essa definição.

Sendo assim, fizeram-se testes para diferentes parâmetros que variavam entre 0 e 1 e como melhores resultados, encontrou-se para o café arábica α=0,9 e β=0,3 e para o Robusta α=0,9 e β=0,6.

Tendo os parâmetros de SE de Holt definidos, o próximo passo foi verificar a variação da série de dados. Para isso, fez-se uma análise de estatística descritiva (Tabela 1).

Tabela 1 – Análise estatística descritiva dos dados antes e após a diferenciação

Estatística Descritiva	**Antes da diferenciação**		**Após a diferenciação**	
	Café Arábica	Café Robusta	Café Arábica	Café Robusta
Desvio padrão	96,747	123,976	12,369	19,266
Variância da amostra	9360,035	15369,929	152,999	371,167

Segundo Morettin e Bussab (2010) uma série de dados com comportamento homogêneo apresenta desvio padrão e variância próximo de zero e, quanto menor for a variância, mais próximos da média eles estarão.

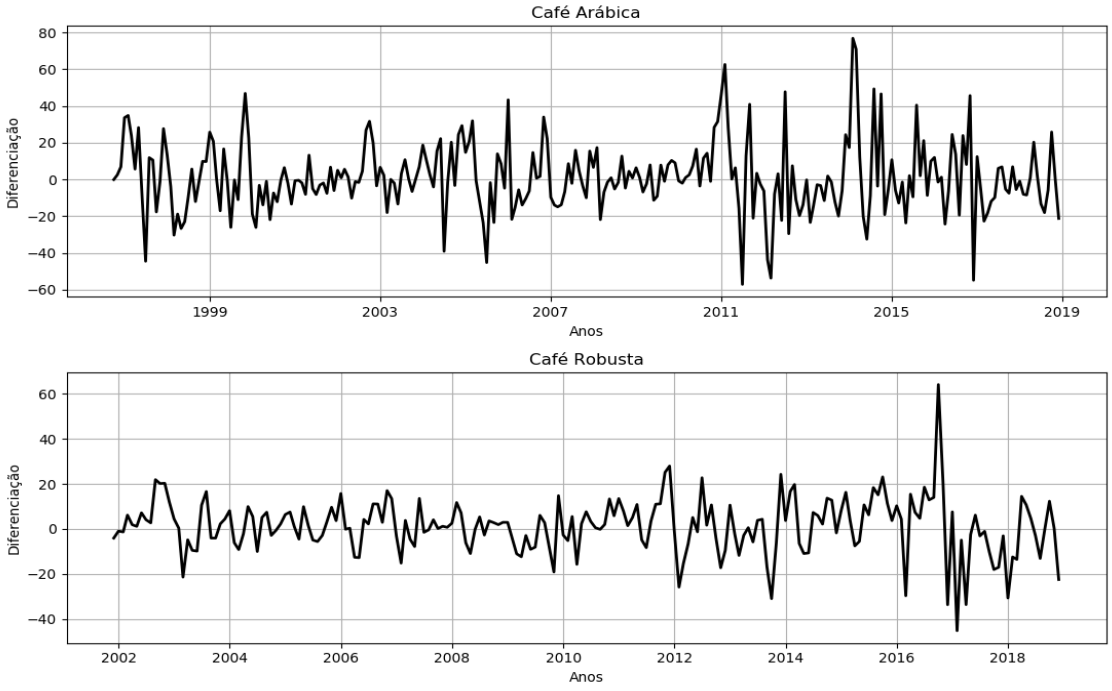
Percebe-se que o valor do desvio padrão em relação aos dados observados foi de 96,747 e 123,976, e a variância foi 9360,035 e 15369,929, respectivamente, em relação as séries Arábica e Robusta, indicando a existência de variabilidade nos dados. Portanto, para torná-las estacionárias, aplicou-se uma diferenciação (Equação 5), a qual reduziu os valores de desvio padrão e variância para a série Café Arábica em 87,21% e 98,37% respectivamente, enquanto que para a série Café Robusta ouve uma redução do desvio padrão de 84,46% e de 97,59% da variância. Graficamente, essa redução de variabilidade dos dados após a diferenciação de ambas as séries pode ser visualizada na Figura 2.

Em seguida, analisaram-se os dados pela função de autocorrelação parcial (FACP) (SIQUEIRA et al., 2018), a qual possibilitou identificar que os preços do café Arábica possuem coeficiente autocorrelação significativo até o segundo atraso, enquanto o Robusta apresenta autocorrelação até o quarto atraso. Fizeram-se testes considerando esses atrasos e constatou-se que o modelo AR (1) foi o que chegou aos menores erros de teste para ambas as séries analisadas.

Com os parâmetros dos modelos definidos, fez-se a previsão de preços de café para 1, 3, 6 e 12 passos à frente, referentes aos 24 meses dos anos de 2017 e 2018. A Tabela 2 mostra os resultados obtidos dos conjuntos de testes pelas medidas de erro MSE, MAE e MAPE.

Considerando o desempenho geral dos modelos para a série de café Arábica, o modelo AR obteve os melhores desempenhos para *P*=1, 3, 6 e 12 passos à frente, exceto para o MAPE e *P*=1. Como esperado, previsão para 1 passo à frente obteve a maior acurácia entre todas.

Figura 2 – Dados diferenciados de preços praticados para o café Arábica e Robusta, referente ao período de setembro de 1996 até dezembro de 2018 e novembro de 2001 até 2018, respectivamente.



Fonte: Autoria própria

Para a série Robusta, a previsão do modelo SE com um passo à frente obteve o melhor desempenho de previsão se considerado o MSE. Entretanto, para todos os demais horizontes e demais métricas de desempenho, o modelo AR apresentou resultados superiores em relação ao SE. Uma observação importante é que na medida em que o horizonte se eleva, é esperado que as métricas de erro sejam majoradas. Entretanto, observa-se uma irregularidade neste caso, já que para *P*=6 e *P*=12 isso quase não se observa. Este fator mostra que a complexidade do problema é elevada. As razões para isso podem estar relacionadas à volatilidade do preço e ao comportamento cíclico ao longo do ano.

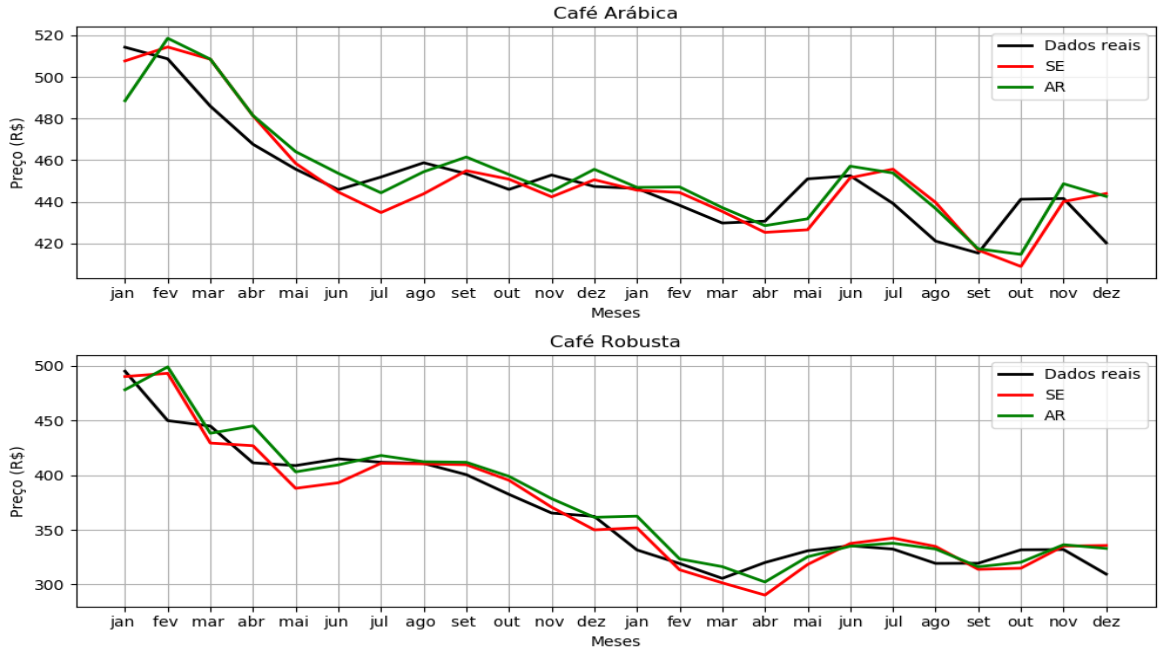
Tabela 2 – Erro Quadrático Médio (MSE), Erro Absoluto Médio (MAE) e Erro Percentual Absoluto Médio (MAPE) para as previsões dos modelos SE, AR e ARIMA.

Número de pas-	**Café Arábica**	**Café Robusta**

sos à frente	1	3	6	12	1	3	6	12
MSE								
SE	181,553	594,331	1227,368	934,625	**274,795**	792,153	1257,205	2584,559
AR	**172,816**	**180,721**	**186,648**	**185,004**	288,633	**291,546**	**295,766**	**284,433**
MAE								
SE	**10,071**	19,787	29,080	23,955	13,124	22,187	29,551	45,506
AR	10,924	**10,857**	**11,070**	**10,930**	**12,418**	**13,307**	**13,287**	**13,035**
MAPE								
SE	**2,244**	4,458	6,535	5,312	3,592	6,324	8,182	12,450
AR	2,407	**2,405**	**2,449**	**2,418**	**3,344**	**3,633**	**3,612**	**3,536**

Na Figura 3 é exposto cada previsão feita com os modelos descritos para 1 passo à frente, o qual apresentou os melhores resultados de previsões para as duas séries analisadas.

Figura 3 – Gráfico da comparação dos dados reais com os dados previstos para 1 passo à frente referentes aos 24 meses dos anos de 2017 e 2018, utilizando os modelos SE e AR.



Fonte: Autoria própria

CONCLUSÕES

Esse estudo investigou a aplicação de modelos lineares de Suavização Exponencial (SE) e Auto Regressivo (AR), para a previsão dos preços praticados de café Arábica e Robusta, usando dados reais do CEPEA-USP.

Para todos os modelos considerou a previsão dos últimos 24 meses de ambas as séries, com horizontes de previsão de 1, 3, 6 e 12 passos à frente. Os parâmetros de entrada para os modelos foram definidos conforme as características das séries e seus atrasos e com base na

função de autocorrelação parcial. As previsões multi-passos foram realizadas de forma recorrente.

Como resultado, o modelo AR foi superior em praticamente todos os 8 cenários propostos, com exceção ao caso de P=1 considerando o MSE para o café Robusta e MAPE para o Arábica. Ademais, a elevação do horizonte, em geral, acarretou em uma piora no desempenho dos preditores. Entretanto, em alguns casos, considerando 6 e 12 passos adiante, o segundo caso levou a resultados um pouco melhores, algo pouco usual para este tipo de problema.

Para trabalhos futuros, recomenda-se a utilização de outros modelos de previsão lineares como o modelo Auto Regressivo Integrado de Médias Móveis (ARIMA) e também modelos computacionais como as Redes Neurais Artificiais, os quais já são aplicados à previsão de séries de outras naturezas (SIQUEIRA et al., 2018; SILVA et al., 2018). Na tentativa de obter melhores resultados, outras variáveis que tem influências sobre o preço do café também podem ser testadas como entradas para os modelos de previsão, como fatores climáticos e econômicos.

AGRADECIMENTOS

Os autores agradecem à Universidade Tecnológica Federal do Paraná (UTFPR), à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), projeto Universal do processo nº 405580/2018-5, pelo apoio financeiro ao projeto de pesquisa.

REFERÊNCIAS

BACCI, L. A.; MELLO, L. G.; INCERT, T.; PAIVA, A. P.; BALESTRASSI, P. P. (2019). Optimization of combined time series methods to forecast the demand for coffee in Brazil: A new approach using Normal Boundary Intersection coupled with mixture designs of experiments and rotated factor scores. *International Journal of Production Economics.* 212: 186–211.

BELOTTI, J. T.; LUNA, I.; SIQUEIRA, H. (2019). Avaliação de abordagens alternativas de previsão no estudo de séries de vazões. Journal of Applied Instrumentation and Control, v. 6, p. 19-26, 2019.

BOX, G. E. P.; JENKINS, G.M.; REINSEL G. (1994). Time Series Analysis: Forecasting & Control. Prentice Hall, New Jersey.

CEPEA - Centro de Estudos Avançados em Economia Aplicada (ESALQ/USP). Disponível em: < http://cepea.esalq.usp.br>. Acesso em: 4 fev. 2019.

CUARESMA, C. J.; HLOUSKOVA, J.; OBERSTEINER, M. (2018). Fundamentals, speculation or macroeconomic conditions? Modelling and forecasting Arabica coffee prices. *European Review of Agricultural Economics*. 45: 583–615.

FOUSEKIS, P. (2017). Price co- movement and the hedger's value-at-risk in the futures markets for coffee. *Agricultural Economics Review.* 18: 35–47.

GARDNER JR., E.S., MCKENZIE, E., 1989. Seasonal exponential smoothing with damped trends. Manag. Sci. 35, 372–376.

HONG, T. T. K. (2016). Effects of exchange rate and world prices on export price of Vietnamese coffee. *International Journal of Economics and Financial Issues*. 6: 1756–1759.

HYNDMAN, R. J.; ATHANASOPOULOS, G. (2018). Forecasting: principles and practice. Melbourne, Australia.

ICO - International Coffee Organization. coffe Brazil. Disponível em: <http://www.ico.org/>. Acesso em: 4 abr. 2019.

MENEGHINI, M.; ANZANELLO, M.; KAHMANN, A.; TORTORELLA, G. L. (2018). Ajuste de previsão de demanda quantitativa com base em fatores qualitativos: estudo de caso em um restaurante fast food. *Sistemas & Gestão*. 13: 68-80.

MILAS, C.; OTERO, J.; PANAGIOTIDIS, T. (2004). Forecasting the spot spices of various coffee types using linear and non-linear error correction models. *International Journal of Finance and Economics*. 9: 227–288.

MONTGOMERY, D.; JOHNSON, L.; GARDINER, J. (1990). Forecasting and Time Series Analysis. New York: McGraw-Hill, 1990.

MORETTIN, P. A.; BUSSAB, W. O. (2010). Estatística básica. São Paulo: Saraiva.

SILVA, N.; SIQUEIRA, I.; OKIDA, S.; STEVAN JR., S. L.; SIQUEIRA, H. (2018). Neural Networks for Predicting Prices of Sugarcane Derivatives. *Sugar Tech*. DOI: 10.1007/s12355-018-0648-5.

SIQUEIRA, H.; BOCCATO, L.; LUNA, I. ; ATTUX, R. ; LYRA, C (2018). Performance analysis of unorganized machines in streamflow forecasting of Brazilian plants. Applied Soft Computing, v. 68, p. 494-506, 2018.

TAYLOR, J. W. (2003). Exponential smoothing with a damped multiplicative trend. *International Journal of Forecasting*. 19: 715– 725.

TRATAR, L. F., MOJŠKERC, B., TOMAN, A., 2016. Demand forecasting with four-parameter exponential smoothing. *International Journal of Production Economics.* 181, 162–173.

WINTERS, P. R. (1960). Forecasting sales by exponentially weighted moving averages. *Management Science.* 6: 324-342.

YASHAVANTH, B. S.; SINGH, K. N.; PAUL, A.; PAUL, R. K. (2017). Forecasting prices of coffee seeds using Vector Autoregressive Time Series Model. *Indian Journal of Agricultural Sciences*. 87: 754-758.



Computer Vision and Image Processing for Detecting and Quantifying whiteflies: A Systematic Review

João Gabriel Junqueira da Silva[1], *Marcelo Gonçalves Narciso*[2], *Cristhiane Gonçalves*[1]

[1]Instituto de Informática, Universidade Federal de Goiás
Goiânia, Goiás, Brasil
joaogabrieljunq@gmail.com, cristhianega@gmail.com

[2]Embrapa Arroz e Feijão
Santo Antônio de Goiás, Goiás, Brasil
marcelo.narciso@embrapa.br

RESUMO

A evolução da agricultura de precisão impulsionou o aumento da produção agrícola e proporcionou o surgimento de grandes desafios em termos de eficiência, segurança alimentar, sustentabilidade e impacto ambiental. Especificamente, o controle de pragas é de suma importância no que tange a saúde das plantações. Este artigo apresenta uma revisão sistemática da literatura, abordando técnicas e métodos de processamento digital de imagens em plantas para a detecção e classificação de pragas, em específico para a praga Mosca Branca. Nesta revisão foram considerados artigos publicados entre 2008 e 2018 indexados por cinco bases de dados científicas. A revisão identificou que as técnicas são bem exploradas, diversificadas e possuem alto desempenho em termos de acurácia. No entanto há uma dificuldade no que tange a detecção de moscas em estágio ninfal, e principalmente técnicas aplicadas a campo aberto e dispositivos portáteis, o que daria maior autonomia à agricultores e cientistas. Ideias futuras são discutidas no final do artigo com base nas dificuldades relatadas e trabalhos futuros propostos.

PALAVRAS-CHAVE: Visão computacional, classificação de imagens, agricultura de precisão, mosca-branca.

ABSTRACT

The evolution of precision agriculture increased agricultural production and provided major challenges in terms of efficiency, food safety, sustainability and environmental impact. Specifically, the pest management is extremely important healthy crop. This paper presents a systematic review of literature, addressing techniques and methods of digital image processing in plants for the detection and classification of pests, specifically for whitefly pest. In this review,

articles published between 2008 and 2018 indexed by five scientific databases were considered. The review identified that the techniques are well explored and have high performance in terms of accuracy. However, there is a problem with the detection of whiteflies in the nymphal stage, mainly techniques applied to the open field and portable devices, which would give greater autonomy to farmers and scientists. Future ideas are discussed at the end of the article based on the difficulties reported and future works were proposed.
KEYWORDS: Computer vision, image classification, precision agriculture, whiteflies.

INTRODUCTION

Precision agriculture can be considered as an exponent of general future development through the Internet of Things (IoT) and Big Data (WOLFERT et al., 2017). Modern technologies have given human society the ability to produce enough supplies to feed the demand of more than 7 billion people (MOHANTY; HUGHES; SALATHé, 2016). However, with the increase in production capacity, great challenges have arisen in terms of efficiency, food safety, sustainability and environmental impact. To address these challenges, the complex, multivariate and unpredictable agricultural ecosystems need to be better understood by monitoring, measuring and analyzing continuously various physical aspects and phenomena (KAMILARIS; PRENAFETA-BOLDú, 2018).

Image processing techniques (IPTs) can be used to enhance agricultural practices, by improving accuracy and consistency of processes while reducing farmers manual monitoring. Often, it offers flexibility and effectively substitutes the farmers visual decision making (SAXENA; ARMSTRONG, 2014). During the last several years, spectacular advances in the fields of artificial intelligence, computer vision, and deep learning have resulted in remarkable performance in image classification and vision tasks. (SCHWARTZMAN et al., 2016).

Linking this need to the emergence of new techniques, this paper aims to explore methods and tools of IPTs through a systematic review of the bibliography, with the objective of presenting the state of the art in the use of IPTs to identify the whitefly pest in images of plants.

WHITEFLY

The whitefly (*Bemisia tabaci*) is a major pest to agriculture. Adults are able to fly for long distances and to colonize staple crops, herbs and ornamentals, and to vector viruses belonging to several important taxonomic groups (BAR et al., 2019). If affects several cultures like pumpkin, zucchini, cotton, eggplant, broccoli, cauliflower, peas, beans, papaya, cassava, watermelon, melon, cucumber, pepper, cabbage, soybean and many others. The easiest way to control the pest infection is the use of pesticides. But the more use of pesticide are hazardous and it not only kill pest in plant , but also affect the health of human, animal and plant. To overcome this problem it is necessary to control the use of pesticide, pest detection is the most important process for an effective cultivation (JIGE; RATNAPARKHE, 19-2).

Figure 1: Leaf infested with whiteflies



Fonte (Embrapa, 2019)

Barbedo (2014) explains the most direct way of measuring whitefly infestation is to manually identify and count the insects inside a selected region. In general, this approach does not require sophisticated apparatus and, more importantly, relies on the remarkable human ability to resolve ambiguities and unclear situations, even under far from ideal conditions. On the other hand, human beings are susceptible to physiological and psychological phenomena that may be important sources of error: fatigue, visual illusions, boredom, among others. Also, humans are usually much slower than machines in performing simple tasks like counting.

Image processing has been proved to be effective tool for analysis in various fields and applications (VIBHUTE; , 2012). In this paper we explore some of these techniques to detecting and counting the whitefly pest.

MATERIAL AND METHODS

In this research, we performed a Systematic Review of the Literature (SRL), since it offers a organized and well-defined process. Mallett et al. (2012) defines SRL as a rigorous and transparency form of literature review. Three main stages were taken following the SRL guideline defined by Kitchenham e Charters (2007): (I) defining a review protocol; (II) conducting the review of the studies; and (III) reporting the review.

The review protocol definition include the definition of the research questions, keywords and search strategy, inclusion and exclusion criteria for the selection of studies and a checklist for study quality assessment. These elements is detailed next. To support this process and provide a collaborative way between the authors a software platform were used: Parsifal (https://parsif.al/).

Research questions

The main objective of this SRL was to answer the following question: “What is the state of the art in the application of digital image processing techniques for acquisition, detection and counting of Whitefly pest in images?”.

Some more specific questions that unfolded the previous one were formulated: (RQ.1) What are the existing methods for acquisition and preprocessing images for Whitefly pest? (RQ.2) What are the techniques used for detection and classification of pests in plants, specifically for the Whitefly pest? (RQ.3) What are the results obtained so far for detection and classification for the white fly pest? (RQ.4) What measures are used to evaluate the effectiveness of existing methods?

Keywords, synonyms and search strategy

The following keywords were chosen for driving the SRL, considering the usually techniques therms and pest scientific and popular name: (K.1) computer vision; and (K.2) deep learning; (K.3) neural networks; (K.4) image processing; (K.5) image classification; (K.6) plague detection; (K.7) whiteflies; (K.8) bemisia tabaci; (K.9) plant disease.

The general logic search string was defined as (”Computer vision” OR ”Deep learning” OR ”Neural networks” OR ”Image processing” OR ”Image classification”) AND (”Plague detection” OR ”Whiteflies” OR ”Bemisia tabaci” OR ”Plant disease”).

This string was adjusted for every electronic database in order to meet its search syntax. We searched on five databases: (1) ACM Digital Library; (2) IEEE Digital Library; (3) ISI Web of Science; (4) Science Direct; and (5) Scopus, since they gather most of the publications on the area of computer science and applications.

Inclusion/Exclusion criteria

Inclusion and exclusion criteria were defined for guiding the selection of primary studies relevant to the SRL. A study was included if it was: (1) available in full-text; (2) published in a journal or in the annals of a conference; (3) a technical report, including surveys; or (4) a master dissertation or a doctoral thesis. On the other hand, it was excluded if it: (1) was not written in English; (2) was published before 2008; (3) present evaluations without describing the used method; (4) do not use digital image processing techniques; (5) do not use digital image processing techniques for the whitefly pest detection.

A set of Yes/No/Partially quality assessment questions was also defined based on a refinement of the main research questions (RQs). For instance, they asked if the classification methods or techniques have been objectively reported or if there is an application or final product, has it been described in detail.

Conducting the review

The proper search strings, with the corrected search conditions, were applied to the databases. Next, as a Study Selection step, the bibliographic references were evaluated according to the inclusion and the exclusion criteria (with the help of the SRL software platforms). When necessary, the full text of the studies was consulted. Table 1 summarizes the number of papers

recovered when searching in each database, and the amount of accepted (selected), reject and duplicate studies during the Selection step. In Total, 25 studies were accepted.

Table 1: Number of studies in the SRL.

Database	Total	Accepted	Rejected	Duplicated
ACM DL	103	4	82	17
IEEE DL	94	3	89	2
WebOfScience	129	5	69	55
ScienceDir	81	9	64	8
Scopus	31	4	12	15

The studies were included according to the inclusion and exclusion criteria, and in addition, the following quality criteria were applied: (1) Have classification methods or techniques been objectively reported?; (2) Were the methods of acquisition and preprocessing explicitly cited?; (3) Was the technique used compared to other existing techniques?; (4) Is there an evaluation of this technique?; (5) If there is an application or final product, has it been described in detail?.

These criteria were elaborated based on the research questions and for each criterion a value was assigned in conjunction with a concept: Yes (S) = 1 point, No (N) = 0 and Partially (P) = 0.5, where the junction of all the criteria form an index that was classified as follows, works with index between 0 and 1 were considered very bad, indexes between 1.5 and 2 were considered bad, works between 2.5 and 3 were considered good, between 3.5 and 4 work was considered very good, and the works that obtained between 4.5 and 5 were considered excellent. The value of the index indicated the priority for reading the articles and was also a factor for the exclusion of articles if the work has a very low index and there is no specific contribution, which would justify the low index. After the quality assessment, 21 studies were selected for full read.

Next, we describe the main concepts in some of the studies selected by the SRL.

RESULTS AND DISCUSSION

Barbedo (2014) explains there are two main strategies for automatically counting whiteflies found in the literature, one using sticky traps, and the other using plant leaves directly. In sticky traps images, only specimens capable of flying are considered, in leaves images early stages like nymphs can be considered.

Sticky traps images

Sticky traps consist of glue cards where the specimens are attracted by the color of the paper and are caught by the sticky substance, the smooth and neutral surface make the counting easier.

Cho et al. (2008) proposed a method to explores size and color features to identify and count whiteflies, aphids and thrips, using color transformations, simple mathematical morphology operations, and thresholding. Boissard, Martin e Moisan (2008) proposed a cognitive vision system that combines image processing, neural learning and knowledge-based techniques to detect and count pests. Qiao et al. (2008) described a image processing system to estimate the population of adult specimens of whiteflies collected on sticky traps based on size and color. Bechar et al. (23-2) uses techniques like Mixture-of-Gaussians and Principal Component Analysis to extract small spots present on the scene and detect whiteflies in sticky traps in a real-time system. Solis-Sánchez et al. (2011) developed a new algorithm capable to detect and identify five species, including whitefly, using geometrical characteristics of insects that affect greenhouse. Barbedo (2014) presented a new method for counting whiteflies on soybean leaves based on conventional digital image, the algorithm is capable of detecting and quantifying not only adult whiteflies, but also nymphal specimens using region-of-interest (ROI) delineation, application of color transformations, threshold-based segmentation, and detection of young nymphs. Xia et al. (2015) used watershed algorithm and Mahalanobis distance on YCrCb color space, to propose a pest identification procedure, for low resolution images in sticky traps using portable devices. Espinoza et al. (2016) proposed a novel approach for the detection and monitoring of adult-stage whitefly (*Bemisia tabaci*) and thrip (*Frankliniella occidentalis*) in greenhouses based on the combination of an image-processing algorithm and artificial neural networks and described a new range of possibilities with the improvement of neural networks in the last years. Sun et al. (2017) proposed a technique rather than directly counting the pests captured on the traps, the concept is to treat trapped pests as noise in a two-dimensional (2D) image with 2DFT serving as a specific noise collector.

Plant leaves images

These approach consists on take pictures directly on leaves surface in greenhouse or open fields, the main difficult is that leaves imperfections and open field climatic disturbances make counting harder.

Huddar et al. (26-2) proposed a method capable of counting insects not only in greenhouses, but also in open farms, the algorithm has four steps: color conversion, segmentation based on the relative difference in pixel intensities, noise reduction by erosion, and counting based on Moore neighbor tracing and Jacob's stopping. Li, Xia e Lee (2015) proposed new application of multifractal analysis for the detection of small-sized pests (e.g., whitefly) from leaf surface images *in situ*. Multifractal analysis was adopted for segmentation of whitefly images based on the local singularity and global image characters with the regional minima selection strategy. Maharlooei et al. (2017) presented the use of digital imaging technique to detect and count soybean aphids on soybean leaves. Lino et al. (18-2) made a comparison between the performances of LBP and Haar features in a Cascade Classifier, which aims automatically to count immature whiteflies (nymphs). Jige e Ratnaparkhe (19-2) presented an algorithm for automatic detection and counting of whiteflies in images of cotton plantations, the algorithm uses color space conversion, background subtraction, threshold operation, morphological operation and counting of the whitefly number on the leaf with centroid algorithm and region properties. Fuentes et al. (2017) proposed a detector based on neural networks for real-time tomato diseases and pest recognition, including the whitefly pest in tomato plant images. The objective was to find the neural network architecture best suited for the task and three types of detectors were considered: one using Faster Region-based Convolutional Neural Network (Faster R-CNN), another using Region-based Fully Convolutional Network (R-FCN), and the latest Single Shot Multibox Detector (SSD). Chen et al. (2018) presented an image processing method using CNN to segment and count aphid nymphs on leaves. Wang et al. (2018) proposed a cognitive segmentation approach to pest images. The method works in the follow way. First, a pest image is divided into blocks via an image block processing method. Second, an adaptive learning algorithm is used to accurately select the initial cluster centers. Third, preliminary segmentation results are achieved using K-means clustering. Finally, three digital morphological features of an ellipse are adopted to remove leaf veins. Gutierrez et al. (2019) presented the development and comparison of two different approaches for vision based automated pest detection and identification, using learning strategies. A solution that combines computer vision and machine learning is compared against a deep learning solution. The main focus of his paper is on the selection of the best approach based on pest detection and identification accuracy.

CONCLUSIONS

Based on the SRL, we found that in the last ten years a few studies was conducted on whitefly pest detection, average two papers per year are published for this purpose. In the sticky traps approach the main difficulty related are detecting whiteflies caught at the edge of the trap, where it is relatively difficult for the image processing system to detect them. In the leaves images, the most related difficulty is detecting early stages nymphs and classify the different stages.

As future work, it is suggested the use of nymphs detection techniques used in other

species, such as aphids for the detection of whiteflies nymphs and could be interesting to see how these techniques could be improve in open field in images acquired in mobile devices. An automatic solution for use by farmers and field researchers could be proposed, deep learning techniques has become popular and is also an alternative in embedded technologies. There is no open dataset available and those reported in papers, show the scarcity of images and data for this field.

Despite these challenges, the field presents promising results and has followed the evolution of digital image processing techniques in agriculture.

ACKNOWLEDGEMENTS

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

REFERENCES

BAR, L. et al. Downregulation of dystrophin expression in pupae of the whitefly bemisia tabaci inhibits the emergence of adults. *Insect Mol Biol*, John Wiley & Sons, Ltd (10.1111), v. 0, n. 0, jul. 2019. ISSN 0962-1075. Disponível em: <https://doi.org/10.1111/imb.12579>.

BARBEDO, J. G. A. Using digital image processing for counting whiteflies on soybean leaves. *Journal of Asia-Pacific Entomology*, v. 17, n. 4, p. 685–694, dez. 2014. ISSN 1226-8615. Disponível em: <http://www.sciencedirect.com/science/article/pii/S1226861514000855>.

BECHAR, I. et al. On-line video recognition and counting of harmful insects. In: *2010 20th International Conference on Pattern Recognition*. [S.l.: s.n.], 23–2. p. 4068–4071. ISSN 1051-4651.

BOISSARD, P.; MARTIN, V.; MOISAN, S. A cognitive vision approach to early pest detection in greenhouse crops. *Computers and Electronics in Agriculture*, v. 62, n. 2, p. 81–93, jul. 2008. ISSN 0168-1699. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0168169907002256>.

CHEN, J. J. et al. Automatic segmentation and counting of aphid nymphs on leaves using convolutional neural networks. In: . [S.l.: s.n.], 2018.

CHO, J. et al. Automatic Identification of Tobacco Whiteflies, Aphids and Thrips in Greenhouse Using Image Processing Techniques. In: Anninos, P and Pham, T and Grebennikov, A (Ed.). *MABE'08: PROCEEDINGS OF THE 4TH WSEAS INTERNATIONAL CONFERENCE ON MATHEMATICAL BIOLOGY AND ECOLOGY*. [S.l.: s.n.], 2008. (Mathematics and Computers in Science and Engineering), p. 74+. ISBN 978-960-6766-32-9. ISSN 1792-4308. 4th WSEAS International Conference on Mathematical Biology and Ecology, Acapulco, MEXICO, JAN 25-27, 2008.

ESPINOZA, K. et al. Combination of image processing and artificial neural networks as a novel approach for the identification of bemisia tabaci and frankliniella occidentalis on sticky traps in greenhouse agriculture. *Computers and Electronics in Agriculture*, v. 127, p. 495–505, set. 2016. ISSN 0168-1699. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0168169916304823>.

FUENTES, A. et al. A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition. In: *Sensors*. [S.l.: s.n.], 2017.

GUTIERREZ, A. et al. A benchmarking of learning strategies for pest detection and identification on tomato plants for autonomous scouting robots using internal databases. *Journal of Sensors*, Department of Chemistry and Biochemistry, Faculty of AgriSciences, Mendel University, Brno, Czech Republic, v. 2019, n. 5219471, 2019. Disponível em: <https://www.hindawi.com/journals/js/2019/5219471/>.

HUDDAR, S. R. et al. Novel algorithm for segmentation and automatic identification of pests on plants using image processing. In: *2012 Third International Conference on Computing, Communication and Networking Technologies (ICCCNT'12)*. [S.l.: s.n.], 26–2. p. 1–5.

JIGE, M. N.; RATNAPARKHE, V. R. Population estimation of whitefly for cotton plant using image processing approach. In: *2017 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)*. [S.l.: s.n.], 19–2. p. 487–491.

KAMILARIS, A.; PRENAFETA-BOLDú, F. X. Deep learning in agriculture: A survey. *Computers and Electronics in Agriculture*, v. 147, p. 70–90, abr. 2018. ISSN 0168-1699. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0168169917308803>.

KITCHENHAM, B.; CHARTERS, S. *Guidelines for performing Systematic Literature Reviews in Software Engineering*. 2007.

LI, Y.; XIA, C.; LEE, J. Detection of small-sized insect pest in greenhouses based on multifractal analysis. *Optik - International Journal for Light and Electron Optics*, v. 126, n. 19, p. 2138–2143, out. 2015. ISSN 0030-4026. Disponível em: <http://www.sciencedirect.com/science/article/pii/S003040261500409X>.

LINO, A. F. S. et al. Performance of haar and lbp features in cascade classifiers to whiteflies detection and counting. In: *2017 CHILEAN Conference on Electrical, Electronics Engineering, Information and Communication Technologies (CHILECON)*. [S.l.: s.n.], 18–2. p. 1–6.

MAHARLOOEI, M. et al. Detection of soybean aphids in a greenhouse using an image processing technique. *Computers and Electronics in Agriculture*, v. 132, p. 63–70, jan. 2017. ISSN 0168-1699. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0168169916310791>.

MALLETT, R. et al. The benefits and challenges of using systematic reviews in international development research. *Journal of Development Effectiveness*, Routledge, v. 4, n. 3, p. 445–455, set. 2012. ISSN 1943-9342. Disponível em: <https://doi.org/10.1080/19439342.2012.711342>.

MOHANTY, S. P.; HUGHES, D. P.; SALATHé, M. Using deep learning for image-based plant disease detection. *Frontiers in Plant Science*, v. 7, p. 1419, 2016. ISSN 1664-462X. Disponível em: <https://www.frontiersin.org/article/10.3389/fpls.2016.01419>.

QIAO, M. et al. Density estimation of bemisia tabaci (hemiptera: Aleyrodidae) in a greenhouse using sticky traps in conjunction with an image processing system. *Journal of Asia-Pacific Entomology*, v. 11, n. 1, p. 25–29, mar. 2008. ISSN 1226-8615. Disponível em: <http://www.sciencedirect.com/science/article/pii/S1226861508000071>.

SAXENA, L. P.; ARMSTRONG, L. A survey of image processing techniques for agriculture. In: . [S.l.: s.n.], 2014.

SCHWARTZMAN, A. et al. Image processing, computer vision, and deep learning: new approaches to the analysis and physics interpretation of lhc events. *Journal of Physics: Conference Series*, IOP Publishing, v. 762, p. 012035, out. 2016. ISSN 1742-6596. Disponível em: <http://dx.doi.org/10.1088/1742-6596/762/1/012035>.

SOLIS-SáNCHEZ, L. O. et al. Scale invariant feature approach for insect monitoring. *Computers and Electronics in Agriculture*, v. 75, n. 1, p. 92–99, jan. 2011. ISSN 0168-1699. Disponível em: <http://www.sciencedirect.com/science/article/pii/S016816991000195X>.

SUN, Y. et al. A smart-vision algorithm for counting whiteflies and thrips on sticky traps using two-dimensional fourier transform spectrum. *Biosystems Engineering*, v. 153, p. 82–88, jan. 2017. ISSN 1537-5110. Disponível em: <http://www.sciencedirect.com/science/article/pii/S1537511016303439>.

VIBHUTE, A.; , S. K. B. . *Application of Image Processing in Agriculture: A Survey*. [S.l.: s.n.], 2012. 34–40 p.

WANG, Z. et al. A cognitive vision method for insect pest image segmentation. *IFAC-PapersOnLine*, v. 51, n. 17, p. 85–89, jan. 2018. ISSN 2405-8963. Disponível em: <http://www.sciencedirect.com/science/article/pii/S2405896318311820>.

WOLFERT, S. et al. Big data in smart farming - a review. *Agricultural Systems*, v. 153, p. 69–80, maio 2017. ISSN 0308-521X. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0308521X16303754>.

XIA, C. et al. Automatic identification and counting of small size pests in greenhouse conditions with low computational cost. *Ecological Informatics*, v. 29, p. 139–146, set. 2015. ISSN 1574-9541. Disponível em: <http://www.sciencedirect.com/science/article/pii/S1574954114001228>.

TIPOLOGIA DE MUNICÍPIOS PRODUTORES DE BOVINOS DE CORTE NO BIOMA AMAZÔNIA

Maria do Carmo Ramos Fasiaben[(1)], *Stanley Robson de M. Oliveira*[(1)], *Maxwell Merçon Tezolin Barros Almeida*[(2)1], *Octávio Costa de Oliveira*[(2)1], *Helano Póvoas de Lima*[(1)]

[(1)] Embrapa Informática Agropecuária - CNPTIA, Campinas, SP, Brasil

[(2)] Instituto Brasileiro de Geografia e Estatística - IBGE, Rio de Janeiro, RJ, Brasil

{maria.fasiaben, stanley.oliveira, helano.lima}@embrapa.br,
{maxwell.almeida, octavio.oliveira}@ibge.gov.br

RESUMO

Conhecer a diferenciação dos sistemas de produção praticados pelos produtores é fundamental para propor linhas de pesquisa, desenvolvimento, transferência de tecnologia e políticas públicas apropriadas. O objetivo deste artigo é diferenciar e caracterizar os sistemas de produção de bovinos de corte praticados no Bioma Amazônia. Foram empregados dados do Censo Agropecuário 2006 de estabelecimentos agropecuários produtores de bovinos de corte, agregados por município. Trata-se do último censo com informações disponíveis no âmbito dos estabelecimentos agropecuários. A análise de cluster permitiu separar 463 municípios em seis agrupamentos (clusters), segundo características da produção. Os resultados mostram que alguns clusters não apresentaram o padrão esperado entre uso de tecnologia e produtividade, o que evidencia a necessidade de estudos mais aprofundados sobre esse tema na Amazônia.

PALAVRAS-CHAVE: Análise de cluster, tecnologia, sistemas de produção pecuários.

ABSTRACT

Knowing the differentiation of the production systems practiced by farmers is fundamental to propose lines of research, development, transfer of technology and appropriate public policies. The objective of this paper is to differentiate and characterize the production systems of beef cattle practiced in the Amazon Biome. Data from the Agricultural Census 2006 of

[1] O IBGE está isento de qualquer responsabilidade pelas opiniões, informações, dados e conceitos emitidos neste artigo, que são de exclusiva responsabilidade do autor.

holdings producing beef cattle, aggregated per municipality, were used. The cluster analysis allowed for separating 463 municipalities in six clusters, according to production characteristics. The results show that some clusters did not present the expected pattern between technology use and productivity, which evidences the need for further studies on this topic in the Amazon Biome.

KEYWORDS: Cluster analysis, technology, livestock production systems.

INTRODUÇÃO

A Amazônia é o maior bioma do Brasil, ocupando quase metade do território nacional. O desenvolvimento da pecuária na Amazônia é frequentemente apontado como a causa do desmatamento na região. Entretanto, para Barbosa et al. (2015) as causas do desmatamento na Amazônia são diversas e frequentemente relacionadas. Para esses autores, seria mais exato dizer que a geopolítica expansionista para ocupação da Amazônia iniciada nos anos de 1970, os créditos subsidiados, a crescente demanda nacional e internacional por carne, os investimentos em infraestrutura e a eterna busca humana pela posse da terra foram fatores que contribuíram para que o desmatamento ocorresse. Para esses autores, a pecuária seria mais uma "companheira" do desmatamento do que a sua "causa", já que é a forma mais barata de se ocupar a terra desmatada e assegurar sua posse, que de outro modo poderia estar ameaçada.

Embora a produção de carne na Amazônia tenha crescido, historicamente, graças à expansão de pastagens, isso começa a mudar com a intensificação dos sistemas de produção, induzida tanto pelo mercado e exigências ambientais quanto pelo maior controle do desmatamento na região.

Conhecer a diversidade dos sistemas de produção praticados pelos produtores é um passo importante para orientar trabalhos de pesquisa, assistência técnica e extensão rural na região, bem como para o apoio à formulação de políticas públicas relacionadas ao setor. Para Alvarez et al. (2018), as tipologias agrícolas ajudam a entender a complexidade dos sistemas agrícolas, fornecendo uma representação simplificada da diversidade de sistemas, organizando-os em grupos homogêneos.

O objetivo deste trabalho foi realizar uma tipificação dos municípios do Bioma Amazônia segundo características da criação de bovinos de corte. O trabalho exigiu tabulações especiais de dados do Censo Agropecuário 2006 - o último com informações disponíveis no âmbito dos estabelecimentos agropecuários -, filtrados para estabelecimentos produtores de bovinos de corte e agrupados por município.

MATERIAL E MÉTODOS

Foram analisados dados oriundos do Censo Agropecuário 2006 referentes a estabelecimentos agropecuários produtores de bovinos de corte do Bioma Amazônia, agrupados em 463 municípios. A seleção de variáveis para a tipificação da produção de bovinos de corte desses municípios (variáveis classificatórias) foi realizada a partir de consultas a especialistas na pecuária de corte da Região Amazônica.

As variáveis classificatórias empregadas no presente estudo foram:

1) Quantidade total de bovinos
2) Área total de pastagem/área explorada[2]
3) Área total de lavouras/área explorada
4) Área de silvicultura/área explorada
5) Taxa de lotação
6) Porcentagem de estabelecimentos especializados em cria
7) Porcentagem de estabelecimentos especializados em recria
8) Porcentagem de estabelecimentos especializados em engorda
9) Porcentagem de estabelecimentos especializados em cria-recria
10) Porcentagem de estabelecimentos especializados em cria-engorda
11) Porcentagem de estabelecimentos especializados em recria-engorda
12) Porcentagem de estabelecimentos especializados em cria, recria e engorda
13) Porcentagem de bovinos confinados
14) Porcentagem de vacas com dois anos ou mais
15) Porcentagem de vacas inseminadas
16) Porcentagem de estabelecimentos que fizeram adubação de pastagem
17) Porcentagem de estabelecimentos que fizeram suplementação com sal e ração ou somente ração – incluídos aqui ração, grãos e subprodutos agroindustriais
18) Porcentagem de estabelecimentos que recebem orientação técnica
19) Porcentagem do valor total da produção que provém de bovinos

Foi utilizada tabulação especial do Censo Agropecuário 2006, realizada pelo IBGE (IBGE, 2012). A tabulação consistiu de dados agregados por município de estabelecimentos agropecuários do bioma Amazônia com mais de 50 cabeças de bovinos e que declararam ter o

[2] Área explorada corresponde à soma das áreas de lavouras (temporárias e permanentes), pastagens (naturais e plantadas), florestas plantadas e sistemas agroflorestais.

corte como principal finalidade do rebanho[3]. Por questão de sigilo estatístico foram descartados os municípios com menos de três estabelecimentos.

Após o tratamento dos dados, foi realizado o processo essencial de extração dos padrões de desenvolvimento tecnológico para o bioma Amazônia. Para tanto, empregou-se a análise de clusters, por meio do algoritmo EM (do inglês, *Expectation-Maximization*) (WU, 1983). A análise de cluster procura definir grupos similares de observações, de maneira que as diferenças entre integrantes de um mesmo grupo sejam mínimas e as diferenças entre os grupos sejam máximas. O algoritmo EM atribui uma distribuição de probabilidade a cada instância (município) que indica a probabilidade deste pertencer a cada um dos clusters. O EM pode decidir quantos clusters podem ser gerados por validação cruzada, ou pode-se especificar quantos clusters se deseja gerar a priori.

Neste estudo, número de clusters sugerido pelo algoritmo por validação cruzada foi condizente com o esperado a partir de trabalho de Fasiaben et al. (2016). A validação cruzada (VC) executada para determinar o número de clusters foi feita nas seguintes etapas:

a) o número de clusters é definido como 1;

b) o conjunto de treinamento é dividido aleatoriamente em 10 partes;

c) o algoritmo EM é executado 10 vezes usando as 10 partes da validação cruzada;

d) computa a estimativa de máxima verossimilhança (*log likelihood*) sobre os 10 resultados;

e) Se o *log likelihood* tiver aumentado, o número de clusters será aumentado em 1 e o programa continuará na etapa 2.

O número de partes é fixada em 10, desde que o número de instâncias (municípios) no conjunto de treinamento não seja menor que 10. Se esse for o caso, o número de partes será igual ao número de instâncias. O algoritmo EM está disponível no ambiente de software Weka (FRANK et al., 2016), uma coleção de algoritmos de aprendizado de máquina para tarefas de mineração de dados.

Uma vez definidos os clusters a partir das variáveis classificatórias, variáveis adicionais foram obtidas do Censo para aprimorar a descrição desses agrupamentos.

RESULTADOS E DISCUSSÃO

Segundo o Censo Agropecuário 2006 do IBGE (IBGE, 2012), o Bioma Amazônia detinha um rebanho de bovinos de corte[4] de cerca de 32 milhões de cabeças, que ocupava uma área de

[3] O Censo Agropecuário 2006 teve como particularidade a aplicação de questionário detalhado, incluídas variáveis que denotavam o perfil tecnológico da produção de bovinos, apenas em estabelecimentos que tinham mais de 50 cabeças na data de referência (31/12/2006).

pastagens próxima aos 25 milhões de ha e envolvia 76,6 mil estabelecimentos. A Tabela 1 apresenta informações gerais acerca dos municípios produtores de bovinos de corte, por Cluster e para o Bioma Amazônia.

A tipologia separou os 463 municípios produtores de bovinos de corte da Amazônia em seis grupos (Clusters), sendo o Cluster 5 o mais representativo, concentrando 32% dos municípios, 67% dos estabelecimentos, 68% do efetivo bovino e 65% da área de pastagens do bioma. A Tabela 2 mostra o uso da terra nos estabelecimentos produtores, agregados por município, por cluster e por bioma.

Os estabelecimentos dedicados a bovinos de corte na Amazônia apresentam uma área média de pastagens de 326 ha, com variações importantes entre os clusters. O Cluster 2 caracteriza-se pelo maior grau de diversificação dos estabelecimentos, nos quais a área de lavouras supera, em média, 27% da área explorada. Nesse Cluster, o valor da produção da bovinocultura de corte representa, em média, 10% do valor total da produção dos estabelecimentos, valor muito abaixo que o encontrado nos demais clusters (Tabela 4).

Ainda quanto ao uso da terra, os Clusters 1 e 3 destacam-se pelo percentual da área ocupada com sistemas agroflorestais, acima de 5% da área explorada dos estabelecimentos em ambos os casos, superando em mais que o dobro a média do bioma.

Na Tabela 3 são apresentadas as características do rebanho nos estabelecimentos produtores de bovinos de corte da Amazônia, agregados por município, por cluster e por bioma. O tamanho médio do rebanho de corte da Amazônia é de 421 cabeças por estabelecimento agropecuário, variando conforme os clusters. Os maiores rebanhos são encontrados nos estabelecimentos do Cluster 4, e os menores, no Cluster 3.

Os clusters da Amazônia também se diferenciam conforme as fases da criação conduzidas nos estabelecimentos, com predomínio de Ciclo Completo nos Clusters 2, 4 e 6; de sistemas de Cria no Cluster 1 (o que é corroborado pelo maior percentual de vacas com 2 anos e mais no rebanho); de sistemas de Cria e Recria no Cluster 5 - também corroborado pelo percentual de vacas no rebanho (nesse cluster, os sistemas a Ciclo Completo também se destacam) e de Engorda no Cluster 3.

Na Tabela 4 são apresentados aspectos gerais da tecnologia empregada nos estabelecimentos produtores de bovinos de corte, agregados por município, por cluster e para todo o bioma. Na Tabela 5 o conjunto de variáveis é apresentado em valores absolutos.

[4] Estabelecimentos com mais de 50 cabeças de bovinos e que declararam ter o corte como principal finalidade do rebanho.

Tabela 1 – Informações gerais sobre municípios produtores de bovinos de corte do Bioma Amazônia (por cluster e bioma)

Variável / Cluster	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Amazônia
Número de ocorrências (municípios)	50	27	79	77	150	80	463
Porcentagem de ocorrências em relação ao total do Bioma	11%	6%	17%	17%	32%	17%	100%
Número de estabelecimentos produtores de bovinos de corte (unidades)	5.366	1.298	1.639	10.901	51.555	5.842	76.601
Porcentagem de estabelecimentos em relação ao total do Bioma	7%	2%	2%	14%	67%	8%	100%
Quantidade total de bovinos de corte (em cabeças)	1.485.837	589.004	374.834	6.384.425	22.053.165	1.397.161	32.284.426
Porcentagem de cabeças em relação ao total do Bioma	5%	2%	1%	20%	68%	4%	100%
Área ocupada com pastagens nos estabelecimentos (ha)	1.182.495	547.537	339.947	5.354.959	16.293.059	1.270.871	24.988.868
Porcentagem da área de pastagens em relação ao total do Bioma	5%	2%	1%	21%	65%	5%	100%

Tabela 2 – Uso da terra em estabelecimentos produtores de bovinos de corte no Bioma Amazônia, agregados por município (por cluster e bioma)

Variável / Cluster	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Amazônia
Área total (média) dos estabelecimentos (ha)	417	1031	455	793	579	434	592
Área média explorada dos estabelecimentos (ha)	250	608	272	521	334	258	352
Área média de pastagem dos estabelecimentos (ha)	220	422	207	491	316	218	326
Porcentagem da área ocupada com lavoura por área explorada dos estabelecimentos	5,2%	27,2%	18,2%	4,0%	3,4%	11,8%	5,0%
Porcentagem da área ocupada com pastagens por área explorada dos estabelecimentos	88,2%	69,4%	76,2%	94,3%	94,7%	84,2%	92,6%
Porcentagem da área ocupada com silvicultura por área explorada dos estabelecimentos	1,3%	1,5%	0,0%	0,4%	0,1%	0,1%	0,3%
Porcentagem da área ocupada com sistemas agroflorestais por área explorada dos estabelecimentos	5,3%	1,9%	5,6%	1,3%	1,9%	4,0%	2,1%

Tabela 3 – Características do rebanho em estabelecimentos produtores de bovinos de corte do Bioma Amazônia, agregados por município (por cluster e bioma)

Variável / Cluster	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Amazônia
Porcentagem de cabeças em relação ao total do Bioma	5%	2%	1%	20%	68%	4%	100%
Número médio de UA por estabelecimento	184	323	158	406	296	160	291
Número médio de cabeças por estabelecimento	277	454	229	586	428	239	421
Porcentagem de vacas com 2 anos ou mais em relação ao rebanho total	31%	25%	27%	23%	31%	28%	29%
Porcentagem de bois e garrotes para corte com 2 anos e mais em relação ao rebanho total	15%	20%	20%	24%	18%	19%	19%
Estabelecimentos que declararam fazer Cria (%)	27%	14%	25%	13%	19%	21%	19%
Estabelecimentos que declararam fazer Recria (%)	16%	5%	8%	6%	8%	5%	8%
Estabelecimentos que declararam fazer Engorda (%)	17%	22%	32%	16%	13%	20%	15%
Estabelecimentos que declararam fazer Cria e Recria (%)	14%	15%	9%	18%	27%	7%	23%
Estabelecimentos que declararam fazer Cria e Engorda (%)	7%	6%	5%	3%	3%	14%	4%
Estabelecimentos que declararam fazer Recria e Engorda (%)	4%	7%	3%	12%	5%	2%	5%
Estabelecimentos que declararam fazer Cria, Recria e Engorda (%)	15%	31%	18%	32%	25%	31%	25%

Tabela 4 – Aspectos gerais da tecnologia empregada nos estabelecimentos produtores de bovinos de corte no Bioma Amazônia, agregados por município (por cluster e bioma)

Variável / Cluster	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Amazônia
Taxa de lotação (UA/ha)	0,83	0,77	0,76	0,83	0,94	0,74	0,89
Porcentagem de estabelecimentos especializados em criação de bovinos de corte	57%	47%	51%	65%	64%	56%	63%
Porcentagem de estabelecimentos que fazem rotação de pastagem	55%	52%	49%	57%	63%	48%	60%
Porcentagem de estabelecimentos que fizeram adubação em pastagens	1%	10%	7%	4%	1%	2%	2%
Porcentagem de estabelecimentos que usaram ração	5,8%	15,3%	17,4%	13,1%	6,9%	12,3%	8,5%
Porcentagem de bovinos confinados	0,7%	10,6%	1,4%	1,3%	0,7%	0,6%	1,0%
Porcentagem de vacas inseminadas	3%	12%	9%	6%	5%	1%	5%
Porcentagem de vacas que receberam embriões	0,1%	0,5%	0,2%	0,4%	0,2%	0,1%	0,2%
Porcentagem do valor da produção advindo dos bovinos em relação ao valor total da produção	84%	10%	54%	83%	83%	63%	71%

Tabela 5 – Variáveis para caracterização dos Clusters e do Bioma Amazônia (em valores absolutos)

Variável / Cluster	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Amazônia
Número de ocorrências (municípios)	51	28	69	79	148	88	463
Número de estabelecimentos produtores de bovinos de corte (unidades)	5.366	1.298	1.639	10.901	51.555	5.842	76.601
Área total dos estabelecimentos (ha)	2.238.037	1.337.870	745.338	8.641.541	29.869.116	2.533.674	45.365.576
Área explorada total dos estabelecimentos (ha)	1.340.795	789.033	446.356	5.678.601	17.209.988	1.509.439	26.974.213
Área ocupada com pastagens nos estabelecimentos (ha)	1.182.495	547.537	339.947	5.354.959	16.293.059	1.270.871	24.988.868
Área ocupada com lavouras nos estabelecimentos (ha)	69.851	214.611	81.428	227.024	576.623	177.609	1.347.146
Área ocupada com silvicultura nos estabelecimentos (ha)	17.569	11.918	-	22.252	20.748	1.332	73.820
Área ocupada com sistemas agroflorestais nos estabelecimentos (ha)	70.879	14.966	24.981	74.366	319.559	59.627	564.379
Quantidade total de bovinos de corte (em UA - INCRA)	986.637	419.816	258.815	4.428.355	15.270.179	935.387	22.299.189
Quantidade total de bovinos de corte (em cabeças)	1.485.837	589.004	374.834	6.384.425	22.053.165	1.397.161	32.284.426
Cria (CAB)	355.337	29.894	66.773	435.151	2.559.039	226.820	3.673.014
Recria (CAB)	182.417	19.785	28.958	234.482	1.061.873	59.130	1.586.645
Engorda (CAB)	292.302	118.000	120.914	977.059	3.093.204	291.937	4.893.416
Cria e Recria (CAB)	175.873	56.875	23.545	640.845	3.928.535	101.546	4.927.219
Cria e Engorda (CAB)	99.282	17.357	15.514	161.966	641.048	163.511	1.098.678
Recria e Engorda (CAB)	67.842	60.453	20.997	1.104.863	1.616.234	69.435	2.939.824
Cria, Recria e Engorda (CAB)	312.784	286.640	98.133	2.830.059	9.153.232	484.782	13.165.630
Estabelecimento Cria (número)	1.448	180	409	1.457	9.813	1.232	14.539
Estabelecimento Recria (número)	850	70	138	691	3.885	273	5.907
Estabelecimento Engorda (número)	916	290	525	1.717	6.897	1.171	11.516
Estabelecimento Cria e Recria (número)	755	194	151	1.995	14.151	411	17.657
Estabelecimento Cria e Engorda (número)	387	73	80	293	1.702	800	3.335
Estabelecimento Recria e Engorda (número)	227	94	43	1.257	2.432	134	4.187
Estabelecimento Cria, Recria e Engorda (número)	783	397	293	3.491	12.675	1.821	19.460

Entre as tecnologias analisadas, observa-se que a mais consolidada no Bioma Amazônia era a rotação de pastagens[5], adotada, em média, em 63% dos estabelecimentos do bioma e com variações entre os clusters.

De um modo geral, os dados evidenciaram que a produção de bovinos de corte na Amazônia se dava com baixo uso de tecnologia à época do Censo Agropecuário 2006. Merece destaque o baixo emprego de adubação de pastagens, que era praticada em média por 2% dos estabelecimentos no Bioma, com marcada diferença entre os clusters, atingindo, como o maior valor, a cifra de 10% dos estabelecimentos no Cluster 2.

A suplementação (com uso de ração) foi empregada em maior grau pelo Cluster 3, mesmo assim, por 17% dos estabelecimentos. O percentual de estabelecimentos que usavam ração no Bioma Amazônia era de 8% em 2006.

O percentual de bovinos confinados girava em torno de 1% nos clusters, sendo também essa a média do bioma. Aqui a exceção foi o Cluster 2, no qual o percentual de animais confinados era próximo a 10% em relação ao total do rebanho.

Em relação ao manejo reprodutivo, o maior percentual de vacas inseminadas ocorria no Cluster 2 (12% do total de vacas), lembrando que nesse cluster os sistemas de produção predominantes eram de ciclo completo. Já a prática da transferência de embriões nos estabelecimentos era muito baixa em todos os clusters.

Um fato que merece destaque neste estudo foi o certo descolamento entre o nível tecnológico expresso em alguns dos clusters e as taxas de lotação médias identificadas. Os casos mais notórios foram o dos Clusters 2 e 3, ambos com valores para as variáveis relacionadas à tecnologia bem acima das médias do Bioma Amazônia, porém com taxas de lotação inferiores às médias do Bioma. Já o Cluster 5, o mais frequente, apresentou taxa de lotação acima da média da Amazônia, com percentuais do uso das diferentes tecnologias próximos à média do Bioma. Esse tipo de resultado aponta para a necessidade de estudos mais aprofundados relacionados à produtividade da pecuária de corte na Amazônia. Uma suposição é que haja um efeito da localização espacial nas médias municipais, podendo estar, os clusters mais tecnificados e com taxas de lotação menores, localizados em áreas de solos menos férteis e/ou em locais de pastagens mais degradadas e em possível estágio de recuperação. Nesse sentido, há que ressaltar, ainda, a limitação de se trabalhar com os dados agregados no nível municipal, pois não se considera a heterogeneidade dos sistemas de produção dentro de um

[5] Segundo o Manual do Recenseador (IBGE, 2007, pág. 99), "Rotação de pastagens - é a prática de dividir a pastagem em piquetes, de modo a possibilitar o pastoreio rotativo".

mesmo município. Tipologias elaboradas a partir de microdados (dados dos estabelecimentos agropecuários) devem permitir identificar, de forma mais precisa, a real diferenciação dos sistemas de produção praticados pelos produtores.

A Figura 1 mostra a localização espacial dos seis clusters encontrados no Bioma Amazônia.

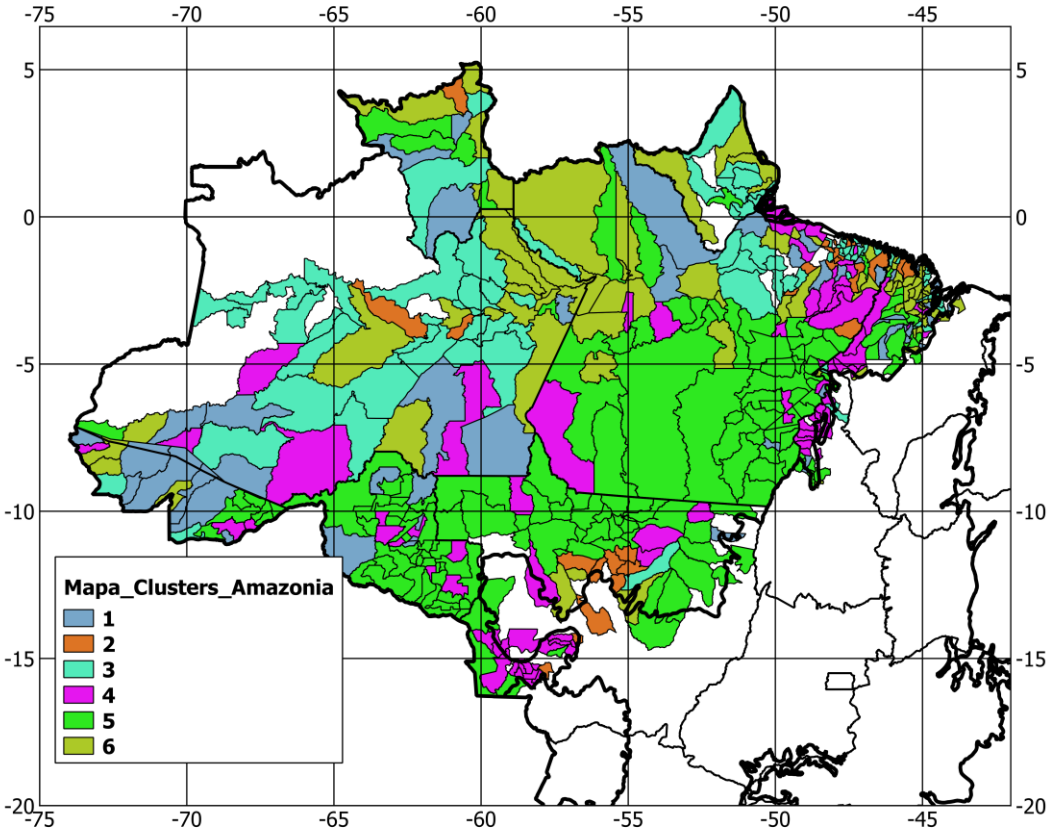


Figura 1- Localização espacial dos tipos (clusters) de municípios produtores de bovinos de corte no Bioma Amazônia.

CONCLUSÕES

A análise de cluster separou os municípios produtores de bovinos de corte do Bioma Amazônia em seis agrupamentos, baseados nas características dos sistemas de produção praticados pelos produtores, com dados agregados por município. Entretanto, no caso da Amazônia, em alguns clusters a relação entre as variáveis que expressam o nível tecnológico e as taxas de lotação não se deu dentro do padrão esperado (maior uso de tecnologia, maior produtividade). Esse fato pode estar relacionado a efeitos da localização espacial sobre os clusters (tipos de solos onde estão localizados, nível de degradação das pastagens), mas

também ao fato de se estar trabalhando com dados agregados no âmbito municipal, pois não se considera a heterogeneidade dos sistemas de produção dentro de um mesmo município. Futuros trabalhos, com tipologias elaboradas a partir de dados dos estabelecimentos agropecuários devem permitir identificar, de forma mais precisa, a real diferenciação dos sistemas de produção praticados pelos produtores.

AGRADECIMENTOS

Os autores agradecem a Santiago Vianna Cuadra pela elaboração do mapa com a localização espacial dos clusters de municípios produtores de bovinos de corte no Bioma Amazônia.

REFERÊNCIAS

ALVAREZ, S.; TIMLER, C.J.; MICHALSCHECK, M.; PAAS, W.; DESCHEEMAEKER, K.; TITTONELL, P.; ANDERSSON, J.A.; GROOT, J.C.J. Capturing farm diversity with hypothesis-based typologies: An innovative methodological framework for farming system typology development. PLOS ONE | https://doi.org/10.1371/journal.pone.0194757 May 15, 2018

BARBOSA, F.A. et al. Cenários para a pecuária de corte amazônica. 1.ed.- Belo Horizonte: Ed. IGC/UFMG, 2015. 146 p. Disponível em: < http://csr.ufmg.br/pecuaria/wp-content/uploads/2015/07/relatorio_cenarios_para_pecuaria_corte_amazonica.pdf> Acesso em: 14 mai. 2018.

FASIABEN, M. do C. R.; ALMEIDA, M. M. T. B.; MAIA, A. G.; OLIVEIRA, O. C. de; COSTA, F. P.; BARIONI, L. G.; DIAS, F. R. T.; MOREIRA, J. M. M. A. P.; ABREU, U. G. P. de; SENA, A. L. dos S.; SANTOS, J. C. dos; LAMPERT, V. do N.; OLIVEIRA, P. P. A.; PEDROSO, A. de F. Typifying beef cattle producers in Brazilian biomes. In: INTERNATIONAL SYMPOSIUM ON GREENHOUSE GASES IN AGRICULTURE, 2., 2016, Campo Grande, MS. Proceedings... Brasília, DF: Embrapa, 2016. p. 404-409. (Embrapa Gado de Corte. Documentos, 216).

FRANK, E.; HALL, M. A.; WITTEN, I.H. (2016). The WEKA Workbench. Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques", Morgan Kaufmann, Fourth Edition, 2016.

IBGE. Censo agropecuário 2006 : Brasil, grandes regiões e unidades da federação : segunda apuração. Rio de Janeiro, 758 p, 2012. ISBN: 2408107300

IBGE. Censo Agropecuário 2006: Manual do Recenseador CI - 1.09 A. Rio de Janeiro, 189 p, 2007.

WU, C.F.J. (1983). On the Convergence Properties of the EM Algorithm, The Annals of Statistics, 11, 95-103.

Requirements identification and proposal of a domain model for implementing IoT in the sugar supply chain

Roberto Fray da Silva[1], Gustavo Marques Mostaço[1], Carlos Eduardo Cugnasca[1]

[1] Departamento de Engenharia de Computação e Sistemas Digitais, Escola Politécnica da Universidade de São Paulo, São Paulo, São Paulo, Brasil, roberto.fray.silva@gmail.com, gmostaco@usp.br, carlos.cugnasca@usp.br

RESUMO

A implementação do paradigma de IoT na cadeia de suprimentos do açúcar é essencial para melhorar a rastreabilidade dos produtos e o monitoramento de processos. Esta cadeia de suprimentos gera uma quantidade considerável de dados heterogêneos, além de possuir sérios problemas de interoperabilidade. Existem poucos modelos de arquiteturas de IoT disponíveis na literatura que possam ser aplicados neste contexto, permitindo o desenvolvimento de aplicações. No entanto, estes não focam em requisitos específicos da cadeia. Neste trabalho, a cadeia de suprimentos de açúcar no Brasil é analisada, identificando seus principais requisitos e como os principais modelos de referência de IoT podem atender a estes requisitos. Também é descrito o desenvolvimento de um modelo de domínio baseado no modelo de referência IoT-A, com foco nas entidades físicas e virtuais. Observou-se que o modelo proposto possui o potencial para suprir os requisitos identificados, endereçando os principais problemas advindos da adoção dos outros modelos de arquitetura analisados.

PALAVRAS-CHAVE: framework, IoT, cadeia de suprimentos, açúcar

ABSTRACT

The implementation of the IoT paradigm in the sugar supply chain is essential to allow for better product traceability and process monitoring. This supply chain generates a considerable amount of heterogeneous data and has serious interoperability issues. There are few IoT architectural models available in the literature that could be applied in this context allowing for application development. Nevertheless, these do not focus on its specific requirements. We studied the sugar supply chain in Brazil, identifying its requirements and analyzing how the main IoT reference models can fulfill these requirements. We then developed a domain model based on the IoT-A reference model, focusing on the physical and virtual entity view. It was observed that the model could fulfill the identified requirements, addressing the main issues that arise from adopting the other IoT architectural models analyzed.

KEYWORDS: framework, IoT, supply chain, sugar.

INTRODUCTION

A supply chain (SC) can be defined as the group of companies responsible, directly or indirectly, for fulfilling the demands of a customer segment (CHOPRA, MEINDL, 2013). These companies are also called agents, and the groups of agents responsible for a specific activity are called links. They generate a high quantity of heterogeneous data that come from different sources and is stored in different formats, mostly inside local databases (DB).

The internet of things (IoT) can be defined as the presence of pervasive objects that have a unique identification scheme and are able to interact with each other and with users to reach established goals through the Internet (ATZORI, IERA, MORABITO, 2010). These objects can be related to identification (such as RFID), monitoring (such as sensors in a wired or wireless network) and devices (mobile phones, tablets, computers, among others). It also encompasses networking, privacy, security, and management issues. The implementation of these technologies will result in an enormous quantity of data (GUBBI et. al, 2013).

Currently, the solutions implemented in this context utilize different types of technologies, sensors, readers, tags, and commercial systems that do not necessarily interact with each other. This raises significant interoperability problems. Also, the new technologies will have to coexist with the current systems of traceability, product, and process monitoring.

The use of an IoT reference model for the implementation of these technologies is necessary because it allows for better interoperability between the traceability, product, and

process monitoring systems throughout the whole SC. We could not find in the literature a systematic identification of requirements for agroindustrial SCs, nor architectural models that allowed for the development of applications. Most of the models proposed are: (i) too generic, without the description of system's components or their interactions, such as in Verdouw et. al (2016); or (ii) too narrow, focusing on a single system, such as in Jiang et. al (2014).

Therefore, the objectives with this paper are to identify the main requirements to implement IoT in the Brazilian sugar SC and to develop a domain model for this context, focusing on the physical and virtual entities view.

THE SUGAR SUPPLY CHAIN

According to Silva et. al (2015), the Brazilian sugar SC is composed of nine links: (1) farm, (2) raw sugar industry, (3) warehouse, (4) harbor in Brazil, (5) harbor in the destination country, (6) sugar refining plant, (7) food industry (if it will be used to make other products), (8) retailer, and (9) consumer. From a technology point of view, it is important to notice that the product being tracked changes its form a couple of times among the SC, affecting how data should be collected and processed.

At the farm, which is the first link of the SC, it is important to trace the truck that will transport the sugarcane. However, as the product will be refined later, a few traceability problems may occur at the farm level for this product. From links number two (raw sugar industry) to six (sugar refining plant), the product to be tracked is the raw sugar in bulk. Relative humidity (RH) is an important environmental variable that should be monitored in order to minimize the impact of quality problems (SILVA et. al, 2015).

From links number six (sugar refining plant) to nine (consumer), the product to be tracked is the sugar package, or the food product package, in case it is used by some food industry as an ingredient for other products. In this phase, all packages must be traced to fulfill the consumers' traceability requirements. Also, the environmental variables that should be monitored are temperature and RH. Only trucks are used to transport product in this form.

METHODOLOGY

The methodology used in this paper can be divided into five steps:

1. Identification of the main requirements for IoT implementation in the sugar SC, based on a literature review of the sugar SC area and the models developed to implement IoT-related technologies in this context;

2. Analysis of five reference models for IoT implementation: Casagras, IEEE P2413, IoT-A, ITU-T, and OneM2M;
3. Selection of the most relevant reference model for this context;
4. Study of the chosen model's documentation in order to identify its main components and the methodology recommended for generating the domain model;
5. Follow the methodology to develop the domain model, considering the requirements identified in step 2.

RESULTS

This section contains the main results of the paper and is divided into the following subsections: requirements identification; IoT reference models analysis; physical and virtual entity view; and the IoT domain model.

Requirements identification

The requirements were identified by analyzing the models, which were selected by the authors considering the most important ones available in the literature for this context. This is mainly related to models that consider: all the SC links; interoperability problems; demands of agents in the different links; operational and planning aspects; and finally product quality monitoring.

We identified 14 requirements. They were separated into three categories and are illustrated in Table 1. The first category, named Allow, is related to the model functionalities, and it contains seven items that should be allowed by the model. These are related to legacy and actual technologies coexistence, communication aspects, SC virtualization, product traceability, and quality control.

Table 1. Requirements identification divided into the three main categories

Category	Requirement
1 - Allow	Technologies coexistence
	Creation of virtual SCs (VERDOUW et. al, 2016)
	Distributed traceability approach (VERDOUW et. al, 2016)
	Asynchronous communication (VERDOUW et. al, 2013)
	Cloud and local interfaces for application services (VERDOUW et. al, 2013)
	Hybrid routing of data - P2P and centralized (VERDOUW et. al, 2013)
	Object quality control using virtual objects (VERDOUW et. al, 2013; VERDOUW et. al, 2016; PANG et. al, 2015)
2 - Consider	Structured and unstructured data (SILVA et. al, 2015; PANG et. al, 2015)
	Data from the three main systems
	Evolution of the system - technologies and processes (BAUER et. al, 2013)
	Virtual objects must provide multiple views for different users (VERDOUW et. al, 2013)
3 - Provide	Capability to monitor, manage events, optimize processes, and provide autonomy for the objects (VERDOUW et. al, 2016)
	Information access - authorization and reliability (VERDOUW et. al, 2016)
	Information fusion on-site (sensors - pre-processing), in-system (ERPs - processing, decision making), and in-cloud (processing, certifiers, SC replanning) (PANG et. al, 2015)

The second category, named Consider, is related to general characteristics of the model and contains four items that should be considered by the model. It encompasses the aspects of the data generated in this context, the characteristics of the technologies and processes, and the need to provide multiple views for the different users based on their needs.

The last category, named Provide, corresponds to the functionalities that should be provided by the model, and are related to information access, data fusion, and the capabilities to monitor and act on events.

IoT reference models analysis

The IoT reference models provide standard notation and a group of concepts that form the basis for the development of architectural models. In this subsection, we analyze the five main IoT reference models that were identified as possible to implement on the sugar supply chain.

The IoT-A reference model was a project developed by the EU between 2010 and 2013 (BAUER et. al, 2013). It provides a high level of abstraction for developing architectural reference models. Its documentation is very detailed, providing examples of use cases. It includes models for different levels of abstraction: the **domain model** describes high-level applications; the **context view** describes the main components of the system in the context; the **physical-entity view** describes the relations between the physical entities (PE) and the virtual entities (VE); the **communication model** describes the interaction between the devices and objects; the **information model** describes how data is generated, processed and transmitted; the **functional model** describes the main functionalities of the system and how they interact with each other; in addition, the **deployment and operation view** describes the main characteristics and implementation aspects to be considered (BAUER et. al, 2013).

The ITU-T reference model was proposed by the International Telecommunication Union (ITU). It is composed of four layers: application, service and application support, network and device. These interact with two groups of capabilities: security and management (ITU, 2012). It is less comprehensive than the IoT-A model and does not specify a clear functional model. Also, there are fewer guidelines for architectural model development.

Another very important IoT reference model is presented by the Casagras project (SMITH, 2012), it is an initiative between the EU, Japan, China, USA, and Korea. The model is composed of four layers: physical interface; interrogator-gateway; information management, application, and enterprise; wider area communications and Internet. Although

this model describes the layers in detail, it is less comprehensive than the other two, increasing the difficulty on using it for developing an architectural model.

Unlike the models cited above, the oneM2M model is driven by industries and has the objective to standardize machine-to-machine (M2M) communications (ONEM2M, 2015). It also considers the multiple views of the different stakeholders and has three layers: application, services, and network. Nevertheless, it does not provide a detailed description of the baselines to develop an architectural model.

The IEEE P2413 standard is in development and aims at building architectural models for the implementation of IoT (WEYRICH, EBERT, 2016). However, this model is only a generic framework and still has no implementation or further documentation.

Table 2 contains the results of the reference models analysis. In this analysis, we evaluated four criteria: (i) Comprehensiveness: functionalities of the model related to the requirements; (ii) Requirements: if it fulfilled the identified requirements; (iii) Indication of adoption: is it possible to be implemented for the sugar SC; and (iv) Documentation: how well documented in the model.

Table 2. Analysis of the identified IoT reference models

Reference model	Comprehensiveness	Requirements	Indication of adoption	Documentation
IEEE P2413				
ITU-T	X		X	X
Casagras	X			X
OneM2M	X	X		X
IoT-A	X	X	X	X

The IoT-A reference model is the most comprehensive model in terms of description of its functionalities, domain, and information. Besides, it fulfills all the requirements for our context, and have very detailed documentation. We also believe that there are indications of its future adoption, such as its application on different research projects, and the continuous investment of the EU on projects related to the IoT-A.

On the other hand, the ITU-T model is less extensive and documented than the IoT-A, and it does not provide sufficient description to fulfill all the identified requirements. The Casagras model has a similar problem, but it is less detailed in terms of functionalities and the information modeling process.

The main concerns related to the OneM2M standards are that they are less used in the literature than the IoT-A and ITU-T models. For this reason, there are few guarantees that

these models will continue to be used. Lastly, the IEEE P2413 provides a very interesting idea, but it is still poorly documented, not being capable to fulfill any of our criteria.

Physical and virtual entity view

The physical-entity view describes the physical entities (PEs) and the virtual entities (VEs) that are part of the model, alongside with resources, devices, services, and users. The physical entities can be single or hierarchical. The single PEs are the vehicles that transport the products. The hierarchical PEs are the raw product batches and the finished products' pallets. The VEs are divided into single and multiple. The single VEs are the vehicles transporting raw sugar. The multiple VEs are the raw sugar batch and the finished products truck, which contains pallets of product packages.

The main resources of the model are: the alarm system, which is an on-device resource on the gateway of the SC link; the basic data processing, that occurs on the gateway of the SC link or the truck's sink node; on-network processing, which occurs in the cloud; data processing on the agents' sink nodes. The services that use these resources are: measurement service; process information providing (which depends on the agent); smartphone app and web interfaces for consulting information in-loco or remotely; and system status.

The devices that are part of the context are: sensors in vehicles transporting raw sugar (RH), in industry's production lines (temperature, RH), in the DCs and warehouses (temperature, RH) and in trucks transporting finished products (temperature, RH); identification tags on trucks, pallets, boxes and packages; gateways and sink nodes on each SC agent. The main users of the model are the SC agents, consumers, Government and certification authorities.

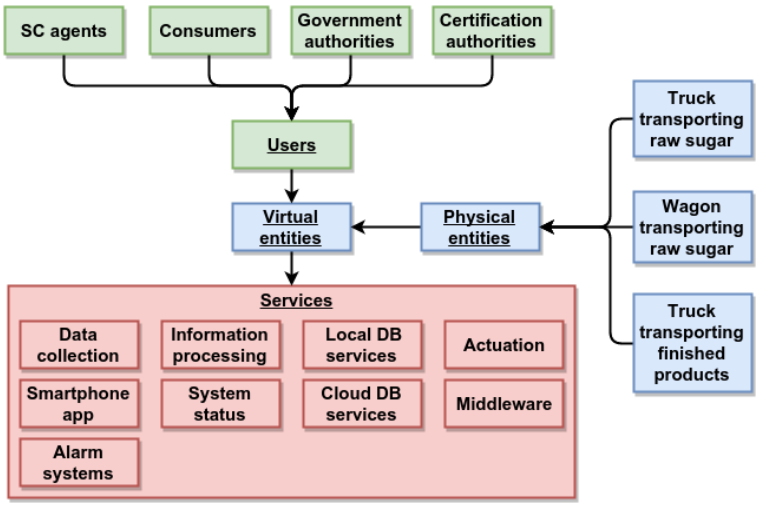
IoT domain model

The IoT domain model is composed of the users, PEs and VEs, and services. Its goal is to describe the system at a higher level. It contains the main elements of the system and how they interact with each other. Figure 1 illustrates this model.

The VEs will be able to access nine services using our proposed model: (i) Data collection: gather data from sensors, tags, and processes for the different agents; (ii) Smartphone app: used by the transporters to monitor the data collected by the sensors in the vehicle; (iii) Alarm systems: warn the responsible in the SC link if any problem occurs; (iv) Information processing: extract information from the collected data; (v) System status:

routines to check the system for errors; (vi) Local DB services: insert and modify data on the agents' DBs; (vii) Cloud DB services: insert and modify data on the cloud DB; (viii) Actuation: act on the PEs' environment; and (ix) Middleware: standardize all the data collected to insert it on the cloud DB. These are enough to fulfill the SCs requirements and address the shortcomings of the other reference models.

Figure 1. IoT domain model for the sugar supply chain



DISCUSSIONS

A domain model was proposed for the implementation of IoT in the sugar SC, describing all its functionalities, considering the IoT-A reference model. This is important to reduce the interoperability problems that may happen if a reference model is not used.

We provided a requirements list for the sugar SC, which can be adapted for other agroindustrial SCs. An important aspect to consider while building domain models, according to Bauer et. al (2013), is that it should consider the evolution of technologies in the context. We addressed this aspect by providing a generic framework that encompasses all available technologies in communication, data collection (considering product traceability and product and process monitoring), data processing and data storage.

The balance between implementation and operational costs, and also the benefits of implementing the proposed model are important topics for further investigation. Mainly because there are agents or links that can have great influence over others. These are called focal links and, in the sugar SC, the food industry and the retailers have that kind of power. Both can force the implementation of the abovementioned technologies on their suppliers. This generates a problem in terms of investments division among the chain links.

Based on this research, we believe that both the Government (of the producing country) and the SC agents should share the infrastructure costs. The Government can participate in developing loans that have low taxes and longer payback periods, making the implementation more attractive for these agents. On the other hand, agents themselves should cover the operational costs, since among the benefits of implementing the proposed model there are financial gains, that will be able to mitigate those costs and increase the agents' competitiveness. However, there should be a third-party organization in order to regulate how these costs will be shared among the agents, otherwise, the focal companies may force other SC links to pay for most of the costs.

The main limitations of the proposed domain model are: the need for validation with the SC agents; the lack of a simulation model to analyze its impacts in comparison to the current situation; the need to analyze aspects of security and privacy; and the need to further describe the information and communication models.

CONCLUSIONS

The implementation of IoT in agroindustrial SCs is important in order to fulfill the consumers' traceability needs and product and process monitoring along the SCs. We found that most models in the literature are not comprehensive enough to allow the development of applications, because they lack important components or functionalities. We identified the main requirements for the implementation of IoT in the Brazilian sugar SC and proposed a model that fulfills these requirements. The IoT-A reference model was used as a basis.

Future work is related to: validation with the SC agents; evaluation of its implementation through simulation; and extension to other agroindustrial SCs. We will also consider aspects related to: evolution of the technologies and their impact; security and privacy factors and how they shape the data; and information exchange in our model.

ACKNOWLEDGMENTS

This work was supported by Itaú Unibanco S.A. through the Itaú Scholarship Program, at the Centro de Ciência de Dados (C2D), Universidade de São Paulo, Brazil, by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brazil (CAPES) - Finance Code 001, and also by the National Council for Scientific and Technological Development (CNPq).

REFERENCES

ATZORI, L.; IERA, A.; MORABITO, G. The Internet of Things: A survey. Computer Networks, v. 54, n. 15, p. 2787-2805, 2010.

BAUER, M. et. al. Internet of Things – Architecture IoT-A Deliverable D1.5 – Final architectural reference model for the IoT v3.0. 2013.

CHOPRA, S.; MEINDL, P. Supply Chain Management: Strategy, Planning and Operation, 5th ed. New Jersey: Pearson Education, 2013.

GUBBI, J.; BUYYA, R.; MARUSIC, S.; PALANISWAMI, M. Internet of Things (IoT): A vision, architectural elements, and future directions. Future Generation Computer Systems, v. 29, n. 7, p. 1645-1660, 2013.

ITU - International Telecommunication Union. ITU-T Y.4000/Y.2060, Recommendation: Overview of the Internet of things. 2012. Accessed on: 01/06/2019. Available on: https://www.itu.int/rec/T-REC-Y.2060-201206-I.

JIANG, L.; XU, L.D.; CAI, H.; JIANG, Z.; BU, F.; XU, B. An IoT-Oriented Data Storage Framework in Cloud Computing Platform. IEEE Transactions on Industrial Informatics, v. 10, n. 2, p. 1443-1451, 2014.

ONEM2M. oneM2M - The interoperability enabler for the entire M2M and IoT ecosystem - White paper. OneM2M, Technical Report, 2015. Accessed on: 01/06/2019. Available on: http://www.onem2m.org/images/files/oneM2M-whitepaper-January-2015.pdf.

PANG, Z.; CHEN, Q.; HAN, W.; ZHENG, L. Value-centric design of the internet-of-things solution for food supply chain: Value creation, sensor portfolio and information fusion. Information Systems Frontiers, v. 17, n. 2, p. 289–319, 2015.

SILVA, R.F.; PRAÇA, I.; YOSHIZAKI, H.T.Y.; CUGNASCA, C.E. Proposal of a traceability model for the raw Brazilian sugar supply chain using RFID and WSN. In: Proceedings of the 2015 POMS Conference. Washington, DC, USA: Production and Operations Management Society (POMS), 2015, p. 1–10.

SMITH, I.G. The Internet of things 2012 - New horizons, A. Smith, I.G., Vermesa, On., Friess, P., Furness, Ed., Halifax, UK, 2012.

VERDOUW, C.N.; VUCIC, N.; SUNDMAEKER, H.; BEULENS A. Future Internet as a Driver for Virtualization, Connectivity and Intelligence of Agri Food Supply Chain Networks. International Journal of Food System Dynamics, v. 4, p. 261–272, 2013.

VERDOUW, C.N.; WOLFERT, J.; BEULENS, A.J.M.; RIALLAN, A. Virtualization of food supply chains with the internet of things. Journal of Food Engineering, v. 176, p. 128-136, 2016.

WEYRICH, M.; EBERT, C. Reference Architectures for the Internet of Things. IEEE Software, v. 33, n. 1, p. 112–116, 2016.

Difundindo a Agricultura Digital na Agricultura Familiar da Baixada Fluminense

Larissa Paiva[1], Pedro Vieira Cruz[1],
Sérgio Manuel Serra da Cruz[1], Ana Cláudia de Macedo Vieira[2]

[1] Departamento de Computação, Universidade Federal Rural do Rio de Janeiro, Seropédica, Rio de Janeiro, Brasil, serra@pet-si.ufrrj.br

[2] Departamento de Farmacobotânica, Universidade Federal do Rio de Janeiro, Ilha do Fundão, Rio de Janeiro, Brasil, vieira@ufrj.br

RESUMO

Até meados do século XXI a demanda global de alimentos crescerá 70% devido ao crescimento populacional, tal situação requererá a adoção de novas técnicas e o uso racional dos recursos naturais para assegurar uma produção eficiente de alimentos para fornecer uma dieta balanceada para a população. O objetivo deste trabalho é apresentar uma plataforma de agricultura familiar digital denominada Magé Agro Familiar que pode ser utilizada em arranjos produtivos locais para apoiar a comercialização direta de produtos orgânicos e agroecológicos e a difusão dos conceitos de agricultura digital entre os pequenos agricultores familiares da Baixada Fluminense (Rio de Janeiro, Brasil).

PALAVRAS-CHAVE: Fazendas Inteligentes, Agricultura 4.0, Aplicativos móveis.

ABSTRACT

Until the middle of the 21st century the global food demand will grow about 70% due to population growth, it will require the adoption of new techniques and the rational use of natural resources to ensure efficient food production to the population. The objective of this work is to present a novel digital agriculture named Magé Agro Familiar that can be used in local productive arrangements to support the direct marketing of organic and agroecological products and the diffusion of the concepts of digital agriculture among the small family farmers of Baixada Fluminense (Rio de Janeiro, Brazil).

KEYWORDS: Smart Farms, Agriculture 4.0, mobile apps.

INTRODUÇÃO

Até meados do século XXI a demanda global de alimentos crescerá devido ao crescimento populacional, tal situação requererá a adoção de técnicas e novos sistemas inovadores além do uso racional dos recursos naturais para assegurar uma produção eficiente de alimentos para fornecer uma dieta balanceada para a população (ONU, 2018). O Brasil poderá desempenhar um papel central na mitigação dessas demandas. Segundo o Censo Agropecuário do IBGE, 84,4% dos estabelecimentos rurais do Brasil são propriedades familiares e ocupam 74,4% da mão de obra que atua no campo (IBGE, 2006). De acordo com Bittencourt (2018), tais propriedades compreendem apenas 24,3% de toda a área rural do país e produzem cerca de 70% da comida que chega às mesas das nossas residências.

Esses números revelam que a agricultura familiar (AF) tem relação direta com a segurança alimentar e nutricional da população brasileira. A AF impulsiona economias locais e contribui para o desenvolvimento rural sustentável ao estabelecer uma relação íntima e vínculos duradouros da família com seu ambiente de moradia e produção (GUANZIROLI, CARDIM, 2000). No entanto, a AF não é homogênea, varia de acordo com as regiões brasileiras e vivencia diversos problemas, é cercada por um conjunto de preconceitos, por exemplo, de que o agricultor busca, basicamente, a subsistência e é avesso à inovação tecnológica (PEDROSO, 2017).

Atualmente existem inúmeras dificuldades para que o agricultor familiar estabeleça boa comunicação com os seus consumidores, por exemplo, dificuldades de escoamento da produção; dificuldades planejamento de produção de acordo com as demandas; desconhecimento das cotações nos mercados, entre outras. Adicionalmente, os produtores recebem novas demandas dos consumidores que estão se tornando mais exigentes em suas decisões de compra. Agora, os consumidores consideram um amplo conjunto de fatores além de apenas preço e qualidade; consideram rastreabilidade, procedência, sustentabilidade e relação do produtor com o meio ambiente. Essas dificuldades podem culminar em prejuízos por não conseguirem escoar a produção, contribuindo para o aumento do desperdício de alimentos, de esforço humano e desperdícios de recursos naturais ao longo da cadeia produtiva.

Tanto o relatório *World Urbanization Prospects* de 2018 das Nações Unidas quanto diversos autores relatam a inovação e a defesa da AF são fundamentais e devem caminhar conjuntamente (ONU, 2018 e HLPE, 2019). De acordo com Sousa e Silva (2006), propostas

para o apoio a AF devem envolver desafios de caráter tecnológico e inovações pontuais e compartilhadas, como condições de mercado, educação, saúde, água, solo e logística. Segundo Canevesi et al. (2017), as inovações na AF devem considerar os princípios da autonomia, da superação de assimetrias geradas por modelos excludentes e concentradores de renda; elas devem criar condições para a manutenção da viabilidade econômica das propriedades familiares e ampliar sua capacidade de se reproduzir como unidade social familiar, além de poder contribuir para a modernização do setor. Massruhá e Leite (2017) apontam que a modernização da AF passa pela capacitação, uso adequado de insumos, utilização de sistemas e equipamentos apropriados ao segmento e adoção de algumas técnicas da agricultura digital como, por exemplo, computação ubíqua baseada em dispositivos móveis, sensores, métodos analíticos para processar grandes massas de dados, sistemas de suporte à tomada de decisões baseados em computação em nuvem e métodos computacionais de alto desempenho. A modernização na AF pode elevar os índices de produtividade, eficiência do uso de insumos, da redução de custos com mão de obra, melhorar a qualidade do trabalho e a segurança dos trabalhadores e reduzir os impactos ao meio ambiente.

O objetivo deste trabalho é contribuir com a difusão de inovações na AF no estado do Rio de Janeiro através da oferta de ferramentas da agricultura digitais de baixo custo que podem contribuir com a mitigação das dificuldades supracitadas. A plataforma *Magé Agro Familiar* é uma ferramenta colaborativa que pretende agregar valor aos produtos da AF e contribuir com o aumento da conectividade, sustentabilidade e a racionalização da cadeia produtiva de produtores que praticam AF.

Conceitualmente, a plataforma é classificada como um Sistema Agrícola de Próxima Geração (SAPG) (ANTLE et al., 2017), ela tem como propósito disponibilizar informações em tempo real e apoiar a negociação direta entre os produtores de regiões rurais ou peri-urbanas e o crescente mercado consumidor (restaurantes, escolas, creches, sacolões, feiras, consumidores finais, entre outros) de alimentos frescos (orgânicos e agroecológicos) e produtos artesanais de qualidade produzidos pelos agricultores que adotem aspectos éticos, étnicos relacionados com a sustentabilidade.

MATERIAL E MÉTODOS

Esta seção apresenta materiais, métodos e principais tecnologias adotadas na condução da pesquisa e na elaboração deste manuscrito.

Materiais

Segundo Antle et al., (2017) e Janssen et al., (2017), os SAPG são definidos como o encadeamento de sistemas de informação modulares apoiados tanto em fluxos de dados

quanto por fluxos de requisições à micros serviços. Os SAPG requerem uma infraestrutura computacional composta por softwares distribuídos com baixo acoplamento baseados no modelo de camadas, bancos de dados distribuídos, APIs e nuvens de computadores. Os SAPG podem ser executados em dispositivos móveis e plataformas virtualizadas, facilitando o reuso e compartilhamento de dados e componentes, a colaboração e a tomada de decisão por parte dos diferentes tipos de atores que atuam ao longo das três fases da cadeia agroalimentar.

Os SAPG são compostos por *frontend* e *backend* dispostos em múltiplas camadas, a saber: (i) *camada de dados* - são representados por repositórios de dados digitais de natureza semiestruturada utilizados para armazenar os dados da plataforma; (ii) *camada de modelos* - representam os problemas da cadeia agroalimentar a serem equacionados; (iii) *camada de síntese* - representa, por exemplo, os algoritmos, rotinas, subsistemas, serviços, scripts de transformações de dados ou *workflows* que serão codificados para contribuir com a solução dos problemas identificados na camada de modelagem. Por fim, (iv) *camada de interfaces* - representam os pontos de interação entre os diferentes atores da cadeia agroalimentar e os demais módulos do SAPG.

Em geral, os SAPG serão constituídos por módulos de baixo acoplamento hospedados em ambiente de nuvem de computadores (JANSEN et al., 2017). Os fluxos de dados entre os módulos de um SAPG podem ser simples ou complexos e podem incluir, por exemplo, sequências de coleta, extração, transformação de dados e operações de enriquecimento (anotações, proveniência e ligação) de dados; utilização de um ou vários modelos preditivos; integração de dados de diferentes modelos matemáticos; visualização de dados produzidos pelos modelos, geração de mapas e relatórios digitais, entre outros.

Métodos

Do ponto de vista metodológico, delinearam-se as seguintes etapas: Após realizar entrevistas com especialistas nas áreas de agronomia e economia com o intuito de identificar as oportunidades de mercado; conduziu-se uma revisão sistemática da literatura (não apresentada neste texto por limitações de espaço) para avaliar trabalhos relacionados. Além disso, realizaram-se várias sessões de entrevistas presenciais com aproximadamente 30 pequenos produtores familiares na Cooperativa de Produtores de Magé (COPAGÉ). Adotou-se o mapeamento de requisitos baseado no processo de ideação do *Design Thinking* (BROWN, 2017) onde detectaram-se as oportunidades e desafios para projetar a plataforma.

Devido as possíveis limitações de Internet no campo, a camada de interface (*frontend*) da plataforma foi desenhada para operar com cobertura limitada de redes de dados, a camada está baseada no conceito *off-line first* e adota tecnologias (*Web Storage, App Cache, React*

Nativo, *Javascript* e *SQLite*) para melhorar a forma como os usuários utilizam os protótipos na Web. As demais camadas da plataforma (*backend*) estão baseadas no paradigma *serverless computing* (AMAZON, 2019), no banco NoSQL MongoDB (MONGODB, 2019) e em APIs de troca de mensagems do Whatsapp, de geolocalização do Google Maps e de troca de dados de solos da plataforma *OpenSoils* (CRUZ et al, 2018). Além disso, configurou-se uma nuvem de computadores na Amazon Web Services (EC2/S3) e o gerenciador de API WSO2 Manager para documentar as APIs e facilitar sua conectividade com outros sistemas.

A plataforma foi concebida para intermediar a negociação de mais de 40 produtos distintos produzidos por 30 agricultores familiares certificados por organismos credenciados e, que atuam em sistemas orgânicos de produção ou oriundos de processo extrativista sustentável e não prejudicial aos ecossistemas locais. As interfaces dos protótipos adotam design responsivo, se reorganizam e se adaptam a diferentes tamanhos de telas, apresentando seu conteúdo da melhor forma possível de acordo com os dispositivos dos usuários.

A prototipação e avaliação das funcionalidades foi baseada nos círculos ágeis descritos por Ries (2012), onde o autor descreve-os em três etapas (aprender, construir e medir). O método maximiza os aprendizados adquiridos pelos desenvolvedores junto aos usuários da plataforma por meio de ações incrementais e iterativas.

Todos os artefatos da plataforma são *open-source* e seu desenvolvimento ocorre nas dependências do PET-SI da UFRRJ (r1.ufrrj.br/petsi/mageagrofamiliar). Os primeiros experimentos de campo ocorrem na COPAGÉ, utilizando equipamentos com processador Intel Core i7, 1TB de disco rígido e 16GB de RAM e sistemas operacionais Unix-like. O aplicativo está sendo avaliado em *smartphones* e *tablets* equipados com sistema operacional Android versão 5.0 ou superior.

RESULTADOS E DISCUSSÃO

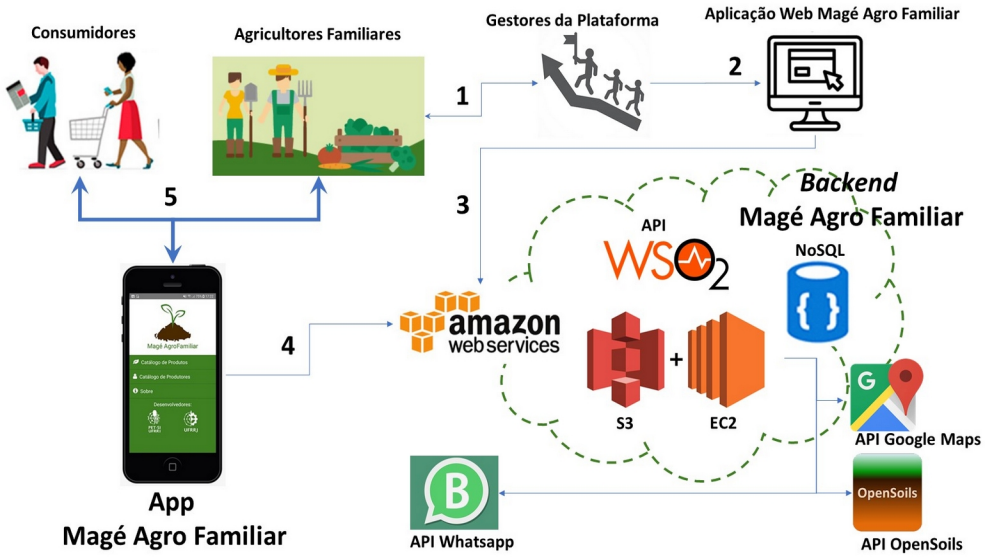
A região de Magé e Santo Aleixo, localizada na Baixada Fluminense do estado do Rio de Janeiro, foi escolhida como sendo o ponto de partida deste estudo devido ao projeto de extensionista conduzido por Vieira et al. (2016). Além disso, a região possui um expressivo quantitativo de agricultores familiares e um grande potencial de crescimento agrícola; a região tem disponibilidade de água e áreas próprias para cultivo de orgânicos, apresenta facilidade de escoamento por meio da boa malha viária e, principalmente, proximidade com o amplo mercado consumidor na Capital do estado do Rio de Janeiro e demais municípios da Baixada.

Do ponto de vista social, a plataforma tem como finalidade de ser utilizada de modo colaborativo por clientes e agricultores que tenham poucos conhecimentos dos fundamentos de computação, devendo operar em localidades com limitada cobertura de telefonia e utilizar

tecnologias de baixo custo operacional. Segundo levantamentos na literatura, a plataforma Magé Agro Familiar é uma das primeiras plataformas sociais de agricultura digital com foco na rede de pequenos agricultores da Baixada Fluminense.

A Figura 1 ilustra, de modo simplificado a operação conceitual da arquitetura proposta, suas camadas e resume o ciclo de interações entre os usuários da plataforma.

Figura 1. Visão conceitual da utilização da plataforma Magé Agro Familiar e ciclo com as principais interações entre os usuários (numeradas de 1 até 5).



Fonte: Elaborado pelos autores.

Conceitualmente ocorre o seguinte ciclo de interações: (1) a plataforma Magé Agro Familiar permite que o próprio produtor estabeleça o que ele irá disponibilizar através de lista semanal de produtos, além dos quantitativos para a comercialização fornecida aos gestores da plataforma; (2) Os produtores também podem consultar dados históricos de produtividade e sobre a qualidade dos solos de suas propriedades através das APIs da plataforma *OpenSoils*. A plataforma não requer que o produtor se preocupe com a logística e todos os problemas de entregar produtos, em centenas de pontos na cidade, apenas com a planejamento, produção e negociação; (3) Os gestores mantêm as listas de produtores, produtos e seus quantitativos, cadastro dos produtores e suas certificações, *feedbacks* emitidos pelos consumidores em notificações. Além disso, eles monitoram a operação do aparato computacional na nuvem da AWS, fazem as divulgações da ferramenta em diversos canais de comunicação ou personalizam a plataforma para atender outros APL; (4) Os consumidores se conectam a plataforma através do aplicativo, selecionam os produtos/produtores do *marketplace* e iniciam a negociação; (5) A negociação entre produtor e cliente é direta, dispensando a presença de intermediários, ao término da negociação ocorre atualização dos quantitativos de produtos

disponíveis na plataforma. Caso haja demanda por um produto cujo quantitativo seja maior que o disponibilizado por um produtor a plataforma poderá sugerir outros produtores para suprir a demanda.

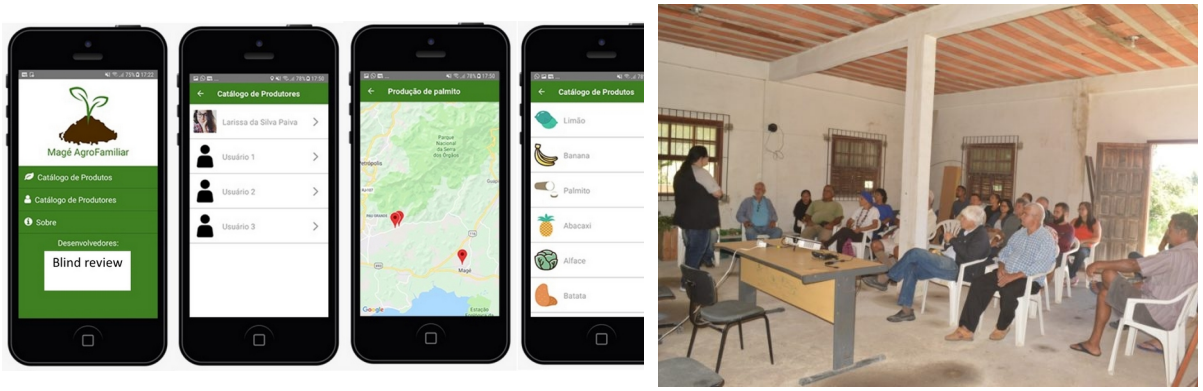
A plataforma permite utilizar o GPS dos dispositivos móveis e facilitar a localização dos produtores e o georeferenciamento das propriedades através do uso de APIs do Google Maps e promover planejamento da produção e marketing direto, ou seja, os alimentos disponíveis no *marketplace* nem sequer foram colhidos quando os clientes (restaurantes, sacolões, minimercados, creches etc.) fizerem suas escolhas. No momento, a plataforma não atua na área distribuição logística porta-a porta, que poderá incluir desde a colheita diária de produtos nas propriedades rurais até entrega para os clientes.

Outros diferenciais da plataforma são a inclusão digital e a auto-organização da rede agricultores familiares da região. O grupo, durante o desenvolvimento do projeto, definiu que apenas os produtores credenciados e com boa reputação profissional e apoiadores dos compromissos socioambientais e de boas práticas de qualidade poderão fazer parte da rede de produtores que terá sua visibilidade assegurada na plataforma Magé Agro Familiar. Periodicamente, os membros da rede avaliam os *feedbacks* emitidos pelos clientes para arbitrar pela continuidade dos produtores ou a inclusão de novos parceiros na rede. Por fim, destacamos que a plataforma é totalmente parametrizável e escalável, sendo facilmente adaptável para outros APL no Estado.

Os primeiros impactos da plataforma já estão sendo alcançados, são essencialmente de natureza inclusiva, ou seja, como resultados preliminares, ocorre paulatinamente a inclusão digital dos pequenos agricultores, trata-se de uma força de trabalho que frequentemente era subvalorizada e marginalizada. Durante as reuniões de projeto para avaliar as funções, projetos de telas (Figura 2) e avaliação de protótipos houve frequência de aproximadamente 20-30 agricultores e agricultoras familiares por reunião. Além disso, observa-se o amadurecimento da rede, durante as reuniões houve aumento da conscientização da necessidade de legalizar as propriedades junto à prefeitura e demais órgãos de controle. No entanto, um dos resultados preliminares mais animadores foi a constatação da ampliação da participação do quantitativo de jovens agricultores, interessados na temática da agricultura digital, na utilização do aplicativo ou explicando aos mais idosos as vantagens de sua utilização. Informes sobre as reuniões estão disponíveis no sítio da prefeitura de Magé[1].

[1] http://canalmage.rj.gov.br/agricultura-sustentavel/2018/09/17/agricultores-de-mage-fazem-parceria-com-ufrrj-e-criam-app-para-vender-producao/

Figura 2. |Telas do aplicativo Magé Agro Familiar (esquerda) exibindo as interfaces de produtores, geolocalização das propriedades e lista de produtos produzidos por agricultura orgânica e (direita) uma imagem de reunião de avaliação do protótipo junto ao grupo de agricultores familiares de Magé.



Fonte: Elaborado pelos autores.

CONCLUSÕES

No mundo contemporâneo, cada vez mais os avanços das TIC e a economia mais justa e sustentável terão um caráter estratégico na produção de alimentos. Nesse contexto, computação ubíqua poderá atuar como uma das molas propulsoras da inovação da cadeia produtiva da AF brasileira. No entanto, o acesso à tecnologia e recursos são os fatores que mais influenciam a persistência da exclusão entre os agricultores familiares.

Este trabalho adotou o como referenciais teóricos o SAPG e o *Design Thinking* para o desenvolvimento da plataforma Magé Agro Familiar que tem como missão contribuir com a difusão da agricultura digital entre os pequenos AF, apoiar a conectividade na comercialização direta de produtos orgânicos e agroecológicos e possivelmente reduzir os desperdícios de alimentos, ajudando a capacitar novos gestores no campo.

Atualmente, a plataforma está sendo desenvolvida e avaliada juntamente aos agricultores e agrônomos da EMATER-Rio e pesquisadores da UFRJ e UFRRJ. Dentre os trabalhos futuros, estima-se finalizar a implantação da plataforma, apresentar aos agricultores e demais parceiros para em seguida suas funcionalidades sob um maior número de variáveis e, em seguida, treiná-los para a utilização antes da divulgação para o público em geral.

AGRADECIMENTOS

O presente trabalho foi parcialmente realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001, do Conselho Nacional de Pesquisas (CNPq) – Código de Financiamento produtividade DT -II (315399/2018-0), do Programa MEC/PET, da EMATER-Rio, da Prefeitura de Magé e da Cooperativa COPAGÉ. Por fim e não menos importante, agradecimentos a eng. Agrônomo

Edison Rodrigues e a todos os agricultores que colaboraram no desenvolvimento da plataforma.

REFERÊNCIAS

AMAZON – Amazon Web Services. 2019. Disponível em: < https://aws.amazon.com/pt/>. Acesso em: 03 jun. 2019.

ANTLE J.M., JONES J.W., ROSENZWEIG C.E., Next generation agricultural system data, models and knowledge products: Introduction. Agricultural Systems.; 2017, pp. 155:186-190.

BROWN, T. Design Thinking: Uma metodologia poderosa para decretar o fim das velhas ideias. 1. Ed. eBook Kindle, 2017. 240 p.

CANAVESI, F. C., BIANCHINI, V., SILVA. H. B. C. Inovação Na Agricultura Familiar No Contexto Da Extensão Rural E Da Transição Agroecológica. In: A política nacional de agroecologia e produção orgânica no Brasil: uma trajetória de luta pelo desenvolvimento rural sustentável. 1. Ed. EMBRAPA, 2017. 20p.

CRUZ, S. M. S et al. OpenSoils: e-Science em Segurança de Solos II. Encuentro Latinoamericano de e-Ciencia (TICAL), Cartagena da Índias. Anais da Octava Conferência de Directores de Tecnología de Información y Comunicación en Instituciones de Educación Superior, y II Encuentro Latinoamericano de e-Ciencia 2018. 8p.

DAVIS, F. D. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of lnformation Technology, MIS Q., vol. 13, no. 3, 1989. p. 319–340.

GUANZIROLI, C.E; CARDIM, S.E. C. S. Novo retrato da agricultura familiar: o Brasil redescoberto. Brasília: Ministério do Desenvolvimento Agrário, INCRA, Instituto Nacional de Colonização e Reforma Agrária, 2000.

HLPE - The High Level Panel Of Experts On Food Security And Nutrition. Agroecological and other innovative approaches for sustainable agriculture and food systems that enhance food security and nutrition. June 2019. Disponível em: < http://www.fao.org/fileadmin/user_upload/hlpe/hlpe_documents/HLPE_Briefs-Flyers/HLPE_Report_14-Executive-Summary-A5_WebRes_DEF.pdf >. Acesso em: 08 jul. 2019.

IBGE - INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. Censo agropecuário: agricultura familiar. Brasília, DF, 2006. Disponível em: <http://www.ibge.gov-.br/home/estatistica/economia/agropecuaria/censoagro/agri_familiar_2006/familia_censoagro2006.pdf>. Acesso em: 03 jun. 2019.

JANSSEN. S. J. C. et al. Towards a new generation of agricultural system models, data, and knowledge products: building an open web-based approach to agricultural data. Agricultural Systems. Vol. 155, 2017. p. 269-288.

MASSRUHÁ, S. M. F. S., LEITE, M. A. DE A. Agro 4.0 – Rumo À Agricultura Digital. 2017. JC na Escola Ciência, Tecnologia e Sociedade: Mobilizar o Conhecimento para Alimentar o Brasil. Disponível em: < https://ainfo.cnptia.embrapa.br/digital/bitstream/item/166203/1/PL-Agro4.0-JC-na-Escola.pdf>. Acesso em: 01 jun. 2019.

MONGODB. The database for modern applications. Disponível em: <https://www.mongodb.-com/>. Acesso em: 10 jun. 2019.

ONU, United Nations, Department of Economic and Social Affairs, Population Division (2019). World Urbanization Prospects: The 2018 Revision (ST/ESA/SER.A/420). New York: Disponível em: <https://population.un.org/wup/Publications/Files/WUP2018-Report.pdf/>. Acesso em: 20 set. 2019.

PEDROSO, M. T. M. Acesso à tecnologia: a verdadeira questão social no campo. Cienc. Cult., São Paulo, v. 69, n. 4, p. 48-51, Oct. 2017. Disponível em: <http://cienciaecultura.b-vs.br/scielo.php?script=sci_arttext&pid=S000972520170004000015&lng=en&nrm=iso>. Acesso em: 03 jun. 2019.

RIES, E. A Startup Enxuta. 1. Ed. LeYa, 2012. 288 p.

SOUSA, I. S. F. de; SILVA, J. de S. Agricultura familiar e tecnologia. In: SOUSA, I. S. F. de (Ed.). Agricultura familiar na dinâmica da pesquisa agropecuária. Brasília, DF: Embrapa Informação Tecnológica, p. 25-36, 2006.

VIEIRA, A. C. M. et al. Conhecendo, conservando e comendo Plantas Alimentícias Não Convencionais (PANC) nos municípios de Magé e Guapimirim (RJ). In: 7o Congresso de Extensão Universitária, Ouro Preto, 2016.

Algoritmo genético aplicado no refinamento de adubação em experimento com alfaces

Vinícius Tessele[1], Augusto Vaghetti Luchese [2], Arnaldo Candido Junior[3]

[1] Departamento de Informática, Universidade Tecnológica Federal do Paraná, Medianeira, Paraná, Brasil, vinicius_tessele@hotmail.com, arnaldocan@gmail.com

[2] Departamento de Agronomia, Universidade Federal do Paraná, Palotina, Paraná, Brasil, aluchese@gmail.com

RESUMO

Algoritmos de Inteligência artificial (IA) tem sido aplicado com sucesso nas mais diversas áreas, dentro das ramificações da IA, os algoritmos genéticos se caracterizam por possuírem inspirações na biologia evolutiva e apresentam uma alternativa para solução de problemas complexos como mineração de dados, busca e otimização entre outras. Buscou-se aprimorar o processo de recomendação da adubação em uma área especifica com a ajuda de um algoritmo genético chamado como sistema imunológico artificial (SIA). Uma adubação correta proporciona uma melhora desenvolvimento das plantas e um aumento da produção. No experimento o algoritmo SIA, pode ser utilizado para gerar recomendações de N-P-K de modo a verificar sua eficiência na agricultura. Foi utilizada o cultivo da alface, por se tratar de um cultivar de desenvolvimento rápido o que permitiu a realização de sete iterações para a coleta dos dados desta pesquisa. De posse dos dados gerados, observou-se um aumento significativo da produtividade mesmo com a redução na aplicação dos nutrientes.

PALAVRAS-CHAVE: Inteligência artificial, Agricultura, Adubação.

ABSTRACT

Artificial Intelligence (AI) algorithms have been successfully applied in the most diverse areas, within the ramifications of AI, the genetic algorithms are characterized by having inspirations in evolutionary biology and present an alternative for solving complex problems such as data mining, search and optimization among others. We sought to improve the

fertilization recommendation process in a specific area with the help of a genetic algorithm called the artificial immune system (SIA). Correct fertilization provides improved plant development and increased yield. In the experiment the SIA algorithm can be used to generate N-P-K recommendations in order to verify its efficiency in agriculture. It was used the cultivation of lettuce, because it is a fast growing cultivar which allowed the accomplishment of seven iterations for the collect of the data of this research. With the data generated, a significant increase of productivity was observed even with the reduction in nutrient application.
KEYWORDS: Artificial intelligence, Agriculture, Fertilization.

INTRODUÇÃO

Algoritmos genético faz parte da inteligência artificial que visa desenvolver ferramentas que contenham comportamento e características dos seres vivos para seu funcionamento. Atualmente empregam a IA em várias áreas do conhecimento estes algoritmos são empregados em muitos casos na resolução de problemas complexos como mineração de dados, busca e otimização, aprendizagem de máquina, detecção de falhas. (PILA, 2015).

Na área de otimização são muitos os benefícios nos últimos anos, e tem surgido diversas propostas de algoritmos que possam ser aplicados em problemas reais e que obtiveram resultados encorajadores. A otimização pode apresentar diferentes formas como uni-objetivo no qual existe somente um critério a ser otimizado, outros como multi-objetivo, que possuem muitos critérios a serem otimizados, isso quer dizer que podem possuir apenas um único ótimo global ou apresentar múltiplos ótimos. (COELHO; ZUBEN, 2009).

Os sistemas imunológicos artificiais (SAI) é parte da computação evolutiva dos algoritmos genéticos que tem como princípio a teoria da evolução das espécies, descrita por Darwin, com isso o objetivo do algoritmo é avaliar a capacidade dos indivíduos da população a sobreviver e a de se reproduzir e espera-se que características dos indivíduos mais adaptados sejam transferidas para gerações futuras. (ARTERO, 2009). Segundo Coppin (2013) é possível explicar os algoritmos genéticos em comparação a evolução natural e que pequenas alterações nos genes podem ocorrer em uma população e juntamente com a reprodução, tendem a aprimorar a aptidão da população, ao longo do tempo. Característica principal da SIA é utilizar conceitos do sistema imunológicos dos seres vivos, para Abbas, Lichtman e Pillai (2012) o sistema imunológico (SI) representa a defesa do organismo contra agentes infecciosos, são mecanismos de proteção que podem eliminar as substancias estranhas que por ventura possam se estabelecer no organismo. Ao identificar um antígeno as células dos

organismos sofrem o processo de expansão clonal das células especifica a combater o antígeno.

As SIA se apropriam das características da SI para realizar a expansão clonal, que é a possibilidade das células se proliferarem, após serem expostas a um agente infeccioso, outra característica é a especialização, as células em contato com o agente infeccioso, respondem especificamente a cada contato, assim maximizando a sua eficiência, outro fator importante são as células de memória, que tem por finalidade manter por longos períodos as características de reconhecimento do antígeno, para que em nova infecção as respostas imune sejam rápidas.

O objetivo é estudar o método de implementação do algoritmo SIA na agricultura para que a reposição dos nutrientes do solo sejam mais eficiente, utilizando o algoritmo de rede imunológica artificial *Artificial Imune NETwork* (aiNet) proposto pode De Castro e Timmis (2002) a fim de refinar a recomendação de adubação da área estudada, para que isso aconteça o desenvolvimento de uma software que possa receber os dados de entrada referente a produção e efetue o processamento e que no final apresentem os melhores resultados. O algoritmo se destina a otimização e busca de resultados ótimos dentro de um conjunto de possibilidades.

MATERIAL E MÉTODOS

Para o experimento foi necessária uma área de 200m², no distrito de Silveirópolis (- 24° 17’38.1” S – 53° 34’10.7” W), na cidade de Assis Chateaubriand oeste paranaense. Foram construídos quatro canteiros de 23m de comprimento por 1m de largura, que foram divididos em 1,2m de comprimento por 1m de largura totalizando 92 parcelas de 1,2m². Foram selecionadas, 12 parcelas para serem utilizados como testemunhas, que puderam ser utilizadas para confronto dos resultados da produção. Os canteiros tiveram como identificação placas, A. B, C e D e as parcelas de 1 a 20, formando então uma matriz de 4 linhas e 20 colunas.

Após o preparo do local do experimento, foi definido o sequenciamento da pesquisa;

Análise de solo inicial, para verificar as necessidades da área estudada. Os compostos de N.P.K para a aplicação foram selecionados. Ureia, Superfosfato simples e Cloreto de Potássio.

Foram plantados 12 pés de alface-crespa por parcela, totalizando 1.104 pés por iteração. As mudas foram adquiridas do mesmo fornecedor e com no mínimo de 30 dias de germinação. Espaçamento mínimo de 25cm entre plantas, este espaçamento foi respeitado em todas as interações do experimento. A colheita efetuada a 45 dias após o transplante. Ao final

dos 45 dias, foram colhidas as 4 plantas centrais, cortados-s rente ao solo, a massa fresca das plantas pode se determinada utilizando-se uma balança com 0,01g de precisão. A massa fresca bruta (MFB), obtida a partir da parte aérea separada das raízes, a segunda pesagem foi realizada após a limpeza de folhas e caules danificados, denomina massa fresca comercial (MFC).

A análise de solo foi realizada no dia 10 de agosto de 2016, na cidade de Palotina – PR, pela Universidade Federal do Paraná – UFPR. De posse dos dados da análise de solo, foi prescrita a recomendação de adubação com base no Manual de adubação e de Calagem dos Estados do Rio Grande do Sul e de Santa Catarina

Foram aplicados, Ureia na porção de 26g, Superfosfato Simples 27g e de cloreto de Potássio de 19g, em 80 parcelas do experimento. Teve como objetivo, obter dados de referência, como média de produção de alfaces e o desvio padrão, para servir de dados iniciais e comparativa de resultados.

Os canteiros atribuídos testemunhas foram sorteados, e foram adubados da seguinte maneira; nenhum nutriente, outro de 50% da recomendação e 100% da recomendação inicial. A finalidade é observar o comportamento das parcelas testemunhas em relação as parcelas cuja a adubação foi calculada pelo algoritmo.

No desenvolvimento e implementação do algoritmo foi utilizado a linguagem de programação *object pascal* com a ferramenta Delphi Xe10, escolhida por proporcionar um desenvolvimento multiplataforma através dos seus frameworks *FireMonkey*, conforme descrito por Duarte (2015).

Teve como principal característica do software inclusão e alteração de dados de parâmetro, identificação dos tratamentos, lançamento de produção, geração de novas recomendações com base e avaliação do algoritmo.

O Algoritmo aiNet foi proposto por Castro e Zuben com o propósito de encontrar um modelo reduzido de dados pela eliminação de redundância em um conjunto de dados (STIBOR; TIMMIS, 2007). É um modelo inspirado na teoria da rede imunológica proposta por Jerne (1974). Foi desenvolvido para uso de valores no espaço de forma euclidiana, entretanto, nada impede o uso de outras formas. O objetivo do algoritmo aiNet é encontrar anticorpos a partir de sua afinidade, em uma concentração de antígenos no modelo computacional; o qual representa um modelo matemático inspirado na estrutura neural e um organismo inteligente. Estes adquirem conhecimento por meio da experiência (COELHO; ZUBEN, 2009).

A expressão de mutação por afinidade aplicada foi a seguinte:

$$c' = c + \alpha N(0,1),$$

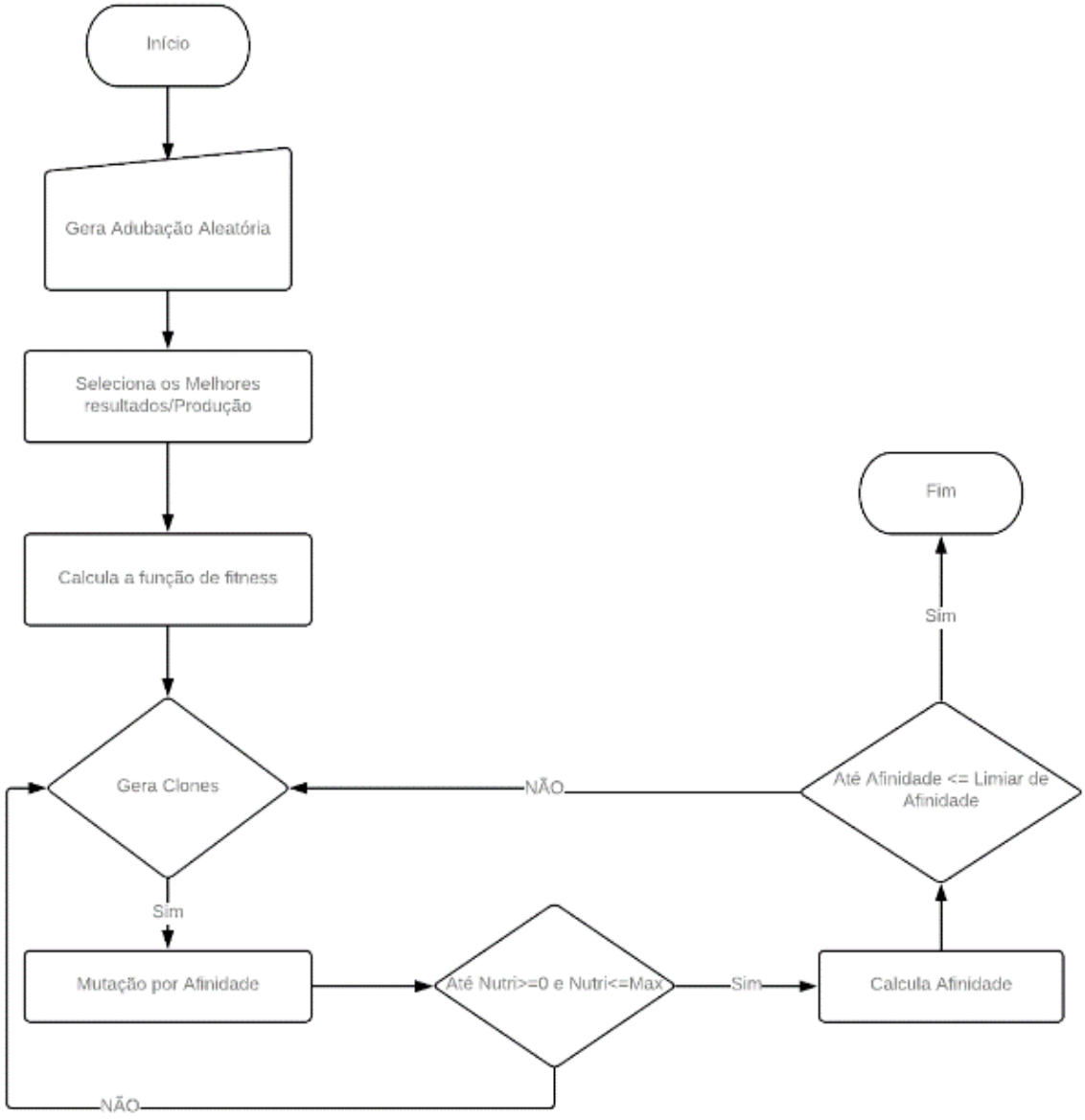
$$\alpha = (1 / \beta) \exp(- f^*),$$

C é a célula de origem que será clonada, c' é célula após a mutação, N (0,1) sendo N uma variável ao acaso de média 0 e desvio padrão $\sigma = 1$, β é o parâmetro que controla o decaimento da

inversa exponencial da função, f * é a aptidão dos indivíduos normalizada da célula clone. (DE CASTRO; TIMMIS. 2002)

Foram realizadas alterações pontuais no algoritmo proposto por De Castro & Timmis (2002) para atender as necessidades do experimento conforme a Figura 1.

Figura 1 – Fluxograma do algoritmo adaptado de De Castro & Timmis (2002) – Problema de Otimização.



Fonte: (autoria própria)

O algoritmo recebe valores de parâmetro inicialmente pré-definidas, gerando inicialmente um conjunto aleatório de recomendações, ao final da colheita e pesagem das alfaces de todas as parcelas, o algoritmo seleciona em ordem decrescentes as parcelas com os melhores resultados obedecendo a função de custos (F), estas parcelas são selecionadas para que sejam clonadas e gerarem novos descendentes. Para a função de custos é considerado o valor do produto em Kg, menos os valores dos

nutrientes por Kg, valores estes observados no mercado atual. Outro parâmetro utilizado para o algoritmo é valores menores que 0 e maiores que 2,5 vezes a recomendação padrão dos nutrientes são suprimidos da rodada. Após a geração dos clones é iniciado o cálculo de aptidão média da população gerada. A Figura 2 exemplifica o cadastro dos parâmetros utilizados no experimento.

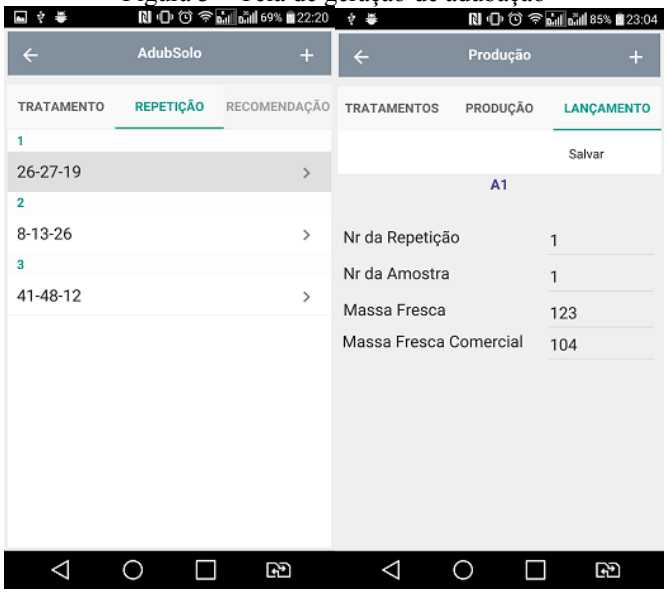
Figura 2 – Tela de Parâmetros



Fonte: (autoria própria)

Na Figura 3 são apresentadas as telas de recomendações de N-P-K geradas em cada uma das rodadas, na figura ao lado a identificação da parcela colhida e o peso da produção como Massa fresca e Massa fresca comercial.

Figura 3 – Tela de geração de adubação



Fonte: (autoria própria)

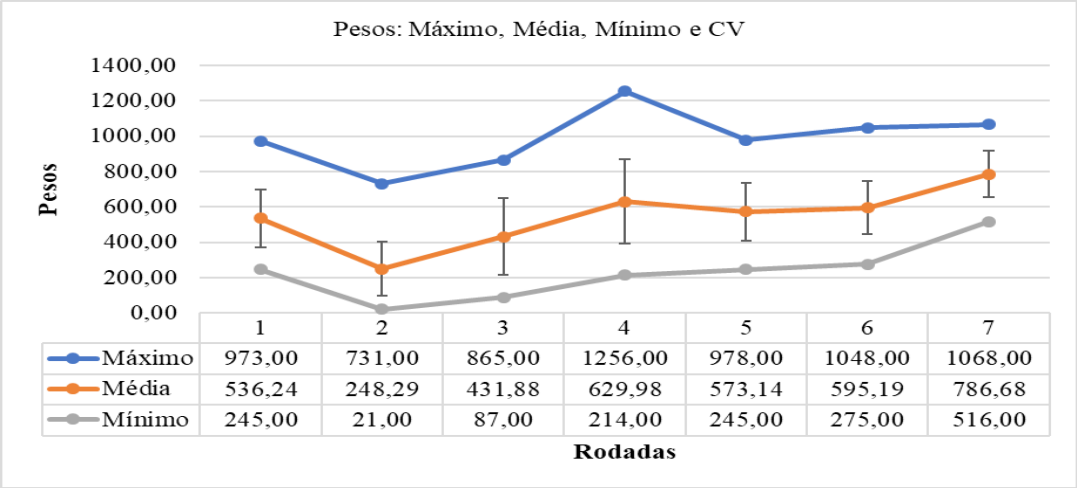
RESULTADOS E DISCUSSÃO

O experimento tem se mostrado resultado satisfatório a partir da sexta rodada, sugere que o algoritmo do SIA apresenta um melhor resultado na recomendação de adubação, possibilitando o aumentado da

quantidade produzida e consequentemente um maior lucro por reduzir a quantidade de adubação aplicados.

Na Figura 4 são apresentados os valores obtidos em cada uma das iterações, é importante verificar que a linha de média do gráfico apresenta uma melhora do coeficiente de variação (CV) das amostras em relação as rodadas 3 e 4, que apresentam uma maior dispersão dos pesos.

Figura 4. Pesos máximos, média, mínimo e CV.



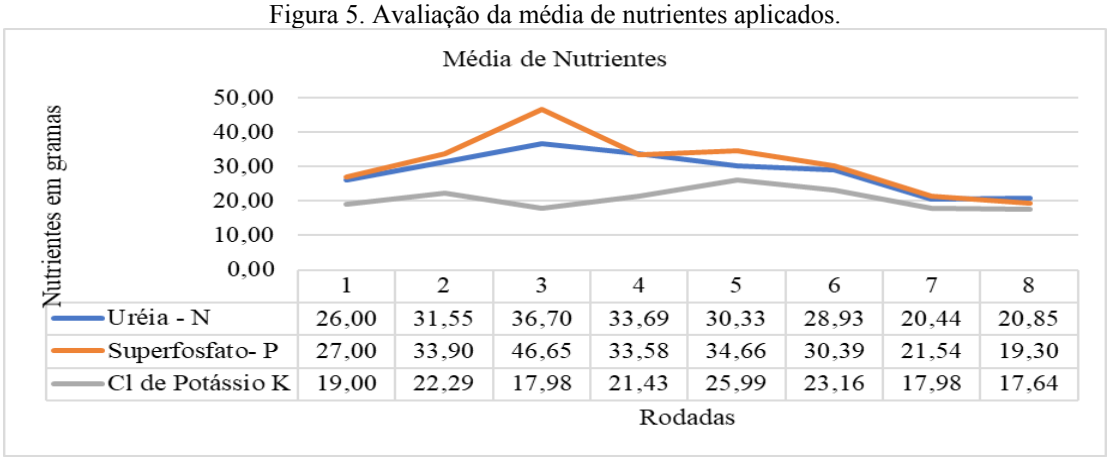
	1	2	3	4	5	6	7
Máximo	973,00	731,00	865,00	1256,00	978,00	1048,00	1068,00
Média	536,24	248,29	431,88	629,98	573,14	595,19	786,68
Mínimo	245,00	21,00	87,00	214,00	245,00	275,00	516,00

Fonte: (autoria própria)

Dentre os diversos parâmetros, o algoritmo deveria selecionar parcelas que possuíam produção maior ou igual 800g, os valores que ultrapassarem este parâmetro, são selecionados para a clonagem e a maturação por afinidade.

Na Figura 5 percebemos outro importante resultado, no qual a média dos nutrientes aplicados no experimento, diminuem a partir da terceira rodada, essa informação sugere que o algoritmo tenha encontrado valores mais equilibrados para a aplicação e consequente aprendizado tenha sido efetivo.

Figura 5. Avaliação da média de nutrientes aplicados.



	1	2	3	4	5	6	7	8
Uréia - N	26,00	31,55	36,70	33,69	30,33	28,93	20,44	20,85
Superfosfato- P	27,00	33,90	46,65	33,58	34,66	30,39	21,54	19,30
Cl de Potássio K	19,00	22,29	17,98	21,43	25,99	23,16	17,98	17,64

Fonte: (autoria própria)

CONCLUSÕES

Para este experimento o algoritmo mostrou que oferece resultados positivos, o que sugere que nas próximas rodadas o aumento da produtividade deverá ser continuo, percebe-se que é possível obter resultados positivos no processo de otimização para a recomendação de adubação utilizando os algoritmos genéticos, entretanto existe a necessidade de um números de interações e coleta de dados de produção maiores, para confirmar tal tendência e a possibilidade de realizar testes estatísticos apropriados para verificar a significância do resultado.

REFERÊNCIAS

ABBAS, A. K.; LICHTMAN , A. H.; PILLAI, S. H. I. V. Imunologia celular e molecular. 7. ed. Rio de Janeiro: Elsevier, 2012.

ARTERO, A. O. Inteligência artificial: teoria e pratica. São Paulo. Ed Livraria da Física 2009.

COPPIN, Bem. Inteligência artificial. Rio de Janeiro. LTC. 2013.

COELHO, G. P., ZUBEN, F. J. V. Uma proposta imuno-inspirada para Omni-Otimização

DE CASTRO, L, N; TIMMIS, J. Atificial Immune Systems: a new computational intelligence approach. London, Uk: Springer, 2002.

DUARTE, W. Delphi para Android e iOS: Desenvolvendo Aplicativos Móveis. Rio de Janeiro: Brasport, 2015.

STIBOR, T.; TIMMIS, J. An investigation on the compression quality of aiNet. In: Foundations of Computational Intelligence, 2007. FOCI 2007. IEEE Symposium on. IEEE, 2007. p. 495-502.

PILA, A. D. História e terminologia a respeito da computação evolutiva. Revista Ciências Exatas e Tecnologia. Faculdade Comunitária de Santa Bárbara. 2006.

Impactos da IoT na Avicultura: um Mapeamento Sistemático

Felipe Dias de Oliveira[1], *Alixandre Santana*[1], *Luis Filipe Alves Pereira*[1]
[1]Universidade Federal Rural de Pernambuco, Garanhuns, Pernambuco, Brasil,
alixandre.santana@ufrpe.br, luis.filipe@ufrpe.br

RESUMO

O Brasil é hoje o maior exportador e o segundo maior produtor mundial de carne de frango. Apesar dessa posição de destaque, a ausência de dados sobre a utilização de ferramentas baseadas em *Internet of Things* (IoT) na indústria de avicultura nacional reforça a hipótese de que a maior parte dos avicultores brasileiros ainda usam métodos convencionais de operação em detrimento de uma maior produtividade. Nesse estudo, foram mapeados 17 trabalhos da literatura internacional que apresentam soluções baseadas em IoT que perpassam desde a segurança alimentar, fatores ambientais de produção, rastreabilidade e saúde dos animais. Espera-se, portanto, contribuir para discussão entre a comunidade científica da computação e do agronegócio por meio de uma revisão sistemática.

PALAVRAS-CHAVE: avicultura, IoT, revisão da literatura, automação

ABSTRACT

Brazil is today the largest exporter and the second largest producer worldwide of chicken meat. Despite such prominent position, the lack of data about the use of tools based on Internet of Things (IoT) at national poultry farming leads to the hypothesis that the most Brazilian poultry farmers still use conventional methods to the detriment of higher productivity. In this work, we mapped 17 publications from the international literature that show IoT-based solutions that pass through food safety, environmental production factors, traceability, and animal health. This way, we expect to contribute to a discussion between the computer and agribusiness scientific communities through a systematic review.

KEYWORDS: poultry farming, IoT, systematic review, automation

INTRODUÇÃO

O rápido avanço tecnológico observado nas últimas décadas vem embarcando dispositivos eletrônicos com conexão remota a objetos comuns do nosso dia-a-dia. Dessa forma, cria-se objetos capazes de receber, manipular e enviar informações. Nesse contexto, o termo *Internet-of-Things* (IoT) refere-se a rede global que é criada interconectando *objetos inteligentes* (MIORANDI et al., 2012). Novas aplicações da IoT surgem a todo momento nas mais diversas áreas do conhecimento humano, tais como: saúde (DOUKAS; MAGLOGIANNIS, 2012), transporte e logística (SICARI; RIZZARDI; A COEN-PORISINI, 2019), automação residencial (PARK et al., 2017), dentre outros.

Na agroindústria, aplicações da IoT estão equipando produtores rurais com ferramentas de suporte à decisão e soluções automatizadas que permitem um avanço na produtividade, qualidade da produção, e no lucro da propriedade (OLAKUNLE et al., 2018). Soluções como essas podem (i) estimar o surgimento de pragas na lavoura e, consequentemente, reduzir o uso frequente de inseticidas e fungicidas (HEAMIN et al., 2017); (ii) controlar as condições ideais de temperatura e umidade em instalações para o armazenamento da produção (DEEPAK et al., 2017); (iii) disponibilizar para produtores e consumidores um registro atualizado e seguro do caminho percorrido pelos alimentos em todas as etapas da cadeia produtiva (REGATTIERI; GAMBERI; MANZINI, 2007); (iv) gerenciar dados técnicos, estratégicos e operacionais de toda a propriedade rural para planejamento, documentação, e otimização de processos (KALOXYLOS et al., 2012); dentre outros.

No Brasil, uma dos setores de maior destaque da agroindústria é a avicultura. O país é atualmente o maior exportador de carne de frango do mundo e detém a segunda maior produção mundial (USDA, 2019). Não há dúvidas sobre a importância desse setor para balança comercial brasileira. No entanto, devido a ausência de dados sobre a utilização de IoT na indústria de avicultura no Brasil, nós partimos do pressuposto que - assim como ocorre na Malásia (MANSOR et al, 2018) - a maioria dos avicultores ainda usam métodos convencionais de operação onde todas as rotinas de monitoramento e controle de suas granjas são feitas por trabalhadores humanos em detrimento de uma maior produtividade.

Acreditamos ainda, conforme apontado no plano de ação para internet das coisas, que a agricultura brasileira é um setor estratégico que pode se beneficiar significativamente da

implementação de IoT (BNDES, 2019). Dessa forma, o presente trabalho busca mapear soluções baseadas em IoT para problemas encontrados na avicultura, elicitando o estado da arte para discussão na comunidade científica da informática e do agronegócio por meio de uma revisão sistemática.

MATERIAL E MÉTODOS

O foco da presente revisão sistemática está em estudos que propõem soluções de software e hardware baseadas em IoT que foram desenvolvidas nos últimos anos para otimizar processos relacionados à avicultura. Foram utilizados os *guidelines* de Kitchenham e Charters (2007) como método de pesquisa. Procurou-se então responder à seguinte pergunta de pesquisa: *Quais são as soluções baseadas em IoT encontradas na literatura voltadas para o contexto da avicultura e quais seus objetivos?*

Os critérios de seleção de estudos, bem como as estratégias de busca adotadas, são descritas a seguir:

- **Critérios de seleção de estudos:** a Tabela 1 apresenta os cinco critérios de inclusão e os cinco critérios de exclusão que foram estabelecidos para guiar a seleção de estudos para a presente revisão sistemática.

Tabela 1 - critérios utilizados para incluir ou excluir estudos da revisão sistemática proposta.

Critérios de Inclusão	Critérios de exclusão
• Publicação entre 2010 e Janeiro de 2019	• Apresentação de soluções usando IoT para outras áreas da agroindústria
• Apresentação de soluções para avicultura usando IoT	• Trabalhos incompletos
• Trabalhos completos	• Trabalhos disponíveis apenas como resumo
• Trabalhos não duplicados	• Capítulo de livros
	• Publicações no formato de livros

- **Estratégias de busca:** o protocolo de busca e revisão foi definido em três etapas: (i) execução de *string* de busca em indexadores de artigos; e revisões aplicando os critérios de inclusão e exclusão aos (ii) títulos; (iii) resumos; e (iv) seções de

introdução e conclusão dos artigos. Nesse contexto, a seguinte *string* de busca foi replicada nas bases *IEEE Xplore*, *Web of Science* e *Scopus*:

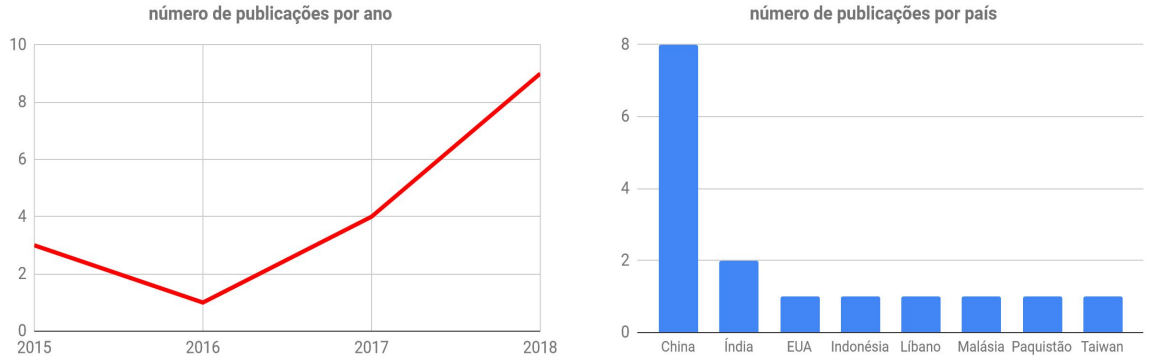
(*aviculture* OR *poult** OR *chicken** OR *fowl* OR *broile** OR *egg*) AND (*iot* OR *internet of things*)[1]

Dentre 404 estudos capturados pela *string* de busca nos indexadores de artigos, apenas 17 atenderam as condições de interesse definidas no protocolo.

RESULTADOS E DISCUSSÃO

Foram identificados 17 estudos que apresentam soluções baseadas em IoT para otimização de processos na avicultura. O gráfico apresentado na Figura 2 (a) mostra como o número dessas publicações cresceu nos últimos anos, em especial em 2018 em relação aos anos anteriores. Este fato permite-nos antever uma presença cada vez maior da IoT na avicultura nos próximos anos. Ao mesmo tempo, os dados coletados também nos revela que países asiáticos concentram a maior parte dos esforços nesse sentido: a Figura 2 (b) apresenta a distribuição das publicações selecionadas pelos países de origem de seus autores.

Figura 2 - métricas sobre os estudos selecionados nesta revisão sistemática: (a) crescimento do número de publicações que propõem soluções de IoT para avicultura ao longo dos últimos anos; e (b) número de publicações em função dos países de origem de seus autores.



(a) (b)

Fonte: produzida pelos autores

[1] o operador '*' indica a possibilidade de variações nas terminações dos radicais das palavras.

A incorporação de tecnologias de automação em qualquer domínio contribui , de uma forma geral, para melhorar o processo produtivo, a atualizar as funcionalidades do sistema e a incrementar o sistema de controle de qualidade. No contexto específico da avicultura, os objetivos das soluções desenvolvidas mapeadas foram divididos em três categorias **(i) sistemas para rastreabilidade, (ii) sistemas para monitoramento e controle de ambientes, (iii) sistemas de controle de produtividade e saúde de animais**. Os sistemas para rastreamento oferecem aos produtores, transportadores, e consumidores segurança e transparência quanto à origem e o processo produtivo de cada alimento individualmente. Por outro lado, os sistemas para monitoramento e controle coletam dados em tempo real sobre as instalações produtivas para gerar relatórios/alertas, ou acionar atuadores eletromecânicos com capacidades de intervir em situações que podem causar perdas ou adoecimento das aves. Existiram ainda as soluções preocupadas com a saúde e produtividade dos animais. Na sequência, as soluções encontradas são detalhadas.

Sistemas de rastreabilidade de produtos

Sistemas de rastreabilidade são úteis para garantir a origem e a transparência de um processo produtivo. Na indústria de alimentos, tem uma importante função na segurança alimentar[2]. No contexto da avicultura, Liu et al. (2015) propõem uma arquitetura de um sistema de **rastreio para ovos** baseado em vídeo e em redes sem fio. Nesse sistema um consumidor poderia consultar todos os **dados, incluindo os vídeos gerados durante o processo de produção dos ovos**. Pode-se questionar possíveis problemas de privacidade que os trabalhadores da cadeia produtiva possam ter, uma vez que suas imagens estariam expostas na web. Das dificuldades apresentadas nos projetos de rastreabilidade como descritas em (LIU et al., 2018), estão a grande quantidade de conexões que ocorrem entre os fornecedores, produtores, indústrias de transformação e vendedores, o que dificulta a correta coleta de informações sobre o produto que estão sendo manipulado.

Outro sistema similar e proposto por Zhu and Tang (2015), que **rastreia embalagens de ovos fazendo processamento de um identificador disponível na embalagem com um aplicativo android**. Esse sistema apresenta como vantagem seu baixo custo, uma vez que não precisa de outros sensores além da câmera do *smartphone*.

[2] http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1676-56482002000200008

Yu et al. (2018) apresentam um modelo de rastreabilidade segura baseada em tecnologias de IoT como RFID (do inglês, radio-frequency identification), QR code (do inglês, quick response code), tecnologia de sensores e comunicação wireless. Tal modelo abrange "*toda a cadeia da indústria, provendo clientes, empresas e o departamento de governo responsável de um sistema multicanal de rastreabilidade que tem importância significativa na redução do risco de segurança alimentar e de recall de produtos com defeitos"* (YU et al. 2018).

Sistemas de monitoramento e controle de fatores ambientais

Os fatores ambientais como elevação da temperatura e umidade, quantidade de luz no ambiente e qualidade do ar influenciam diretamente na produtividade das aves, como destacado por (GHAZAL et al. 2017). Há algumas variáveis que podem ser controladas e manipuladas para produzir uma condição adequada de confinamento na qual a menor diferença poderá afetar a produtividade da fazenda (MANSOR et al. 2018), daí a importância de sistemas para esse propósito.

Para proprietários que possuem quantidades significativas de granjas, realizar este monitoramento nem sempre é possível. Elevados custos para manter mão de obra suficiente para o trabalho inviabilizaram a produção. Como forma de solucionar esse problema, Lashari et al. (2018) propõem um modelo de sistema para monitoramento de **métricas ambientais como temperatura, umidade, e concentração de gases** para granjas inteligentes. O modelo proposto utiliza sensores e um Raspberry Pi[3] para capturar as informações de granjas geograficamente distribuídas e enviá-las para um servidor na nuvem. Os autores destacam que o proprietário tem a facilidade de acessar as informações de qualquer local e a qualquer tempo, podendo prevenir situações que possam causar perdas ou adoecimento das aves.

Em YuQun et al. (2018), os resultados experimentais mostram que o sistema proposto podem automaticamente coletar e analisar parâmetros ambientais como **temperatura do ar, umidade, intensidade da luz, velocidade do vento e qualidade do ar** em ambientes de produção avícola, de forma que o ambiente pode ser razoavelmente controlado a partir de um balanço dinâmico e automatizado de seus parâmetros.

Na mesma direção, Ghazal et al. (2017) propõem uma rede de sensores sem fio para monitorar e medir diversas variáveis ambientais de granjas como **temperatura, qualidade do**

[3] https://en.wikipedia.org/wiki/Raspberry_Pi

ar e quantidade de luz. O sistema proposto oferece diretrizes convenientes para diferentes cenários e desta forma, permite a adaptação para todas **as configurações de clima** e de parâmetros, promovendo condições favoráveis para o ambiente de confinamento para os galináceos.

Já Choukidar e Dawande (2017) apresentam um sistema de automação para monitoramento de granjas pelo qual também são coletados **níveis de ração e água**. O sistema alerta o administrador para o momento em que ele deve repor níveis dos reservatórios. O autor mostra que houve melhoras na redução do desperdício de ração e queda da mortalidade quando comparado com granjas convencionais.

Zhou et al (2017) propõem uma solução que consiste em dois subsistemas, incluindo uma rede local de sensores sem fio que é usada para coletar dados ambientais como **temperatura, CO_2 e controle de ventilação** em cada celeiro. A solução usa controladores de Arduino, módulos ZigBee[4], Raspberry Pi e *motor drivers*. A solução ainda é composta de um subsistema de controle monitoramento baseado em nuvem usado para visualizar dados, transmitir dados para a nuvem e transmitir comandos dos usuários para os controladores mecânicos.

Similarmente, Wicaksono et al (2017) e Goud e Sudharson (2015) propõem uma rede de sensores sem fio para monitoramento **de temperatura, umidade e níveis de água** em fazendas de aves. In Choukidar and Dawand (2017) , uma combinação de sensores sem fio e GPRS[5] (do inglês, *general packet radio service*) são usados para controlar temperatura, nível de água fumaça, gás e níveis de ração. O sistema foca em manter as melhores condições ambientais possíveis com o mínimo de esforço humano.

Sistemas de controle de produtividade e saúde dos animais

Um dos grandes desafios da avicultura é identificar uma doença no seu estágio inicial e evitar que ela se se espalhe entre os animais. Soluções mostram que é possível estimar a **saúde das aves de acordo com o seu comportamento (**FEIYANG et al. 2016, RAJ e JAYANTHI, 2018). Para estimar o estado de saúde das aves, Feiyang et al. (2016) utilizam RFID para monitorar os animais e mapear seus comportamentos. A plataforma dos autores é composta de: (1) sensores de peso e de torres coletoras de tags RFID para monitorar o comportamento

[4] https://pt.wikipedia.org/wiki/ZigBee
[5] https://pt.wikipedia.org/wiki/General_Packet_Radio_Service

das aves; (2) Tags RFID colocadas em cada uma das aves de forma que o movimento do animal entre duas torres possa ser registrado. A partir de então, características como **velocidade, habilidade de coletar comida e tempo de descanso** são extraídas; (3) as informações do passo anterior são entradas para classificadores que utilizam o método *K-means*. Os resultados do sistema foram similares aos obtidos a partir da classificação manual dos animais. Com o mesmo propósito, utilizando processamento de áudio e vídeo, Raj e Jayanthi (2018) conseguem realizar o mesmo tipo de classificação. Os autores criam um arquitetura que utiliza sensores de distância, de gás, de temperatura, de umidade, câmera RGB, câmera térmica e microfone.

A **produtividade** de postura de galinhas é tema da pesquisa de Chien and Chen (2018). A plataforma proposta executa a análise da produtividade individual de ovos de galinhas em tempo real, desta forma, possibilitando a substituição de animais com produtividade abaixo de um nível pré-determinado para que a taxa média de produtividade seja mantida.

CONCLUSÕES

O presente trabalho identificou o impacto de soluções tecnológicas baseada em IoT na avicultura. Tais soluções foram segmentadas em três categorias, a saber: sistemas para rastreabilidade, sistemas para monitoramento e controle de ambientes, sistemas de controle de produtividade e saúde de animais.

Os sistemas para rastreabilidade da produção atuam para assegurar a confiança que os clientes possuem na marca. Os sistemas de monitoramento controlam parâmetros ambientais das granjas visando garantir condições favoráveis de confinamento para os animais. Por fim, os sistemas de controle de produtividade e saúde de animais são capazes de identificar doenças em estágios iniciais e evitar que as mesmas se espalhem para outros animais.

Vemos que o impacto da IoT na avicultura é perceptível em várias frentes, todas elas são complementares e contribuem para a melhoria da produtividade das indústrias e o fortalecimento das suas marcas.

REFERÊNCIAS

BNDES. 2019. Internet das Coisas: um plano de ação para o Brasil. Disponível em <https://www.bndes.gov.br/wps/portal/site/home/conhecimento/pesquisaedados/estudos/estud

o-internet-das-coisas-iot/estudo-internet-das-coisas-um-plano-de-acao-para-o-brasil> Acesso em 20 jan 2019.

CHOUKIDAR, G. A. and DAWANDE, N. Smart poultry farm automation and mo-nitoring system. International Conference on Computing, Communication,Control and Automation (ICCUBEA), 2017.

CHIEN, Y. R.; CHEN, Y. X. An rfid-based smart nest box: An experimental study of laying performance and behavior of individual hens. Sensors, 2018.

DEEPAK, V., et al. Cold storage management system for farmers based on IoT. Int. J. Recent Trends Eng. vol. 3, no. 5, pp. 556–561, 2017.

DOUKAS, C.; MAGLOGIANNIS, I. Bringing IoT and cloud computing towards pervasive healthcare. IEEE Sixth International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing, 2012.

FEIYANG , Z. et al. Monitoring behavior of poultry based on RFID radio frequency network, International Journal of Agricultural and Biological Engineering, vol. 9, no. 6, pp. 139–147, 2016.

GHAZAL, B. et al. A poultry farming control system using a zigbee-based wireless sensor network. International Journal of Control and Automation, 2017.

GOUD, K.; SUDHARSON, A. Internet based smart poultry farm.Indian Journal of Science and Technology, 2015.

HEAMIN, L. et al. Disease and pest prediction IoT system in orchard: A preliminary study. Ninth IEEE International Conference on Ubiquitous and Future Networks (ICUFN), 2017.

KALOXYLOS, A. et al. Farm management systems and the Future Internet era. Computers and electronics in agriculture, pp. 130-144, 2012.

KITCHENHAM, B.; CHARTERS, S. Guidelines for Performing Systematic Literature Reviews in Software Engineering, versão 2.3. Relatório Técnico, Evidence-Based Software Engineering (EBSE), 2007.

LASHARI, M. H. et al. IoT based poultry environment monitoring system. IEEE International Conferenceon Internet of Things and Intelligence System (IOTAIS), 2018.

LIU, F. et al. The egg traceability system based on the video capture and wireless networking technology. International Journal of Sensor Networks, 2015.

LIU, L. et al. Information collection system of duck products based on IoT. EURASIP Journal on WirelessCommunications and Networking, 2018.

MANSOR, H. et al. Development of smart chicken poultry farm. Indonesian Journal of Electrical Engineering and Computer Science, pp. 498–505, 2018.

MIORANDI, D., et al. Internet of things: Vision, applications and research challenges. Ad Hoc Networks, p. 1497–1516, 2012.

OLAKUNLE, E., et al. An overview of Internet of Things (IoT) and data analytics in agriculture: Benefits and challenges. IEEE Internet of Things Journal p. 3758-3773, 2018.

PARK, E., et al. Comprehensive approaches to user acceptance of Internet of Things in a smart home environment, IEEE Internet Things Journal, vol. 4, no. 6, pp. 2342–2350, 2017.

RAJ, A.; JAYANTHI, J. IoT-based real-time poultry monitoring and health status identification. Institute of Electrical and Electronics, pp. 1–7, 2018.

REGATTIERI, A., GAMBERI, M., MANZINI, R. Traceability of food prod-ucts: General framework and experimental evidence. J. Food Eng.,v. 81, no. 2, pp. 347–356, 2007.

SICARI, S., RIZZARDI, A., COEN-PORISINI. Smart transport and logistics: A Node-RED implementation. Internet Technology Letters 2, no. 2, 2019.

USDA (UNITED STATES DEPARTMENT OF AGRICULTURE), Foreign Agricultural Service. Disponível em: <https://apps.fas.usda.gov/psdonline/app/index.html#/app/ advQuery >. Acesso em: 07 jul. 2019.

WICAKSONO, D. et al. Design and analysis automatic temperature control in the broiler poultry farm based on wireless sensor network. International Conferences on Information Technology, Information Systems and Electrical Engineering (ICITISEE), 2017.

YU, et al. Materials, processes,and facile manufacturing for bioresorbable electronics: A review. Advanced Materials, 2018.

ZHOU, J. et al. Develop a low-cost remote monitoring and control system for poultry barn. American Society of Agricultural and Biological Engineers, ASABE Annual International Meeting, 2017.

ZHU, S.; Tang, P. The design and implementation of eggs' traceability system based on mobile qr code. Advance Journal of Food Science and Technology, 2015.

Agricultura Digital: Levantamento junto ao Produtor Rural na Região Metropolitana de Campinas

Maria Angelica de Andrade Leite[1], *Daniel Azevedo Duarte*[2], *Silvia Maria Fonseca Silveira Massruhá*[1]

[1] Embrapa Informática Agropecuária, Campinas, São Paulo, Brasil,

angelica.leite@embrapa.br, silvia.massruha@embrapa.br

[2] Dados Assessoria de Imprensa e Comunicação, Campinas, São Paulo, Brasil,

daniel@dadoscomunicacao.com.br

RESUMO

Este trabalho apresenta um levantamento das principais necessidades do produtor rural da Região Metropolitana de Campinas com relação ao uso de tecnologias digitais na sua propriedade. O estudo foi realizado pelo Agrohub Campinas, uma iniciativa que visa fomentar a inovação das tecnologias digitais a partir das necessidades do produtor. Essa iniciativa vem ao encontro do Ecossistema de Agricultura Digital que está sendo proposto pela Embrapa Informática Agropecuária. Nesse ecossistema prevê-se uma interação entre os diversos atores comprometidos com a implantação da agricultura digital representados pelos produtores, *startups* e instituições públicas e privadas incluindo centros de pesquisa, universidades, empresas do agronegócio, órgãos reguladores e instituições governamentais. Esse levantamento é um estudo inicial e visa contribuir para aproximar as pesquisas que estão sendo desenvolvidas com as demandas dos produtores rurais.

PALAVRAS-CHAVE: Agricultura 4.0, Agricultor, Tecnologia Digital.

ABSTRACT

This paper presents a survey of the main needs of rural producers in the Metropolitan Region of Campinas regarding the use of digital technologies in their property. The study was conducted by Agrohub Campinas which is an initiative that aims to foster the innovation of digital technologies based on the needs of the producer. This initiative is related to the Digital

Agriculture Ecosystem that is being proposed by Embrapa Informática Agropecuária. In this ecosystem there is an interaction between the various actors involved in the implementation of digital agriculture represented by producers, startups and public and private institutions including research centers, universities, agribusiness companies, regulatory agencies and government institutions. This survey is an initial study and aims to contribute to approximate the research that is being developed with the demands of rural producers.
KEYWORDS: Agriculture 4.0, Farmer, Digital Technology.

INTRODUÇÃO

Diante dos desafios apresentados na agricultura, principalmente o de aumentar a produção agrícola sem ampliar a área plantada significativamente, torna-se premente o uso cada vez mais intenso de novas tecnologias para permitir os ganhos de produtividade de forma sustentável.

Ao longo do tempo a agricultura tem passado por várias transições que podemos classificar como três ondas principais. Na primeira onda ocorreu a revolução verde que consistiu em usar recursos tecnológicos para produzir mais alimentos no mesmo espaço de terra. Deste modo, se desenvolveram sementes de plantas geneticamente modificadas que produziam mais, tinham melhor resposta a fertilizantes e maior resistência a pragas. Além disso, começou-se a utilizar métodos modernos para a mecanização do campo. Na segunda onda, a intensificação agrícola se fortalece e a monocultura cede lugar aos sistemas integrados e rotacionados de produção como os sistemas de integração-lavoura-pecuária-floresta (ILPF). São sistemas mais demandantes de conhecimentos envolvendo múltiplas disciplinas. As pesquisas se tornam sistêmicas pois há a necessidade de entendimento de todo o sistema produtivo. Despontam os sistemas guiados e a agricultura de precisão. Na terceira onda tem-se a agricultura de base biológica e os sistemas se tornam complexos envolvendo várias variáveis. Existe a preocupação em desenvolver alimentos mais nutritivos e funcionais e que sejam mais resistentes tanto a doenças quanto a condições adversas diminuindo a necessidade de uso de insumos (PILLON, 2017).

A terceira onda da agricultura vai atender um consumidor mais exigente demandando alimentos de melhor qualidade, rastreáveis e produzidos de forma a respeitar o meio ambiente, em condições de trabalho éticas e considerando o bem estar animal. Além disso, a agricultura torna-se uma das protagonistas para o atingimento dos objetivos de desenvolvimento sustentáveis da ONU (NAÇOES UNIDAS BRASIL, 2019) como o de

“Acabar com a fome, alcançar a segurança alimentar e melhoria da nutrição e promover a agricultura sustentável”.

Para responder a todos esses desafios a terceira onda conta com a transformação digital que, por meio das tecnologias disruptivas como a computação em nuvem, big data, inteligência artificial, Internet das Coisas e robótica avançada, entre outras, será possível coletar e analisar um imenso volume de dados que serão produzidos pelas áreas de biotecnologia, geotecnologia e nanotecnologia com o uso de sensores, drones, satélites e máquinas agrícolas inteligentes culminado na chamada agricultura digital ou agricultura 4.0. (SHEPHERD et al. 2018; MASSRUHÁ; LEITE, 2018; LEITE; MASSRUHÁ, 2014; TRENDOV; VARAS; ZENG, 2019).

Vários levantamentos têm sido realizados cujas informações contribuem para a visão das condições atuais relativas à implantação da agricultura digital. Esses levantamentos incluem estudos sobre as *startups* do agronegócio, as AgTechs, as condições de conectividade da área rural e os hábitos do produtor rural.

De acordo com o 2o. Censo Agtech Startups Brasil (AGTECHGARAGE, 2019), realizado pelo AgTechGarage e a Esalq/USP, existe uma estimativa de pouco mais de 300 *startups* AgTechs no Brasil sendo que 184 participaram do censo. Esse número mais que dobrou em dois anos dado que no último Censo foram consideradas 75 Agtechs. Das 184 *startups*, foi verificado que 6% estão localizadas na Região Centro-oeste. A região Sudeste abriga a maioria das *startups* com 60% do total. A maioria delas atua em suporte à decisão, Internet das Coisas, gestão agrícola e agricultura de precisão.

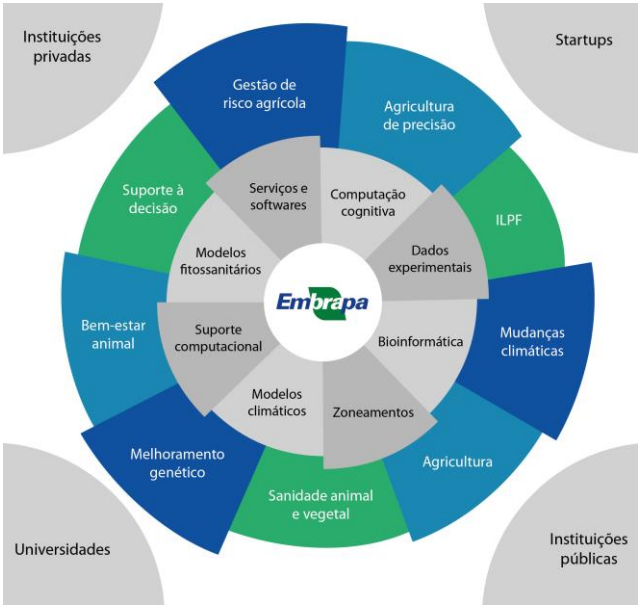
A 7ª Pesquisa Hábitos do Produtor Rural (ABMRA, 2017) aponta que cada vez mais o produtor adota tecnologias e está conectado. A pesquisa, realizada pela Associação Brasileira de Marketing Rural e Agronegócio, foi divulgada no final de 2017, onde foram entrevistados 2835 agricultores e pecuaristas (distribuídos nos 15 principais estados produtores). Ela mostra que 96% dos entrevistados possuem celular, 96% dos entrevistados usam WhatsApp, 67% o Facebook, 24% o YouTube, 8% o Instagram e 5% tem o Skype.

Essa nova dinâmica do agronegócio brasileiro também está se confirmando com os resultados do Censo Agro 2017, realizado pelo Instituto Brasileiro de Geografia e Estatística (IBGE, 2017). De acordo com o censo, 1.425.323 produtores declararam ter acesso à internet, sendo que 659 mil (46,2%) através de banda larga, e 909 mil (63,77%), via internet móvel. Em 2006, o total de estabelecimentos agropecuários que tinha acesso à internet era de apenas 75 mil, o que representa um aumento de 1.790,1% em 2017. O acesso ao telefone também cresceu. Pode-se destacar a evolução observada na existência do aparelho nos

estabelecimentos, passando de 1,2 milhão para 3,1 milhões de propriedades com acesso a telefone entre 2006 e 2017, um acréscimo de 1,9 milhão ou 158%.

Ciente de que a implantação da agricultura digital no Brasil requer a superação de muitos desafios, a Embrapa Informática Agropecuária propõe o Ecossistema de Agricultura Digital, conforme ilustrado na Figura 1. Nesse ecossistema a Embrapa Informática Agropecuária vai atuar como facilitadora tendo o papel de fomentar o trabalho colaborativo entre os vários atores englobando os produtores rurais, *startups*, instituições públicas e privadas para que possam compartilhar expertise e conhecimentos.

Figura 1: Ecossitema de Agricultura Digital da Embrapa Informática Agropecuária



Fonte: (EMBRAPA INFORMÁTICA AGROPECUÁRIA, 2018).

Uma das iniciativas alinhadas com esse ecossistema é o Agrohub Campinas (DADOS COMUNICAÇÃO, 2019) que tem como missão facilitar o acesso dos agricultores a soluções inovadoras para problemas reais de seu cotidiano fomentando a troca de informações e o encontro de diferentes *players* do agro a partir da necessidade e da base da cadeia: o produtor rural. Visando compreender o estado atual das necessidades do produtor rural da Região Metropolitana de Campinas (RMC), o Agrohub Campinas realizou uma pesquisa cujo objetivo foi o levantamento de informações relativas às necessidades desses produtores bem como sua visão sobre o uso das tecnologias digitais no ambiente rural. Esse trabalho apresenta a pesquisa realizada assim como uma análise das respostas obtidas na pesquisa.

MATERIAL E MÉTODOS

A Região Metropolitana de Campinas constitui um polo que reúne várias instituições e empresas atuantes no agronegócio. Conta com importantes instituições de pesquisa na agricultura como Instituto Agronômico de Campinas (IAC), Embrapa Informática Agropecuária, Embrapa Territorial, Embrapa Meio Ambiente, Faculdade de Engenharia Agrícola da Unicamp, Instituto de Tecnologia de Alimentos (ITAL), Agência Paulista de Tecnologia dos Agronegócios (APTA), Coordenadoria de Desenvolvimento Rural Sustentável (CDRS), Instituto Zootecnia (IZ), entre outras. Além disso, sedia várias instituições de pesquisa e tecnologia como o Polo Tecnológico de Campinas, CPQD, Laboratório Nacional de Luz Síncrotron (LNLS), Faculdade de Engenharia Elétrica e da Computação e Instituto de Computação da Unicamp, Pontifícia Universidade Católica de Campinas, Universidade Paulista (UNIP), entre outras.

Segundo dados da CDRS, Campinas possui 1012 propriedades rurais e 23 cooperativas em um raio de 70 km. Já a Região Metropolitana de Campinas possui cerca de 3 mil propriedades rurais e a Região Administrativa de Campinas mais de 10 mil propriedades com produção voltada a variadas culturas desde frutas e hortaliças até grãos, pecuária e eucalipto. Dado o ambiente propício para o desenvolvimento de inovações na área agrícola, ela foi escolhida como o foco para o levantamento das condições do produtor rural visando ter informações sobre a sua visão com relação ao uso de tecnologias no campo e suas necessidades.

Para isso, foram realizadas pesquisas quantitativas e qualitativas junto a produtores dos Sindicatos Rurais de Campinas e Valinhos nos meses de maio e junho de 2019. Foram definidas 17 perguntas que visaram compreender o tipo de negócio dos produtores, margem de lucro, existência de marca, embalagem dos produtos, aspectos de venda e comercialização, expectativa de aumento de produtividade; custos e uso de tecnologia. As 17 perguntas realizadas são mostradas na Tabela 1.

Tabela 1: Conjunto de perguntas aplicadas nas entrevistas e opções de respostas

Perguntas realizadas	**Opções de respostas**
a) O que o senhor produz?	Frutas; hortaliças; Grãos; Pecuária; Flores; Outros
b) Como o senhor faz a administração da sua produção?	Não faço; No papel; No computador; Software de gestão administrativa; Software de gestão administrativa e agronômica; Software de gestão administrativa e agronômica em tempo real
c) Como o senhor identifica e combate a pragas e ervas daninhas?	Não me preocupo; Quando vejo que está prejudicando, uso os "remédios" no meu jeito; Consulto um agrônomo; Tenho dados no computador, vejo o cultivo de vez em quando e consulto um agrônomo; Atualizo um software agronômico periodicamente e consulto um agrônomo; Software acompanha o cultivo em

	tempo real e alerta para usarmos só o necessário.
d) Como o senhor mede e administra os efeitos do clima e fertilização em seu cultivo?	Não me preocupo; Quando está seco, irrigo ou rego tudo. É tudo que posso fazer; Vejo as necessidades hídricas e de fertilizantes em cada área do cultivo; Usamos instrumentos para medir necessidade real de água e fertilização; Atualizo um software periodicamente com dados e uso só o que preciso; Software acompanha o cultivo em tempo real.
e) Como o senhor faz a classificação dos seus produtos (frutas, verduras e legumes)?	Não classifico. É tudo igual; No olho; Minha equipe separa por peso, tamanho e características visuais manualmente; Usamos balanças e réguas que ajudam a classificação; Automatizada com maquinário com sensores modernos.
f) O senhor acredita que a tecnologia pode ajudar?	Sim; Não; Talvez.
g) Quais tecnologias te interessam para sua produção?	Conexão à internet na propriedade; Software de administração; Software de monitoramento agrícola; Sensores de monitoramento agrícola e de irrigação; Equipamentos (robôs, drones,irrigação mecanizada, câmara fria,etc)
h) Qual o maior desafio para melhorar seus resultados?	Gestão da propriedade; Compra de Insumos; Identificação e combate a pragas, doenças ou ervas daninhas; Clima; Margem de lucro; Comercialização.
i) O senhor acredita que pode aumentar a produtividade de seus cultivos?	Sim; Não; Talvez
j) O senhor acredita que pode reduzir custos com uma gestão mais eficiente?	Sim; Não; Talvez.
k) Como o senhor organiza a compra de insumos?	Compro quando lembro, não planejo; Quando sei que outros produtores começaram a comprar; Compro com antecedência sempre na mesma época; Compro em grupo com outros produtores mas não acompanho o mercado; Acompanho a cotação dos insumos e compro quando vejo as melhores oportunidades; Acompanho as cotações e faço negociação coletiva.
l) Como o senhor sabe quanto precisa de insumos?	Compro sempre a mesma quantidade; Avalio apenas olhando a plantação e decido a quantidade; Consulto um agrônomo; Tenho os dados no computador, vejo o cultivo de vez em quando e consulto um agrônomo; Atualizo um software periodicamente com dados sobre o cultivo para comprar só o que preciso; Software acompanha o cultivo em tempo real.
m) Você tem alguma diferenciação de embalagem ou câmara fria?	Não; Só uma embalagem simples; Temos embalagem e marca própria; Temos embalagem, marca própria, *packing* com câmara fria; Temos embalagem, marca própria, *packing* com câmara fria e publicidade; Caixa fornecida por outra firma, e tenho a marca própria; Câmara fria.
n) Como o senhor trabalha sua marca?	Não tenho marca própria; Tenho só um logo e um nome; Tenho nome, logo e nós mesmos divulgamos nas mídias sociais; Trabalhamos o *branding* (nossa marca) profissionalmente; Procuramos sempre diferenciais como certificações, ações sociais e ambientais.
o) O senhor acredita que pode melhorar sua comercialização?	Sim; Não; Talvez.
p) Como o senhor vende seus produtos?	Vendo a produção a um intermediário; Vendo à Ceasa e intermediários; Vendo à Ceasa, intermediários e redes varejistas; Temos um monitoramento estruturado sobre preços em diferentes compradores; Vendemos como uma cooperativa, em conjunto; Usamos inteligência de mercado para obter melhores retornos.
q) O senhor acredita que pode aumentar a margem de lucro dos seus produtos?	Sim; Não; Talvez.

Para a realização da pesquisa foi necessária a definição do tamanho da amostra relativa ao número de respondentes que fosse relevante. A definição do tamanho da amostra foi realizada com base em cálculo amostral (COCHRAN, 1977) que considera fatores como

população, amostra, erro amostral, heterogeinedade da população e nível de confiança, segundo a fórmula $n = (N Z^2 p (1-p))/((N-1) e^2 + Z^2 p (1-p))$, na qual:

n = tamanho da amostra obtido por meio do cálculo;
N = total da população pertencente à pesquisa;
Z = desvio indicado ao valor médio aceitável para que o nível de confiança seja atingido;
e = margem de erro máxima que a pesquisa permite;
p = proporção que se deseja encontrar no cálculo.

Como existem cerca de 3 mil produtores na RMC, a amostra necessária para garantir 80% de confiança dentro de uma margem de erro de 10% é de 41 pesquisados. Na pesquisa efetuada houve um retorno de 42 pesquisados contabilizando a somatória das pesquisas quantitativas, realizadas por meio de questionário, e qualitativas em entrevistas pessoais.

A etapa quantitativa da pesquisa foi realizada a partir do compartilhamento, via *WhatsApp* e e-mail dos respectivos Sindicatos Rurais, de questionário de perguntas desenvolvido na plataforma *Google Forms*, no qual cada pergunta poderia ter apenas uma resposta de forma a destacar o tópico mais "urgente' ou a solução mais interessante para cada um dos produtores. Os questionários foram enviados para os produtores cadastrados nos Sindicatos Rurais de Campinas e Valinhos perfazendo um total de 300 questionários enviados Esses questionários tiveram 32 respondentes.

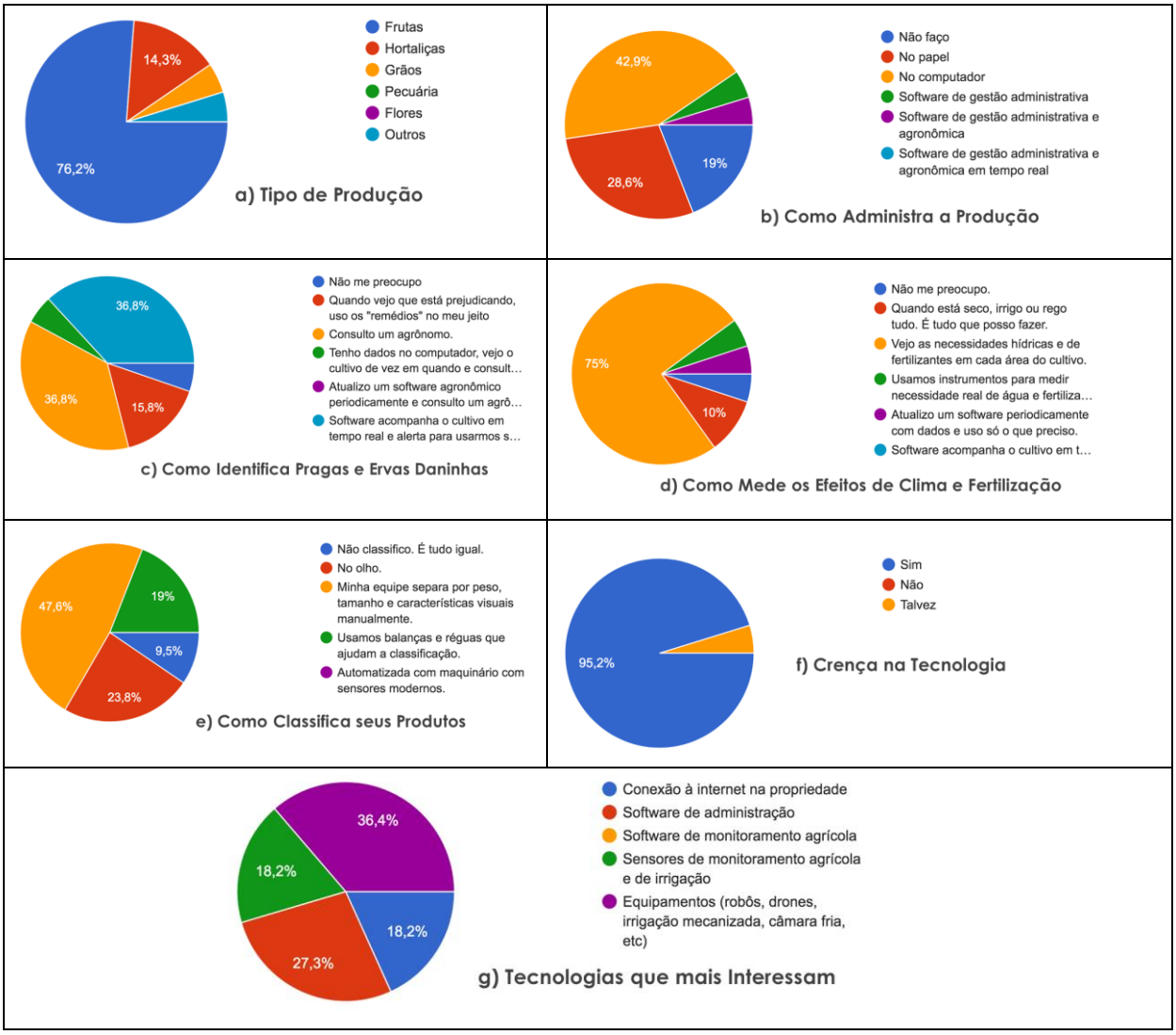
A etapa qualitativa da pesquisa constituiu-se em entrevistas pessoais realizadas na sede dos referidos sindicatos rurais. Essas entrevistas se basearam nas mesmas perguntas da pesquisa quantitativa, foram gravadas e alcançaram 10 produtores privilegiando o diálogo com objetivo de captar nuances e impressões mais sutis. No total, nesta primeira etapa, foram obtidas mais de 3 horas de gravações com relevantes *feedbacks* e *insights* dos produtores.

O processamento das perguntas foi realizado de modo a totalizar as respostas ao questionário de prospecção quantitativa e qualitativa em gráficos do tipo "pizza" para demonstrar os percentuais das alternativas escolhidas para cada pergunta. O cálculo das porcentagens foi realizado automaticamente pela plataforma do Google Forms.

RESULTADOS E DISCUSSÃO

A pesquisa apresentada nesse trabalho focou na visão dos produtores rurais da Região Metropolitana de Campinas. A Figura 2 apresenta os gráficos relativos às respostas às perguntas relacionadas com o uso de tecnologias digitais e software pelos produtores.

Figura 2 – Gráficos resultantes do levantamento junto aos produtores da RMC.



De acordo com o levantamento, pode-se ver que os produtores utilizam softwares para acompanhar diversas fases na sua produção embora em diferentes níveis. A atividade de administração da produção e identificação de pragas e plantas daninhas são as mais informatizadas. Ainda há bastante oportunidade de implantar o uso de softwares e tecnologias digitais em várias áreas do processo de trabalho rural. No gráfico 2(g) os produtores apontaram as tecnologias que mais interessam e, dentre os pesquisados, de acordo com o gráfico 2(f), 95,2% creem que a tecnologia pode ajudar na solução de seus problemas. Essa crença abre o caminho para que se estabeleça o diálogo maior entre os produtores, as cooperativas e as empresas ofertantes de tecnologias seja por meio de soluções prontas e gratuitas para alguns casos ou soluções adquiridas no mercado ou, ainda, soluções específicas que podem ser desenvolvidas junto aos centros de pesquisa.

CONCLUSÕES

Esse trabalho apresentou um levantamento com relação ao uso das tecnologias digitais e ferramentas de software nas propriedades rurais da Região Metropolitana de Campinas. Trata-se de um levantamento inicial para balizar futuras pesquisas e o desenvolvimento de tecnologias para o setor rural. Os próximos passos para o aprimoramento desse levantamento incluem: i) realizar encontros com os presidentes de todos os sindicatos rurais da região visando ampliar a base de pesquisa sobre as demandas dos produtores e engajar lideranças; ii) participar em um Stand no 2º Encontro dos Produtores Rurais de Campinas e Região para apresentar a agricultura 4.0 e elencar soluções gratuitas, prontas e acessíveis para percepção de valor imediata; iii) promover o primeiro encontro entre produtores e solucionadores na busca de soluções a serem desenvolvidas; iv) dar continuidade às entrevistas para ter uma visão mais abrangente da situação dos produtores da Região Metropolitana de Campinas; v) estratificar e classificar os produtores respondentes para verificar se eles representam os diversos perfis dos agricultores da região de Campinas visando aprimorar a amostra nas futuras entrevistas; vi) publicação de um trabalho mais completo considerando todas as questões realizadas.

O objetivo é aproximar cada vez mais os diversos atores envolvidos no Ecossistema de Agricultura Digital permitindo promover a implantação da agricultura 4.0 nas propriedades rurais. Nesse cenário, o Agrohub Campinas está dando um passo no sentido de conhecer as demandas e necessidades dos produtores da Região Metropolitana de Campinas para que possa atuar em parceria aproximando e promovendo a troca de informações de modo a fomentar a adoção e o desenvolvimento da agricultura 4.0 a partir dos problemas dos produtores da região. O AgroHub Campinas também se propõe a ser uma plataforma de testes, consolidação e oferta de soluções capazes de resolver demandas de agricultores e pecuaristas podendo considerar as demandas de outras regiões do Brasil e, futuramente, considerar sua expansão para outros países. A Embrapa Informática Agropecuária se coloca como um dos atores desse ecossistema articulando, junto às instituições parceiras, um polo de empreendedorismo em agricultura digital para o desenvolvimento e compartilhamento de serviços para os usuários do setor produtivo agrícola brasileiro.

REFERÊNCIAS

ABMRA. 7ª Pesquisa Hábitos do Produtor Rural. 2017. Disponível em: <http://www.abmra.org.br/2016/index.php/pesquisa-abmra/>. Acesso em: 08 jul. 2019.

AGTECHGARAGE. 2o. Censo Agtech Startups Brasil. Disponível em: <https://www.agtechgarage.com/censo/>. Acesso em: 08 jul. 2019.

BRITO, D. 40% das aeronaves não tripuladas no Brasil são usadas no agronegócio. Globo Rural. 2018. Disponível em: < https://revistagloborural.globo.com/Noticias/Pesquisa-e-Tecnologia/noticia/2018/09/40-das-aeronaves-nao-tripuladas-no-brasil-sao-usadas-no-agronegocio.html>. Acesso em: 10 jun. 2019.

COCHRAN, W. G. Sampling techniques. 3.ed. New York : John Wiley & Sons, 1977. 428p.

DADOS COMUNICAÇÃO. AgroHub Campinas. Disponível em <https://dadoscomunicacao.com.br/2019/06/22/agrohub-campinas-visa-fomentar-inovacao-a-partir-das-dores-do-agricultor/>. Acesso em: 05 jul 2019.

EMBRAPA INFORMÁTICA AGROPECUÁRIA. Relatório de gestão 2015-2018: pesquisa e inovação em tecnologia da informação e comunicação para a agricultura. Campinas, 2018. 67 p.

IBGE. Censo Agro 2017. Disponível em: < https://censos.ibge.gov.br/agro/2017/>. Acesso em: 02 jul. 2019.

LEITE, M. A. de A.; MASSRUHA, S. M. F. S.; EVANGELISTA, S. R. M.; SOUZA, K. X. S. de. Tecnologias emergentes - futuro e evolução tecnológica das AgroTIC. In: MASSRUHÁ, S. M. F. S.; LEITE, M. A. de A.; LUCHIARI JUNIOR, A.; ROMANI, L. A. S. (Ed.). Tecnologias da informação e comunicação e suas relações com a agricultura. Brasília, DF: Embrapa, 2014. Cap. 17. p. 335-353.

MASSRUHÁ, S. M. F. S.; LEITE, M. A. de A. Agro 4.0 - rumo à agricultura digital. In: MAGNONI JÚNIOR, L.; STEVENS, D.; SILVA, W. T. L. da; VALE, J. M. F. do; PURINI, S. R. de M.; MAGNONI, M. da G. M.; SEBASTIÃO, E.; BRANCO JÚNIOR, G.; ADORNO FILHO, E. F.; FIGUEIREDO, W. dos S.; SEBASTIÃO, I. (Org.). JC na Escola Ciência, Tecnologia e Sociedade: mobilizar o conhecimento para alimentar o Brasil. 2. ed. São Paulo: Centro Paula Souza, 2017. p. 28-35. il. Publicado também em: Controle & Instrumentação, ano 21, n. 235, p. 56-59, 2018.

NAÇOES UNIDAS BRASIL, Transformando Nosso Mundo: A Agenda 2030 para o Desenvolvimento Sustentável. Disponível em: <https://nacoesunidas.org/pos2015/agenda2030/>. Acesso em: 07 jul. 2019.

PILLON, C. N. Dos pós de rocha aos remineralizadores: passado, presente e desafios. In: CONGRESSO BRASILEIRO DE ROCHAGEM, 3., 2017, Pelotas. Anais... Assis: Triunfal Gráfica e Editora, 2017. 455 p. Editado por Adilson Luis Banberg, Carlos Augusto Posser Silveira, Éder de Souza Martins, Magda Bergmann, Rosane Martinazzo e Suzi Huff Theodoro. p. 16-23.

SHEPHERD, M.; TURNER, A. J.; SMALL.B.; WHEELER, D. Priorities for science to overcome hurdles thwarting the full promise of the ‘digital agriculture’ revolution. Journal of the Science of Food and Agriculture. DOI:10.1002/jsfa.9346. 2018

TRENDOV , N. M.; VARAS , S.; ZENG, M. Digital Technologies In Agriculture And Rural Areas – Status Report. 1. ed. Food and Agriculture Organization of the United Nations: Rome, 2019. 152 p.

Dispositivo para Captura e Processamento de Imagens Térmicas para Detecção de Estresse Hídrico

Júlio Anderson de Oliveira Júnior[1], *Marcelo Gonçalves Narciso*[2]

[1] Estagiário em eletrônica, Embrapa Arroz e Feijão, Goiânia, Goiás, Brasil, julio12310@gmail.com

[2] Departamento de Biotecnologia, Embrapa Arroz e Feijão, Goiânia, Goiás, Brasil, marcelo.narciso@embrapa.br

RESUMO

Na tentativa de reduzir o tempo da coleta de dados de temperatura foliar, em fenotipagem de plantas em casa de vegetação, um dispositivo foi criado a partir de uma câmera térmica de baixo custo (Flir One Lepton). Foram coletadas imagens térmicas de folhas de arroz (dossel) para detecção de algum nível de estresse hídrico. O dispositivo captura as imagens térmicas juntamente com uma série de informações do ambiente onde a planta está inserida. Posteriormente, um programa do dispositivo processa estes dados fornecendo um resumo das medidas encontradas e compara os valores entre a amostra estudada e a planta de controle. Logo, partindo-se de uma câmera térmica relativamente barata, em comparação as demais no mercado, foi possível obter um dispositivo para detectar estresse hídrico da planta e fornecer outras análises a partir da temperatura do dossel da planta, como a detecção da presença de organismos patogênicos nas folhas.

PALAVRAS-CHAVE: Lepton. Processamento. Temperatura.

ABSTRACT

In an attempt to reduce the time of data collection of leaf temperature, in plant phenotyping in a greenhouse, a device was created from a low-cost thermal camera (Flir One Lepton). Thermal images of rice leaves (canopy) were collected to detect some level of water stress. The device captures the thermal images along with a series of information from the environment where the plant is inserted. Subsequently, a device program processes these data by providing a summary of the measures found and compares the values between the sample

studied and the control plant. Therefore, starting from a relatively inexpensive thermal camera, in comparison to the others in the market, it was possible to obtain a device to detect water stress of the plant and to provide other analyzes from the temperature of the plant canopy, such as the presence of organisms pathogens in the leaves.
KEY WORDS: Lepton. Processing. Temperature.

INTRODUÇÃO

Com o aumento da população mundial e as mudanças climáticas em várias regiões do mundo, a demanda mundial por alimentos tende a crescer nos próximos anos. Além disso, o avanço da tecnologia e a propagação de informações fizeram com que um número maior de indivíduos buscasse uma maior qualidade nos produtos consumidos. Logo, uma produção agrícola sustentável é vital para assegurar que todas essas exigências sejam atendidas. Para isso, várias pesquisas e experimentos foram desenvolvidos para alcançar uma melhor taxa de produção e qualidade das técnicas de cultivo. Dentre os estudos realizados, o melhoramento genético vegetal ganha bastante destaque por apresentar soluções tecnológicas quanto à resistência das plantas aos desafios ambientais que impactam negativamente o crescimento das plantas, como a falta e/ou excesso de água, nutrientes e luminosidade (SOUZA, 2015).

Para que o melhoramento genético supere e avance em seu programa, é necessário um conhecimento maior sobre as características fisiológicas que determinam o fenótipo das plantas. O fenótipo é a expressão do genótipo, em função de um ambiente em complexa transformação. Atualmente, sabe-se que a grande diferença entre genótipo e fenótipo é particularmente relacionada às interações entre plantas e ambiente. Portanto, a fenotipagem de plantas faz parte do estudo integral de um determinado sistema biológico e, atualmente, é considerado um fator limitante exatamente por não ter acompanhado o avanço na área de genotipagem de planta. A fenotipagem utiliza técnicas tradicionais e requer métodos destrutivos, os quais demandam tempo, recursos financeiros e mão-de-obra.

Sabe-se que a busca por conhecimentos mais refinados exige a formação de equipe multidisciplinar para que experimentos sejam conduzidos com o objetivo de detectar respostas das plantas mediante as diferentes condições de cultivo. Nesse sentido, o avanço na área de fenotipagem de plantas prevê a utilização de novos métodos (não invasivos/não destrutivos) e estes estão sendo considerados cruciais para o avanço dos programas de melhoramento,

fisiologia e a integração com a bioquímica, biologia molecular, dentre outras áreas. Esses novos métodos requerem uma abordagem inter, multi e transdisciplinar, envolvendo fisiologistas, melhoristas, engenheiros, físicos, químicos, bioinformatas, e ecofisiologistas para atuar no desenvolvimento e aplicação de sensores, biosensores, métodos espectroscópicos, dentre outros para quantificar as propriedades da planta e do ambiente no tempo e no espaço.

O uso de imagens digitais auxilia na obtenção de dados fenotípicos e podem fornecer informações distintas, conforme a banda do espectro eletromagnético utilizado. Para detectar informações como a temperatura foliar/dossel, que é uma das características que varia de forma expressiva e é altamente dependente do nível de estresse da planta, são utilizadas imagens térmicas que captam a radiação infravermelha distante (8 a 15 μm). Ao captar as temperaturas da superfície das folhas da planta, é possível perceber mudanças no estado hídrico da mesma em função da restrição de água imposta. Logo, a informação obtida está ligada, diretamente, a abertura ou fechamento dos estômatos, os quais são responsáveis pelas trocas gasosas entre a planta e o ambiente. Outros exemplos de aplicações de imagens térmicas são a sua utilidade na detecção de regiões das folhas infectadas por organismos patogênicos, detecção de danos mecânicos e no acompanhamento do comportamento dos estômatos (COSTA; GARCÍA-TEJERO; CHAVES, 2009).

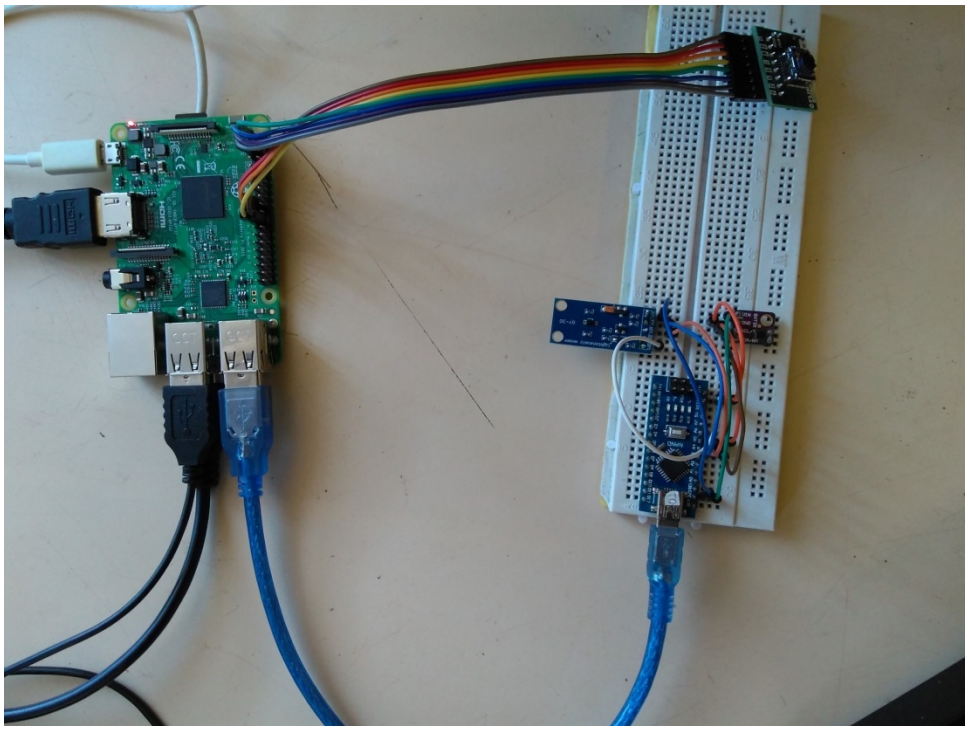
Por ser portátil, a captura de imagens térmicas torna-se mais dinâmica e prática, economizando tempo na coleta de dados. Além disso, como recurso adicional, o dispositivo foi projetado de modo que, após a fase de coleta de dados, um processamento das imagens fosse realizado. Como saída, o dispositivo fornece as informações da planta de forma sucinta e de fácil analise. Com estas informações, o objetivo deste trabalho foi a criação de um dispositivo portátil, de baixo custo, que pudesse auxiliar nos experimentos de fenotipagem baseados em imagens térmicas para obtenção de dados de temperatura foliar para monitoramento do *status* hídrico da planta, e assim discriminar genótipos mais tolerantes à deficiência hídrica nos programas de melhoramento vegetal.

MATERIAL E MÉTODOS

Para controlar todos os componentes utilizados, foi usado um minicomputador desenvolvido pela fundação Raspberry Pi, o minicomputador Raspberry Pi 3 (ver www.raspberrypi.org).

Além disso, todos os softwares produzidos foram feitos neste minicomputador. Na captura de imagens térmicas, foi utilizada uma câmera contendo um sensor infravermelho de longo comprimento de onda (LWIR) da empresa FLIR (ver https://www.flir.com/flirone), o sensor FLIR Lepton 3. Este sensor é destinado para dispositivos móveis e foi necessário o uso de um módulo (adaptador) para conectá-lo ao minicomputador. Para coletar dados do ambiente em que as plantas estão submetidas, foi conectado ao minicomputador um Arduino Nano (ver https://store.arduino.cc/usa/arduino-nano), usado para alimentar e controlar um sensor de temperatura e umidade (SHT3x) e um sensor de luminosidade (GY-30). Com este aparato (figura 1), além da imagem térmica da amostra avaliada, foi possível obter informações como a temperatura ambiente, a umidade relativa do ar e a luminosidade. Para a coleta de dados, o dispositivo foi posicionado acima das amostras de modo a captar uma área maior da superfície das folhas.

Figura 1 – Dispositivo contendo câmera térmica e sensores



Fonte: (Própria, 2019)

Implementação

Como o sensor (câmera Lepton) foi colocado em módulo para poder ser conectado às portas GPIO da Raspberry Pi 3, foi necessário o uso de um programa escrito em linguagem C para que o minicomputador pudesse controlar a captura de imagens. Este programa foi

desenvolvido na própria Embrapa Arroz e Feijão, e foi baseado em repositórios online para o sensor Lepton. Ao ser acionado, o programa cria uma janela onde é emitido um vídeo do que o sensor capta do ambiente. Nesta mesma janela, existem dois botões referentes à captura de uma imagem e a calibragem do sensor. Ao apertar o botão de captura de imagem, o programa emite uma foto e um arquivo texto com as temperaturas ponto a ponto da imagem. Por conta da resolução do sensor, a imagem de saída possui 160x120 pixels e consequentemente, o arquivo texto apresenta temperaturas distribuídas em 120 linhas e 160 colunas sendo cada elemento referente aos pixels da imagem.

Para tornar este processo mais dinâmico, foi desenvolvido um programa na linguagem Python para acionar o programa do sensor térmico e pressionar o botão da janela em intervalos de tempo periódico. Desse modo, no teste realizado posteriormente a cada cinco minutos, uma nova foto era obtida juntamente com seu arquivo texto. Quanto aos sensores conectados no Arduino Nano, um pequeno programa foi desenvolvido para que o Arduino captasse as três medidas do ambiente (temperatura, umidade relativa do ar e luminosidade) continuamente e logo após enviaria via serial para o minicomputador Raspberry Pi 3. Posteriormente, foram adicionadas ao programa em Python algumas linhas de código onde a informação da porta serial era lida e armazenada juntamente com a data e hora da captura da imagem, em um segundo arquivo texto.

Portanto, o dispositivo, ao acionar o programa de captura, obtém uma imagem térmica da região avaliada com 160x120 pixels e dois arquivos textos, um contendo as temperaturas ponto a ponto da imagem e outro com a temperatura ambiente, umidade relativa do ar, luminosidade do ambiente e data e hora da imagem. Para manter uma organização durante a análise dos dados obtidos, todos os arquivos foram salvos com nomes referentes às amostras, e esse sistema foi projetado de modo que torne mais prático o acesso posterior a estas informações. Como passo final, foi criado um programa em Python com uma interface para acesso das informações. Neste programa, é possível acessar o banco de fotos e selecionar a imagem desejada. Ao selecionar a foto, os arquivos textos referentes também são carregados, bem como a imagem e arquivos da planta de controle para comparar as informações.

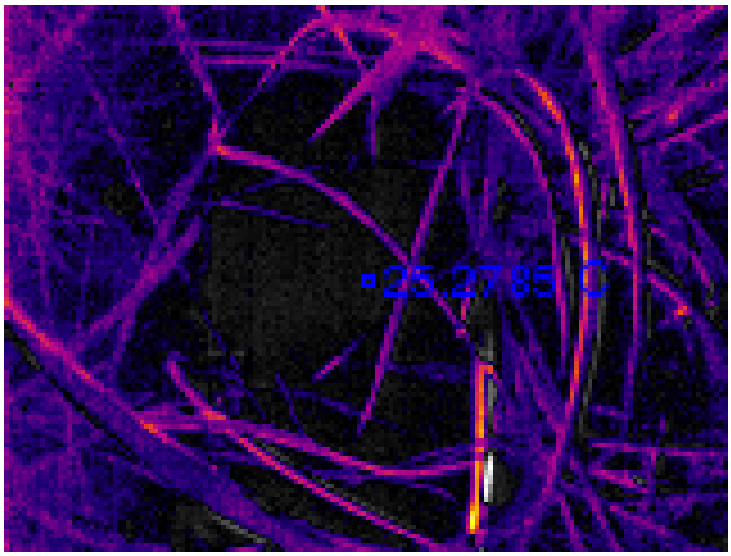
Quando estes dados são carregados, uma janela é criada onde as imagens da planta de controle e a amostra são mostradas, e para facilitar a visualização, as fotos são redimensionadas para um tamanho maior. Posteriormente, a interface fornece a opção de processar as imagens inteiras ou apenas uma região selecionada pelo mouse. Na parte de processamento, as imagens passam por um filtro para diferenciar os valores dos pixels referentes à planta dos pixels do ambiente. Dessa forma os pixels da imagem passam a

apresentar apenas dois valores distintos: zero para pontos não referentes as folhas e um para pontos das folhas. Em seguida, a imagem é transformada em uma matriz, assim como o arquivo texto com as temperaturas. Multiplicando-se a matriz da imagem com a matriz das temperaturas, as temperaturas referentes aos pontos da planta são obtidas. Com estes dados, o programa realiza os cálculos para obter dados como a temperatura média, temperatura máxima e mínima.

RESULTADOS E DISCUSSÃO

Para testar o programa de captura de imagens, foi desenvolvido um experimento onde a câmera foi posicionada entre algumas colunas de arroz em uma casa de vegetação. Como o programa capturava uma imagem a cada cinco minutos, as mesmas colunas foram observadas durante variados horários do dia. Devido à natureza do teste ser apenas para verificação do funcionamento do dispositivo, foram feitas comparações apenas com relação ao posicionamento da câmera (JONES, 2009). Como resultado deste procedimento, foram obtidas algumas fotos e seus respectivos arquivos textos. Ao capturar imagens das colunas de arroz (Figura 2), é possível perceber visualmente a diferença entre os pixels referentes às folhas das plantas com o resto do ambiente.

Figura 2 – Imagem térmica de colunas de arroz



Fonte: (Própria, 2019)

O sensor, capta a radiação infravermelha refletida e emitida pela planta, e de acordo com uma lei da Física (Stefan-Boltzmann) que afirma que a quantidade de radiação emitida depende diretamente da temperatura do objeto, e produz uma imagem que diferencia a

coloração dos pixels de acordo com a radiação recebida. Logo, o padrão de cores da imagem se deve por conta da radiação infravermelha e consequentemente pela temperatura. Ao detectar a diferença de temperatura entre as folhas das plantas e o resto do ambiente, cria um padrão de cor que depende da quantidade de radiação infravermelha que cada objeto emite. Nesse caso, a temperatura cresce de acordo com a cor passando de cores mais escuras para mais claras. Juntamente com a imagem, um arquivo texto (Figura 3) é criado tomando como base os dados captados pelo sensor. A partir destes dados, a temperatura correspondente é calculada e armazenada para cada ponto da imagem.

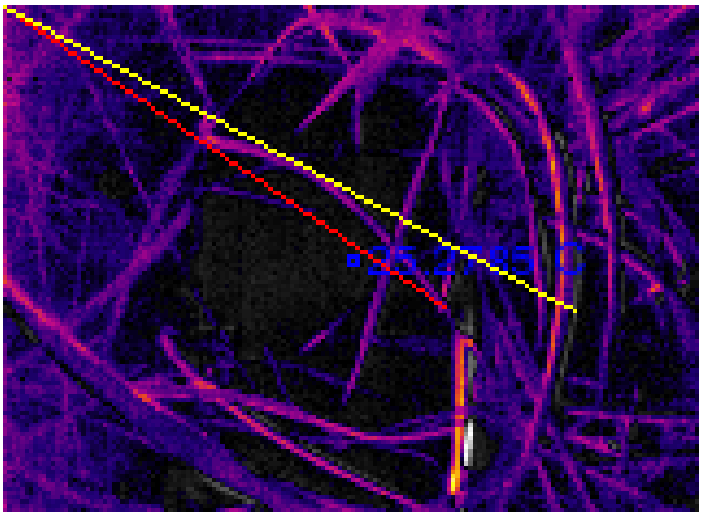
Figura 3 – Corte do arquivo texto com as temperaturas

```
25,907810 25,821003 25,669102 25,821003 25,821003 25,690800 25,625706 25,582310
25,669102 25,625706 25,604008 25,712498 25,582310 25,604008 25,647404 25,690800
25,690800 25,604008 25,560596 25,647404 25,473804 25,560596 25,560596 25,560596
25,625706 25,647404 25,604008 25,517200 25,690800 25,755909 25,799305 25,951206
25,495502 25,582310 25,538898 25,452106 25,473804 25,495502 25,669102 25,604008
25,560596 25,517200 25,495502 25,517200 25,560596 25,538898 25,538898 25,604008
25,517200 25,538898 25,517200 25,495502 25,560596 25,495502 25,517200 25,517200
26,146503 25,712498 25,452106 25,517200 25,517200 25,560596 25,495502 25,517200
25,647404 25,712498 25,842701 25,712498 25,842701 25,842701 25,821003 25,842701
25,864399 25,821003 25,755909 25,842701 25,842701 25,886097 25,907810 25,799305
25,755909 25,734196 25,625706 25,604008 25,473804 25,452106 25,582310 25,690800
25,712498 25,669102 25,669102 25,625706 25,538898 25,538898 25,560596 25,712498
25,647404 25,625706 25,582310 25,582310 25,560596 25,712498 25,712498 25,604008
25,821003 25,777607 25,799305 25,604008 25,799305 25,734196 25,755909 25,647404
25,495502 25,538898 25,669102 25,604008 25,625706 25,712498 25,690800 25,669102
25,560596 25,582310 25,647404 25,886097 25,734196 25,821003 26,059696 25,864399
25,907810 25,777607 25,755909 25,777607 25,712498 25,777607 25,669102 25,582310
25,755909 25,712498 25,647404 25,647404 25,582310 25,582310 25,690800 25,712498
25,582310 25,560596 25,495502 25,517200 25,538898 25,560596 25,560596 25,517200
```

Fonte: (Própria, 2019)

Apesar da quantidade de casas após a vírgula, a precisão do sensor e da natureza dos experimentos, apenas as três primeiras casas foram consideradas. Essa quantidade deve-se por conta dos cálculos efetuados durante o programa de captura de imagens que armazena os valores de cada pixel e depois os transforma em temperaturas na escala Celsius. Posteriormente, a foto é processada pelo programa em linguagem Python para fornecer as informações mais gerais sobre as folhas avaliadas. Como resultado, foi entregue uma nova imagem com os marcadores de temperatura máxima e mínima e dois arquivos textos: um deles com apenas os valores de temperatura referentes aos pontos das folhas presentes na imagem, e o outro com os dados estatísticos calculados e os dados coletados pelos sensores acoplados ao Arduino. Na imagem processada (Figura 4), o marcador em vermelho, inicia a partir do canto superior esquerdo e vai até o ponto de temperatura máxima encontrado nas folhas. Já o marcador em amarelo, parte do canto superior esquerdo e vai ao ponto de temperatura mínima.

Figura 4 – Imagem final com marcadores



Fonte: (Própria, 2019)

Para esta imagem, a temperatura máxima detectada nos pontos referentes às folhas foi 26,04ºC, e a temperatura mínima foi de 24,95ºC. Logo, a amplitude de temperatura ao longo dos pontos avaliados foi de 1,09ºC. Levando em conta a precisão da câmera térmica usada, aproximadamente 0,1ºC, a variação de temperatura pode ser aproximada para 1ºC. Quanto à temperatura média dos pontos das folhas, o valor encontrado foi de 25,47ºC, o que é uma medida aceitável por estar próximo do valor médio da amplitude de temperatura. Para questão de comparação, a temperatura do ar no momento da captura da imagem era de aproximadamente 23,9ºC.

CONCLUSÕES

Observando os resultados obtidos e as especificações dos componentes, a câmera térmica usada apresenta diferentes campos de visão dependendo da posição em que é colocada (JONES, 2009). Para a aplicação usada, a melhor posição para obter uma maior área das folhas da copa é de um ponto superior à planta, pois como as folhas do arroz possuem uma espessura pequena, uma imagem lateral não traria uma quantidade significativa de pontos da superfície das folhas. Dessa forma, além de aumentar a quantidade de pontos da imagem referentes às folhas, pode-se alterar a altura do dispositivo em relação às plantas, fazendo com que a foto seja tirada de uma região com mais de uma amostra.

Este dispositivo foi construído a partir de uma câmera térmica de baixo custo e de algumas peças eletrônicas de fácil acesso, logo o uso desse tipo de equipamento torna-se bastante viável para os mais variados grupos de pesquisa nacionais. Por basear-se no uso de

imagens térmicas o dispositivo pode fornecer informações em vários assuntos referentes a agricultura como alguns fenótipos de plantas, a presença de pragas ou doenças e a presença de falhas de irrigação em plantações.

Foi mostrado o seu uso em casa de vegetação para obter a temperatura de cada folha e assim ter a temperatura média do dossel da planta e poder avaliar cada uma se tem algum estresse hídrico ou se está sendo injuriada por alguma praga. Ajustando alguns parâmetros e posicionando o dispositivo em um drone ou em algum suporte móvel, ele pode também ser aplicado para obter informações no campo. Uma forma de tornar o dispositivo ainda mais portátil e de fácil acesso, é transformar os programas desenvolvidos em um aplicativo para smartphones e usar como câmera térmica o próprio equipamento FLIR ONE (ver https://www.flir.com/flirone). Dessa forma, a resolução das imagens obtidas será muito maior e o uso dessa ferramenta se tornará mais acessível. Com essa ferramenta presente em um telefone celular, todo o processo de captura de imagens e processamento de dados se tornaria mais rápido e intuitivo, pois com o aplicativo essas duas etapas seriam feitas de forma simultânea. Logo, o produto final pode ser modelado pelo próprio usuário mediante a sua necessidade.

REFERÊNCIAS

SOUSA, C. A. F. et al. Nova abordagem para a fenotipagem de plantas: conceitos, ferramentas e perspectivas. **Revista Brasileira de Geografia Física**, v. 8, p. 660-672, 2015.

JONES, Hamlyn G. et al. Thermal infrared imaging of crop canopies for the remote diagnosis and quantification of plant responses to water stress in the field. **Functional Plant Biology**, v. 36, n. 11, p. 978-989, 2009.

MUNNS, Rana et al. New phenotyping methods for screening wheat and barley for beneficial responses to water deficit. **Journal of experimental botany**, v. 61, n. 13, p. 3499-3507, 2010.

FLIR Proprietary-Confidential e FLIR LEPTON. Long Wave Infrared (LWIR) Datasheet.

Arduino. Arduino Nano. Disponível em: https://store.arduino.cc/usa/arduino-nano. Acessado em: 19 de mar. de 2019.

FLIR. FLIR ONE | FLIR Systems. Disponível em: https://www.flir.com/flirone. Acessado

em: 06 de fev. de 2019.

Costa, Miguel & García-Tejero, I.F. & Chaves, M.. (2013). O uso da termografia na agricultura moderna. Revista da Associação Portuguesa de Horticultura (APH). 113. 30-34.

Identificação de padrões de produção e intensificação tecnológica de municípios produtores de bovinos de corte no bioma Pampa

Vinicius do Nascimento Lampert[1], Maria do Carmo Ramos Fasiaben[2], Urbano Gomes Pinto de Abreu[3], Elisa Kohler Osmari[1], Stanley Robson de Medeiros Oliveira[2], Maxwell Merçon Tezolin Barros Almeida[4,5], Octávio Costa de Oliveira[4,5], Helano Póvoas de Lima[2]

[1] Embrapa Pecuária Sul - CPPSUL, Bagé, RS, Brasil
[2] Embrapa Informática Agropecuária - CNPTIA, Campinas, SP, Brasil
[3] Embrapa Pantanal - CPAP, Corumbá, MS, Brasil
[4] Instituto Brasileiro de Geografia e Estatística - IBGE, Rio de Janeiro, RJ, Brasil

{vinicius.lampert@embrapa.br, maria.fasiaben@embrapa.br, urbano.abreu@embrapa.br, elisa.osmari@embrapa.br, stanley.oliveira@embrapa.br, maxwell.almeida@ibge.gov.br, octavio.oliveira@ibge.gov.br, helano.lima@embrapa.br}

RESUMO

O objetivo deste trabalho foi diferenciar e caracterizar a produção de bovinos de corte praticada nos municípios do bioma Pampa, segundo seus padrões de produção e intensificação tecnológica. Para tanto, empregaram-se análises de cluster e discriminante canônica. Foram selecionadas 24 variáveis oriundas de tabulações especiais de dados do Censo Agropecuário 2006, filtrados para estabelecimentos com bovinos de corte e agregados em 116 municípios do Pampa. As variáveis reuniram informações socioeconômicas e tecnológicas dos estabelecimentos, tais como uso da terra, tamanho do rebanho, área de pastagens (naturais e plantadas), taxa de lotação, manejo das pastagens, uso de insumos e participação dos bovinos de corte no valor da produção. As análises identificaram três grupos de municípios produtores de bovinos de corte no Pampa: terminação com silvicultura (TERMSILVI), tecnificado com agricultura (TECAGRI) e tradicional com ciclo completo (TRADICI). Conhecer essa diferenciação é importante para propor linhas de pesquisa, desenvolvimento, transferência de tecnologia e apoiar a formulação de políticas públicas apropriadas ao setor.

PALAVRAS-CHAVE: Tipologia, Clusterização, Sistemas de produção.

[5] O IBGE está isento de qualquer responsabilidade pelas opiniões, informações, dados e conceitos emitidos neste artigo, que são de exclusiva responsabilidade do autor.

ABSTRACT

The objective of this work was to differentiate and characterize beef cattle production practiced in the municipalities of the Brazilian Pampa, according to their production patterns and technological intensification. Cluster analysis and canonical discriminant were used and twenty-four variables were selected from special tabulations of data from the Census of Agriculture 2006, filtered for agricultural holdings with beef cattle, in 116 municipalities of Pampa. The variables included socioeconomic and technological information of the holdings, such as land use, herd size, pasture area (natural and planted), stocking rate, pasture management, input use and beef cattle participation in production value. The analysis identified three groups of municipalities producing beef cattle in Pampa: finishing beef cattle with forestry (TERMSILVI), technified with agriculture (TECAGRI) and traditional with full cycle (TRADICI). Knowing this differentiation is important to propose lines of research, development, transfer of technology and support the formulation of appropriate public policies for the sector.

KEYWORDS: Typologies, Clustering, Production systems

INTRODUÇÃO

Os sistemas de produção de gado de corte baseados em pastagens naturais têm sido a principal atividade econômica no bioma Pampa. A diversidade de espécies nativas caracteriza uma boa oferta de forragem na estação quente e reduzida oferta na estação fria. Segundo Boldrini (2008), em levantamentos botânicos realizados nessas pastagens, registraram-se 523 espécies de gramíneas e 250 espécies de leguminosas, a maioria delas perene e de crescimento estival. Entretanto, mudanças vêm sendo observadas em função da pressão por lucratividade, com o avanço da agricultura nas fronteiras agrícolas (OLIVEIRA et al., 2017).

Distinguir o perfil tecnológico de criação de bovinos de corte dominante de cada município é importante para aprimorar políticas públicas a fim de evitar riscos iminentes de degradação da biodiversidade. O objetivo da pesquisa foi realizar uma tipificação dos municípios com pecuária de corte do bioma Pampa a partir da identificação de padrões de produção e intensificação tecnológica com dados do Censo Agropecuário 2006, agregados por município a partir de estabelecimentos com mais de 50 bovinos de corte.

MATERIAL E MÉTODOS

A região analisada abrange o bioma Pampa brasileiro, subtropical, restrito a 63% do Rio Grande do Sul e a 2,07% do território nacional (Ministério do Meio Ambiente, 2019). Para

caracterizar a produção de bovinos de corte do Pampa - e diferenciar os sistemas de produção predominantes no bioma -, no presente trabalho foram empregadas informações do Censo Agropecuário 2006 (IBGE, 2012), o último disponível com informações detalhadas em nível de estabelecimentos agropecuários. Vale destacar que, no âmbito da bovinocultura, o Censo Agropecuário 2006 teve como particularidade a aplicação de questionário detalhado apenas em estabelecimentos que tinham mais de 50 cabeças de bovinos na data de referência. Em estabelecimentos com 1 a 50 cabeças de bovinos não foi realizado esse nível de detalhamento da criação de bovinos (FASIABEN et al., 2013). Para atender ao presente trabalho, foi utilizada uma tabulação especial do Censo Agropecuário 2006 (IBGE, 2012) com dados agregados por município de estabelecimentos agropecuários com mais de 50 cabeças de bovinos de corte. Por questão de sigilo estatístico, foram suprimidos os dados dos municípios com menos de três estabelecimentos.

O processo de extração dos padrões de desenvolvimento tecnológico da bovinocultura de corte do bioma Pampa foi realizado utilizando a técnica de análise de clusters, por meio do método de particionamento. A seleção de variáveis para a tipificação da produção de bovinos de corte desses municípios foi definida a partir de consultas a especialistas na área, conhecedores da região de estudo (Tabela 1). Para determinar o número de clusters, foram testados vários valores de k (k = 2, ..., 16). Em seguida, para minimizar a subjetividade na escolha do valor de k, foi utilizado o método do cotovelo (Han et al, 2011). O algoritmo K-Means e o método do cotovelo foram implementados no ambiente RStudio.

Tabela 1: Variáveis selecionadas na identificação de padrões de desenvolvimento tecnológico dos sistemas de produção de pecuária de corte no bioma Pampa em escala municipal, 2006

Sigla	Descrição da variável
D_CRIA	Estabelecimentos que declararam fazer cria (%)
D_RECRIA	Estabelecimentos que declararam fazer recria (%)
D_ENGORDA	Estabelecimentos que declararam fazer engorda (%)
D_CRIA_RECRIA	Estabelecimentos que declararam fazer cria e recria (%)
D_CRIA_ENGORDA	Estabelecimentos que declararam fazer cria e engorda (%)
D_RECRIA_ENGORDA	Estabelecimentos que declararam fazer recria e engorda (%)
D_CRIA_RECRIA_ENG	Estabelecimentos que declararam fazer cria, recria, engorda (ciclo completo (%)
PVACAS2ANOS	Total de vacas de 2 anos e mais/ total do rebanho bovino
TX_LOT_UA	Taxa de lotação - total de UA / área total de pastagens
RACAO	Porcentagem de estabelecimentos que fizeram suplementação com sal e ração ou somente ração (inclui ração, grãos, subprodutos agroindustriais)
ADUBPASTAGEM	Estabelecimentos que declaram adubar pastagem (%)
RECEBEORIENT	Estabelecimentos que declaram receber orientação técnica (%)
PCONFINADOS	Animais confinados / total de animais (nº de bovinos em 31.12.2006)
PVACASINSEM	Vacas inseminadas/ total de vacas
P_PAST	Área com pastagens / área explorada* com a produção

P_SILV	Área com silvicultura / área explorada* com a produção
PVTP_BOV	Valor da produção de bovinos / total do valor da produção
UA_BOV[6]	Número total de unidades animais (UA)
PPASTNATURAL	Área de pastagem natural/ área de pastagem total
PPASTDEGRADADA	Área de pastagem plantada degradada/ área de pastagem plantada total
PMACHOS2MAIS	Total de machos de 2 anos ou mais/total de bovinos
PLAVPERM	Área de lavouras permanentes / área explorada com a produção
PLAVTEMP	Área de lavouras temporárias / área explorada com a produção
PFORRAGEIRA	Área com forrageiras para corte / área explorada com a produção

*Área total explorada com a produção corresponde ao somatório de área de lavouras permanentes; área de lavouras temporárias; área com forrageiras para corte; área com cultivo de flores, casa de vegetação e viveiro de mudas; área de pastagem natural; área de pastagem plantada degradada; área de pastagem plantada não-degradada; área de silvicultura; e área de sistemas agroflorestais (em hectares).

A análise de agrupamentos foi utilizada como suporte à MANOVA (do inglês *multivariate analysis of variance*) sobre esses 24 atributos da pecuária de corte de 2006 (IBGE, 2012). Onde a análise discriminante canônica (ADC) foi a técnica utilizada (PROVOST e FAWCETT, 2016), de modo a facilitar a distinção de padrões tecnológicos.

Além das variáveis que serviram para a classificação dos municípios nos clusters, outras variáveis foram obtidas a partir do Censo Agropecuário 2006. Elas foram extraídas diretamente do censo ou foram combinadas entre si gerando informações adicionais. Essas variáveis serviram para completar a caracterização do bioma com o objetivo de enriquecer a visão do sistema produtivo como um todo sem prejudicar o poder discriminatório pelo excesso de variáveis. Cabe destacar, que a aderência dos resultados à realidade foi analisada por especialistas do bioma Pampa.

RESULTADOS E DISCUSSÃO

O presente estudo contou com informações de 116 municípios, que agregaram informações de 19.240 estabelecimentos que declararam possuir mais de 50 cabeças de bovinos de corte representando 14,4% do total de 134.068 estabelecimentos agropecuários. Esta quantidade representa 8.256.387 hectares (66,6%) do total de 12.396.061 hectares ocupados pelo total de estabelecimentos agropecuários no bioma Pampa.

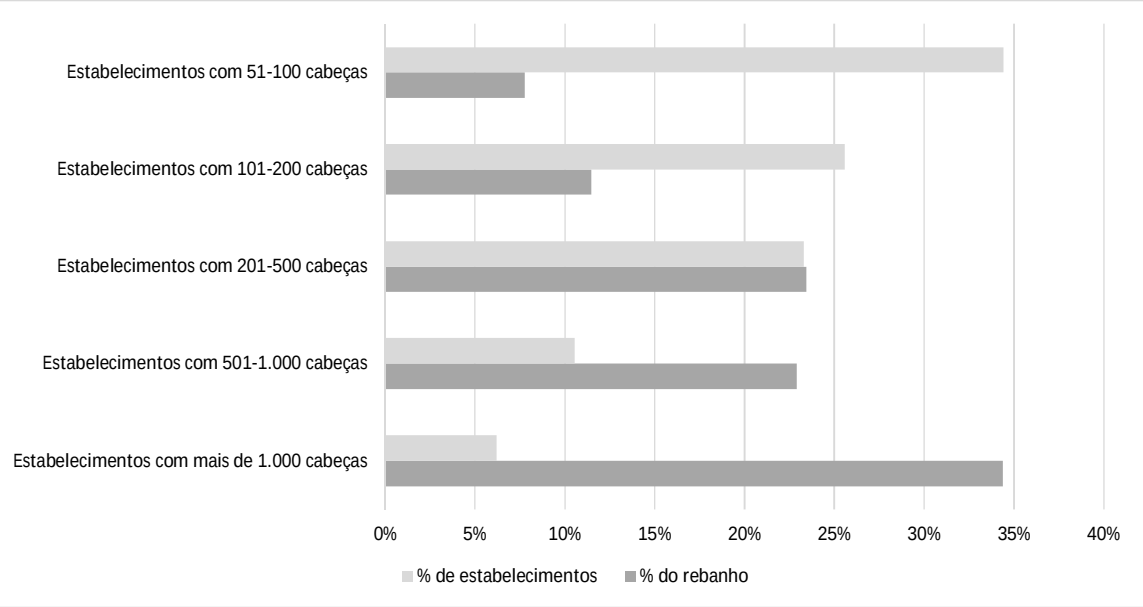
Destaca-se que os estabelecimentos com 51-100 cabeças representam 34,42% dessa população e detêm 7,77% do rebanho enquanto os estabelecimentos com mais de 1.000 cabeças representam 6,19% do total de propriedades, mas detêm 34,38% do rebanho de corte.

[6] A informação sobre "Unidades Animais" não foi colhida no Censo Agropecuário 2006. Ela se baseia em cálculo, feito pelo IBGE, conforme metodologia estabelecida na INSTRUÇÃO NORMATIVA Nº 11, de 4 de abril de 2003 (BRASIL, 2003).

Este fato evidencia a maior concentração de estabelecimentos nas faixas com rebanhos menores (Figura 1).

A análise de agrupamentos não supervisionada permitiu a separação dos municípios produtores de bovinos de corte em três agrupamentos (clusters) distintos, que puderam ser diferenciados por seus perfis de produção e tecnológicos: terminação com silvicultura (TERMSILVI), tecnificado com agricultura (TECAGRI) e tradicional com ciclo completo (TRADICI).

Figura -1 Concentração e distribuição dos estabelecimentos por tamanho do rebanho no bioma Pampa



Fonte: elaborada pelos autores a partir de dados do Censo Agropecuário 2006 (IBGE, 2012).

Verificou-se que no cluster TERMSILVI (25 municípios) encontram-se os municípios com predominância de sistemas de produção dedicados às fases de cria, recria e engorda individualizadas, comparativamente aos demais clusters (Tabela 2). Também em relação aos demais, esse cluster apresenta menor participação do sistema ciclo completo, de percentual de fêmeas no rebanho, além de maior participação de machos, em sistema de engorda ou terminação, com presença marcante de silvicultura e maiores áreas de matas e florestas naturais. É importante frisar que, embora o sistema de ciclo completo esteja presente em 29% dos estabelecimentos do cluster TERMSILVI (sistema de criação de maior ocorrência nesse cluster), ele é menos representativo comparativamente aos demais clusters. Exemplos de municípios representantes: Butiá, Charqueadas e Triunfo.

Tabela 2 – Informações gerais sobre perfis tecnológicos dos municípios e distribuição dos produtores de bovinos de corte do bioma Pampa por tipo de sistema de produção (dados agregados por município a partir de estabelecimentos com mais de 50 cabeças).

	CLUSTER TERMSILVI	CLUSTER TECAGRI	CLUSTER TRADICI	BIOMA PAMPA
Número de municipios	25	36	55	116
Área total dos estabelecimentos (ha)	168.397	1.125.956	6.962.034	8.256.387
Estabelecimentos com bovinos de corte	602	2.424	16.214	19.240
Animais em estabelecimentos (>50 bovinos de corte)	122.606	720.878	5.451.827	6.295.311
Número total de unidades animais (UA)	91.095	530.462	4.080.294	4.701.851
Valor da produção de bovinos / total do valor da produção *	17,8%	22,6%	43,8%	37,8%
Estabelecimentos que adubam as pastagens*	37,0%	37,8%	30,2%	31,4%
Estabelecimentos que recebem orientação técnica*	50,0%	61,8%	49,5%	51,1%
Bovinos confinados (%)*	3,7%	3,8%	1,2%	1,6%
Vacas inseminadas (%)*	12,3%	15,0%	12,7%	12,9%
Estabelecimentos que usam ração*	30,9%	29,1%	23,3%	24,3%
D_Cria*,**	16,3%	12,5%	14,5%	14,3%
D_Recria*,**	10,5%	3,8%	4,5%	4,6%
D_Engorda*,**	25,9%	22,6%	12,3%	14,0%
D_Cria_Recria*,**	6,0%	12,5%	20,9%	19,4%
D_Cria_Engorda	4,2%	6,0%	4,3%	4,5%
D_Recria_Engorda	8,1%	5,1%	3,9%	4,2%
D_Cria_Recria_Engorda*,**	29,1%	37,5%	39,5%	39,0%

* variáveis usadas no modelo; ** fases de produção significativas estatisticamente para distinguir os agrupamentos municipais do bioma

O cluster TECAGRI (36 municípios) foi assim denominado pela maior participação em lavouras, percentual de estabelecimentos com orientação técnica, percentual de animais confinados, maior taxa de lotação por hectare (Tabelas 3 e 4). Essas características revelam que este cluster apresenta uma melhor eficiência no uso do solo e manejo de pastagens. O cluster TECAGRI foi representado principalmente por municípios da divisa com a fronteira noroeste do Estado do bioma Mata Atlântica, próximos de locais tradicionais de lavoura temporária (soja, milho, trigo) como, por exemplo, Júlio de Castilhos, Tupanciretã e São Luiz Gonzaga, ou localizados nas denominadas terras baixas ou alagadas, tradicionais produtoras de arroz irrigado, que abrangem municípios como Camaquã e Palmares do Sul.

O cluster TRADICI (55 municípios) agrega 47% do total de municípios e 86% do total de animais, e, consequentemente, o maior uso da terra no bioma Pampa. Caracteriza-se por seu nível tecnológico mais extensivo a pasto, com menor uso de adubação, menor taxa de lotação por hectare e maior percentual de vacas de cria. Alguns exemplos de municípios desse sistema são Alegrete, Caçapava do Sul, Cacequi, Lavras do Sul, Pinheiro Machado, Sant'ana do Livramento e Quaraí.

Tabela 3 - Características do rebanho em municípios produtores de bovinos de corte do bioma Pampa (dados agregados a partir de estabelecimentos com mais de 50 cabeças).

	CLUSTER TERMSILVI	CLUSTER TECAGRI	CLUSTER TRADICI	BIOMA PAMPA
Tamanho médio dos estabelecimentos (ha)	280	465	429	429
Tamanho médio das pastagens (ha/estab.)	175	244	327	312
Tamanho médio do rebanho (cab)	204	297	336	327
Taxa de Lotação (UA/hectare)*	0,86	0,90	0,77	0,78
Peso médio dos animais (kg)	334	331	337	336
% Mortalidade 0-12 meses	4,0%	3,9%	3,0%	3,1%
% Mortalidade >=12 meses	3,7%	1,8%	1,2%	1,4%
Fêmeas 0-12 meses (%)	9,6%	10,3%	10,5%	10,4%
Machos 0-12 meses (%)	9,4%	10,4%	10,2%	10,2%
Fêmeas 13-24 meses (%)	14,6%	13,1%	12,5%	12,6%
Machos 13-24 meses (%)	13,9%	13,1%	11,0%	11,3%
Fêmeas >=24 meses (%)*	25,9%	35,3%	38,7%	38,1%
Machos >=24 meses (%)* - corte	21,8%	15,0%	14,8%	15,0%
Machos >=24 meses (%) - trabalho	0,9%	0,7%	0,3%	15,3%
Touros >= 24 meses (%)	3,8%	2,1%	1,9%	2,0%

* Variáveis usadas no modelo;

Tabela 4 - Uso do solo em municípios produtores de bovinos de corte do bioma Pampa (dados expressos em hectares, agregados a partir de estabelecimentos com mais de 50 cabeças).

		CLUSTER TERMSILVI		CLUSTER TECAGRI		CLUSTER TRADICI		BIOMA PAMPA	
Lavouras Permanentes*	(1)	836	0,6%	2.996	0,3%	7.904	0,1%	11.736	0,2%
Lavouras Temporárias*	(2)	14.266	10,0%	344.176	34,3%	733.353	11,7%	1.091.796	14,7%
Forrageiras para corte*	(3)	5.781	4,0%	33.983	3,4%	101.467	1,6%	141.231	1,9%
Pastagem Natural**	(4)	86.379	60,3%	499.444	49,8%	4.797.359	76,6%	5.383.181	72,7%
Pastagem Plantada	(5)	19.164	13,4%	91.277	9,1%	504.272	8,1%	614.713	8,3%
Silvicultura*	(6)	14.592	10,2%	16.453	1,6%	50.218	0,8%	81.263	1,1%
Sistemas Agroflorestais	(7)	2.297	1,6%	14.194	1,4%	66.608	1,1%	83.099	1,1%
Total explorado na produção		143.315	100%	1.002.631	100%	6.261.782	100%	7.407.729	100%
Toral de Lavouras[1]		20.883	12,4%	381.264	33,9%	843.326	12,1%	1.245.473	15,1%
Total de Pastagens[2]		105.543	62,7%	590.721	52,5%	5.301.631	76,2%	5.997.894	72,6%
Matas e Florestas Plantadas[3]		16.889	10,0%	30.647	2,7%	116.826	1,7%	164.361	2,0%
Matas e/ou Florestas Naturais	(8)	16.112	9,6%	61.576	5,5%	424.717	6,1%	502.406	6,1%
Outros	(9)	8.970	5,3%	61.748	5,5%	275.535	4,0%	346.252	4,2%
Total dos Estabelecimentos[4]		168.397	100%	1.125.956	100%	6.962.034	100%	8.256.387	100%

* Variáveis usadas no modelo; ** TERMSILVI, TECAGRI e TRADICI apresentam 81,8%, 84,5% e 90,5% de área de pastagem natural em relação ao total de pastagens, respectivamente; [1](1)+(2)+(3); [2](4)+(5); [3](6)+(7); [4](1)+(2)+(3)+(4)+(5)+(6)+(7)+(8)+(9)

As funções discriminantes canônicas foram estimadas por meio de combinações lineares das variáveis originais, e possuem a propriedade de maior poder discriminante, dentro do conjunto de variáveis agrupadas *a priori*. Ou seja, são as combinações lineares com maior

poder de separação dos grupos, ou de maneira equivalente as que fornecem o menor valor da probabilidade de significância (MINGOTI, 2007). Os coeficientes das duas primeiras discriminantes canônicas (Tabela 5) expressaram os valores de ponderação para cada variável dos municípios, e podem ser interpretados como o coeficiente de um modelo de regressão múltipla. Esses valores servem para identificar as variáveis que mais contribuem para discriminação entre os agrupamentos (CAMPOS et al., 2016).

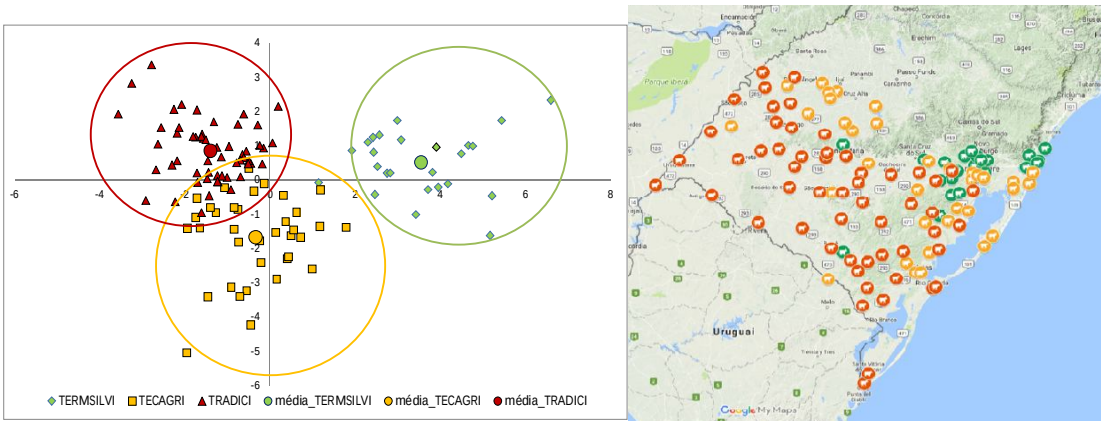
Tabela 5 - Procedimento de Análise Discriminante Canônica (ADC), ordenadas por ordem de importância

Variável	Canônica 1	Canônica 2
PVACAS2ANOS	-0,75	0,14
D_CRIA_RECRIA	-0,66	0,19
P_SILV	0,57	0,08
D_ENGORDA	0,50	-0,27
D_RECRIA	0,48	0,29
UA_BOV	-0,48	0,46
PMACHOS2MAIS	0,44	-0,03
PLAVPERM	0,36	-0,10
PPASTNATURAL	-0,35	0,37
PFORRAGEIRA	0,35	-0,14
D_CRIA_RECRIA_ENGORDA	-0,34	-0,01
PLAVTEMP	-0,24	-0,77
RACAO	0,22	-0,28
PVTP_BOV	0,22	0,62
PCONFINADOS	0,20	-0,36
ADUBPASTAGEM	0,13	-0,33
TX_LOT_UA	0,13	-0,29
RECEBEORIENT	-0,12	-0,46
D_CRIA	0,08	0,27

As distâncias espaciais de Mahalanobis, entre as médias canônicas dos três agrupamentos (*a priori* determinados) foram significativas ($p<0,01$), o que nos permite inferir que os grupos possuem padrões tecnológicos diferentes.

A primeira função discriminante canônica, também chamada de primeira função discriminante linear de Fisher (FDLF), foi responsável pela explicação de 74% das variâncias das características significativas apresentadas, e a segunda função discriminante por 26%. Ocorreu maior dissimilaridade entre os grupos TERMSILVI e TRADICI, sendo intermediária entre os grupos TERMSILVI e TECAGRI. Os grupos TECAGRI e TRADICI foram mais próximos entre si. As duas primeiras combinações lineares canônicas foram utilizadas para estabelecer os eixos cartesianos apresentados na Figura 2.

Figura 2 - Gráfico bidimensional da análise discriminante canônica (P<0,01, Distância de Mahalanobis) e localização espacial dos municípios, segundo os três perfis tecnológicos da pecuária de corte do bioma Pampa.



Fonte: elaborada pelos autores

A FDLF separou municípios com sistemas com presença de vacas (PVACAS2ANOS), sistemas de cria mais recria (D_CRIA_RECRIA) e números de unidades de bovinos (UA_BOV), pois apresentam coeficientes altos (com valores negativos). Em oposto, estão os municípios com sistemas com percentual da área com silvicultura em relação com a área explorada com qualquer atividade de produção (P_SILV), percentual de estabelecimentos que fizeram engorda (D_ENGORDA) e percentual de estabelecimentos que fizeram recria (D_RECRIA), que tiveram coeficientes mais altos, porém positivos. A segunda função discriminante canônica discriminou municípios com sistemas com maior percentual de lavoura temporária (PLAVTEMP) e os que recebem orientação técnica (RECEBEORIENT) com coeficientes negativos altos, em relação aos municípios com sistemas com percentuais de maior valor de produção com bovinos em relação ao valor da produção total (PVTP_BOV) e de unidades de bovinos (UA_BOV), variáveis com coeficientes positivos altos.

CONCLUSÕES

O uso sequencial dos métodos de agrupamento e discriminação canônica permitiu identificar três padrões de produção e intensificação tecnológica nos municípios produtores de bovinos de corte no bioma Pampa. As variáveis selecionadas foram adequadas para realizar essa distinção, considerando o uso de dados agregados por município a partir de estabelecimentos com mais de 50 bovinos de corte. Os resultados foram considerados aderentes à realidade da pecuária de corte regional. Embora os dados empregados no estudo datarem de 2006, último censo agropecuário com informações disponíveis no âmbito de estabelecimentos agropecuários, os resultados são relevantes, pois podem servir como uma referência importante para análises comparativas da dinâmica espacial da pecuária ao longo do tempo com os novos censos,

inclusive considerando a interação com outras atividades agropecuárias. Ressalta-se que esses estudos futuros estarão condicionados às variáveis presentes nos próximos censos realizados pelo IBGE. Espera-se, que com esses novos estudos, obtenha-se conhecimento novo e relevante para a pecuária capaz de contribuir com a conservação da biodiversidade no bioma Pampa.

REFERÊNCIAS

BOLDRINI, I. I.; LONGHI-WAGNER, H. M.; BOECHAT, S. C. **Morfologia e taxonomia de gramíneas sul-rio-grandenses**. 2. ed. Porto Alegre: UFRGS, 2008. 88 p.

BRASIL. INSTRUÇÃO NORMATIVA Nº 11, de 4 de abril de 2003, aprovada pela Resolução/CD nº 7/2003 - DOU 74, de 16/04/03 seção 1, p. 101 – B.S. 16, de 21/04/03.

CAMPOS, K. A.; MORAIS, A. R. de; PAIXÃO, C. A. Alternative for the evaluation of coffee seedlings using Fisher's discriminant analysis**. Revista Ciência Agronômica**, v. 47, n. 2, p. 299-306, 2016.

FASIABEN, M.C.R.; SANTUCCI, J.M.; MAIA, A.G.; ALMEIDA, M.M.T.B.; OLIVEIRA, O.C.de; BARIONI, L.G. **Tipificação de municípios produtores de bovinos no Brasil**. Campinas: Embrapa Informática Agropecuária, 2013. 38 p. (Embrapa Informática Agropecuária. Boletim de pesquisa, 33).

HAN, J.; KAMBER, M.; PEI, J. **Data mining**: concepts and techniques. 3rd ed. San Francisco: Morgan Kaufmann Publishers, 2011. 770 p.

IBGE. **Censo agropecuário 2006**: Brasil, grandes regiões e unidades da federação: segunda apuração. Rio de Janeiro, 758 p, 2012. ISBN: 2408107300

MINGOTI, S. A. **Análise de dados através de métodos de estatística multivariada: uma abordagem aplicada.** 1° ed. Belo Horizonte: Editora UFMG, p. 297, 2007.

Ministério do Meio Ambiente. **Pampa**. Disponível em: <http://www.mma.gov.br/biomas/pampa>. Acesso em: 28 mar. 2019.

OLIVEIRA, T. E.; FREITAS, D. S.; GIANEZINI, M. RUVIARO, C. F.; ZAGO, D.; MÉRCIO, T. Z.; DIAS, E. A.; LAMPERT, V. N.; BARCELLOS, J. O. J. Agricultural land use change in the Brazilian Pampa Biome: the reduction of natural grasslands. **Land Use Policy**, v. 63, p. 394-400, 2017.

PROVOST, F; FAWCETT, T. **Data science para negócios: o que você precisa saber sobre mineração de dados e pensamento analítico de dados.** 1° ed. Rio de Janeiro: Editora Alta Books, p. 404, 2016.

Aplicação de técnicas de classificação em imageamento aéreo com RPA para determinação de classes de produtividade do trigo

Jonathan de Matos Diniz[1], *Alessandro Stachak*[2], *Alaine Margarete Guimarães*[3], *Eduardo Fávero Caires*[4]

Setor de Ciências Agrárias e de Tecnologia, Universidade Estadual de Ponta Grossa, Paraná, Brasil

jonathan.lucas.m@gmail.com[1], alessandro.sbk@gmail.com[2], alainemg@uepg.br[3], efcaires@uepg.br[4]

RESUMO

A agricultura de precisão baseia-se na determinação de diferentes áreas de manejo, permitindo a redução de custos e de impactos ambientais. Uma das maneiras de avaliar tais áreas é pela produtividade, classificando-as em faixas de baixa, média e alta. Para determinar essas classes, ferramentas de sensoriamento remoto podem ser utilizadas, tais como, imageamento aéreo com satélites ou aviões. O presente trabalho apresenta uma alternativa acessível com a utilização de aeronaves remotamente pilotadas (RPAs) e técnicas de mineração de dados para determinar as classes de produtividade da cultura do trigo. Os experimentos foram realizados em julho de 2016, em uma área localizada em Piraí do Sul – PR, onde foi semeado o trigo (cultivar Tbio Sinuelo). Foram realizados voos mensais com uma RPA modelo Ebee da Sensefly, com diferentes resoluções espaciais (3,4, 10 e 20 cm/pixel). Os mapas de refletância para cada mês foram gerados e submetidos a correção geométrica para extração dos dados. Com os dados mensais e a produtividade obtida por meio de coletas de campo, técnicas de balanceamento de classes e o algoritmo Suport Vector Machine (SVM) foram utilizados para realizar a classificação, obtendo altos índices kappa, a medida em que o grau de balanceamento foi ajustado.

PALAVRAS-CHAVE: VANT, Agricultura de precisão, Sensoriamento remoto.

ABSTRACT

Precision agriculture is based on the determination of different areas of management, allowing thereduction of costs and environmental impacts. One of the ways to evaluate such areas is by productivity, classifying them into low, medium and high productivity ranges. To determine these classes, remote sensing tools can be used, such as aerial imagery with satellites or airplanes. The present paper presents an accessible alternative with the use of remotely piloted aircraft (RPAs) and data mining techniques to determine wheat crop productivity classes. The experiments were carried out in July 2016, in an area located in Piraí do Sul - PR, where wheat (Tbio Sinuelo) was sown. Monthly flights were carried out with a Sensefly RBE Ebee model, with different spectral resolutions (3.4, 10 and 20 cm / pixel). Thereflectance maps for each month were generated and submitted to geometric correction for extraction of the data. With monthly data and productivity obtained through field collections, class balancing and the Support Vector Machine (SVM) algorithm could be used for classification, obtaining high kappa indexes, the extent to which the degree of balancing is adjusted.

KEYWORDS: UAV, Precision agriculture, Remote sensing.

INTRODUÇÃO

Sistemas para gerenciamento de informações no campo, bem como inovações na interface de trabalho para usuário, são ferramentas poderosas e essenciais para os agricultores, uma vez que permitem o manuseio eficiente de recursos possibilitando redução de gastos (LEWIS, 1998; SØRENSEN et al., 2011). Além disso, esses sistemas fornecem um leque de decisões tanto para tratamento da cultura quanto para o comportamento diante do mercado econômico. A agricultura de precisão visa disponibilizar tais informações para que auxiliem o agricultor durante as tomadas de decisões em campo.

O sensoriamento remoto (SR) junto com a mineração de dados são grandes aliados na agricultura de precisão para a obtenção dessas informações, pois são uma maneira prática e não destrutiva comparada aos métodos comuns de obtenção de dados de campo (CHANG, 2017). Um recurso de fácil acesso e boa relação custo-benefício para o SR é a utilização de Aeronaves Remotamente Pilotadas (RPA – Remotely Piloted Aircraft), pois além de possuirem baixo custo em relação as demais técnicas ainda contam com a vantagem de sua alta resolução espacial e temporal.

Aliado a isso, a mineração de dados pode ser utilizada para encontrar padrões através do processo de Descoberta de Conhecimento em Bases de Dados (KDD - Knowledge Discovery Database) como já utilizado por Murthy et al. (2003) para classificação da colheita de trigo

em imagens multitemporais, bem como em Melgani & Bruzzone (2004) na classificação de imagens de sensoriamento remoto com o uso do algoritmo de Máquina de Vetores de Suporte (SVM).

Em geral, a SVM possui vantagens sobre outros classificadores, por conta de sua capacidade de generalização a partir de um conjunto limitado de amostras de treinamento (MOUNTRAKIS, 2011). Além disso, já demonstrou melhor desempenho sobre outros algoritmos, como comprovado por Pal & Mather (2005) que realizaram a comparação entre SVM, algoritmo de Máxima Verossimilhança e Redes Neurais Artificias, em que obtiveram maior kappa (0.87) com a SVM.

Por mais que existam diversas pesquisas para predição de produtividade, poucas delas utilizam métodos de classificação associados ao sensoriamento remoto, e essa restrição aumenta quando se trata do uso de RPAs. Sendo assim, o presente trabalho teve como objetivo determinar classes de produtividade do trigo utilizando técnicas de mineração de dados e imageamento com RPA.

MATERIAL E MÉTODOS

Área de estudo

Os estudos foram realizados em área de aproximadamente 7,2 ha da Fazenda Paiquerê, situada no município de Piraí do Sul, região Centro-Sul do Paraná, com o cultivo do trigo (cultivar Tbio Sinuelo).

A cultura foi semeada em 28 de julho de 2016 com densidade de 70 sementes por metro e espaçamento entre linhas de 0,17 cm. O experimento foi dividido em quatro parcelas e cada uma foi submetida a um sistema distinto de tratamento de nitrogênio (N) afim de obter diferenças no crescimento da planta desde o início do plantio.

Todas as parcelas receberam, na semeadura, 200 kg ha^{-1} da fórmula 18-36-00($N - P_2O_5 - K_2O$). Na segunda parcela, durante a pré-semeadura, foi aplicado 50 kg N ha^{-1} e na terceira parcela 100 kg N ha^{-1} no mesmo período. Para a quarta parcela houve aplicações aos 40 (27 kg N ha^{-1}) e 50 (100kg N ha^{-1}) dias após a semeadura.

Coleta de dados em campo

A coleta de imagens foi realizada mensalmente, em cinco datas, a partir do primeiro mês de plantio utilizando a RPA eBee (SenseFly). Em todas as datas, a resolução espacial 3,4 cm/pixel foi empregada, entretanto, no quarto mês após a semeadura, as resoluções de 10 e 20

cm/pixel também foram usadas para possíveis comparações. A Tabela 1 apresenta o número de imagens obtidas em cada voo, com as diferentes resoluções utilizadas.

A fim de garantir uma melhor precisão nas imagens, Pontos de Controle em Solo (GCPs) foram posicionados ao redor da área de interesse e suas posições coletadas com receptor GNSS geodésico, o GPS Trimble R4 (NTech Industries Incorporation, EUA), para posterior correção através do Software QGIS [1](Open Source).

Número de imagens obtidas por voo						
	Resolução 3,4cm/px		Resolução 10cm/px		Resolução 20cm/px	
Datas	RGB	NIR	RGB	NIR	RGB	NIR
28/07/2016	154	192	-	-	-	-
28/08/2016	126	158	-	-	-	-
27/092016	127	143	-	-	-	-
10/10/2016	120	146	39	46	18	21
01/11/2016	113	123	38	42	-	-

Tabela 1 - Datas e quantidade de imagens obtidas em cada voo realizado com as diferentes cameras e respectivas resoluções espaciais

Os mapas de produtividade foram gerados com software GreenStart APEX 3.7.1.0 a partir dos dados obtidos por duas colhedoras, a John Deere 9670 STS com monitor de colheita GS2 2630 e a John Deere 9650 STS com monitor GreenStar Original. Ao final, o software forneceu uma malha poligonal com valores na faixa de 0 a 28.000ton/ha, totalizando 3.115 instâncias, incluindo ruídos de leitura.

Com base nos dados disponibilizados pela Secretaria da Agricultura e Abastecimento do Paraná (SEAB) sobre a safra de 2016-2017 os valores de produtividade abaixo de 2.500 ton/ha foram atribuídos a classe Baixa, valores entre 2.500 e 4.500 ton/ha a classe Média e acima de 4.500 ton/ha a classe Alta. Valores maiores que 9.000 ton/ha e iguais a zero foram considerados como ruídos.

Processamento

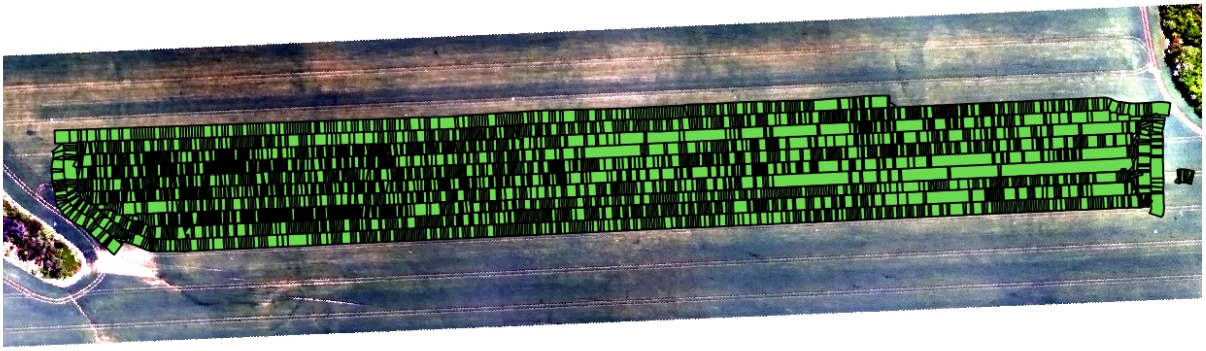
Utilizando o software eMotion2[2] (SenseFly), as imagens foram georreferenciadas com os dados do GPS da própria RPA, atribuindo a cada imagem as informações de latitude, longitude e altitude referentes ao momento de sua captura. Após isso, os ortomosaicos e mapas de refletância foram gerados a partir do Software Pix4D (Pix4D SA) e, utilizando os

[1] Equipe de Desenvolvimento QGIS (2009). Sistema de Informação Geográfica QGIS. Projeto Geoespacial Open Source. URL http://qgis.osgeo.org.

[2] Software comercial, disponibilizado juntamente com a RPA eBee para controle da aeronave, mais informações em: https://www.sensefly.com/software/emotion/.

pontos de GPS coletados para os GCPs, a correção dos mapas de refletância a partir do software QGIS foi realizada.

Figura 1 – Malha poligonal de produtividade sobre a área experimental da Fazenda Paiquerê (1:2500)



Ainda dentro do ambiente QGIS, a partir da malha poligonal obtida com a máquina no momento da colheita (Figura 1), os dados estatísticos de cada polígono, tais como média, variância e mediana das bandas espectrais, foram extraídos em todos os mapas de refletância para o processamento.

Na etapa de pré-processamento, utilizando a Linguagem R, a remoção de atributos desnecessários e de ruídos existentes em cada uma das bases, bem como a discretização do atributo meta em três classes de produtividade (Alta, Média e Baixa), foram realizadas.

Ao final do pré-processamento todas as bases possuíam 3.090 instâncias, contudo, as classes estavam desbalanceadas, o que prejudica o resultado da classificação. Para suprir este problema optou-se por fazer o balanceamento de classes utilizando o filtro Resample disponível no software Weka.

Através da distribuição normal, o filtro Resample cria novas instâncias para as menores classes substituindo por outras de classes majoritárias a fim de manter a mesma quantidade de instâncias. O parâmetro “biasToUniformClass” recebe valores de 0 a 1 para controlar o balanceamento, sendo que, quanto mais próximo de zero, mais parecida ao original e quanto mais perto de 1, mais balanceada a nova base será. Nesse trabalho adotou-se a variação de 0 a 0,5 para o referido parâmetro.

Para a estimativa das classes de produtividade, correlacionando os dados espectrais com a produtividade em campo, o algoritmo Máquina de Vetores de Suporte (SVM) presente na biblioteca LibSVM (CHANG & LIN, 2001)foi executado, com validação cruzada de 10 folds. A avaliação do resultado da classificação foi feita com base na estatística Kappa (LANDIS; KOCH, 1977), que corresponde a uma métrica que compara uma acurácia observada com a acurácia esperada.

Por se tratar do ápice do estágio fenológico da cultura, as imagens utilizadas foram as de outubro, seguido das imagens de setembro e agosto de maneira a tentar antecipar os resultados.

RESULTADOS E DISCUSSÃO

Os resultados da classificação foram avaliados com base nos valores de índice Kappa obtidos, os quais aumentaram a medida em que o nível de balanceamento também aumentou. Ao aplicar o menor balanceamento possível com o filtro, o valor Kappa obtido foi de aproximadamente 0,66 para imagens NIR do mês de outubro em todas as resoluções, chegando a 0,80 quando submetidos ao filtro com bias 0,4, classificando corretamente em média 89,4% das instâncias. Esses resultados foram superiores aos obtidos por Pantazi et al. (2016), que atingiram uma porcentagem média de acerto de classificação de 81,65% utilizando Redes Supervisionadas de Kohonen (SKN) para estimar classes de produtividade a partir de NDVI e dados de solo.

Tabela 2 – Índices Kappa para as bases de dados utilizadas com e sem balanceamento de classes, referentes ao mês de outubro.

	Índice Kappa – Imagens de outubro de 2016					
Bias	NIR 3,4	NIR 10	NIR 20	RGB 3,4	RGB 10	RGB 20
-	0	0	0	0,0024	0,005	0
0	0,6652	0,6652	0,6652	0,6526	0,6273	0,6229
0.1	0,7385	0,7385	0,7393	0,7272	0,6867	0,6974
0.2	0,762	0,762	0,7613	0,7561	0,7212	0,7265
0.3	0,796	0,796	0,7961	0,7896	0,7658	0,7587
0.4	0,8081	0,8081	0,8082	0,8003	0,7745	0,7824
0.5	0,822	0,822	0,822	0,815	0,79	0,796

Com as imagens RGB do mês de outubro os resultados foram pouco menores do que a NIR (Tabela 2), porém ainda seguem o mesmo padrão de crescimento com o aumento do parâmetro bias.

As imagens dos meses de setembro e agosto (Tabela 3) produziram resultados similares aos do mês de outubro. Em ambos os meses os melhores resultados foram também para os dados da câmera NIR (3,4cm/pixel), ainda que por pouca diferença. Quando aplicado bias 0,4 para o filtro, os dois meses atingiram índice de avaliação 0,80 nas imagens NIR e 0,79 nas imagens RGB, como evidenciado na Tabela 3.

Tabela 3 – Índices Kappa para as bases de dados utilizadas com e sem balanceamento de classes, referentes aos meses de agosto e setembro.

	Índice Kappa com as diferentes câmeras			
	Agosto de 2016		Setembro de 2016	
Bias	NIR 3,4	RGB 3,4	NIR 3,4	RGB 3,4
-	0	0,0016	0	-0,0049
0	0,6662	0,6444	0,6652	0,6462
0.1	0,7379	0,7189	0,7385	0,707
0.2	0,7634	0,7473	0,762	0,7413
0.3	0,796	0,7773	0,796	0,7778
0.4	0,8081	0,7975	0,8081	0,7994
0.5	0,8231	0,8151	0,822	0,8113

CONCLUSÕES

A partir dos testes, é visível que o algoritmo Suport Vector Machine consegue, com um bom nível de confiabilidade, determinar classes de produtividade do trigo. É possível observar ainda, que não houve mudanças significativas nos resultados entre as resoluções e entre os meses de setembro e outubro o que demonstra que não são necessárias altas resoluções para chegar a um resultado satisfatório e que é possível determinar antecipadamente a classificação das áreas de manejo.

São necessários ainda outros trabalhos para que se possa validar definitivamente o método e também possíveis testes com mais classes, contudo, deve-se considerar as variáveis que afetam a cultura no momento da discretização do atributo meta, uma vez estas podem alterar a produtividade anual.

AGRADECIMENTOS

Ao CNPq pelo apoio financeiro, à professora Alaine Margarete Guimarães pela oportunidade da participação na pesquisa científica e à equipe do Mestrado em Computação Aplicada da UEPG pela orientação e ajuda durante o processo da pesquisa.

REFERÊNCIAS

CHANG, A.; JUNG, J.; MAEDA, M. & LANDIVAR, J. Crop height monitoring with digital imagery from Unmanned Aerial System (UAS). **Computers and Eletronics in Agricultures**, v.141, p. 232 - 237, 2017.

CHANG CC.; LIN CJ. LIBSVM - A Library for Support Vector Machines (2001). Disponível em: <http://www.csie.ntu.edu.tw/~cjlin/libsvm/> Acessado em: 01 out. 2019.

CHAUDHARY, A.; KOLHE, S. & KAMAL, R. An improved random forest classifier for multi-class classification. **Information Processing in Agriculture**, Elsevier, v.3, p. 215-222, 2016.

CHAVES, A. A.; LA SCALEA, R. A. Uso de VANTs e processamento digital de imagens para a quantificação de áreas de solo e de vegetação. **Anais XVII Simpósio Brasileiro de Sensoriamento Remoto** - SBSR, João Pessoa-PB, Brasil, v. 25, 2015.

KUMAR, P.; GUPTA, D. K.; MISHRA, V. N.; PRASAD, R. Comparision of support vector machin, artificial neural network, and spectral angle mapper algorithms for crop classification using LISS IV data. **International Journal of Remote Sensing**, Taylor & Francis, v.36, p. 1604-1617, 2015.

LANDIS, J. Richard; KOCH, Gary G. The measurement of observer agreement for categorical data. **Biometrics**, p. 159-174, 1977.

LEWIS, T. Evolution of farm management information systems. **Computers and Eletronics in Agriculture**, v. 19, n. 3, p. 233-248, 1998. Disponível em: <https://www.sciencedirect.-com/science/article/pii/S0168169997000409> Acessado em: 13 ago. 2018.

MELGANI, Farid; BRUZZONE, Lorenzo. Classification of hyperspectral remote sensing images with support vector machines. **IEEE Transactions on geoscience and remote sensing**, v. 42, n. 8, p. 1778-1790, 2004.

MOLIN, J. P. Definição de unidades de manejo a partir de mapas de produtividade. **Engenharia Agrícola**, v. 22, n. 1, p. 83-92, 2002.

MOUNTRAKIS, G.; IM, J.; OGOLE, C. Support Vector machines in remote sensing: A review. **Journal of Photogrammetry and Remote Sensing** - ISPRS, Elsevier, v.66, p. 247-259, 2011.

MURTHY, C.; RAJU, P. & BADRINATH, K. Classification of wheat crop with multi-temporal images: performance of maximum likelihood and artificial neural networks. **International Journal of Remote Sensing**, Taylor & Francis, v.24, p. 4871-4890, 2003.

PAL, M.; MATHER, P. Support vector machines for classification in remote sensing. **International Journal of Remote Sensing**, Taylor & Francis, v. 26, p. 1007-1011, 2005.

PANTAZI, Xanthoula Eirini et al. Wheat yield prediction using machine learning and advanced sensing techniques. **Computers and Electronics in Agriculture**, v. 121, p. 57-65, 2016.

SØRENSEN, C. G. et al. Functional requeriments for a future farm management information system. **Computers and Eletronics in Agriculture**, v. 76, n. 2, p. 266-276, 2011. Disponível em: <http://www.sciencedirect.com/science/article/pii/S016816991100055X>. Acessado em 13 ago. 2018.

Proposta de uma estação meteorológica utilizando uma rede de longo alcance: LoRa

Guilherme Henrique Alves da Silva[1,3], *Marco Antonio Franzin Leite*[1,3], *Renan Guides Diniz de Barros*[1], *Paulo Roberto Carvalho dos Santos*[1,3], *Juliana Saragiotto Silva*[2]

[1] Departamento de Área de Eletroeletrônica, Instituto Federal do Mato Grosso, Cuiabá, Mato Grosso, Brasil, guilhermehenriquee9@gmail.com, marcofranzin1999@gmail.com, renanguides@hotmail.com, paulorobcss@gmail.com

[2] Departamento de Área de Informática, Instituto Federal do Mato Grosso, Cuiabá, Mato Grosso, Brasil, juliana.silva@cba.ifmt.edu.br

[3] Programa de Educação Tutorial AutoNet, Instituto Federal do Mato Grosso, Cuiabá, Mato Grosso, Brasil

RESUMO

O monitoramento climático é essencial para as atividades humanas, desde atividades do dia-a-dia (como a decisão sobre qual roupa vestir, em função da temperatura ambiente) até decisões sobre o processo produtivo de uma cultura agrícola. Com as evoluções tecnológicas vigentes na atualidade, surgem as necessidades da implementação de modelos computacionais que auxiliem a análise das informações coletadas neste processo de monitoramento climático. Desta forma, observando essa realidade, o objetivo desse trabalho é apresentar a proposta de uma estação meteorológica utilizando uma rede de longo alcance (*Long Range* – LoRa), por meio do monitoramento dos dados, em tempo real, sem a necessidade de conexão com a internet. Este processo de monitoramento pode ser visualizado por meio de uma página *web*, que fornece o horário e a sequência da obtenção dos dados. A montagem do circuito elétrico consistiu na utilização dos sensores DHT22 e pluviômetro, conectados ao microcontrolador *ESP32 LoRa*. Para a criação do banco de dados e da página *web* foram empregados os *softwares Apache*, *phpMyAdmin* e *MySQL*. Os resultados mostraram-se satisfatórios, tendo em vista que houve a visualização das informações do monitoramento climático por meio de uma conexão com a página *web*, bastando o acesso à rede *wireless*.

PALAVRAS-CHAVE: Agricultura de Precisão, Mecanismos, Previsão do Tempo, Rede de Longo Alcance.

ABSTRACT

Climate monitoring is essential for human activities, from everyday decisions (such as which clothes to wear, based on temperature) to decisions about an agricultural culture productive process. With today's operating technology evolution, comes the need for implementation of computational models that assist the collected data analysis in the climate monitoring process. Therefore, observing this reality, this work's objective is to present a metereological station proposal employing a long range network (Long Range - LoRa), through real time data monitoring, unnacompanied by the need of internet connection. This monitoring process can be viewed through a web page, which provides the time and data retrieval sequence. The electrical circuit assembly consisted in utilizing the DHT22 and pluviometer sensors, connected to the ESP32 LoRa microcontroller. For the database and web page development, the used software were Apache, phpMyAdmin and MySQL. The results were satisfactory, given that there was climate monitoring data visualized by a web page connection, just accessing the wireless network.

KEYWORDS: Precision Agriculture, Mechanism, Weather forecast, Long Range Network.

INTRODUÇÃO

Há muito tempo a temática “crescimento populacional” versus “produção de alimentos” vem preocupando vários estudiosos. Com isso, teorias vêm sendo levantadas e derrubadas ao passo que os avanços tecnológicos apresentam alguns caminhos possíveis para estes desafios decorrentes da crescente população mundial. Uma das teorias sobre crescimento demográfico, que já foi superada, é a Malthusiana. Preocupado com os problemas socioeconômicos (desemprego, fome, êxodo rural, rápido aumento populacional) decorrentes da Revolução Industrial e que afetavam seriamente a Inglaterra, o pastor Thomas Robert Malthus (1766-1834) acreditava que, enquanto o crescimento populacional obedeceria a uma progressão geométrica, o aumento da produção agrícola teria um ritmo apenas aritmético (ABRAMOVAY, 2010). Desta forma, entendia que a população mundial tinha potencial de crescimento ilimitado e, a natureza, inversamente, não teria o mesmo crescimento de recursos para alimentá-la (LUCCI et al., 2005).

Ademais, é importante mencionar que a população mundial vive de acordo com as condições do ambiente, em termos de alimentação (FONTANA et al., 2015). Com isso, tem-se a necessidade pela busca de mecanismos que contribuam com o objetivo de acelerar a produção, como o monitoramento das variáveis climáticas do ambiente, que são fundamentais para o êxito da produção (SOUSA; DRUMOND; NALDI, 2015).

É justamente neste contexto que está inserida a Agricultura de Precisão, que tem o objetivo de melhorar o manuseio do solo. A lavoura é dividida em parcelas menores, sendo observadas, com maior rigor, todas as informações da área demarcada. Essa inspeção contribuiu para o aumento da produtividade e, por conseguinte, possibilita a diminuição na utilização de insumos agrícolas (como agrotóxicos, adubos e fertilizantes) (MOLIN, 2003). Do mesmo modo, a implementação de métodos automatizados possibilitou o estudo mais abrangente das variáveis climáticas e, com isso, houve o aumento da capacidade de manuseio das informações coletadas. Entretanto, no transcorrer das atividades, existem imprevistos na execução, caracterizando-se pela perda de conexão, interferência, problemas elétricos e mecânicos, dentre outros (STRASSBURGER et al., 2011). A perda de conexão é exemplificada como a falta de internet em todos os pontos das propriedades. Com isso, há necessidade de tecnologias que consigam trabalhar sem a conexão com a rede.

Face ao exposto, o objetivo desse trabalho é apresentar a proposta de uma estação meteorológica utilizando uma rede de longo alcance (*Long Range – LoRa)*, por meio do monitoramento dos dados, em tempo real, sem a necessidade de conexão com a internet.

Desta forma, este artigo está estruturado em 5 seções, incluindo esta introdução. Na seção 2 são apresentados os termos essenciais para a compreensão do tema. Na seção 3 é descrita a metodologia utilizada na elaboração do trabalho. Na sequência, na seção 4 são mencionados os resultados e as etapas para a elaboração do trabalho. Na seção 5 realizou-se a análise do experimento, ressaltando o seu funcionamento. Por fim, na última seção, são abordadas as conclusões e as perspectivas de pesquisas futuras.

REVISÃO BIBLIOGRÁFICA

Esta seção tem como objetivo descrever, de forma sucinta, os termos essenciais para a compreensão desta pesquisa, como Agricultura de Precisão, Microcontroladores e Rede de Longo Alcance.

Agricultura de Precisão

A Agricultura de Precisão tem como intuito a análise de variáveis que participam do processo de produção. O objetivo dessa inspeção é a busca da melhor utilização dos recursos do ambiente (MOLIN, 2003). A sua implementação existe desde a Revolução Industrial, onde foi empregada a estratégia de divisão das grandes propriedades em pequenos lotes, no qual eram captados dados para realizar estudos do solo, com o objetivo de desenvolver técnicas de acordo com as necessidades da área analisada (LAMPARELLI, 2008).

Nos dias atuais, nota-se a presença da utilização dos conceitos de Agricultura de Precisão em atividades que se utilizam de métodos que visam o estudo da variabilidade da produção, com o intuito da implementação de mecanismos que busquem auxiliar os produtores, de modo a melhorar o desempenho da produção, juntamente com a preocupação com o meio ambiente (BERNARDI, 2014).

Os avanços tecnológicos trouxeram, para a Agricultura de Precisão, a implementação de vários dispositivos eletroeletrônicos, com o objetivo do monitoramento das variáveis presentes na produção (TERUEL, 2012), tais como temperatura, umidade, índice de precipitação, entre outras. Um exemplo desses dispositivos são os microcontroladores – como descrito a seguir.

Microcontroladores

Um microcontrolador consiste em um único chip, cuja finalidade é desempenhar o papel de um microcomputador, que em sua composição já venham embutidos acessórios (processador, memória e periféricos de entrada/saída) necessários para o seu funcionamento. A programação para a utilização desses periféricos ocorre por meio da linguagem C (OLIVEIRA JÚNIOR; DUARTE, 2010).

Atualmente, existem vários modelos de microcontroladores no mercado; cada um possui características peculiares e sua aplicação depende da necessidade do usuário. Dentre os vários modelos, destacam-se: Arduino, *ESP*, *PIC* e *CLP*. O *ESP* possui vários modelos, sendo os principais: *ESP01, ESP8266* e *ESP32*. O *ESP32* é o modelo que se destaca, possuindo características que são aplicáveis no cenário da Internet das Coisas (*Internet of Things* - IoT), tendo o *wifi* integrado, o que possibilita a conexão com uma rede (SINGH; KAPOOR, 2017). Essas placas de desenvolvimento (microcontroladores) permitem aos usuários a conexão de módulos, algo que aumenta o leque de aplicações do equipamento; um desses módulos é o *LoRa* – que é descrito a seguir.

Rede de Longo Alcance

A rede de longo alcance (*Long Range* - LoRa) consiste em uma tecnologia de comunicação sem fio criada pela *Semtech Corporation* e organizada pela *Lora Alliance*. Desta forma, para o seu funcionamento necessita de um protocolo específico, neste caso, o *LoraWan* (TEIXEIRA; ALMEIDA, 2014). A rede LoRa tem como característica a comunicação entre o dispositivo e o meio de transmissão, sendo responsável pelo teste de conexão e a verificação da transferência de dados. No que se refere à comunicação sem fio, existem várias tecnologias que podem ser empregadas, sempre levando em consideração os seguintes requisitos: velocidade na taxa de transmissão, protocolo, frequência e técnica de modulação de pulso (FOROUZAN, 2007).

Uma característica das tecnologias de transmissão é a frequência de operação. A rede *LoRa* tem uma faixa de rádio frequência específica, entre 10^2 e 10^6 Hertz. A utilização da união da modulação *LoRa* e do protocolo *LoRaWan* permitem aos usuários o envio de informações a longas distâncias, em áreas urbanas, de 2 a 5 km. (TEIXEIRA; ALMEIDA, 2014).

Uma vez que alguns conceitos basilares do trabalho foram apresentados, a seção seguinte se dedica à apresentação dos materiais e métodos utilizados no desenvolvimento do protótipo.

MATERIAL E MÉTODOS

O presente trabalho compreende a elaboração de um protótipo, que visa a simulação do funcionamento de uma estação meteorológica, a partir da utilização de equipamentos para a medição de variáveis climáticas de uma determinada região.

O desenvolvimento desse trabalho consistiu nas seguintes etapas: (i) Revisão da Literatura; (ii) Identificação dos requisitos do protótipo; (iii) Elaboração do protótipo, empregando recursos de *hardware* e *software*; e (iv) Testes e análise dos resultados.

Na etapa (i) foram realizadas pesquisas em artigos da área, com o objetivo de conceituar termos técnicos, facilitando a compreensão sobre o tema. Na sequência, a etapa (ii) consistiu na pesquisa dos requisitos necessários para o desenvolvimento do protótipo.

Após a verificação dos requisitos necessários ao projeto, a etapa (iii) compreendeu a criação do protótipo, para o monitoramento das variáveis climáticas do ambiente, utilizando sensores e a tecnologia LoRa para o envio dos dados – o que possibilitou a utilização do protótipo em locais sem conexão com a internet. A visualização dos dados foi realizada a partir de uma página *web*. Para tanto, foi necessário a programação e a simulação do protótipo, utilizando os *softwares* (dispostos no Quadro 1) e os componentes de *hardware* (dispostos no Quadro 2) envolvidos na montagem do protótipo. Por fim, a última etapa desta pesquisa (iv), consistiu na realização dos testes e na análise dos resultados, por meio dos quais foi possível relatar os principais benefícios da implementação do protótipo.

Até o momento foram apresentadas as etapas e os recursos materiais utilizados nesta pesquisa. A seção seguinte expõe os procedimentos adotados na elaboração do protótipo.

Quadro 1. *Softwares* utilizados

Software	**Descrição**	**Figura**
Apache	Servidor responsável por disponibilizar páginas *web*. A versão utilizada foi a 2.4.	
MySql	É um sistema de gerenciamento de banco de dados, sendo utilizada a versão 8.0.	
PHP	Linguagem de *script open source* empregada no desenvolvimento de páginas *web*. Foi utilizada a versão 7.3.6.	
IDE Arduino	Ambiente de desenvolvimento integrado; versão utilizada – 1.8.9.	
Sublime Text 3	Editor de texto que permite a programação em HTML; neste caso, empregou-se a versão 3.2.1.	
Raspbian	Sistema Operacional utilizado no *Raspberry Pi*; versão utilizada – 4.14.	
Fritizing	Permite realizar a simulação de conexões elétricas, deixando evidente todas as ligações realizadas. A versão utilizada foi a 0.9.3b.	

Fonte: elaborado pelos autores

Quadro 2. Componentes de *hardware* utilizados

Hardware	**Descrição**	**Figura**
DHT22	Sensor utilizado para medição da temperatura e da umidade.	
Protoboard	Utilizada para a montagem de circuitos elétricos experimentais.	
Raspberry Pi 3 (Modelo *B+)*	Consiste em um microcomputador.	
Resistor	Componente de circuitos elétricos que tem a função de limitar os valores da corrente elétrica.	
Jumpers	Filamentos de metal responsáveis pela condução de eletricidade.	
Pluviômetro de báscula	Equipamento responsável pela medição da chuva.	
ESP32 LoRa	É um microcontrolador com módulo *LoRa*.	

Fonte: elaborado pelos autores

PROTÓTIPO DA ESTAÇÃO METEOROLÓGICA

Para o desenvolvimento do protótipo da estação meteorológica, foram realizados procedimentos, sendo descritos, em detalhes, a seguir.

Configuração do Raspberry Pi 3 B+

Em um primeiro momento, houve a instalação do Sistema Operacional apropriado para o equipamento, o *Raspbian*. Após a configuração do Sistema Operacional, houve a necessidade da atualização dos aplicativos vindos de fábrica, com o intuito de garantir o pleno funcionamento do equipamento. Para tanto, o seguinte comando foi inserido diretamente no terminal de comandos: *sudo apt-get update*. O *Raspberry Pi*, no projeto, tem a função de dar

suporte ao banco de dados, possibilitando o desenvolvimento das atividades de consulta, criação e manutenção das tabelas.

Desenvolvimento do Banco de Dados

O segundo procedimento consistiu no desenvolvimento do banco de dados para o estudo de caso. Desta forma, optou-se pela escolha dos aplicativos *Apache*, *phpMyAdmin* e *MySQL*. Na sequência, realizou-se a instalação destes programas no *Raspberry Pi*, facilitando o gerenciamento do banco de dados, por meio da interface disposta na Figura 1.

Figura 1. Aplicativo utilizado para a manutenção do banco de dados (*phpMyAdmin*)

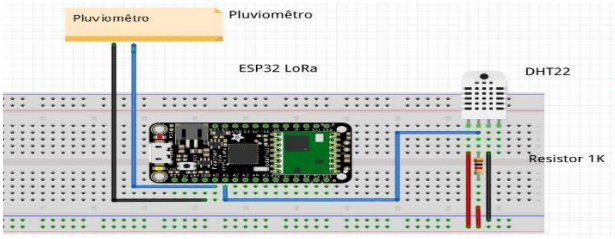


Fonte: elaborada pelos autores

Simulação do Circuito Elétrico

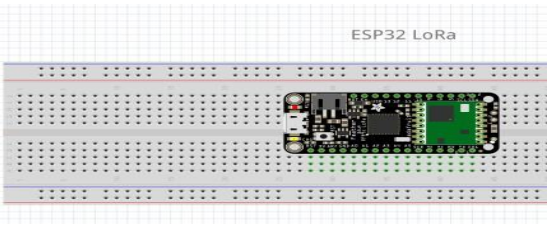
O desenvolvimento da simulação dos circuitos elétricos do projeto foi realizado no software *Fritzing*. Assim sendo, foi necessário a implementação de dois circuitos elétricos, onde cada possuía uma função específica: o primeiro (Figura 2), de coletar e enviar os dados; e, o segundo (Figura 3), responsável por receber e armazenar os dados no banco de dados.

Figura 2. Circuito para coleta e envio dos dados



Fonte: elaborada pelos autores

Figura 3. Circuito que recebe e armazena os dados



Fonte: elaborada pelos autores

Desenvolvimento da Programação

A atividade de programação do protótipo consistiu no desenvolvimento de dois algoritmos: o primeiro, responsável pelo envio das informações, por meio da tecnologia *LoRa;* e, o segundo, com o papel de receber os dados e, em seguida, passar para o banco de dados, de modo a deixar armazenados as leituras do dia. Para tanto, utilizou-se o ambiente de desenvolvimento *Arduino*, cuja função era compilar os comandos estabelecidos, possibilitando aos programadores uma

ampla diversidade de bibliotecas, o que facilitou a execução do trabalho. Os algoritmos desenvolvidos foram passados para os microcontroladores e, estes, tinham a função de interpretar e realizar as operações necessárias para a execução dos respectivos algoritmos.

É importante registrar que, os códigos desenvolvidos neste projeto estão disponíveis no *GitHub*, sendo acessado pelo seguinte link: <https://github.com/guilhermehenrique99/estacao-meteorologica>.

Criação da Rede Wireless

A elaboração de uma rede *wireless* teve como objetivo permitir o acesso dos usuários aos dados provenientes da estação meteorológica.

Desta forma, realizou-se a configuração da rede *wireless* no *Raspberry Pi*. O aparelho possuía a capacidade da geração do sinal, o que permitiu a atribuição de nome e senha – garantindo, assim, a segurança na observação dos dados. O nome da rede e a senha configuradas foram, respectivamente, "EstaçãoMeteorológica" e "12345678", ambas podem ser modificadas, variando de acordo com as especificações do usuário. A Figura 4 explicita o nome da rede criada e, para a sua conexão, basta que a senha seja preenchida corretamente.

Figura 4. Rede *wireless* do *Raspberry Pi*



Fonte: elaborada pelos autores

Criação da página web

Por fim, houve a implementação de uma página *web* (demonstrada na Figura 5), com objetivo de facilitar a amostragem das variáveis captadas pelos sensores. Os dados recebidos foram enviados para o banco de dados, que estava conectado diretamente com a página *web*.

No desenvolvimento da página *web* foi utilizada a linguagem *PHP*, que possui suporte para a criação e a customização da página. Além disso, para a programação do código fonte, empregou-se o editor de texto *Sublime Text 3*, tendo em vista a capacidade de compilação de algoritmos da linguagem *PHP*.

Figura 5. Página *web* desenvolvida



ID	Data	Temperatura	Umidade	Milímetro de Chuva
1	2019-05-17 16:12:33	25.8	56.3	0
2	2019-05-17 16:17:18	24.3	54.4	0.25
3	2019-05-17 16:17:19	24.3	54.3	0.5
4	2019-05-17 16:17:20	24.3	54.2	0.75
5	2019-05-17 16:17:21	24.3	54.3	1

Fonte: elaborada pelos autores

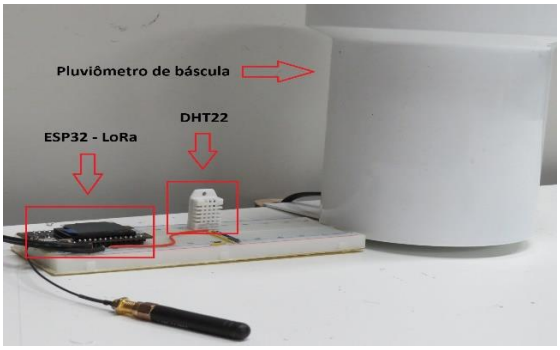
Após a conexão com a rede *wireless* criada (EstaçãoMeteorológica), tem-se a necessidade da utilização de um navegador *web*, para o acesso à página. Em seguida, faz-se necessário digitar o endereço *IP* 10.42.0.1, para que se tenha acesso, em tempo real, aos dados captados pelos sensores – como temperatura, umidade e milímetro de chuva, juntamente com a data e a hora da coleta (Figura 5). Exemplos do funcionamento do equipamento em um determinado período, são relatados na próxima seção.

RESULTADOS E DISCUSSÃO

Esta seção dedica-se a apresentar a validação do protótipo, resultante desta pesquisa, de acordo com os métodos utilizados no seu desenvolvimento e a análise do funcionamento do equipamento proposto.

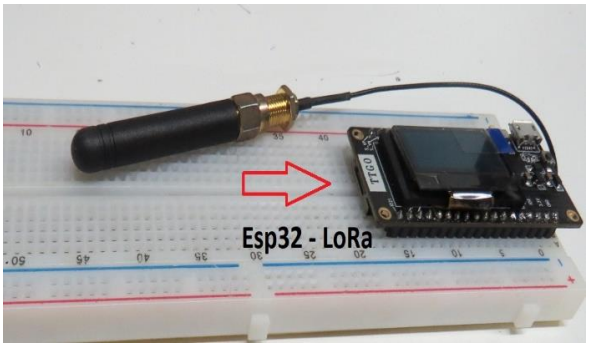
A montagem do protótipo consistiu-se no emprego dos componentes de *hardware* (dispostos no Quadro 2). Desta forma, o primeiro circuito (Figura 6) tinha como função a leitura das variáveis climáticas do ambiente (umidade, temperatura e milímetros de chuva), por meio dos sensores (*DHT22* e pluviômetro), e o envio ao banco de dados. Já o segundo (Figura 7) circuito, tinha a responsabilidade de receber os dados advindos dos sensores e armazená-los no banco de dados.

Figura 6. Circuito responsável por enviar os dados



Fonte: elaborada pelos autores

Figura 7. Circuito responsável por receber os dados



Fonte: elaborada pelos autores

Para o acesso aos dados coletados, houve a necessidade da criação e configuração de uma rede *wireless*. O mecanismo utilizado para o suporte da rede foi o *Raspberry Pi*, tendo em sua configuração a opção de desenvolvimento do *wireless*; além disso, serviu como base para o desenvolvimento do banco de dados do projeto. Após a conexão como banco de dados, foi utilizado o navegador, para a visualização da página, bastando digitar o endereço *IP* correspondente (10.42.0.1). A página era atualizada automaticamente, permitindo o monitoramento das informações, em tempo real. Vale ressaltar que, para otimizar o funcionamento do protótipo, quando não havia registro de chuva, a coleta das variáveis umidade e temperatura ocorria a cada 10 minutos (tempo pré-estabelecido durante a programação). No entanto, durante o período de chuva, o algoritmo foi configurado para enviar os dados assim que forem coletados, de modo a registrar o índice de chuva da região.

Figura 8. Circuito responsável por receber os dados

Instituto Federal do Mato Grosso

Estação Meteorológica

Dados climáticos

ID	Data	Temperatura	Umidade	Milímetro de Chuva
62	2019-06-05 17:17:00	24.6	43.7	1.75
63	2019-06-05 17:17:01	24.7	41.3	2
64	2019-06-05 17:17:02	24.7	43.7	2.25
65	2019-06-05 17:17:03	24.7	43.7	2.5
66	2019-06-05 17:17:05	24.6	43.7	2.75
67	2019-06-05 17:17:16	24.7	43.8	3
68	2019-06-05 17:17:17	24.7	43.8	3.25
69	2019-06-05 17:17:18	24.7	43.8	3.5
70	2019-06-05 17:17:18	24.7	43.8	3.75
71	2019-06-05 17:17:19	24.7	41.3	4
72	2019-06-05 17:17:20	24.6	43.7	4.25

Fonte: elaborada pelos autores

Os dados coletados pela estação são enviados para o banco de dados. O mecanismo utilizado para a transmissão dos dados foi a tecnologia LoRa – que permite uma maior distância de comunicação, podendo chegar até a 3 km, dependendo das obstruções do ambiente (prédio, árvores, carros, entre outros) e da antena utilizada, permitindo, assim, o monitoramento de maiores áreas. Nesta pesquisa, os testes foram realizados no IFMT – Campus Cuiabá – Cel. Octayde Jorge da Silva, com uma distância de 100 metros. Os dados coletados estão disponíveis no seguinte endereço eletrônico: <https://www.youtube.com/channel/UC7pqeI_cZnRBfgz1to_Grxg>

Na literatura, existem trabalhos que abordam o desenvolvimento de estações meteorológicas para áreas com cobertura de sinal de telefone e rede de internet. Silva e Padilha (2011), por exemplo, relatam a utilização dos Serviços Gerais de Pacote por Rádio (*General*

Packet Radio Services – GPRS) como mecanismo responsável pelo envio de dados coletados pelos sensores; neste caso, foi necessário estar na área de cobertura de sinal da telefonia. Já no trabalho de Elias et al. (2014), elaborou-se uma estação meteorológica utilizando arduino, juntamente com o *shield* ethernet, realizando a conexão da estação com a página *web*.

No entanto, o diferencial do protótipo proposto nesta pesquisa, é o fato de se buscar a implementação de tecnologias que trabalham sem a necessidade do equipamento estar conectado a uma rede ou em uma área de cobertura dos sinais de telefonia.

CONCLUSÕES

A partir dos elementos aqui apresentados, pode-se concluir que o desenvolvimento do protótipo mostrou-se satisfatório, uma vez o estudo permitiu a implementação do protótipo de uma estação meteorológica, para o monitoramento das variáveis climáticas, sem a necessidade de conexão com a internet. Assim, por meio dessa proposta, constata-se que é possível realizar o estudo das condições do ambiente de uma determinada região, utilizando-se uma rede LoRa.

É importante registrar, ainda, que, a utilização de estações meteorológicas em propriedades agrícolas, onde não há disponibilidade de conexão com a internet em todos os locais, impossibilita o envio dos dados para uma central – o que dificulta o monitoramento das variáveis climáticas e, assim, necessita-se do deslocamento do proprietário até à região, sempre que houver a necessidade de realizar um monitoramento ambiental. Por isso, vislumbra-se que é necessária a implementação de novas tecnologias, como a proposta neste estudo, como uma alternativa para minimizar este problema.

Ademais, como perspectiva de trabalhos futuros propõe-se a elaboração de uma rede de estações meteorológicas, realizando a conexão entre elas. Para tanto, faz-se necessário a passagem dos dados de uma estação a outra, até que se chegue a uma central de dados, possibilitando o armazenamento dos dados de toda a rede e, assim, o acesso aos dados climáticos de toda a propriedade.

AGRADECIMENTOS

Os autores agradecem ao PET AutoNet, pela utilização do laboratório; ao Prof. Dr. Rupert pelo apoio financeiro; ao Prof. Dr. Manoel Rodrigo Moreira, pelo apoio nos recursos de fundamentação teórica; à Claudia Compagnoni, pelo incentivo no decorrer do trabalho.

REFERÊNCIAS

ABRAMOVAY, R. Alimentos versus População: está ressurgindo o fantasma malthusiano. **Revista Ciência e Cultura**, v.62, n.4, Out. 2010.

BERNARDI, A. C. C.; NAIME, J. M.; RESENDE, A. V.; BASSOL, L. H.; INAMASU, R.Y. **Agricultura de Precisão**: Resultados de um novo olhar, Embrapa, Brasília, DF. 2014.

ELIAS, A. A. A.; SILVA, J. C. P.; GONÇALVES, R. N.; SOUZA, T. S. Ardweather: uma estação meteorológica baseada no arduino e em web services restful. In: SAFETY, HEALTH AND ENVIRONMENT WORD CONGRESS, 14.,2014, Cubatão. **Proceedings...** Cubatão: COPEC, 2014.p. 44-48.

FONTANA, R. L. M.; COSTA, S. S.; SILVA, J. A. B.; RODRIGUES, A. J. Teorias Demográficas e o Crescimento Populacional no Mundo. **Revista Ciência Humanas e Sociais**, Aracaju, v.2, n.3, p.113-124, mar. 2015.

FOROUZAN, B. A. **Comunicação de dados e redes de computadores**. 4. ed. São Paulo: McGraw-Hill, 2007. 1134 p.

LAMPARELLI, R. A. C. **Agricultura de Precisão**, 2016. Agência Embrapa de Informação Tecnológica, 2016.Disponível em: http://twixar.me/XzLn. Acesso em: 12 maio. 2019.

LUCCI, E. A.; BRANCO, A. L.; MENDONÇA, C. R. A. **Território e sociedade**: Geografia geral e do Brasil. 3. ed. São Paulo: Saraiva, 2005.

MOLIN, J. P. Agricultura de Precisão. Parte I: O que é o Estado da Arte em sensoriamento. **Revista de Engenharia Agrícola**, Jaboticabal, v.17, n.2, p.97-107, Dez. 2003.

OLIVEIRA JÚNIOR, M.; DUARTE, R. O. **Introdução ao Projeto com Microcontroladores e Programação de Periféricos**. Universidade de Minas Gerais. 2010. (Apostila de Introdução ao Projeto com Microcontroladores e Programação de Periféricos).

SILVA, J. L.; PADILHA, L. F. O. **Estação Meteorológica para regiões remotas**. 2011. 46 f. Monografia (Bacharel em Engenharia da Computação) – Curso de Engenharia da Computação, Universidade Positivo. 2011.

SINGH, K. J.; KAPOOR, D. S. Create your own Internet of Things: A survey of IoT platforms, **IEEE Consumer Electronics Magazin**e, vol. 6, no. 2, pp. 56-68, April 2017.

SOUSA, M. M.; DRUMOND, L. C. D.; NALDI, M. C. Sistema computacional para aquisição automática e disponibilização de dados meteorológicos. **Revista Engenharia Agrícola**, Jaboticabal, v.35, n.3, p.606-612, maio/jun. 2015.

STRASSBURGER, A. S.; MENEZES, A. J. E. A.; PERLEBERG, T. D.; EICHOLZ, E. D.; MENDEZ, M. E. G.; SCHOFFEL, E. R. Comparação da temperatura do ar obtida por estação meteorológica convencional e automática. **Revista Brasileira de Meteorologia**, v.26, n.2, p.273 – 278, 2011.

TEIXEIRA, G. B.; ALMEIDA, J. V. P. **Rede LoRa® e Protocolo LoRaWan® aplicados na agricultura de precisão no Brasil**. 2014. 77 f. Monografia (Bacharel em Engenharia Eletrônica) – Curso de Engenharia Eletrônica, da Universidade Tecnológica do Paraná, 2014.

TERUEL, B.; ALVES, E. C. S.; ROCCIA, C. J.; ARNOLD. F. J.; BRAVO-ROGER, L. L. Desempenho de rede de sensores sem fio em casa de vegetação. **Revista Brasileira de Engenharia Agrícola e Ambiental**. v.16, n.12, p.1374 – 1380. Campina Grande, PB. 2012.

PisciculTec: um sistema WEB para gerenciamento e monitoramento de atividades para piscicultura

Mateus Resende Rodrigues[1], Igor Medeiros Vanderlei[1], Luis Filipe Alves Pereira[1], Anderson Fernandes de Alencar[2]

[1] Curso de Ciência da Computação, Universidade Federal Rural de Pernambuco, Garanhuns, Pernambuco, Brasil, mateus.resende.r@gmail.com, igor.vanderlei@gmail.com, luisfilipeap@gmail.com

[2] Curso de Pedagogia, Universidade Federal Rural de Pernambuco, Garanhuns, Pernambuco, Brasil, anderson.alencar@ufrpe.br

RESUMO

A piscicultura fatura mais de meio bilhão de reais ao ano no nordeste brasileiro. Apesar disso, pesquisas bibliográficas e entrevistas informais com profissionais da área mostraram que a maior parte dos piscicultores da região gerenciam os dados da sua produção utilizando (*i*) diferentes planilhas que não se comunicam entre si, ou (*ii*) lápis e papel. Nesse contexto, o presente trabalho apresenta uma solução de *software* - disponível a todos os piscicultores gratuitamente e de código aberto - capaz de auxiliar o planejamento e o monitoramento da produção em fazendas de piscicultura. Acreditamos que a adoção desta solução por estes produtores pode contribuir para o aumento da produtividade do setor na região nordeste.

PALAVRAS-CHAVE: Piscicultura, *Software*, Aplicação WEB.

ABSTRACT

Northeastern Brazil fish farming makes more than half a billion Brazilian Reais per year. Nevertheless, our assessment shows that most fish farmers in that region manage their production data using either (i) different spreadsheets that do not communicate with each other, or (ii) simple notebooks. In this context, this paper presents an open source software

solution for the above mentioned problem. Our goal is to help the planning and monitoring of production at fish farms. We believe that the adoption of our solution by the fish farmers can contribute to increasing the productivity of the sector in the northeast region of Brazil.
KEYWORDS: Fish-farming, Software, Web application.

INTRODUÇÃO

O manejo produtivo refere-se às atividades realizadas diariamente para garantir o bom funcionamento e a otimização de sistemas de produção na agroindústria. Na piscicultura, isto envolve a frequência de alimentação dos peixes, a quantidade de ração ofertada, o monitoramento das condições químicas da água utilizada, o controle do tamanho da população em cada tanque de produção, e outros (KUBITZA, 2009; TAVARES-DIAS, 2009). Por fim, a qualidade e a eficiência da produção está intimamente ligada ao devido monitoramento de todos esses fatores.

Na região nordeste do Brasil, a tilápia representa 63% da produção de peixes em cativeiro, gerando um faturamento total de mais de meio bilhão de reais (VIDAL, 2016). Apesar de relevante, esses números estão aquém do esperado quando leva-se em consideração as potencialidades do nosso país (SCHULTER; VIEIRA FILHO, 2017).

Nesse contexto, o presente trabalho apresenta o PisciculTec: uma solução de *software* para auxiliar piscicultores a planejarem e gerenciarem suas pisciculturas de forma ágil e organizada visando atingir maiores níveis de produtividade. Por se tratar de uma aplicação WEB, não há necessidade de nenhuma instalação prévia e as funcionalidades podem ser acessadas através de um *browser* instalado em qualquer computador, *smartphone*, ou *tablet* com acesso à internet.

O PisciculTec permite o acompanhamento ao longo do tempo das propriedades físicas e químicas da água e do crescimento dos peixes nos tanques de produção. Além disso, o sistema também indica o número ideal de tanques para a produção, a quantidade de ração que deve ser ministrada e a população inicial de peixes em cada tanque para que os objetivos de produtividade do piscicultor sejam atendidos.

O trabalho desenvolvido está disponível sob a licença de software livre, com o código-fonte disponível na url <https://github.com/lmts-ufape/psicultec>. O sistema também está disponível para uso, hospedado no servidor da equipe desenvolvedora em <http://app.uag.ufrpe.br/psicultec>.

MATERIAL E MÉTODOS

O *software* proposto foi desenvolvido na plataforma web, utilizando a linguagem de programação PHP 7.3.6, o *framework* Laravel 5.6.39, banco de dados Postgresql 10.9 e servidor Apache 2.4.29. Para auxiliar na reprodução dos gráficos foi utilizada a biblioteca Laravel Charts 6.5.

O trabalho realizado vem sendo orientado pela metodologia Scrum que se caracteriza por ser adaptativa ao invés de preditiva (SOARES, 2004). Dessa forma, foi possível trabalhar em um cenário em que os requisitos poderiam sofrer pequenas modificações ao longo do processo de desenvolvimento do *software*. A metodologia Scrum baseia-se principalmente em três princípios: equipes pequenas, requisitos pouco estáveis e iterações curtas, divididas em *sprints*. O ciclo de vida Scrum possui três fases principais: (i) pré-planejamento, (ii) desenvolvimento, e (iii) pós-planejamento.

Na fase de pré-planejamento, foram realizadas pesquisas bibliográficas e entrevistas informais com profissionais da área, a fim de identificar os requisitos iniciais. Esses requisitos foram documentados e classificados de acordo com a sua importância e organizado em três unidades de trabalho: (i) dados da piscicultura, (ii) planejamento de produção e (iii) manejo. Após esse processo, os dez requisitos apresentados se encontram identificados no Quadro 1:

Quadro 1 - Lista de requisitos implementadas no software em questão.

Dados da Piscicultura	
REQ01	cadastrar novas pisciculturas
REQ02	cadastrar novos tanques de produção
Planejamento de Produção	
REQ03	realizar o planejamento da piscicultura - calcular e exibir a quantidade e o volume dos tanques necessários para atingir a produção desejada
REQ04	computar o escalonamento da produção - calcular e exibir a escala de produção, indicando a data de povoamento, a quantidade de peixes no povoamento e a data de pesca de cada tanque, considerando (i) o peso total da produção desejada, (ii) a periodicidade da produção desejada, (iii) quantidade e o volume dos tanques

	cadastrados, (iv) o peso médio dos peixes produzidos, (v) o tempo necessário para atingir o peso desejado
Manejo	
REQ05	registrar os parâmetros físicos e químicos da água medidos nos tanques
REQ06	registrar o povoamento dos tanques
REQ07	registrar a biometria dos peixes nos tanques
REQ08	visualizar a quantidade ideal de ração que deve ser ministrada em cada tanque
REQ09	visualizar gráficos com as informações de todos os parâmetros físicos e químicos da água nos tanques ao longo do tempo, bem como os valores de referência
REQ10	visualizar gráficos com dados do crescimento dos peixes ao longo do tempo

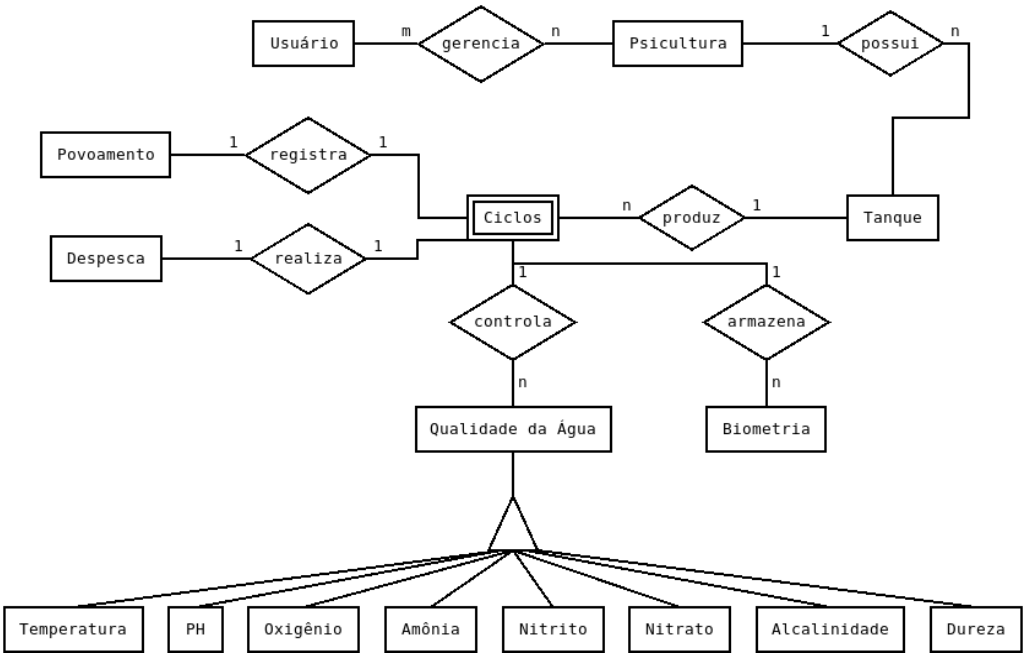
Fonte: dados da pesquisa.

Com os requisitos definidos, o modelo Entidade Relacionamento (Figura 1) foi projetado para modelar os dados gerenciados pelo sistema. Nele, percebe-se as relações entre as principais entidades do modelo construído, a saber:

- uma piscicultura possui **n** tanques;
- cada tanque produz **n** ciclos;
- cada ciclo de produção é caracterizado por um povoamento e uma despesca;
- em cada ciclo de produção existe o controle de **n** fatores de qualidade da água e **n** biometrias.

A fase de desenvolvimento foi organizada em 9 *sprints* com a duração de 15 dias cada. O *sprint* se iniciava com uma reunião de planejamento, que tinha por objetivo escolher as histórias de usuário que seriam implementadas e estimar o esforço da sua efetivação. Em seguida os requisitos eram implementados, testados e, caso fossem validados, seriam integrados à *release* em desenvolvimento. Por fim, ocorria a reunião de encerramento do *sprint*, com a participação de especialistas da área de piscicultura e de desenvolvimento de *software*. Nesta reunião, eram revisadas as histórias de usuário implementadas na *sprint*.

Figura 1 - Modelo Entidade Relacionamento do sistema proposto.



Fonte: dados da pesquisa.

Para a fase de pós-planejamento realizaremos a divulgação do *software* para comunidade acadêmica e produtores da região, através da publicação de artigos acadêmicos, realização de cursos de capacitação e produção de material didático em apostilas e vídeo tutoriais, que serão disponibilizados no site do projeto.

RESULTADOS E DISCUSSÃO

Ao acessar o PisciculTec pela primeira vez, o usuário deve realizar o seu cadastro e informar os dados de sua psicultura (**REQ01**). A partir deste momento, existe a possibilidade de cadastrar os tanques (**REQ02**), caso eles já estejam construídos, ou realizar o planejamento da psicultura (**REQ03**) para calcular a quantidade e volume dos tanques, de acordo com a produção pretendida.

O planejamento da psicultura (**REQ03**) é realizado informando o peso médio desejado para os indivíduos, a duração do ciclo previsto para atingir o peso esperado, a periodicidade da pesca e o peso total esperado, conforme exibido na Figura 2.

Figura 2 - Tela do PisciculTec em que o usuário insere os dados para o planejamento de uma nova piscicultura.



Fonte: dados da pesquisa.

Após a realização do cadastro dos tanques disponíveis, é dado ao piscicultor a opção de realizar o cálculo de escalonamento (**REQ04**) para garantir que ele tenha um fluxo contínuo de produção de acordo com os seus objetivos. O cálculo do escalonamento é feito a partir das informações dadas como entrada na tela exibida na Figura 3.

Figura 3 - Tela do PisciculTec em que o usuário insere dados para calcular o escalonamento da produção com o objetivo de garantir um fluxo contínuo de produção de acordo com os seus objetivos.



Fonte: dados da pesquisa.

Com base no peso de produção esperado e no peso médio do indivíduo será calculada a quantidade de peixes adultos esperada e, a partir desta quantidade, será aplicada a taxa de mortalidade média de forma reversa para chegar à quantidade de alevinos que deverá ser utilizada no povoamento. Para o cálculo das datas de povoamento de cada tanque, leva-se em consideração a duração do ciclo, isto é, a periodicidade de pesca desejada. O resultado de uma simulação de escalonamento pode ser observada na Figura 4. Esta simulação utilizou dados fictícios para um ciclo produtivo de 3 meses com periodicidade de produção mensal. Para o cálculo das datas consideramos um mês composto por 4 semanas, de modo que as ações de povoamento e despesca sempre ocorram no mesmo dia da semana definido pelo produtor. Observa-se ainda que a funcionalidade de escalonamento não possui ligação com as funcionalidades de fluxo de manejo, deixando o produtor livre para realizar diversas simulações de planejamento sem interferir nos dados de sua produção.

Figura 4 - Tela do PisciculTec em que o usuário visualizada o resultado do escalonamento da produção.



Fonte: dados da pesquisa.

O ciclo de manejo da produção é realizado de forma individual em cada tanque. Este ciclo inicia com o registro dos parâmetros físicos e químicos de qualidade da água (**REQ05**), tendo opções para registrar a temperatura, o pH, o nível de oxigênio, amônia, nitrito, nitrato, alcalinidade e dureza da água. Após o registro da primeira aferição dos parâmetros de qualidade da água no tanque, fica disponibilizada a opção de registrar o povoamento do tanque (**REQ06**), informando a data do povoamento, a quantidade de alevinos inseridos e o peso total. Com o povoamento devidamente registrado, o usuário tem disponibilizado a

possibilidade de registrar a biometria dos peixes do tanque (**REQ07**), informando a quantidade da amostra utilizada e o peso total dessa amostra. Também é possível visualizar quantidade ideal de ração que deve ser ministrada em cada tanque (**REQ08**), como exibido na Figura 5.

Figura 5 - Tela dos PisciculTec em que o usuário visualiza a quantidade de ração que deve ser ministrada em cada tanque.



Fonte: dados da pesquisa.

Por fim, o usuário também tem ao seu dispor gráficos com as informações de todos os parâmetros físicos e químicos da água (**REQ09**), conforme ilustrado na Figura 6. Além de poder acompanhar o crescimento dos peixes também de forma gráfica (**REQ10**).

Figura 6 - Tela do Piscicultec em que o usuário visualiza a variação do pH da água auferida ao longo do tempo em um tanque de produção.



Fonte: dados da pesquisa.

CONCLUSÕES

O presente trabalho apresentou o *software* PisciculTec que auxilia o planejamento e monitoramento de pisciculturas. Procurando focar na intuitividade e clareza das informações expostas, o sistema proposto oferece soluções para: (i) calcular o número de tanques necessários para produção; (ii) calcular a quantidade de peixes que deve ser utilizada para povoar cada tanque; (iii) calcular a quantidade de ração que deve ser ministrada em cada tanque; (iv) visualização dos dados de crescimento dos peixes ao longo do tempo; (v) visualização das variações de parâmetro de qualidade da água ao longo do tempo; e outros.

O PisciculTec está disponível de forma gratuita - e com o código aberto - a todos os piscicultores interessados. Espera-se, dessa forma, contribuir para o aumento da produtividade dos pequenos e médios produtores, sobretudo da região nordeste que, em sua maioria, não possuem recursos para investir em soluções tecnológicas e acabam por utilizar planilhas pouco intuitivas ou até lápis e papel para o simples registro de dados.

Como trabalhos futuros, pretendemos introduzir inteligência artificial no PisciculTec, visando um *feedback* mais completo para os produtores sobre o andamento e condições de suas criações. Também pretende-se realizar testes de usabilidade com os produtores locais, com o intuito de captar informações para possíveis melhoras no *software*.

REFERÊNCIAS

KUBITZA, F. Manejo na produção de peixes. Revista Panorama da Aquicultura, Rio de Janeiro. v. 18, n. 110, Novembro / Dezembro, 2009, p. 14-27. Disponível em: <http://www.acquaimagem.com.br/docs/Pan111_Kubitza.pdf>. Acesso em: 9 jul. 2019.

TAVARES-DIAS, M (org.). Manejo e sanidade de peixes em cultivo. Macapá: Embrapa Amapá, 2009. Disponível em: <http://pisceotec.com.br/arquivos/livro_manejo-peixes-cultivo.pdf>. Acesso em: 9 jul. 2019.

SCHULTER, E. P.; VIEIRA FILHO, J. E. R. Evolução da piscicultura no Brasil: Diagnóstico e desenvolvimento da cadeia produtiva de tilápia. Texto para Discussão. Brasília: Rio de Janeiro: Ipea , 1990-. ISSN 1415-4765. No. 2328. Disponível em: <http://repositorio.ipea.gov.br/bitstream/11058/8043/1/td_2328.pdf>. Acesso em: 9 jul. 2019.

SOARES, Michel dos Santos. Metodologias Ágeis Extreme Programming e Scrum para o Desenvolvimento de Software. Revista Eletrônica de Sistemas de Informação, [S.l.], v. 3, n. 1, june 2004. ISSN 1677-3071. Disponível em:

<http://www.periodicosibepes.org.br/index.php/reinfo/article/view/146/38>. Acesso em: 09 jul. 2019. doi:https://doi.org/10.21529/RESI.2004.0301006.

VIDAL, M. F. Panorama da piscicultura no nordeste. Caderno Setorial Escritório Técnico de Estudos Econômicos do Nordeste - ETENE, ano 1, n. 3, novembro, 2016. Disponível em: <https://www.bnb.gov.br/documents/80223/1218176/4_piscicultura.pdf/a281f37f-5929-ede0-8edc-041cc2f46742>. Acesso em: 9 jul. 2019.



Gerência de Dados de Processos de Genotipagem no GeneMaisLab

Eliseu Germano[1], *Luana Rodrigues*[1], *Thiago Souza*[1], *Tereza Borba*[1], *Marcelo Narciso*[1]

[1]Embrapa Arroz e Feijão

Santo Antônio de Goiás, Goiás, Brasil

eliseusilva@inf.ufg.br, {luana.rodrigues, thiago.souza, tereza, marcelo.narciso}@embrapa.br

RESUMO

Um ensaio de genotipagem consiste em um conjunto de informações de amostras e marcadores moleculares para realização de experimentos destinados ao melhoramento genético. O desafio de gerenciar os dados dos ensaios em uma base de dados estruturada e permitir a exploração deles a partir de uma interface acessível via serviços web, é o principal objetivo do GeneMaisLab. Um modelo arquitetural foi elaborado e uma implementação foi realizada. Neste artigo são apresentados os resultados do projeto GeneMais em termos de artefatos produzidos e validação da proposta de desenvolvimento de um software de gestão de dados de processos de genotipagem.

PALAVRAS-CHAVE: Gestão de dados genotípicos, Sistemas de Informação, Arquitetura Orientada a Serviços.

ABSTRACT

A genotyping assay consists of a set of molecular sample and marker information for experiments aimed at genetic improvement. The challenge of managing test data in a structured database and enabling it to be exploited from an accessible interface via web services is the main goal of GeneMaisLab. An architectural model was developed and an implementation was carried out. In this paper we present the results of the GeneMais project in terms of artifacts produced and validation of the proposal of development of data management software for genotyping processes.

KEYWORDS: Genotype data management, Information Systems, Service Oriented Architecture.

INTRODUÇÃO

Um processo de genotipagem consiste em um mecanismo no qual podem ser identificadas as variantes genéticas em indivíduos de uma espécie (KORDROSTAMI; RAHIMI, 2015). Esse mecanismo pode ser representado a partir de um ensaio de genotipagem, que consiste de um conjunto

de amostras e marcadores moleculares para realização de um experimento destinados ao melhoramento genético (HOSBINO et al., 2002). Assim, gerenciar dados de um ensaio consiste em representar os processos envolvidos desde uma etapa de elaboração de um experimento até o momento em que ocorre a obtenção dos resultados do mesmo.

Embora existam demandas para realizar sistematizações de processos de genotipagem, devido a diversos fatores, há uma complexidade em representar a estrutura desses processos. Uma das principais dificuldades em fazer essa representação está relacionada a dinamicidade dos elementos que podem compor um ensaio, como, marcadores moleculares, métodos de extração de DNA, representação dos genes analisados, entre outros.

Nesse sentido, o GeneMaisLab surgiu com o propósito de criar formas de padronizações de elementos de um ensaio que possuem uma menor variabilidade. De forma sintetizada, o GeneMaisLab consiste de um *software* para elaboração e acompanhamento de ensaios, em que os dados de genotipagem são divididos em subconjuntos de dados que possuem uma variabilidade mais baixa (e.g., plataformas de genotipagem). Assim, esses subconjuntos de dados podem ser tratados separadamente, e quando agrupados permitem representações customizadas de um ensaio.

Podendo ser acessado a partir da web, o GeneMaisLab oferece de forma centralizada em um servidor web, mecanismos para gerenciar um conjunto de dados do processo de genotipagem que atende a demanda de pesquisadores e analistas da Embrapa. De modo geral, os pesquisadores podem atuar no sistema como solicitantes de serviços que devem ser executados em uma determinada análise, enquanto analistas e colaboradores podem armazenar informações sobre o andamento dessa análise (e.g., o status de execução, os resultados obtidos, o local em que os experimentos estão sendo realizados).

Neste artigo são apresentadas as principais características do projeto Genemais sob uma perspectiva das funcionalidades do sistema GeneMaisLab e dos benefícios do sistema após o seu desenvolvimento. Dessa forma, são abordadas algumas tomadas de decisões que ocorreram no projeto tanto do ponto de vista computacional, quanto do ponto de vista do domínio no qual ele faz parte. Também são apresentados os resultados em termos de integração de dados do GeneMaisLab com outros sistemas da Embrapa (e.g., Alelo).

O artigo está organizado em outras três seções. A segunda seção apresenta traz os principais conceitos relacionados ao domínio do projeto Genemais. A terceira seção apresenta os resultados do projeto Genemais, incluindo a sua estrutura e funcionamento do GeneMaisLab. Por fim, a quarta seção apresenta a conclusão e trabalhos futuros.

MATERIAL E MÉTODOS

O GeneMaisLab é um sistema desenvolvido para viabilizar o gerenciamento de ensaios de genotipagem e marcadores moleculares na Embrapa. A partir dele é possível realizar uma melhor gestão das atividades relacionadas aos ensaios de genotipagem por meio da centralização dos

dados do processo de genotipagem, do armazenamento adequado desses dados em um servidor e da padronização de alguns desses dados diante de um mecanismo de gestão de padrões do sistema. Nesta seção é feita uma descrição do sistema (com ênfase nas motivações de desenvolvimento do projeto), os seus principais componentes de *software* produzidos e um conjunto de tecnologias envolvidas nessa produção.

Descrição do GeneMaisLab

Diante de um volume grande de informações genotípicas que são geradas continuamente por pesquisadores da Embrapa, surgiu a necessidade do desenvolvimento de um sistema de gestão de ensaios de genotipagem para auxiliar na gerência dessas informações. A ausência de um *software* específico para este fim, propiciava alguns impactos negativos nas atividades de pesquisa, tais como, perdas de dados por serem armazenados de forma inadequada, provocando retrabalho ou duplicidade de informações, perdas de dados por afastamento de pesquisadores, além da falta de padronização das informações coletadas nas análises laboratoriais.

Para lidar com esse cenário, foi desenvolvido um sistema chamado GeneMaisLab, no escopo do projeto chamado Genemais. Entre os vários objetivos desse sistema, o principal deles consiste em fazer a gestão dos dados de processos de genotipagem em um local único, centralizado, de forma estruturada e segura. Isso é feito no sistema por meio do gerenciamento dos dados de execução dos ensaios desde o planejamento até as etapas em que são feitas as considerações finais. Também é proposto no projeto que por meio do sistema seja viável o registro de eventos e o rastreamento das informações durante a execução dos ensaios.

Com o objetivo de sistematizar todo processo de gestão de dados relacionados aos ensaios de genotipagem e aos marcadores moleculares, o GeneMaisLab oferece uma interface *web* personalizada de acordo com as funções de cada perfil de funcionário ou colaborador da Embrapa. Os perfis de usuários definidos no sistema são:

- Solicitante – pesquisadores da Embrapa que são responsáveis por planejar o ensaio e as condições adequadas para a sua execução;
- Executor – analistas da Embrapa que são responsáveis por implantar, conduzir e registrar os resultados dos processos de execução dos ensaio;
- Colaborador – estagiários ou membros externos da empresa;
- Administrador – responsável gerenciar as funcionalidades do sistema e realizar o controle de acesso usuários ao mesmo.

Definição do Fluxo de Solicitação de Atividades dos Ensaios

O fluxo de atividades relacionados aos processos de genotipagem no GeneMaisLab variam de acordo com as funções de cada perfil de usuário no sistema. A Figura 1 apresenta o principal fluxo de atividades relacionado a execução de um ensaio no sistema. Esse fluxo consiste em quatro etapas que caracterizam os possíveis estados de um ensaio, sendo elas, planejamento, aceitação, em análise e executado.

Figura 1: Principal Fluxo de Atividades no GeneMaisLab



A primeira etapa (planejamento) é realizada por um usuário com perfil *solicitante* e as ações realizadas no sistema são relacionadas a elaboração do ensaio. Dessa forma, o solicitante preenche o formulário de cadastro de ensaios, solicita (caso necessário) o cadastro de novos padrões de dados para um gestor de padrões, importa dados de sistemas externos ao GeneMaisLab (a partir de componentes de *software* integrados a *web services* de outros sistemas), e por fim, cadastra o ensaio como um novo serviço laboratorial a ser realizado.

A segunda etapa (aceitação) é realizada por um usuário com perfil *executor* e as ações realizadas no sistema são relacionadas a avaliação do ensaio. Assim, o executor recebe as informações do ensaio como um novo serviço laboratorial solicitado (e.g., experimentos), avaliam as informações do ensaio (e.g., verifica se há recursos disponíveis para realizar os experimentos) e aceita ou rejeita a solicitação. Caso o serviço seja aceito, a terceira etapa (análise) consiste em executar as atividades relacionadas ao ensaio, adicionar observações durante o processo de execução e finalizar o processo anexando os resultados dos experimentos realizados.

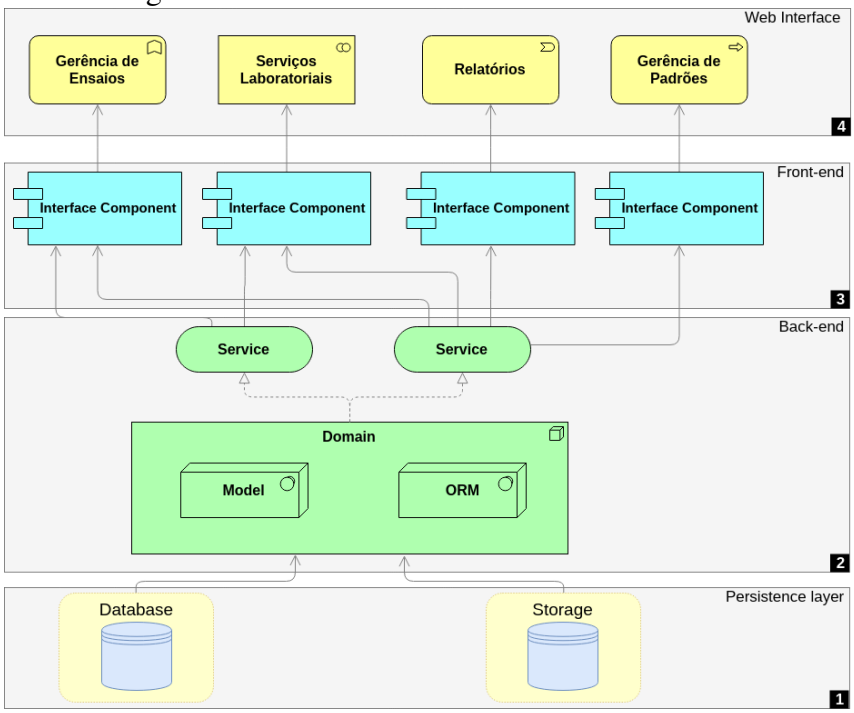
Por fim, a quarta e última etapa do fluxo (executado) está relacionada ao acesso aos ensaios cadastrados no sistema. Todos os ensaios podem ser listados ou consultados por quaisquer usuários do sistema. No entanto, apenas os ensaios públicos podem ser abertos pelos todos usuários (operação que permite acessar os detalhes de um ensaio, como, por exemplo, visualizar como os resultados de uma análise). Alguns ensaios podem ser cadastrados como restritos (nível de visibilidade privado). Esses últimos podem ser abertos apenas por usuários que fazer parte do grupo de pesquisa no qual o ensaio pertence. Tanto a escolha do grupo de pesquisa, quanto o nível de visibilidade do ensaio, é algo definido pelo solicitante durante a etapa de elaboração do ensaio.

Componentes do Sistema

Com acesso a partir de uma interface web, o GeneMaisLab possui um conjunto de elementos computacionais orientados a serviços que seguem o estilo arquitetural REST (FIELDING; TAYLOR, 2000). Esses elementos estão divididos em camadas de software de acordo com o que é apresentado na Figura 2. Essas camadas representam os níveis de abstração dos componentes.

Ainda na Figura 2, no nível mais baixo (nível 1) estão os componentes responsáveis pela persistência de dados e arquivos. No nível acima (nível 2) estão os componentes implementam as rotinas de *back-end* do sistema, ou seja, que fazem o gerenciamento dos recursos do nível

Figura 2: Modelo de Camadas do GeneMaisLab



1 e oferecem interfaces de alto nível para serem utilizadas, por exemplo, na implementação das regras de negócio dos serviços. Uma situação em que isso ocorre são nos componentes do nível 2 são os que fazem Mapeamento objeto-relacional (do inglês *Object-relational mapping* – ORM). Enquanto a camada de *back-end* provê os recursos implementados na forma de serviços, a camada de *front-end* (nível 3) consome esses recursos e apresenta os por meio de componentes de interface. Esses componentes são responsáveis por renderizar os dados do sistema em elementos visuais, que por sua vez, são apresentados no nível 1 a partir interfaces *web*.

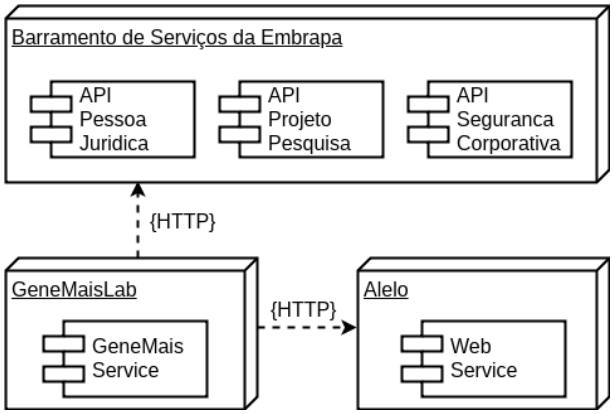
Devido a demanda do GeneMaisLab por dados de outros sistemas corporativos da Embrapa, foi necessário realizar algumas integrações entre esses sistemas. A Figura 3 representa a partir de uma diagrama da UML as principais integrações realizadas. As integrações foram realizadas a partir do Barramento de Serviços da Embrapa[1], exceto a integração realizada com o sistema Alelo Vegetal[2].

As APIs do Barramento de Serviços da Embrapa consumidas pelo GeneMaisLab foram:

- API Pessoa Jurídica – referente os serviços disponíveis para consumos das informações corporativas referente as unidades centrais e descentralizadas que compõe a estrutura organizacional da empresa, das pessoas físicas e jurídicas com vínculo com a Embrapa.
- API Projeto de Pesquisa – referente os serviços disponíveis para consumos das informações corporativas referente aos projetos de pesquisa juntamente com seus planos de ação e atividades.
- API Segurança Coorporativa – refere-se as formas de autenticação dos usuários aos

[1]https://sandbox-api.sede.embrapa.br/store/?tenant=embrapa.sede
[2]http://alelobag.cenargen.embrapa.br/AleloConsultas/Home/index.do

Figura 3: GeneMaisLab no Contexto de Sistemas da Embrapa



sistemas corporativos da Embrapa.

Em relação ao Alelo Vegetal, a integração foi realizada com o objetivo de obtenção de dados Genotípicos armazenados e categorizados no Banco ativo de germoplasma (BAG). O BAG consiste de um conjunto de dados (acessos) conservados a curto e médio prazos, geneticamente representativos de um grupo, família, espécie ou gênero, conservados, regenerados, caracterizados e disponibilizados para uso em pesquisas científicas e intercâmbio, sem a prática de descarte. Essas informações ficam armazenadas em um banco de dados relacional, que por sua vez, foi consultado pelo *web service* desenvolvido como um componente do Alelo.

As tecnologias utilizadas para implementar esses componentes de software estão relacionadas com a linguagem Java e com alguns de seus principais frameworks e bibliotecas. Dentre os utilizados estão o Spring Boot[3], Hibernate[4] e Google Web Toolkit Project (GWTP)[5]. Todos os componentes de foram integrados pela ferramenta de automação de compilação e gerência de dependências Maven[6] e são executados em *containers* da plataforma Docker[7]. Os detalhes computacionais relacionados ao desenvolvimento do GeneMaisLab podem ser encontrados em (GERMANO et al., 2017), (GERMANO; NARCISO, 2018b) e (GERMANO; NARCISO, 2018a).

RESULTADOS E DISCUSSÃO

Neste seção são apresentados os resultados da construção dos principais componentes do GeneMaisLab, sendo eles: Gestão de Ensaios, Gestão de Padrões, Gestão Laboratorial e Gestão do Conhecimento.

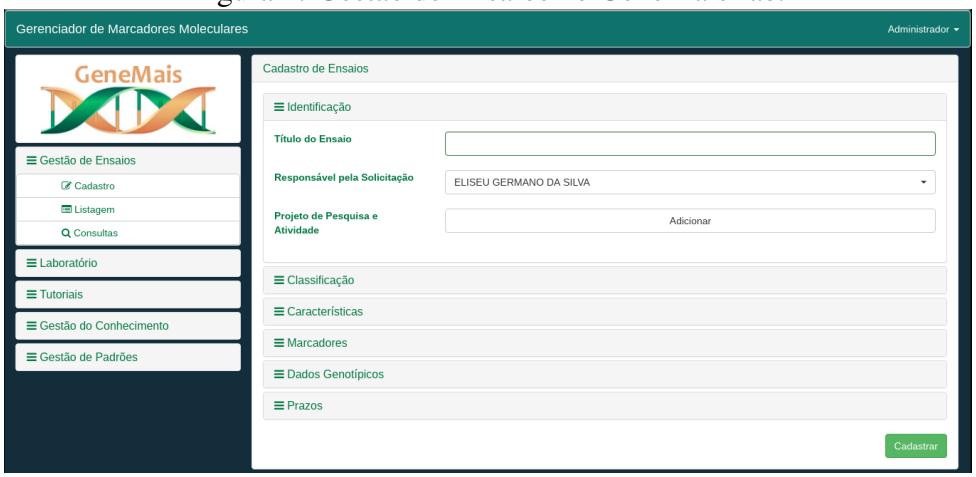
Resultados da Implementação da Gestão de ensaios

Componente responsável por disponibilizar um conjunto de serviços para manipulação dos ensaios de genotipagem, assim como das informações que o compõe, desde as formas de

[3]https://spring.io/projects/spring-boot
[4]https://hibernate.org/
[5]http://dev.arcbees.com/gwtp/
[6]http://maven.apache.org
[7]https://www.docker.com/

identificação do ensaio até suas características mais peculiares, como os dados da equipe de genotipagem, locais de execução, marcadores moleculares utilizados, os mapas de placas com seus conjuntos de amostras, entre outras. A Figura 4 apresenta a tela de elaboração dos ensaios.

Figura 4: Gestão de Ensaios no GeneMaisLab.



Durante a gestão de um ensaio, podem ser definidas informações de identificação de um ensaio (ex., funcionário responsável pela solicitação e o projeto de pesquisa relacionado) e informações administrativas (ex., prazo para entrada das amostras e prazo de entrega dos resultados). O software também permite que, durante o cadastro de um ensaio, sejam importadas informações de outro sistemas da Embrapa, como por exemplo, alguns detalhes de informações genotípicas que podem ser obtidos a partir do Sistema Alelo.

Resultados da Implementação da Gestão de Padrões

Componente que possibilita o gerenciamento dos padrões relacionados aos ensaios de genotipagem (Figura 5), viabilizando adição e modificação dinâmicas desses padrões, como, os métodos de extração de DNA (kit, CTAB, etc.), natureza da amostra (folha, semente, buque, etc.), as plataformas de genotipagem (Dart, Cornel, Ilumina, etc.), cultura (arroz, feijão, milho, etc.), caráter(seca, resistência à doença, etc.) entre outros.

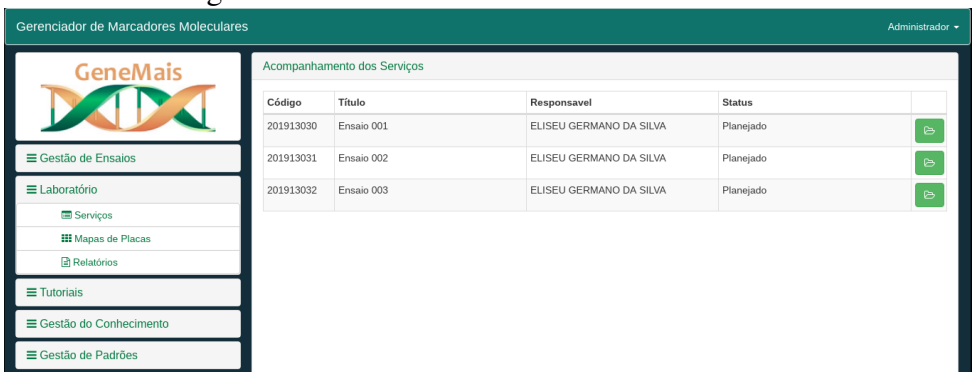
Figura 5: Gestão de Padrões no GeneMaisLab.



Resultados da Implementação da Gestão de Laboratorial

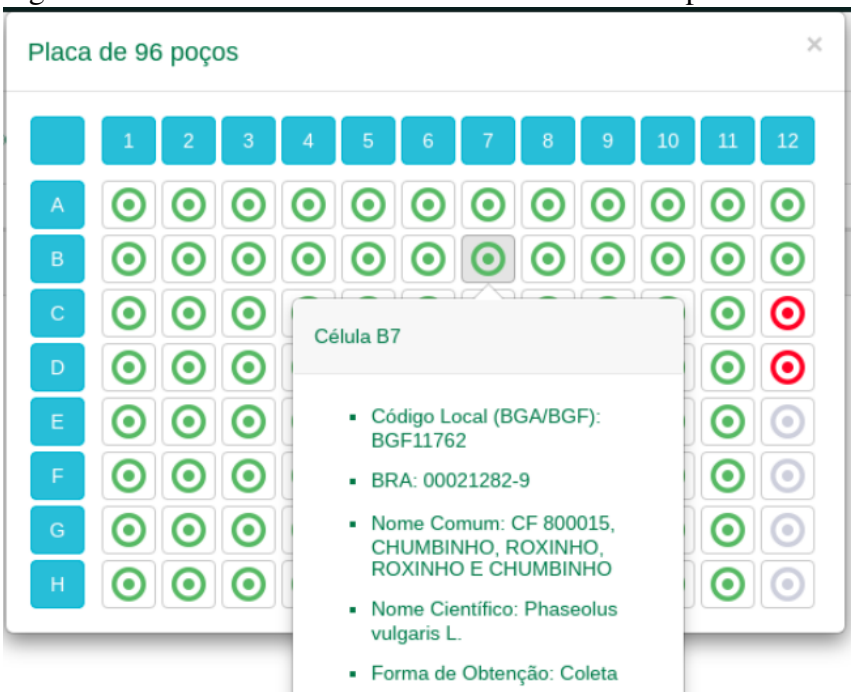
Uma vez cadastrado, um ensaio é colocado em uma lista de serviços (Figura 6) a ser realizado em um laboratório da Embrapa. Esses serviços podem ser acessados por usuários que realizaram a execução dos experimentos nos laboratórios (ex., analistas e colaboradores), que por sua vez, podem armazenar informações sobre o andamento de um ensaio, como o status de execução, os resultados obtidos e o local em que os experimentos estão sendo realizados.

Figura 6: Gestão Laboratorial no GeneMaisLab.



Diante da Gestão Laboratorial também é possível ter acesso a diferentes visões das informações de um ensaio, como por meio dos relatórios que sintetizam as informações dos ensaios de acordo com algum tipo de filtro escolhido, ou pela *interface* de um mapa de placas, que consiste em um desenho esquemático das placas de genotipagem de 96 ou 382 poços (Figura 7)). Assim o usuário poderá verificar a qualquer momento a disposição dos indivíduos, os indivíduos controles (testemunhas) e os marcadores que estarão sendo avaliados no ensaio.

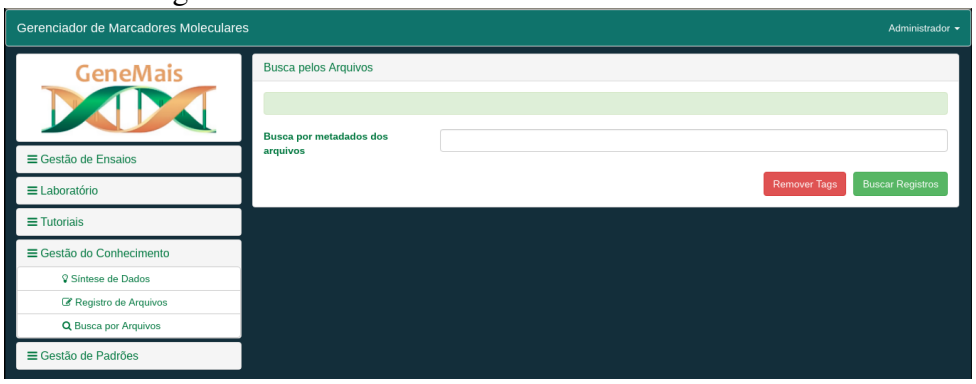
Figura 7: Gestão Laboratorial no GeneMaisLab – Mapa de Placas.



Resultados da Implementação da Gestão do conhecimento

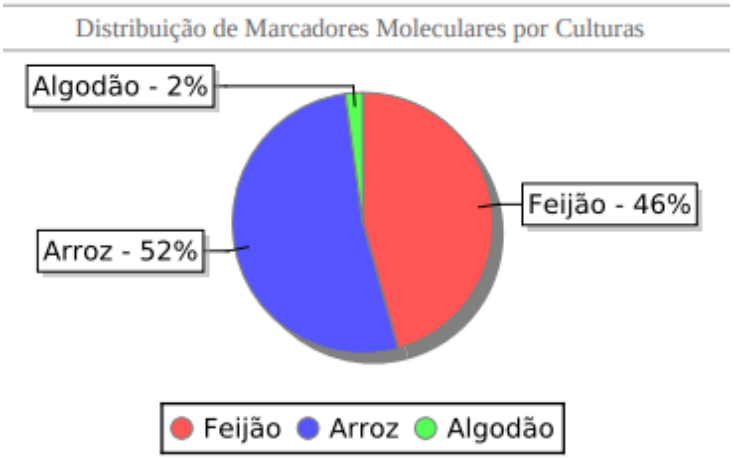
A gestão de conhecimento consiste de recursos do sistema que permite que durante a importação de arquivos no sistema sejam definidos um conjunto de metadados relacionados a esse arquivo. Esse metadados, por sua vez, são utilizados como elementos de busca em um elemento de interface para consulta aos arquivos Figura 8.

Figura 8: Gestão do Conhecimento no GeneMaisLab.



Outra funcionalidade da gestão de conhecimentos são os relatórios gerados a partir de macro visões dos dados registrados no banco de dados do GeneMaisLab. Na Figura 9, por exemplo, é representada a distribuição de marcadores moleculares por culturas. Nesse exemplo em específico, é possível visualizar de forma rápida alguns comportamentos dos dados registrados relacionado aos marcadores, como o fato de que a proporção de marcadores de Arroz e Feijão é consideravelmente maior que dos marcadores de Algodão.

Figura 9: Distribuição de Marcadores Moleculares por Culturas.



CONCLUSÕES

O GeneMaisLab é um sistema que foi desenvolvido para viabilizar o gerenciamento de ensaios de genotipagem e marcadores moleculares na Embrapa. Por meio dele é possível realizar

uma melhor gestão das atividades relacionadas aos ensaios de genotipagem viabilizada pela da centralização dos dados do processo de genotipagem, do armazenamento adequado desses dados em um servidor e da padronização de alguns desses dados em função de um mecanismo de gestão de padrões.

Os benefícios que o Genemaislab trará aos usuários estão relacionados a possibilidade de ter um maior controle sobre os dados produzidos durante os processos de Genotipagem. Esse controle pode ser obtido a partir do registro de eventos durante a execução dos ensaios, do rastreamento das informações que são produzidas durante a execução de um ensaio e da possibilidade de consultar os resultados produzidos após o encerramento de todas atividades de um ensaio. Dessa forma, os usuários envolvidos na realização de um ensaio (e nos serviços em torno dele) podem ter um melhor acompanhamento do andamento das atividades a partir do sistema e das notificações (emitidas por e-mail, por exemplo) que são encaminhadas. Além disso, o sistema viabiliza uma melhor comunicação entre pesquisadores e analistas (executores dos ensaios) a partir da troca de mensagens que envolvem os serviços que são realizados durante o processo de genotipagem.

Para a Embrapa o sistema oferece a partir de suas interfaces a possibilidade de acompanhamento das atividades que são realizadas entre pesquisadores e analistas. Dessa forma, é possível que seja verificado o andamento das tarefas, a execução delas diante dos prazos estabelecidos além do acompanhamento de custos que são gerados em cada um dos ensaios de acordo com os recursos que são necessários para execução dos mesmos.

REFERÊNCIAS

FIELDING, R. T.; TAYLOR, R. N. *Architectural styles and the design of network-based software architectures*. [S.l.]: University of California, Irvine Doctoral dissertation, 2000.

GERMANO, E.; NARCISO, M. Uma abordagem baseada em websocket para comunicação em tempo real no genemaislab. In: *IX EATI–Encontro Anual de Tecnologia da Informação*. [S.l.: s.n.], 2018. p. 110.

GERMANO, E.; NARCISO, M. G. Encapsulando um sistema gerenciador de ensaios de genotipagem e de marcadores moleculares em containers docker. 2018.

GERMANO, E. et al. Gene+: uma solução computacional distribuída para gerenciar ensaios de genotipagem e marcadores moleculares. *Embrapa Informática Agropecuária-Comunicado Técnico (INFOTECA-E)*, Campinas: Embrapa Informática Agropecuária, 2017., 2017.

HOSBINO, A. et al. Marcador microssatélite na conservação de germoplasma vegetal. *Biotec Ciência Desenvol*, v. 29, p. 146–150, 2002.

KORDROSTAMI, M.; RAHIMI, M. Molecular markers in plants: concepts and applications. *G3M*, v. 13, p. 4024–4031, 2015.

Sistema Imageador Infravermelho para mensuração do diâmetro e volume de pêssegos

Edson d'Avila[1], Antonio Carlos Loureiro Lino[2], Inácio Maria Dal Fabbro[3]; Ana Cristina da Silva[1]

[1] Departamento de Elétrica, Instituto Federal de São Paulo - IFSP, São Paulo, Brasil, davila.dre@gmail.com, cristina.silva.rane@gmail.com
[2] Centro de Engenharia e Automação do Instituto Agronômico de Campinas – CEA/IAC, Jundiaí, São Paulo, Brasil, acllino58@gmail.com
3 Faculdade de Engenharia Agrícola – Feagri, Universidade Estadual de Campinas - Unicamp, Campinas, Brasil, imdf14@gmail.com

Resumo

A fruticultura no Brasil vem gradativamente fortalecendo sua cadeia produtiva e implementando uma modernização nos padrões de classificação de frutas. O Brasil desponta como o terceiro maior produtor mundial de frutas. Atento ao potencial da fruticultura, o Ministério da Agricultura Pecuária e Abastecimento – MAPA, em 2017, lançou o plano nacional de desenvolvimento da fruticultura para ampliar a participação do Brasil no mercado exportador; aumentar o consumo interno e regulamentar toda cadeia de modo a garantir a qualidade. Levando em consideração este cenário, o presente trabalho de pesquisa implementou um sistema imageador infravermelho capaz de medir o volume e o diâmetro de um pêssego modelo com o uso de lasers de forma veloz e precisa. O sistema, por extensão, mensura o volume e o diâmetro de pêssegos e os classifica em consonância com as normas propostas em 2008 pelo programa brasileiro para modernização da horticultura e produção integrada de frutas – PBMH. O sistema proposto consiste, basicamente, de duas fontes emissoras de laser infravermelho estruturadas em linhas com comprimento de onda de 850 (nm); uma câmera digital com lente e filtro sintonizado em 850 (nm) para captura de imagens segmentadas; um goniômetro digital e um notebook acer com microprocessador Intel Core i7 e 16 GB de memória RAM. As linhas lasers são projetadas na superfície da fruta modelo em certo ângulo em relação ao eixo óptico do sensor de imagem. As linhas sofrem uma deformação pela superfície do modelo que, após análise de três imagens capturadas, se traduz em diâmetro da fruta com um erro percentual de 3,2% e no volume com um erro percentual de 2,5%, despendendo cerca de 0,1 segundo de processamento.

Palavras-chave: Frutas; Diagnóstico por imagens; Processamento de imagens; Frutas – tecnologia pós-colheita.

Abstract

Fruticulture in Brazil has been gradually strengthening the productive chain and implementing newer and more modern classification standards. Brazil is noteworthy as the third major producer of fruits in the world. Acting in accordance with the potential of the fruticulture, the Ministry of Agriculture, Livestock and Logistics, in 2017, launched the national plan for fruticulture development in an effort to amplify Brazil's reach in the world market; raise internal demand and regulate all the production chain to guarantee quality. In light of this scenario, the present research effort implemented an infrared imaging system capable of measuring volume and diameter of a model peach through the use of lasers in a fast and precise manner. The system, by extension, measures the volume and the diameter of peaches and is able to classify them in accordance with the proposed regulations of 2008 by the brazilian program of modernization of horticulture and integrated fruit production. The proposed system consists, basically, in two infrared laser-emitting sources structured in lines with a wavelength of 850(nm); a digital camera with lens and filters specifically tuned for 850(nm) to capture segmented images; a digital goniometer and a personal computer. The laser lines are projected in the surface of the model fruit in a given angle in relation to the optical sensor. The lines suffer a specific deformation according to the model's surface which, after the analysis of three captured images, can be translated in the fruit's diameter with a percental error of 3,2% and the fruit's volume with a percental error of 2,5%, taking approximately 0,1 second in this task.

Keywords: Fruits; Diagnosis by images; Image processing; Fruits - post-harvest technology.

Introdução

A importância da cultura do pessegueiro no Brasil pode ser avaliada através de dados estatísticos. Em 2017 o Brasil exportou o equivalente a U$ 22 milhões em pêssego. O Brasil ocupa a terceira posição na produção mundial de frutas com 44,1 milhões de toneladas por ano, precedido pela Índia com 71,7 milhões de toneladas por ano e China, em primeiro lugar, exibindo a cifra de 227,4 milhões de toneladas por ano (COSTA; SOUZA, 2018).

Não é de hoje que esforços vem sendo despendidos para especificar características de valor e de interesse para comercialização de frutas no Brasil. Em 1997 a Secretaria de Agricultura e Abastecimento do Estado de São Paulo lançou um programa que em 2000 se apresentou como o Programa Brasileiro para a Modernização da Horticultura – PBMH. Em 2008 foi lançada a norma para classificação de pêssegos, a exemplo de outras frutas. A norma, de particular interesse para a presente pesquisa, enfatizou seu principal objetivo como a classificação e a separação dos pêssegos em lotes visualmente homogêneos e com descrição clara através de características mensuráveis (PBMH, 2008).

A norma proposta, em 2008, pelo PBMH especifica o maior diâmetro transversal em milímetros para classificar o pêssego por tamanho. A Tabela 1 apresenta os diâmetros mínimos e

máximos para classificação do pêssego por faixas de calibre que vão de C_1 a C_8. O calibre 1, por exemplo, acolhe pêssegos com diâmetro maior que 25 (mm) até 35 (mm) inclusive.

Tabela 1. Classificação pelo diâmetro

Calibre	Diâmetro Mínimo (mm)	Faixa	Diâmetro Máximo (mm)
0	25	$< C_0 \leq$	35
1	35	$< C_1 \leq$	45
2	45	$< C_2 \leq$	51
3	51	$< C_3 \leq$	56
4	56	$< C_4 \leq$	61
5	61	$< C_5 \leq$	67
6	67	$< C_6 \leq$	73
7	73	$< C_7 \leq$	80
8	80	$< C_8$	–

Fonte: PBMH, 2008

Em 2017, o Ministério da Agricultura Pecuária e Abastecimento – MAPA, lançou o plano nacional de desenvolvimento da fruticultura que visa ampliar a participação da cadeia produtiva de frutas na pauta de exportações; melhorar condições e procedimentos para aumentar o consumo interno de frutas e regulamentar toda cadeia de modo a garantir a qualidade dos produtos estipulando requisitos mínimos, parâmetros de tolerância e a revogação de dispositivos e padrões defasados ou inexequíveis.

Neste momento em que se discute qualidade, padrões, parâmetros e tolerâncias, a presente pesquisa oferece a possibilidade, para além do diâmetro, de medir, também, o volume de forma veloz e precisa. Quem sabe o volume possa representar melhor o tamanho de um fruto do que seu maior diâmetro. Para Forbes (2000) o volume de frutas, em vez do diâmetro, permite que os supermercados exibam frutas em lotes de tamanho mais consistente. As embalagens para transporte são mais eficientes se projetadas pelo volume das frutas. Essa consistência é também uma demanda cada vez maior dos consumidores de hoje.

Atualmente os fabricantes de equipamentos para classificação de frutas buscam na visão de máquina o diferencial competitivo de apresentar máquinas mais precisas e velozes. Pesquisas recentes apontam que sistemas de classificação de frutas por análise de imagens podem ser precisos, velozes e que possuem baixa interação com as frutas. (LINO, 2008; FRACAROLLI, 2014; SANCHES et al., 2016; COSTA, 2018). É neste quadro em tela que a presente pesquisa está embasada, mas com o diferencial de estruturar a luz no comprimento de onda do infravermelho com formato em linha que está sintonizada com o sensor de imagem por intermédio de um filtro óptico. Com este sistema imageador obtém-se uma imagem em sua totalidade em um mesmo instante que está segmentada da cena em que se encontra. Por segmentação entende-se o processo de separar o objeto de interesse em uma cena (SOLOMON; BRECKON, 2013).

A captura de uma imagem digital de uma superfície 3D com uma luz estruturada, consiste, basicamente, em se projetar um feixe de luz com um formato bem conhecido sobre a cena. Conhecendo-se a estrutura da luz projetada e como a estrutura é deformada pela superfície do objeto, é possível se obter informações da superfície em estudo. O laser infravermelho é uma opção viável e desejável para compor um sistema de visão de máquina porque com grande potência concentrada em pequena área tem um único comprimento de onda preciso, fácil de se sintonizar com um filtro óptico e pode trabalhar simultaneamente com outros sistemas ópticos em uma mesa de classificação de frutas sem interferir e ou receber interferências de outros tipos de iluminação. A luz em forma de linha é uma das estruturas mais utilizadas em visão de máquina e pode ser facilmente obtida a partir de uma lente cilíndrica (BESL, 1989).

Reis e Jorge (2013) detalham um sistema que permite a reconstrução digital de objetos em 3D, proposto para aplicação em estudos de frutas em pós – colheita, utilizando três componentes principais, sendo o primeiro componente um sensor Kinect destinado a obter imagens de profundidade e imagens no padrão RGB, possibilitando a execução do algoritmo de reconstrução tridimensional. O segundo componente é representado por uma plataforma giratória controlada por um micro controlador. O terceiro componente é o software destinado a montar o objeto em 3D através das imagens obtidas pelo sensor Kinect.

Jorge e Biscegli (2007) relatam uma metodologia aplicada à geração de modelos tridimensionais (3D) de imagens de frutas, obtidas por ressonância magnética nuclear (RMN), com precisão na ordem de 97,7 a 99,85%, qualificando-o ao cálculo de volumes de frutas. As imagens foram adquiridas no equipamento de RMN Varian 2 Tesla. O algoritmo de reconstrução 3D utilizado neste trabalho é baseado na reconstrução 3D por meio da triangulação.

Um sistema de classificação de frutas é basicamente um sistema de medição capaz de medir o diâmetro, o volume ou qualquer outro atributo da fruta. Uma vez que o sistema de medição retorna um valor confiável e com uma determinada precisão, é possível aplicar uma determinada norma e fazer a classificação da fruta. Um sistema de medição pode ser definido como um conjunto de elementos, hardware e software, interligados que permite obter o valor de um mensurando, ou seja, uma grandeza que se pretende medir.

O sistema de medição busca encontrar o número de vezes que a unidade de medição está contida dentro do mensurando levando em consideração o meio, o procedimento e método, o agente que efetua a medição e a clara definição do mensurando (ALBERTAZZI; SOUSA, 2018).

O processo de desenvolvimento de um sistema de medição requer muitas medições. Quando se pretende mensurar uma fruta, um modelo pode substituir a fruta com vantagens. Frutas são perecíveis e impermanentes, ou seja, amadurecem rapidamente e os formatos são efêmeros e instáveis. O pêssego é uma fruta que amadure mesmo depois de colhida e em pouco tempo se observa a redução de seu volume (COSTA, 2018). Outro fator que dificulta a pesquisa com frutas *in natura* é a sazonalidade, ou seja, só estão disponíveis no mercado em certas épocas do ano. Para contornar as dificuldades de se trabalhar

com uma fruta *in natura*, pode-se modelar uma fruta em resina para preservar sua forma tridimensional com seus predicados próprios como volume e maior diâmetro transversal. A modelagem com resina de poliéster, por exemplo, guarda precisamente o formato da fruta para inúmeras medições ao longo do tempo. Um modelo permite, ainda, que se tenha um padrão que forneça a medida materializada de maneira permanente, como um bloco-padrão ou um material de referência. O emprego de modelo para substituir corpos vegetais em ensaios ópticos visa manter a geometria do corpo preservando sua forma 3D contra a deterioração para sucessivos ensaios.

O objetivo principal do trabalho, nesta fase, foi o de desenvolver a metodologia, bem como o *setup* experimental. É de interesse futuro levantar a geometria das diversas variedades de pêssego e de diversos tamanhos. Até a presente fase da pesquisa, o sistema proposto demonstrou-se adequado e capaz de atender a norma de para classificação de pêssegos, obtendo o maior diâmetro transversal do modelo de fruta com pequena margem de erro relativo percentual. Os materiais utilizados na construção do sistema imageador infravermelho foram encontrados comercialmente no mercado, isto põe, possivelmente, o uso do sistema ao alcance de pequenos, médios produtores e cooperativas para a classificação de suas frutas.

Apesar do pêssego modelo utilizado no levantamento dos dados ter sido obtido a partir de um pêssego chiripá, qualquer fruta com forma que guarde similitude, pode, a princípio com ajustes, ser mensurado pelo sistema imageador infravermelho.

A utilização de radiação infravermelho ofereceu duas notáveis vantagens durante os trabalhos, a segmentação da imagem para análise computacional foi grandemente facilitada e a medição do sistema imageador infravermelho não sofreu interferências da iluminação ambiente. Por sua vez, se o sistema imageador infravermelho trabalhar, paralelamente, com outras ferramentas ou máquinas, não sofrerá influência e não influenciará.

No que tange a segmentação de imagem, o filtro em frente ao sensor da câmera captura somente a linha laser refletida pelo objeto em cena, permitindo a captura de uma imagem que contém apenas o dado a ser analisado. Este procedimento reduz o esforço computacional de forma significativa.

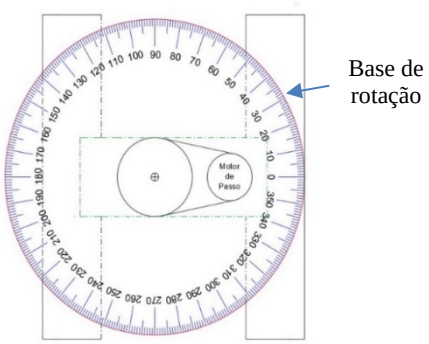
Material e Métodos

O sistema imageador infravermelho, objeto da presente pesquisa, foi projetado e implementado para classificar um pêssego modelo pelo diâmetro a partir de três imagens capturadas de diferentes vistas de sua superfície, espaçadas de 45º. O sistema imageador infravermelho também foi programado para medir o volume do pêssego modelo.

Um pêssego modelo foi elaborado a partir da modelagem com resina de poliéster, tendo como base um pêssego *in natura* do tipo chiripá obtido no mercado. Em seguida, o pêssego modelo foi fixado com plastilina sobre a base de rotação do goniômetro de posicionamento automatizado digital – GPAD, Figura 1. O eixo maior do pêssego modelo foi alinhado com o eixo de rotação da base de rotação. Um módulo eletrônico microcontrolado acionou o motor de passo para girar o pêssego modelo em torno do

seu eixo entre posições angulares precisas e sincronizadas com os disparos da câmera digital – Canon T5i equipada com uma lente 50mm, f/1.4 e um filtro sintonizado rosqueado na lente.

Figura 1. Goniômetro de posicionamento automatizado digital – GPAD



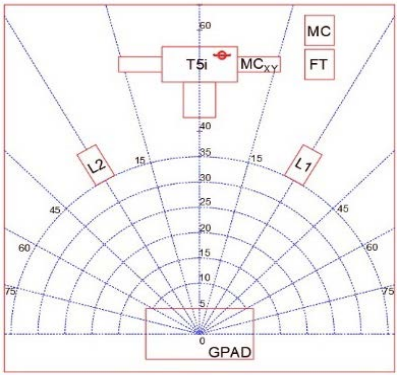
Fonte: Próprio autor

O módulo eletrônico foi programado para rotacionar o pêssego modelo em 180º com intervalos de 2,5 em 2,5 graus e disparar a T5i a cada intervalo, em um total de 72 imagens capturadas. Ao rotacionar o pêssego modelo em 180º obtém-se todas silhuetas de 2,5 em 2,5 graus, as demais, acima dos 180º, não são necessárias, pois são iguais e rebatidas.

Duas fontes lasers infravermelho foram montadas com diodos lasers de 850 (nm), telescópios com alargamento de feixe e convergência de foco, lentes cilíndricas e circuitos para controle da corrente elétrica nos diodos. As fontes lasers L1 e L2 foram posicionadas sobre as linhas de +30º e -30º e distantes 330 (mm) da origem “O”, onde está posicionado o GPAD.

As linhas foram ajustadas para se cruzarem na origem “O” da mesa óptica, Figura 2. Todas as medidas de distâncias e ângulos foram realizadas com a trena laser Bosch GLM 50 C com precisão de ± 1,5 (mm) e inclinações de 0 – 360º ± 0,2º.

Figura 2. Mesa óptica – disposição dos componentes do sistema imageador infravermelho



Fonte: Próprio autor

Uma macro escrita para o *ImageJ*, software livre de análise de imagens digitais, calculou o volume e o diâmetro a partir das áreas obtidas das três imagens capturadas com defasagens de 45º entre si. A_i é a área da i-ésima imagem e A_{med} é a área média calculada pela equação 1.

$$A_{med} = \frac{\sum_{i=1}^{3} A_i}{3} \quad (1)$$

Sabendo-se que existe um círculo cuja a área tem o valor de A_{med}. O raio do círculo pode ser calculado pela equação 2.

$$r = \sqrt{\frac{A_{med}}{\pi}} \quad (2)$$

Sabendo-se que uma esfera pode ser gerada pela revolução de um semicírculo de raio "r" em torno de um eixo, e que seu volume pode ser calculado pela equação 3.

$$V = \frac{4}{3} * \pi * r^3 \quad (3)$$

A linha laser projetada na superfície do pêssego modelo forma um ângulo de 30º com o centro óptico do sensor da câmera fotográfica. Para se obter a silhueta do pêssego modelo em escala 1:1 é necessário considerar que um ponto P(a,b) na superfície do pêssego modelo é projetado no plano β em co-seno de 60º. Logo, a imagem capturada pela câmera fotográfica foi reduzida pela metade no plano β.

Com um densímetro de sólido por diferença de massas com precisão – P de ± 0,13 cm^3 se mediu o volume do pêssego modelo. O valor obtido pelo densímetro foi assumido como o valor verdadeiro convencional do volume – VVC_{vol}. O VVC do mensurando é uma estimativa suficientemente próxima do valor verdadeiro do mensurando e que se assume como o valor verdadeiro.

Com um paquímetro digital Mitutoyo 150 (mm) x 0,01 (mm) se mediu o maior diâmetro transversal do pêssego modelo. O valor obtido pelo paquímetro, foi assumido como VVC_{dia}.

O VVC_{dia} e o VVC_{vol} foram, então, comparados com os valores obtidos pelo sistema imageador infravermelho na configuração para capturar e analisar três imagens 0º, 45º e 90º. O erro relativo percentual – E% foi calculado pela equação 4.

$$E_r\% = \left(\frac{VVC - V}{VVC}\right) * 100 \quad (4)$$

Resultados e Discussões

A Tabela 2 apresenta os cálculos do volume e do diâmetro do pêssego modelo, em várias situações, mas sempre utilizando três imagens espaçadas em ângulos de 45º. Se um pêssego gira em uma esteira de transporte, não é possível precisar em que ângulo será capturada a primeira imagem, considerado o ângulo inicial, mas a partir da primeira, a imagem seguinte será obtida com 45º de defasagem e a terceira com 90º de defasagem em relação a primeira.

Na primeira linha da Tabela 2 encontra-se o grupo de imagens 1 em que o volume e o diâmetro do pêssego modelo foram mensurados a partir das imagens capturadas nos ângulos 0º, 45º e 90º. Neste grupo, o diâmetro mensurado pelo sistema imageador infravermelho é de 56,2 (mm). Se comparado com

o valor de 58 (mm) obtido pelo paquímetro, o erro relativo percentual é de aproximadamente 3,2%. O valor do volume obtido pelo densímetro de sólidos foi de 95,11 (cm^3) e o valor mensurado pelo sistema imageador infravermelho foi de 92,717 (cm^3), com um erro relativo percentual de aproximadamente 2,5%.

Na segunda linha da Tabela 2 encontra-se o grupo de imagens 2 em que o volume e o diâmetro do pêssego modelo foram mensurados a partir das imagens capturadas nos ângulos 2,5º, 47,5º e 92,5º. Neste grupo, o erro relativo percentual é de 2,3% para o volume e 3,1% para o diâmetro. O mesmo raciocínio é seguido sucessivamente até a linha 36. Observa-se que o erro relativo percentual varia entre 2,5% e -1,6% para o volume e varia entre 3,2% e 1,8% para o diâmetro. Logo no pior dos casos o erro máximo não ultrapassa 2,5% para o volume e 3,2% para o diâmetro. O tempo gasto pela macro do *ImageJ* para calcular o volume e o diâmetro foi de 0,097 (s).

Pelas normas do PBMH (2008), o pêssego modelo foi classificado como uma fruta de calibre 4, ou seja, encontrar-se na classe que agrupa os pêssegos com diâmetro entre 56 e 61 (mm) considerando qualquer linha da Tabela 2.

Tabela 2. Diâmetro e volume a partir de três imagens espaçadas de 45º

Grupo de imagens	Intervalos entre imagens	Volume (cm^3)	Er_{vol} %	Diâmetro (cm)	Er_{dia} %
1	0, 45; 90	92,717	2,52	5,62	3,18
2	2,5; 47,5; 92,5	92,923	2,30	5,62	3,11
3	5; 50; 95	93,143	2,07	5,62	3,03
4	7,5; 52,5; 97,5	93,344	1,86	5,63	2,96
5	10; 55; 100	93,598	1,59	5,63	2,88
6	12,5; 57,5; 102,5	93,775	1,40	5,64	2,81
7	15; 60; 105	94,040	1,12	5,64	2,72
8	17,5; 62,5; 107,5	94,184	0,97	5,64	2,67
9	20; 65; 110	94,593	0,54	5,65	2,53
10	22,5; 67,5; 112,5	94,773	0,35	5,66	2,47
11	25; 70; 115	95,018	0,10	5,66	2,39
12	27,5; 72,5; 117,5	95,191	-0,08	5,66	2,33
13	30; 75; 120	95,296	-0,20	5,67	2,29
14	32,5; 77,5; 122,5	95,377	-0,28	5,67	2,26
15	35; 80; 125	95,481	-0,39	5,67	2,23
16	37,5; 82,5; 127,5	95,461	-0,37	5,67	2,24
17	40; 85; 130	95,571	-0,48	5,67	2,20
18	42,5; 87,5 132,5	95,564	-0,48	5,67	2,20
19	45; 90; 135	95,718	-0,64	5,68	2,15
20	47,5; 92,5; 137,5	95,702	-0,62	5,68	2,15
21	50; 95; 140	95,685	-0,60	5,67	2,16
22	52,5; 97,5; 142,5	95,720	-0,64	5,68	2,15
23	55; 100; 145	95,790	-0,71	5,68	2,12

24	57,5; 102,5; 147,5	95,847	-0,78	5,68	2,10
25	60; 105; 150	95,967	-0,90	5,68	2,06
26	62,5; 107,5; 152,5	96,001	-0,94	5,68	2,05
27	65; 110; 155	96,234	-1,18	5,69	1,97
28	67,5; 112,5; 157,5	96,320	-1,27	5,69	1,94
29	70; 115; 160	96,468	-1,43	5,69	1,89
30	72,5; 117,5; 162,5	96,544	-1,51	5,69	1,87
31	75; 120; 165	96,556	-1,52	5,69	1,86
32	77,5; 122,5; 167,5	96,597	-1,56	5,69	1,85
33	80; 125; 170	96,605	-1,57	5,69	1,85
34	82,5; 127,5; 172,5	96,545	-1,51	5,69	1,87
35	85; 130; 175	96,598	-1,56	5,69	1,85
36	87,5; 132,5; 177,5	96,544	-1,51	5,69	1,87

Conclusões

O sistema imageador infravermelho demonstrou-se adequado e capaz de atender a norma do PBMH (2008) para classificação de pêssegos por diâmetro, obtendo o maior diâmetro transversal do pêssego modelo com uma estreita margem de erro relativo percentual – $E_r\%$ = 3,2% e velocidade de aproximadamente 0,1 segundo.

O tempo de resposta do sistema imageador infravermelho, para retornar o volume e o diâmetro, foi de 0,097 segundo, uma velocidade já adequada para o trabalho em esteiras classificadoras. Ajustes na velocidade e na precisão do sistema imageador infravermelho bem como a otimização da macro do *ImageJ* a ser escrita em linguagem compilada, indicam a possibilidade de se obter tempos, ainda, bem menores para o uso prático do sistema imageador infravermelho.

Referências Bibliográficas

ALBERTAZZI, Armando; SOUSA, André Roberto de. *Fundamentos de metrologia científica e industrial*. 2ª ed. Barueri, SP: Manole, 2018. (segunda edição revisada, atualizada e ampliada).

BESL, Paul Joseph. Active optical range imaging sensor. In: SANZ, Jorge L. C. (Ed.). *Advances in machine vision*. 1st edition. New York: Springer-Verlag, 1989. p. 1-63. Springer Series in Perception Engineering, series editor: Ramesh Jain.

COSTA. Daniele Candido da. *Uso de imagens digitais na avaliação da qualidade pós-colheita de frutas de caroço*. 92 f. Dissertação (Mestrado Agricultura Tropical e Subtropical) – Instituto agronômico de Campinas, Programa de pós-graduação, agricultura tropical e subtropical, 2018.

COSTA, José Eduardo Brandão; SOUZA, Jorge Luís de. Cenários do mercado brasileiro e internacional de frutas e do mamão. In: *VII Simpósio do Papaya Brasileiro: Produção e Sustentabilidade Hídrica*, Vitória, ES, 22 a 25 de agosto de 2018. Disponível em: < http://papayabrasil.com.br/cd/dados/palestras/P1.pdf>. Acesso em 12 dez 2017.

FORBES, Keith. *Volume estimation of fruit from digital profile images*. 137f. Master of Science in Engineering (Electrical Engineering) – University of Cape Town, 2000. Disponível em: <

http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.58.8474&rep=rep1&type=pdf>. Acesso em 12 dez. 2017.

FRACAROLLI, Juliana Aparecida. *Levantamento das coordenadas espaciais de corpos vegetais associado ao mapeamento de bioatividade e cor*. 132 f. Tese (Doutorado em Engenharia Agrícola) – Feagri / Unicamp, Campinas, 2014. Disponível em: <http://repositorio.unicamp.br/bitstream/REPOSIP/257126/1/Fracarolli_JulianaAparecida_D.pdf>. Acesso em: 12 abr. 2018.

JORGE, Lúcio André de Castro; BISCEGLI, Clovis Isberto. *Reconstrução 3D, visualização e cálculo de volume de frutas*. São Carlos: Embrapa Instrumentação Agropecuária, 2007. Disponível em: < https://www.infoteca.cnptia.embrapa.br/bitstream/doc/30798/1/DOC312007.pdf>. Acesso em 12 dez. 2018.

LINO, Antônio Carlos Loureiro. *Aplicação de moiré de projeção com deslocamento de fase na construção de modelos digitais topográficos de órgãos vegetais*. 92 f. Tese (Doutorado em Engenharia Agrícola) – Feagri / Unicamp, Campinas, 2008. Disponível em: <http://repositorio.unicamp.br/handle/REPOSIP/256982>. Acesso em: 12 dez. 2017.

MAPA. *Plano nacional de desenvolvimento da fruticultura*. Brasília: Ministério da Agricultura, Pecuária e Abastecimento, 2017.
Disponível em: <http://www.agricultura.gov.br/noticias/mapa-lanca-plano-de-fruticultura-em-parceria-com-o-setor-privado/PlanoNacionaldeDesenvolvimentodaFruticulturaMapa.pdf>. Acesso em 5 maio 2018.

PBMH. *Normas de classificação de pêssego e nectarina*. São Paulo: CEAGESP, 2008. Programa brasileiro para a modernização da horticultura e produção integrada de frutas. (Documentos, 31).

REIS, Vitor Alexandre dos; JORGE, Lúcio André de Castro. *Reconstrução digital de objetos em 3D por nuvem de pontos utilizando o Kinect*. Centro Universitário UNISEB: *Revista Científica Eletrônica UNISEB*, Ribeirão Preto, v.1, n.2, p. 176 191, ag/dez. 2013.

SANCHES, Juliana et al. Classificação de pêssego douradão por calibre e cor através de imagens digitais. In: *XLV Congresso Brasileiro de Engenharia Agrícola*: CONBEA, 2016. Disponível em: < http://docplayer.com.br/41262814-Classificacao-de-pessego-douradao-por-calibre-e-cor-atraves-de-imagens-digitais.html>. Acesso em: 20 jan. 2017.

SOLOMON, Chris; BRECKON, Toby. *Fundamentos de processamento digital de imagens*: uma abordagem prática com exemplos em matlab. Rio de Janeiro: LTC, 2013.

Coletor automático de dados climáticos utilizando redes sem fio e o protocolo MQTT na transmissão das informações

Jonas Chalegra Boechat[1], *Otávio Mendonça Dantas*[1] , *Maurício Fernando Lima Pereira*[1]

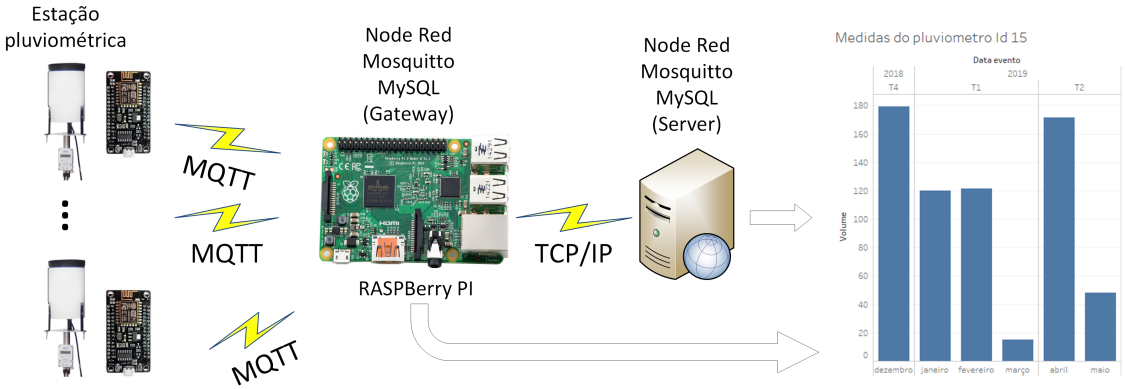
[1] Instituto de Computação - Univers. Federal de Mato Grosso (UFMT)
Av. Fernando Corrêa da Costa, nº 2367 - Bairro Boa Esperança. Cuiabá - MT
CEP 78060-900

jonas_boechat@hotmail.com, otaviomendoncad@gmail.com, mauricio@ic.ufmt.br

Um dos principais instrumentos para a quantificação de precipitação é o pluviógrafo (SENTELHAS; PEREIRA; ANGELOCCI, 2000), onde um dos tipos utilizados é o basculante, no qual uma báscula se movimenta toda vez que um certo volume é atingido. Nesse trabalho utilizou-se esse modelo de pluviógrafo como parte de um coletor de dados climáticos, com o intuito construir um equipamento de baixo consumo de energia, capaz de enviar informações remotamente e que permitisse através dos dados coletados gerar análises que facilitem as tarefas dos agricultores. O equipamento desenvolvido teve por base um sistema estruturado em três níveis (Figura 1). Na primeira camada, um pluviógrafo e outros sensores foram acoplados à um módulo NodeMCU, o qual possui um microcontrolador ESP8266 (MORAIS, 2017) responsável por capturar, armazenar e transmitir dados de precipitação via rede sem fio e usando o protocolo MQTT. Esse protocolo adota o padrão de mensagens *Publisher/Subscriber* e desta forma, os dados são publicados em um *broker*. Na camada intermediária, tem-se um *gateway* baseado no RaspBerry Pi 2 Model B, configurado com o sistema operacional Raspbian e com o *broker* Mosquitto, o serviço Node-RED e o banco de dados MariaDB instalados e configurados para receber dados dos pluviógrafos e persisti-los em memória não volátil, ainda no campo. O Mosquitto funciona como um intermediário, fazendo com que as informações geradas pelos publicadores, nesse caso as estações de coleta, cheguem aos nós do NodeRed. Detalhes mais técnicos a respeito do protocolo MQTT podem ser encontrados em (GREHS, 2016). A terceira camada é composta de um servidor, implementado nesse trabalho através de uma máquina virtual no VMWare, que executa o Ubuntu 18.04 e possui instalado as mesmas ferramentas e serviços instaladas no *gateway*. Para coleta de dados, uma placa de circuíto impresso foi desenvolvida para acoplar o pluviógrafo, o NodeMCU e um cartão de memória. O ESP8266 foi programado para ficar

no modo *DeepSleep*, no qual o sistema de captura e de transmissão ficam desligados em momentos em que não ocorre precipitação. Isso reduz consideravelmente os requisitos de energia demandado pelo sistema, pois nesse modo a corrente consumida pelo equipamento fica aproximadamente em 20μA enquanto não ocorrem precipitações. A cada movimento da báscula do pluviógrafo, gera-se um pulso elétrico indicando que um determinado volume de chuva foi coletado, esse pulso desperta o microcontrolador, o que faz com ele ative sua parte de transmissão, aumentado sua corrente, chegando a picos de 0,08A. Após alguns segundos, o equipamento retorna para o modo *Deep Sleep*. Os testes iniciais do sistema foram realizados em laboratório para validação de toda a arquitetura. Os equipamentos permaneceram consumindo pouca energia enquanto não ocorriam períodos de chuvas e despertaram corretamente quando eventos de oscilação da báscula ocorreram. No Node-Red foram desenvolvidos fluxos de dados que permitiram perceber a chegadas de notificações do *broker*, tendo sucesso na recepção de todos os dados enviados pelas estações. O fluxo projetado conseguiu manipular a informação aderindo nela a data em que foi postada e um número de identificação e salvando toda a informação no banco de dados. Enquanto todo esse processo ocorre, o equipamento toma as ações necessárias para voltar a adormecer e assim economizar energia. Através de uma página Web gerada com uso do Node-RED, o usuário tem a possibilidade visualizar diversos tipos de gráficos volume de chuva capturado por cada pluviógrafo. Destaca-se que essa característica, reduz a necessidade de ir ao campo obter os dados, o que agiliza a análise do cenário em campo.

Figura 1: Estrutura do projeto para coleta automática de dados de pluviosidade



REFERÊNCIAS

GREHS, D. H. *Sistema de irrigação doméstico baseado em internet das coisas*. Dissertação (Trabalho de conclusão de graduação) — Inst. de Informática - UFRGS, Porto Alegre, 2016.

MORAIS, J. *O QUE É ESP8266 - A FAMÍLIA ESP E O NODEMCU*. 2017. Disponível em: <https://portal.vidadesilicio.com.br/o-que-esp8266-nodemcu/>.

SENTELHAS, P. C.; PEREIRA, A. R.; ANGELOCCI, L. R. *Meteorologia agrícola*. [S.l.]: ESALQ-Depto de Física e Meteorologia, 2000.

Um Ambiente para Uso de Outorga Compartilhada de Água

Maria Fernanda Moura[1], Adauto Luiz Mancini [1], Raysa Masson Benatti[1], Lineu Neiva Rodrigues[2], Daniel Victoria[1], Luciana Alvim Romani[1], Stanley R Medeiros Oliveira[1], Leandro Eduardo Annibal Silva[1], Giovani Anhesini Bezerra[1], Gabriel Alves Francioli[1]

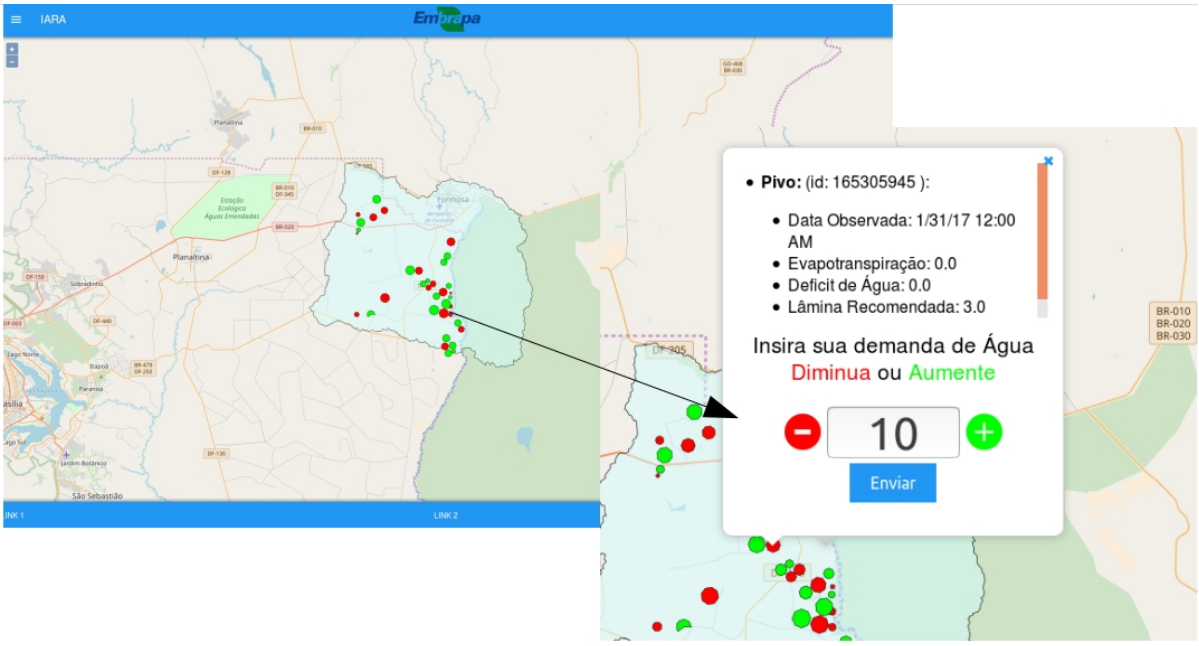
[1] Embrapa Informática Agropecuária, Campinas , SP, *maria-fernanda.moura@embrapa.br,adauto.mancini@embrapa.br, daniel.victoria@embrapa.br, luciana.romani@embrapa.br, colaboradores.cnptia@embrapa.br*

[2] Embrapa Cerrados, Planaltina, DF, *lineu.rodrigues@embrapa.br*

Atualmente vinte e quatro porcento do território brasileiro é ocupado pelo bioma Cerrado, que é a principal fronteira agrícola brasileira e também uma importante fonte de água para a região do semiárido brasileiro. Na região dos cerrados, cada vez mais se observa uma intensificação de conflitos pelo uso da água, principalmente conflito entre irrigantes e geradoras de energia hidrelétrica. Esse crescimento de conflitos pode ser claramente observado na bacia do Rio Preto. Visando a construção de ferramentas para apoiar o uso mais racional da água, entre proprietários que devem fazer o compartilhamento do recurso em pivôs de irrigação, em regime de outorga da microbacia, a Embrapa Cerrados e a Informática Agropecuária, juntamente à Agência Nacional de Águas (ANA), estão desenvolvendo o ambiente IARA - Irrigação: *Analytics* para uso Racional da Água - para auxiliar a minimização do conflito de uso da água na região do Rio Preto (GO).

No ambiente IARA, por meio de um aplicativo móvel de monitoramento de consumo hídrico, e de ferramentas computacionais para estimativa da oferta, da demanda e da área possível de ser irrigada em um dado momento, o agricultor é informado sobre a necessidade de irrigação de suas terras em cada um de seus pivôs. As ferramentas computacionais do IARA calculam a distribuição da água de uma forma otimizada para cada pivô em tempo real, favorecendo a melhor utilização do recurso hídrico para cada agricultor. O aplicativo móvel permite a visualização do uso da água na bacia junto à recomendação diária do uso em cada pivô (Figura 1). Quando o usuário/agricultor discordar da recomendação, pode alterar o valor da

lâmina de água efetivamente irrigada, uma vez que o sistema sugere, mas não obriga o agricultor a aceitar o valor proposto. É importante que ele informe essa alteração, a fim de que esse dado entre no histórico de uso da água na bacia. O aplicativo IARA ainda é um protótipo, porém suas funcionalidades já cobrem aproximadamente 60% do desejado. Ainda, as ferramentas de desenvolvimento utilizadas (Silva et al, 2018; Bezerra et al, 2018) facilitam sua manutenção e evolução para outras plataformas de uso.



Fi

RI

SI
M
M
A
26

BI
L.
se
A
10

Análise comparativa entre Sigfox, LoRa e NB-IoT

Cleiton Anderson Profilio dos Santos[1], *Eli Aparecido Gabilon Fernandes*[1], *Willyan Cândido Silva*[1], *Hana Karina Salles Rubinsztejn*[1]

[1]Faculdade de Computação (FACOM), Universidade Federal de Mato Grosso do Sul
Campo Grande, Mato Grosso do Sul, Brasil
cleitonandersonprofilio@hotmail.com, eligabilon@gmail.com, willyancs@gmail.com,
hana@facom.ufms.br

O objetivo deste trabalho é apresentar um comparativo entre três tecnologias de transmissão de longo alcance para Internet das Coisas (*Internet of Things* - IoT), classificadas como *Low Power Wide Area Network* (LPWAN). Analisando o contexto da agricultura observa-se a integração cada vez maior da tecnologia aos equipamentos e a necessidade de obtenção de dados em tempo hábil é fundamental no processo de tomada de decisão.

Analisando os trabalhos de Rubio-Aparicio et al. (2019), Pandey (2018), Teixeira e Almeida (2017), Ray (2018), Wang et al. (2017) e as informações encontradas no site oficial das instituições responsáveis pelas tecnologias, foi elaborada a Tabela 1 onde é apresentado um comparativo entre as tecnologias Sigfox, LoRa e NB-IoT, para isso foram definidos os seguintes parâmetros de comparação: Consumo de energia (P1); Custo de operação (P2); Alcance (P3); Tolerância à interferência e ruídos (P4); Tamanho dos pacotes (P5); Aplicações (P6); Topologia (P7); Vantagens (P8) e Desvantagens (P9).

Tabela 1: Comparativo entre as tecnologias LPWAN

Parâmetro	Sigfox	LoRa	NB-IoT
P1	Baixo	Baixo	Baixo
P2	Médio	Baixo	Médio
P3	30 - 50 km (campo aberto) 3 - 10 km (zona urbana)	5 - 15 (campo aberto) 2 - 5 (zona urbana)	35 km
P4	Alto	Alto	Baixo
P5	12	243	1500
P6	Número baixo de mensagens enviadas	Número baixo de mensagens enviadas	Baixa latência e comunicação frequente
P7	Estrela	Estrela	Estrela
P8	Baixo custo com infraestrutura Pacotes pequenos Baixo consumo de energia Longo alcance	Gratuita Divisão em classes Baixo consumo de energia Longo alcance	Baixo consumo de energia Longo alcance principalmente em ambientes resistentes a RF Baixa latência
P9	Aquisição de plano Downlink limitado Aquisição de componentes licenciados	Downlink limitado Aquisição de equipamentos em localidades sem cobertura	Mobilidade Interferência Área de cobertura limitada

Fonte: elaborada pelos autores.

Pode-se observar que a tecnologia Sigfox é mais indicada para aplicações que irão transmitir uma quantia pequena de mensagens diariamente, sua vantagem é o fato de que consome pouca energia e que não são necessários investimentos em infraestrutura de rede, sendo que o foco principal está na aplicação final, a operadora Sigfox (no Brasil a WND) oferece toda a infraestrutura, sendo necessário apenas contratar um plano desejado. Além do número limitado de mensagens diárias, a cobertura de sinal ainda é um fator limitante.

Já LoRa é indicada para lugares de difícil acesso, com pouco ou nenhum recurso elétrico fixo e de comunicações. Tem um consumo de energia baixo, podendo ser acoplado a uma bateria ou pilha, é recomendado para uso em campo aberto onde é necessário um longo alcance entre dispositivos, sendo ideal para área rural.

Projetado para atender a requisitos de desempenho da IoT, como extensão de cobertura significativa, baixa complexidade de dispositivos, longa vida útil da bateria e suporte a um grande número de dispositivos IoT, o NB-IoT permite fácil integração e compartilhamento de recursos de rádio com as redes GSM e LTE existentes. Sendo assim, as grandes operadoras que já investiram nessas tecnologias precisarão fazer pouco ou nenhum investimento para disponibilizarem serviços de IoT para seus clientes, já que essa tecnologia utiliza faixas ociosas das redes celulares.

REFERÊNCIAS

PANDEY, P. Comparative study of long range communications systems for iot-cellular, lora sigfox. 04 2018. Disponível em: <https://www.researchgate.net/publication/324557781_Comparative_study_of_long_range_communications_systems_for_IoT-Cellular_LoRA_Sigfox>. Acesso em: 15 jun. 2019.

RAY, B. *NB-IoT vs. LoRa vs. Sigfox*. Junho 2018. Disponível em: <https://www.link-labs.com/blog/nb-iot-vs-lora-vs-sigfox>. Acesso em: 14 jun. 2019.

RUBIO-APARICIO, J. et al. Design and implementation of a mixed iot lpwan network architecture. *Sensors*, v. 19, p. 675, 02 2019. Disponível em: <https://www.researchgate.net/publication/330948068_Design_and_Implementation_of_a_Mixed_IoT_LPWAN_Network_Architecture>. Acesso em: 08 jun. 2019.

TEIXEIRA, G. B.; ALMEIDA, J. V. P. de. *Rede LoRa e Protocolo LoRaWAN Aplicados na Agricultura de Precisão no Brasil*. 77 f. Monografia (bacharelado em Engenharia Eletrônica) — Universidade Tecnológica Federal do Paraná, Ponta Grossa, 2017. Disponível em: <http://repositorio.roca.utfpr.edu.br/jspui/bitstream/1/8436/1/PG_COELE_2017_2_01.pdf>. Acesso em: 15 jun. 2019.

Wang, Y. . E. et al. A primer on 3gpp narrowband internet of things. *IEEE Communications Magazine*, v. 55, n. 3, p. 117–123, March 2017. ISSN 0163-6804. Disponível em: <https://arxiv.org/ftp/arxiv/papers/1606/1606.04171.pdf>. Acesso em: 14 jun. 2019.

Desenvolvimento de metodologia para avaliar a homogeneidade de cobertura de aplicação em folhas de videira

Antonio Carlos Loureiro Lino[1], *Antonio Odair Santos*[1]

[1] Centro de engenharia Automação, Instituto Agronômico de Campinas, Jundiaí, São Paulo, Brasil, lino@iac.sp.gov.br, odairsan@iac.sp.gov.br

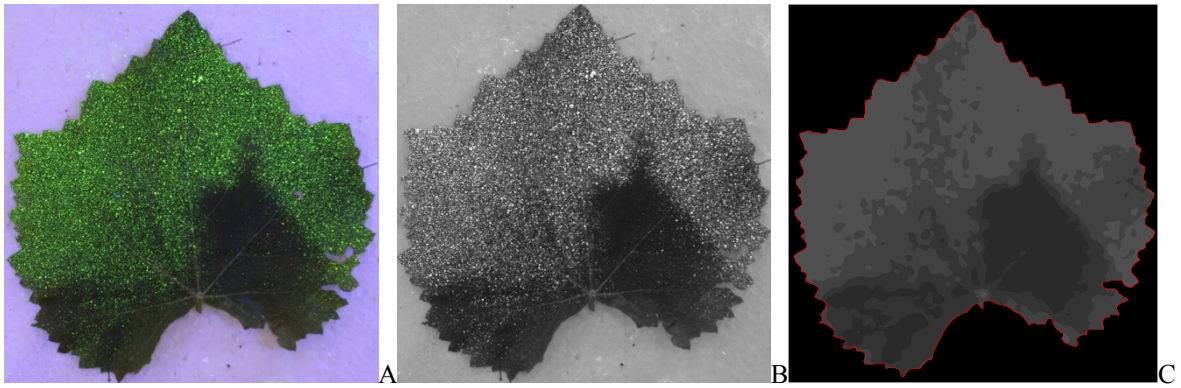
O objetivo deste trabalho foi desenvolver uma metodologia através do uso de tratamento e análise de imagens digitais, capaz de, não somente calcular a porcentagem de cobertura de aplicação de defensivos agrícolas em folhas de videira, como também indicar a uniformidade desta cobertura.

Os experimentos de campo estavam instalados na Escola Agrícola Benedito Storani, localizado em Jundiaí (SP), utilizando-se a cultivar Isabel precoce, estabelecida no espaçamento de 3 m x 2 m, conduzida em cordão bilateral, com sustentação em espaldeira. Na qual foi aplicado um volume de calda de 100 L/há, a a uma pressão de 45 Psi e velocidade de aplicação foi de 3,5 Km/h. Como marcador foi utilizado o corante fluorescente *Saturn Yellow*, na proporção a 0,15% na calda. Amostras foram coletadas nas porções interna e externa, dos terços superior, mediano e inferior das plantas, e então levadas ao laboratório e fixadas sobre uma placa de isopor branco para facilitar a separação das folhas e do "background", evitando interferências no momento das análises. Elas foram fotografadas em ambiente escuro, iluminado somente com luz ultravioleta, para evidenciar as gotas contendo o corante fluorescente. Utilizou-se uma câmera fotográfica digital, da marca Canon (modelo EOS Rebel T5), com lentes Canon EF, 50 mm, posicionada a uma distância de 40 cm. As imagens obtidas foram tratadas e analisadas no software ImageJ, para o qual foi desenvolvida uma rotina computacional (macro) com a finalidade automatizar estas operações.

As imagens (Figura 1ª) foram quebradas nos canais RGB. O canal B (Blue) foi utilizado para calcular a área da folha. Já o canal G (Green), Figura 1(B), que apresentou o melhor contraste entre gotas e folha, foi utilizado para medir a cobertura. O canal G foi ainda utilizado para

fazer a segmentação pela técnica K-means, resultando em uma imagem da folha dividida em 4 regiões chamadas “Clusters” (Figura 1C). Foi então calculada a cobertura, em porcentagem, de cada uma destas regiões. Para cada região, multiplicou-se a cobertura pela área (em percentagem) e calculado Coeficiente de Variação. Quanto menor o CV, mais homogênea é a cobertura de aplicação.

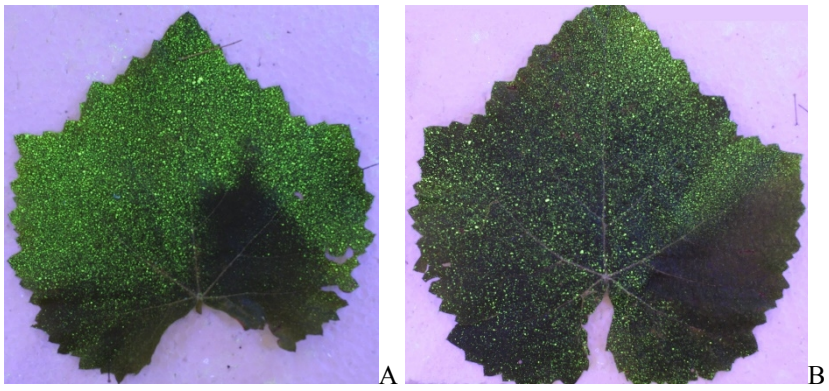
Figura 1 – Folha de videira. A - Original, B - Canal G (Green) e C - Folha segmentada em 4 regiões (clusters).



Fonte: (Próprio autor)

Como se pode observar na Figura 2, as regiões mais claras (verdes), são aquelas que tiveram melhor cobertura, pois é onde houve maior deposição do marcador “Saturn Yellow”. Já as regiões escuras tiveram pior cobertura de aplicação. Esta figura mostra duas folhas de videira com aproximadamente a mesma cobertura, 76,93% (Figura 2A), e 76,26% (Figura 2B). Porém a primeira mostra duas regiões bastante distintas, uma com alta e a outra com baixa cobertura (Figura 2A), enquanto na segunda folha, a cobertura foi mais homogeneamente distribuída. O CV, que foi de 55,91% e 31,75% respectivamente, mostrando que este índice é capaz de indicar a qualidade da cobertura em folhas de videira.

Figura 2 – Folhas com a mesma cobertura de aplicação.



Fonte: (Próprio autor)

A metodologia desenvolvida calculou a porcentagem de cobertura de aplicação em folhas de videira. E o CV, mostrou-se capaz de identificar automaticamente as folhas com cobertura mais uniforme.

CLIMMAPVIEW: PLATAFORMA DE DISSEMINAÇÃO DE DADOS DE PRECIPITAÇÃO PARA O ESTADO DO RIO DE JANEIRO

Gustavo Garcia dos Reis Nunes[1], Sidney Sampaio Alves[1], Gustavo Bastos Lyra[2], Ednaldo Oliveira dos Santos[2], Sergio Manuel Serra da Cruz[3]

[1] Discente do curso de Sistemas da Informação da UFRRJ, Bolsista do grupo PET-SI, Seropédica, Rio de Janeiro, gustavogarcia@pet-si.ufrrj.br, sidney@pet-si.ufrrj.br.

[2] Docente do Departamento de Ciências Ambientais/IF/UFRRJ, Seropédica, Rio de Janeiro, gblyra@gmail.com, edmeteoro@ufrrj.br.

[3] Docente do Departamento de Computação/ICE/UFRRJ, Seropédica, Rio de Janeiro, serra@pet-si.ufrrj.br.

Pesquisas recentes indicam crescimento demográfico mundial, com taxas superiores a 70 milhões de habitantes por ano (INED, 2013). Caso esse aumento continue, a população do Planeta alcançará os 9 bilhões de habitantes em 2050, com previsão de mais de 10 bilhões no final do século.

Portanto, é necessário focar no aumento da produção agrícola, para assim atender a demanda de alimentos da população mundial. Nesse cenário, destacam-se as mudanças e variabilidades climáticas como fatores que mais impactam o setor agrícola. Com isso, o objetivo deste trabalho é apresentar a plataforma *CLIMMAPVIEW (*Figura 1*),* baseada em Sistemas de Informações Geográficas (SIG) de dados abertos. A Plataforma disponibiliza a visualização dos metadados e dados de precipitação brutos e curados de estações pluviométricas para o estado do Rio de Janeiro por meio de uma interface *mobile,* em aplicativos disponíveis em *smartphone* ou em computadores por meio de navegadores, ambos conectados na *Internet*.

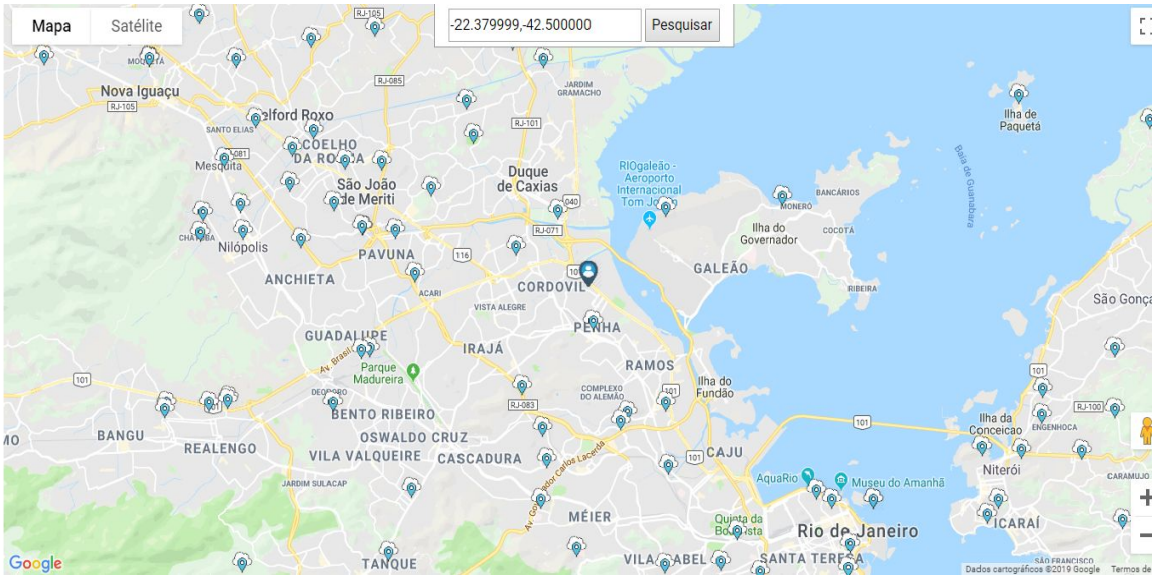


Figura 1: Sistema web de dados abertos, acessado por um navegador.

Para acessar os metadados e dados, a Plataforma recebe as informações da base de dados do sistema existente denominado METEORO (ANDRADE et al, 2015). Na codificação da plataforma, foram usadas as *Application Programming Interface* (APIs) de geolocalização do *Google Maps*. Já no modelo *web* foi usado *javascript* para manipular os dados obtidos pelo sistema METEORO e apresentá-los no mapa virtual de acordo com suas coordenadas, como também utilizou-se a *Hypertext Markup Language (HTML),* para que os navegadores possam interpretar as imagens dos mapas recebidas pela API.

Para desenvolvimento da versão *mobile*, foi usado o *framework* denominado de *react-native*. Ele interpreta códigos em *javascript* como também possui interpretação própria de códigos. Além disso, ele compila aplicações usando componentes nativos para dispositivos que suportam android e IOS. Vale ressaltar, que por ambos os meios de acesso a Plataforma, o sistema usa a geolocalização do dispositivo como meio de orientação de posição para apresentar o mapa.

Ao acessar o *CLIMMAPVIEW* é possível visualizar as estações no mapa indicadas por um marcador de cor azul com uma nuvem para simbolizar que no local marcado existe uma estação, com isso, ao clicar nesse marcador é aberta uma janela com as especificações técnicas da estação. Na versão web é possível realizar o *download* dos dados brutos e curados de acordo com uma data pré determinada pelo usuário, enquanto que na versão *mobile* só é possível visualizar os metadados. Outra funcionalidade é que há um *pin* de cor azul, que referencia a posição do usuário, pois ao clicar nesse *pin*, o Sistema redireciona para a local que se encontra o usuário, dando um zoom até a região.

A viabilidade da visualização das estações meteorológicas em vários dispositivos pode potencialmente demonstrar vantagem competitiva para diversos usuários, com o foco em pequenos e médios agricultores, que têm dificuldade de acesso a *softwares* mais avançados, visto que a única necessidade é ter uma conexão com a Internet, logo, poderão realizar análises e melhorar o seu planejamento agrícola.

REFERÊNCIAS

ANDRADE, L. S. et al. **Aperfeiçoamento do Sistema Meteoro para Gerenciamento de Dados Pluviométricos no Rio de Janeiro**. In: III Reunião Anual de Iniciação Científica da UFRRJ - III RAIC, Seropédica/RJ, 2015.

FAO **– Organização das Nações Unidas para Alimentação e Agricultura**. Disponível em:<http://www.fao.org/fileadmin/user_upload/faoweb/libson/docs/O_Pa%C3%ADs_25_7_2014.pdf /> . Acesso: 25 de julho de 2019.

INED - **Institut National D'étude Démographique.** Disponível em: < https://www.ined.fr/en/everything_about_population/graphs-maps/interpreted-graphs/growth_population_2050/ />. Acesso: 25 de julho de 2019.

Modelagem e simulação de um sistema de subirrigação para cana de açúcar

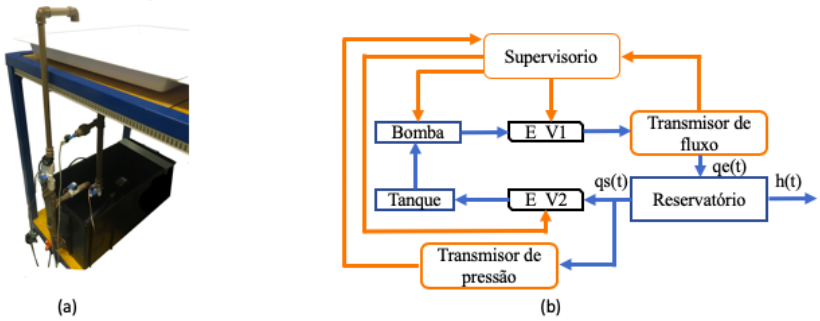
Angel Pontin Garcia[1], *Edna Carolina Moriones Polanía*[2], *Miguel Ángel Salgado Herrera*[1]

[1]Faculdade Engenheria Agrícola, Universidade Estadual de Campinas
Campinas, São Paulo, Brasil
angel.garcia@feagri.unicamp.br, m262340@g.unicamp.br

[2]Faculdade Engenheria Agrícola, Universidade Estadual de Campinas
Campinas, São Paulo, Brasil
e261468@dac.unicamp.br

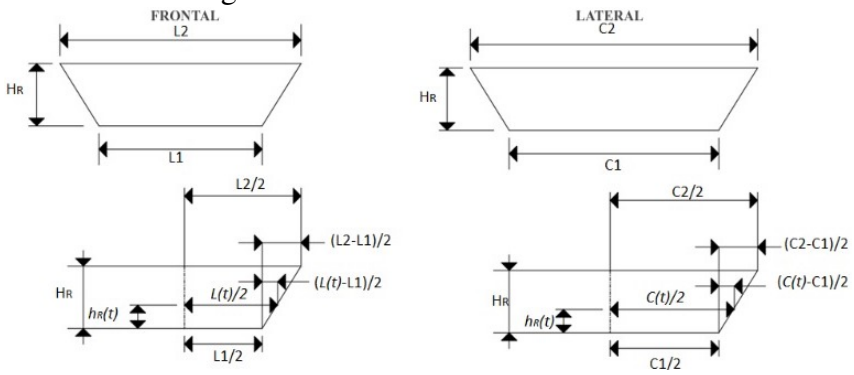
O Brasil é o maior produtor mundial de cana-de-açúcar, e uma das características desta cultura é a necessidade de renovação das plantações para se manter a produtividade do canavial (OECD; FAO, 2018). Uma estratégia que está em estudo para otimizar estas renovações é o sistema de mudas pré-brotadas (MPB) que podem utilizar o método de subirrigação na fase de germinação (FERRAREZI; IERSEL; TESTEZLAF, 2016). O nível da água no reservatório utilizado para o cultivo afeta diretamente a umidade do substrato em que a MPB está plantada (FILHO, 2012). Para a simulação do comportamento dinâmico do nível da água no reservatório das MPB se faz necessário desenvolver um modelo que relaciona o volume de entrada de água com o volume do reservatório (não linear). Dependendo do nível de água no reservatório, os tubetes modificarão proporcionalmente a variação da altura da coluna de água. Com isso, o objetivo deste trabalho é modelar e simular a dinâmica do nível de água no reservatório que é utilizado para a produção de MPB em um sistema de subirrigação. Posteriormente este estudo será utilizado como subsídio para o eleboração de uma estratégia de controle de umidade do substrato. A bancada experimental é apresentada na Figura 1(a). O sistema é composto por um reservatório que aloja as MPBs, sendo que este possui formato de pirâmide truncada invertida. A instrumentação do sistema é composta por eletroválvulas (V1,V2), um relé de acionamento da bomba e uma plataforma NiDAQ para a aquisição de dados. Para a supervisão e gravação dos dados desenvolveu-se um programa em LabView. O sistema a ser modelado é do tipo única entrada e única saída (do inglês, *Single input/Single output- SISO*) e é apresentado na Figura 1(b). A entrada é a vazão fornecida pela bomba (qe(t)) e saída é o nivel do reservatório (h(t)).

Figura 1: Bancada experimental e diagrama de bloques



A vazão do escoamento (qs(t)) está limitado pela resistência (R) do tubo. A equação de volume do reservatório é composta pela área da base S1, constante, e a área superior S2 que depende da altura e varia em função do tempo, Equação 1.

Figura 2: Dimensões do reservatório



$$v_R(t) = \frac{h_R(t)\left(s_1 + s_2(t) + \sqrt{s_1\ s_2(t)}\right)}{3} \tag{1}$$

$$S_2(t) = \frac{h_R^2(t)(L_2 - L_1)(C_2 - C_1) + H_R * h_R(t)(C_1\ L_2 - C_2\ L_1) + L_1\ C_1\ H_R^2}{H_R^2} \tag{2}$$

Com a seção transversal, obtemos uma relação entre a geometria das áreas e a altura, Figura 2. Substituindo a 2 em 1 obtem-se a equação do volume do reservatório, 3.

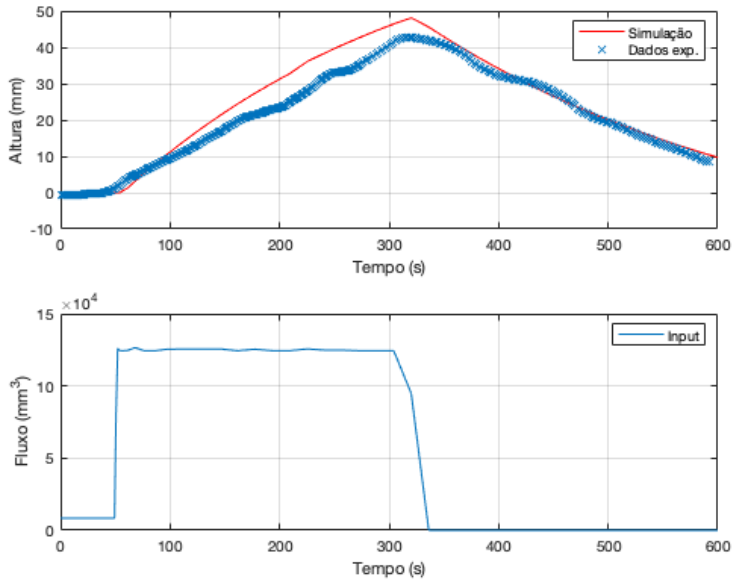
$$\begin{aligned} v_R(t) = 0,0419 \times h_R^3(t) + 13,8345 \times h_R^2(t) + 248640 \times h_R(t) + ... \\ 203,5682 \times \sqrt{0,1258 h_R^4(t) + 41,5034 h_R^3(t) + 372960 h_R^2(t)} \end{aligned} \tag{3}$$

Para o cálculo da altura foi excluído o volume do cone da MPB. Para se obter o modelo final se substituiu a 3 em 4. Para a simulação foi necessário aplicar séries de Taylor para linearizar o sistema e, posteriormente, se aplicou a transformada de Laplace ao modelo.

$$q_e(t) - q_s(t) = \frac{dv}{dt} \tag{4}$$

Para se simular o sistema utilizou-se: $[H_R = 50mm$, $S_1 = 372960mm^2$, $R = (h_R(t) + 16.866)/0.0008)]$. Como resultado se obteve que o tempo de esvaziamento do reservatório foi de $344s$ e o enchimento foi de $243,3s$, para um fluxo de entrada de $[12,5mm^3.s^{-1}]$, a simulação e os dados experimentais são apresentados na Figura 3.

Figura 3: Resposta do sistema com um fluxo de entrada variável



O modelo se mostrou adequado para simular o comportamento dinâmico do nível de água no reservatório de cultivo das MPBs. Os teste mostraram que o modelo se ajustou tanto ao enchimento quanto o esvaziamento do reservatório. Estes modelos serão base para as futuras simulações de desenvolvimento de estratégias de controle para a umidade do substrato que utilizam o método de subirrigação para o cultivo de cana-de-açúcar.

REFERÊNCIAS

FERRAREZI, R. S.; IERSEL, M. W. V.; TESTEZLAF, R. Plant growth response of subirrigated salvia 'Vista Red' to increasing water levels at two substrates. *Horticultura Brasileira*, v. 34, n. 2, p. 202–209, 2016.

FILHO, A. C. et al. X Congreso Latinoamericano y del Caribe de Ingeniería Agrícola e XLI Congresso Brasileiro de Engenharia Agrícola CLIA/CONBEA 2012 Londrina - PR, Brasil, 15 a 19 de julho 2012. p. 1–4, 2012.

OECD; FAO. *Perspectivas Agrícolas 2017-2026- Azúcar*. [S.l.]: Organisation for Economic Co-Operation and Development (OECD), ago. 2018.

Estudos para um controlador de código-aberto em imageador de custo acessível para fenotipagem de plantas

Patric Moreto[1], Luciano Vieira Koenigkan[2], Thiago Teixeira Santos[2]

[1] Bolsista Iniciação Científica, Embrapa Informática Agropecuária, Campinas, São Paulo, Brasil, ptmoreto@gmail.com

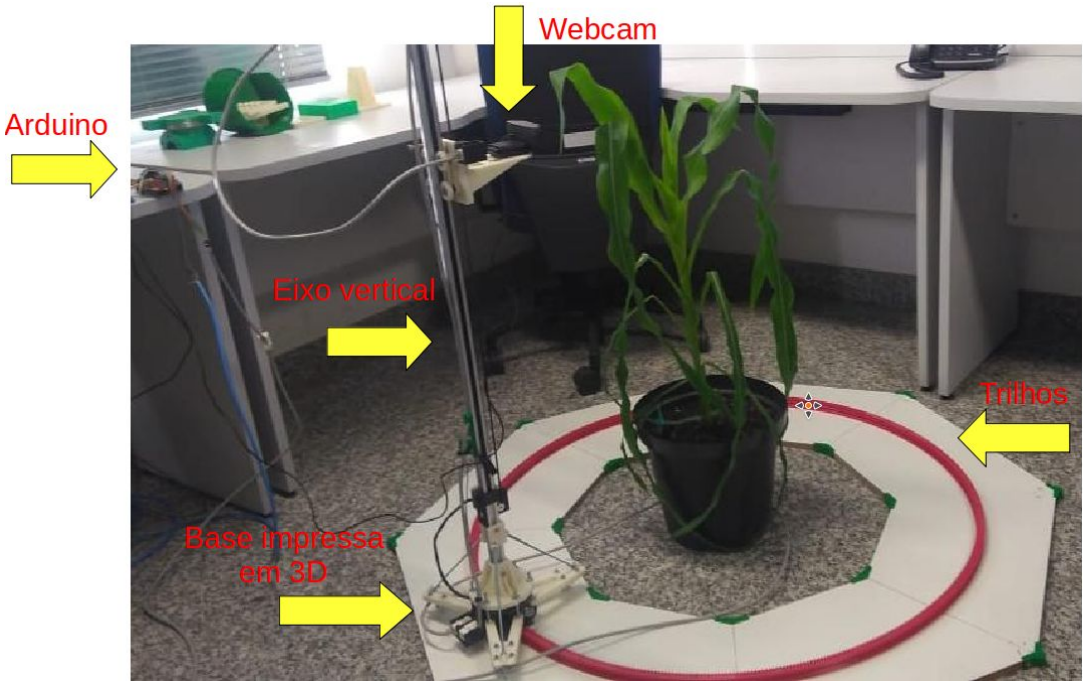
[2] Embrapa Informática Agropecuária, Campinas, São Paulo, Brasil, luciano.vieira@embrapa.br, thiago.santos@embrapa.br

A fenotipagem de plantas por imagem é uma metodologia para caracterização e avaliação do estado de plantas através de imagens capturadas de cada indivíduo em análise por um processo eficiente e não invasivo. Diversas aplicações são possíveis, como a estimação do volume de cachos de uva com alto índice de precisão (SANTOS et al., 2017) e o acompanhamento de medidas como altura e comprimento foliar em gramíneas como milho. Soluções escaláveis para fenotipagem por imagem podem apresentar um custo elevado, então estamos desenvolvendo uma solução baseada nos princípios encontrados nos movimentos "*open source*" e "*do it yourself*": custo acessível, participação de comunidades relacionadas ao projeto e possibilidades de modificações dependendo das necessidades do usuário.

Como parte integrante de um imageador eletromecânico com controle eletrônico, está sendo desenvolvido um *firmware* em linguagens Python e C capaz de movimentar uma câmera digital ao redor de plantas em vasos, visando a produção de modelos 3D das espécimes em análise. Nosso foco é o sistema de controle da movimentação do equipamento e a captação de imagens. Como visto em Tovar et al. (2018), o Raspberry Pi é uma solução comumente empregada em processos de controle como na captura/armazenamento de imagens. Ainda neste projeto, também é realizada uma integração com a placa controladora Arduino para o deslocamento de uma *webcam* em vários eixos. A Figura 1 é uma amostra do equipamento em funcionamento, na qual é possível identificar componentes advindos da

manufatura aditiva (impressão 3D) como os trilhos vermelhos e outros itens de valor acessível e amplamente disponíveis no mercado.

Figura 1 – Imageador de custo acessível.



Fonte: (Elaborada pelo autor, 2019)

Espera-se que o imageador auxilie no procedimento de fenotipagem, obtendo imagens de maneira automatizada, seja para a criação de modelos tridimensionais de cada planta, seja para a captação de imagens 2D para análise (GEHAN et al., 2017). Também é desejado que o modelo seja versátil para atender diferentes especificações de pesquisas na área agrícola (fisiologia e melhoramento vegetal e prospecção gênica, por exemplo).

REFERÊNCIAS

SANTOS, T. T. et al. Automatic grape bunch detection in vineyards based on affordable 3D phenotyping using a consumer webcam. In: CONGRESSO BRASILEIRO DE AGROINFORMÁTICA, 11., 2017, Campinas. *Anais*... Campinas: Editora da Unicamp: Embrapa Informática Agropecuária, 2017, p. 89-98.

TOVAR, J. C. et al. Raspberry Pi-powered imaging for plant phenotyping. *Applications in Plant Sciences*, v. 6, n. 3, 2018.

GEHAN, M. A. et al. PlantCV v2: Image analysis software for high-throughput plant phenotyping. *PeerJ 5*, 2017.



Reconhecimento de Bovino Pantaneiro utilizando visão computacional através da Rede Neural Convolucional: Resultados Premilinares

Fabricio de Lima Weber[1,2], *Vanessa Aparecida de Moraes Weber* [2], *Geazy Menezes*[3], *Adair da Silva Oliveira Junior*[3], *Daniela Arestides Alves*[1], *Marcus Vinicius Morais de Oliveira*[1], *Hemerson Pistori*[2,3], *Urbano Gomes Pinto de Abreu*[1,4]

[1]Programa de Pós-Graduação em Zootecnia, Universidade Estadual de Mato Grosso do Sul, Unidade Universitária de Aquidauana, Mato Grosso do Sul, Brazil
fabrico.weber@gmail.com, daniela.aalvesc@gmail.com,

[2]Universidade Católica Dom Busco
Campp Grande, Mato Grosso do Sul, Brasil
vmoraes@gmail.com, pistori@ucdb.br

[3]Faculdade de Computação, Universidade Federal do MS
Campp Grande, Mato Grosso do Sul, Brasil
geazyme01@gmail.com, adairsojr@gmail.com

[4]Embrapa Pantanal, Empresa Brasileira de Pesquisa Agropecuária
Corumbá, Mato Grosso do Sul, Brasil
urbano.abreu@embrapa.br

O bovino Pantaneiro é um *Bos taurus taurus*, também conhecido como Tucura, que foi introduzido na época da colonização no Bioma Pantanal. Por meio de seleção natural se adaptaram as peculiaridades regionais e adquiriram grande valor adaptativo. Atualmente, no Mato Grosso do Sul há dois principais núcleos de conservação *in situ* com trabalho direcionado a conservação da raça, são os núcleos de Conservação da fazenda Nhumirim (Embrapa Pantanal) e o Núcleo de Bovinos Pantaneiros de Aquidauana (NUBOPAN) da Universidade Estadual do Mato Grosso do Sul (UEMS). A partir de estudos zootécnicos os núcleos buscam, em parceria com produtores, realizar trabalhos de melhoramento genético animal e disponibilizar animais de genética superior para rebanhos do MS. Foi estimado que existem 8.319 indivíduos da raça no Pantanal, sendo categorizada como uma raça vulnerável. A fim de propor um método de identificação

não invasivo para o reconhecimento dos bovinos, este trabalho utilizou as Redes Neurais Convolucionais (CNN), que é uma rede específica para processamento e análise de imagens sendo assim, uma ferramenta da Visão Computacional. As imagens para formar esse conjunto de dados, foram coletadas junto ao NUBOPAN da UEMS. Foram utilizados 51 bovinos Pantaneiros de diversas idades, de ambos os sexos. Para a coleta das imagens o experimento contou com a instalação de um conjunto DVR:MD-1004NS MD-DVR41 da marca MIDI, câmeras com qualidade de imagem de AHD 720p e capacidade de gravação de 1Tb. Os vídeos coletados foram armazenados no DVR e posteriormente transferidos para computadores, em seguida, foram analisados e divididos em 212 vídeos menores onde ocorreu a extração dos quadros que continham imagens dos animais para formação do banco de imagens com 27.849 imagens dos bovinos. Assim, foi criado o banco de imagens RecBov51C em 90:10 para formar uma divisão interna de treinamento e teste. Foi selecionada a arquitetura Resnet50, como os pesos iniciais da ImageNet para execução da rede. A rede foi executada com fine-tuning de 100 e validação cruzada de 10 dobras, durante 50 épocas. Os resultados obtidos pela arquitetura Resnet50, no teste, foi de 99,5534% de acurácia e 0,1507% de desvio padrão. Os resultados mostram que modelos de arquiteturas utilizados na pesquisa alcançaram ótimos resultados de acerto. Os resultados indicam que o modelo pode apoiar pesquisadores e pecuaristas no reconhecimento de bovinos Pantaneiros. Também ficou evidenciado que o modelo de arquitetura de uma rede neural convolucional pode ser aplicado a um sistema de visão computacional para identificação dos animais seja feita automaticamente.

CLASSIFICAÇÃO DE ESPÉCIES DE PLANTAS ATRAVÉS DE REDES NEURAIS CONVOLUCIONAIS

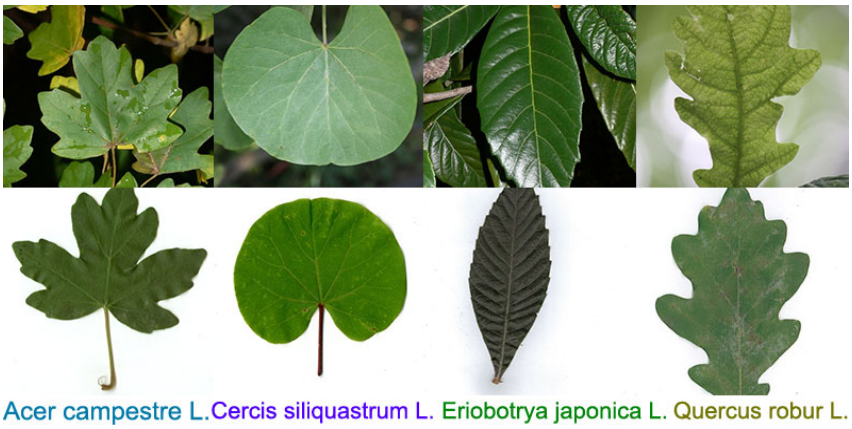
Luís Guilherme Ribeiro[1], *Renann Rodrigues da Silva*[1], *José Carlos Ferreira da Rocha*[1]

[1] Programa de Pós-Graduação em Computação Aplicada, Universidade Estadual de Ponta Grossa, Ponta Grossa, Paraná, Brasil, luisguicomp@gmail.com, renann_r@hotmail.com, jcarlos@uepg.br

De acordo com Christenhusz et al. (2016) existem 317.000 espécies diferentes de plantas, e a cada ano aumenta este número. Isso torna cada vez mais difícil a identificação de espécies vegetais, tendo que a identificação de plantas por chaves convencionais é complexa, demorada e devido ao uso de termos botânicos específicos, se torna frustrante para os não especialistas. Considerando isso, este trabalho tem como objetivo propor o uso de redes neurais convolucionais, do inglês *convolutional neural network* (CNN), para realizar a classificação de espécies de plantas. Os autores Matsugu et al. (2003) afirmam que uma CNN tende a demandar um nível mínimo de pré-processamento quando comparada a outros algoritmos de classificação de imagens, pois a rede aprende os filtros que em um algoritmo tradicional precisariam ser implementados manualmente. Essa independência de um conhecimento a priori e do esforço humano no desenvolvimento de suas funcionalidades básicas pode ser considerada a maior vantagem de sua aplicação.

Então neste trabalho foi utilizado um conjunto de 15.715 imagens (GOËAU et al., 2013), distribuídas em 100 diferentes espécies de plantas (Figura 1), este conjunto foi utilizado para realizar o treinamento da CNN. Além de treinar, a rede é capaz de extrair características que facilitam na identificação das espécies. Estas características foram utilizadas em um experimento que avalia o desempenho, em termos de acurácia, da CNN comparado com os seguintes classificadores: *Multi-Layer Perceptron* (MLP), Logistic Regression (LR), k-Nearest Neighbors (kNN), Support *Vector Machine* (SVM), e *Random Forest* (RF).

Figura 1 – Amostras de espécies de plantas com o fundo uniforme (cima) e natural (baixo).



A rede escolhida para extração de característica da CNN foi a *Inception V3*, esta é uma rede que extrai 2048 características de cada imagem. Então para realizar a análise comparativa, ainda foi necessário aplicar filtros para seleção de características e redução de dimensionalidade, para que os classificadores comuns possam obter melhores desempenhos. Para redução de dimensionalidade foi aplicado a técnica do *principal component analysis* (PCA), reduzindo de 2048 para 695 dimensões, em seguida foi aplicado a técnica de qui-quadrado para seleção dos 30 melhores atributos. Além disso foi aplicado o teste de *Friedman* com post-hoc de *wilcoxon,* considerando um nível de confiança de 95%. Na Tabela 1 é mostrado o desempenho de acurácia média de cada classificador e o *p-value* do pareamento comparado a CNN.

Tabela 1 – Desempenho em termos de acurácia e o *p-value* do teste de *wilcoxon* com relação a CNN.

	CNN	**MLP**	**SVM**	**LR**	**kNN**	**RF**
Acurácia (%)	60,35	24,11	19,15	16,3	16,1	10,82
p-value	---	0,00503	0,00503	0,00501	0,00489	0, 00475

Nos resultados (Tabela 1) é demonstrado que em termos de acurácia média a CNN obteve um desempenho significativamente superior aos demais classificadores, isto justifica a robustez dos filtros que a CNN utiliza para extrair características de imagens.

REFERÊNCIAS

CHRISTENHUSZ, J. M. et al. The number of known plants species in the world and its annual increase. Phytotaxa. v. 3, p. 201-217, 2016.

GOËAU, H. et al. The imageclef 2013 plant identification task. In CLEF. September, 2013.

MATSUGU, M. et al. Subject independent facial expression recognition with robust face detection using a convolutional neural network. Neural Networks, 16(5-6), 555-559, 2003.

MÓDULO SENSOR AUTÔNOMO PARA APLICAÇÃO DE LEITURAS COMPORTAMENTAIS DE CORPOS HÍDRICOS

Camila Cristina Gonçalves Teixeira Leite Moscibrocki[1], *Weliton da Maia*[1], *Luiz Alberto Bordignon* [2]

[1] Aluno(a) do 7º período do curso de Engenharia de Energia, Universidade Positivo, Curitiba, Paraná, Brasil, camilamoscibrocki@gmail.com, welitondamaia@hotmail.com

[2] Professor do curso de Engenharia da Computação, Engenharia Elétrica e Engenharia de Energia, Mestre em Engenharia Elétrica, Universidade Positivo, Curitiba, Paraná, Brasil luiz.bordignon@up.edu.br

O projeto em questão surgiu através da demanda de se obter informações do lago da Universidade Positivo uma vez que o mesmo apresentava problemas de infestações de algas, as quais foram controladas a fim de manter a qualidade aceitável de acordo com os níveis determinados pelo CONAMA (CONAMA, 2005). A partir do momento em que esses dados são compilados em parâmetros de análise, principalmente se as medições forem feitas em curtos períodos e com alta precisão, o manejo das providências acerca da qualidade do corpo hídrico é bastante facilitado a níveis demasiados (TOLMASQUIM, 2016).

O objetivo deste estudo é desenvolver um módulo que colete informações de pH, temperatura, umidade e oxigênio dissolvido. O processo contempla o desenvolvimento de um hardware capaz de coletar essas informações via sensores e mandá-las via comunicação LoRa de longa distância e internet a um banco de dados de uso livre, facilitando o acesso a essas informações. Em um primeiro momento, o desenvolvimento do módulo sensor foi baseado na pesquisa dos principais e mais comuns parâmetros para a determinação da qualidade dos corpos hídricos (ANA, 2018), os quais foram selecionados pela viabilidade de monitoramento em campo sem a necessidade do uso de laboratórios ou procedimentos que implicam análise extensiva.

Em um segundo momento, foi desenvolvido um módulo sensor autônomo com parâmetros de sensoriamento de pH, temperatura da água, temperatura do ar, umidade e oxigênio dissolvido, cumprindo 39% do índice de qualidade dado pela ANA (Agência Nacional de Águas). O módulo sensor é totalmente autônomo, possuindo baterias que são alimentadas por um painel

fotovoltaico. Esse módulo coleta essas informações em tempo real, envia via comunicação LoRa para um receptor conectado à internet, que por sua vez, envia as informações dos sensores para uma plataforma de acesso livre Ubidots. Fornecendo as informações em tempo real, assim como gráficos interativos e um histórico do comportamento desses dados.

Durante 10 dias foram feitos testes coletando dados do lago da Universidade, os quais foram obtidos sem interrupções. Os mesmos aparentam estar em conformidade com os testes feitos em laboratório para a calibração dos sensores (UBIDOTS,2019). O sistema de alimentação também se mostrou eficiente, suprindo a demanda de energia do sistema, e efetuando a carga na bateria.

O módulo funcionou como o planejado, fornecendo os dados do lago em testes efetuados em um período de 10 dias. A partir dos dados coletados de maneira constante nesse período é possível criar um histórico de comportamento destes parâmetros e, tendo conhecimento do comportamento destas variáveis no lago, dá-se a viabilidade de estudo do mesmo.

REFERÊNCIAS

ANA - AGÊNCIA NACIONAL DE ÁGUAS. Indicadores de qualidade da água – 2018a. Disponível em: <http://pnqa.ana.gov.br/indicadores-indice-aguas.aspx>. Acesso em: 20 de out. 2018.

BUZANELLO, E. B. Determinação de Coliformes Totais e Termotolerantes na Água do Lago Municipal de Cascavel, Porto Alegre - Revista Brasileira De Biociência, 2008. Disponível em: <http://www.ufrgs.br/seerbio/ojs/index.php/rbb/article/download/1087/811>. Acesso em 5 jul. 2019.

CONAMA. Resolução nº 20 Classificação de níveis de qualidade de águas. Jun. de 1986a. Disponível em: <www.mma.gov.br/port/conama/res/res86/res2086>. Acesso em: 20 out. 2018.

FUNASA. Manual prático de análise de água. Disponível em: <https://repositorio.ufsc.br/bitstream/handle/123456789/182173/TCC2%20%28corrigido%29.pdf?sequence=1&isAllowed=y>.. Acesso em 11 nov. 2018.

FURTADO, E. W. R., Sensor Eletrônico Para Análise De Qualidade Da Água. Araranguá: UFSC, 2017. Disponível em:< https://repositorio.ufsc.br/bitstream/handle/123456789/182173/TCC2%20%28corrigido%29.pdf?sequence=1&isAllowed=y > Acesso em 06 jul. 2019.

TOLMASQUIM, Mauricio T. EPE - Empresa de Pesquisa Energéticas. Energia renovável: Hidráulica, Biomassa, Eólica, Solar e Oceânica. v.1. Ano de publicação: 2016.

UBIDOTS. Ubidots for Education. Disponível em: https://ubidots.com/education/. Acesso em 20 de out. 2018.

Interdisciplinaridades e aplicações didáticas com uso de maquetes voltadas ao Agronegócio

Reinaldo Toso Júnior[1], Wellington Roque[2], Luis Borges Gouveia[3]

[1] Doutorando em Projetos, RVOE SEDUC D-110499 , Universidad Internacional Iberoamericana (UNINI), Campeche, Campeche, México, reitoso@yahoo.com.br,

[2] Curso Superior de Tecnologia em Redes de Computadores, Fatec Indaiatuba "Dr Archimedes Lammoglia" – Centro Paula Souza, Indaiatuba, São Paulo, Brasil, wellington.roque@gmail.com.

[3] Universidade Fernando Pessoa, Porto, Portugal, lmbg@ufp.edu.pt.

Com a igualdade tecnológica e o advento da sociedade em rede a competição entre os marcados deixou de ser apenas de produtos, sejam acabados, intermediários ou primários, a competição se dá por meio da cadeia de suprimentos mais desenvolvida e está e uma geradora de vantagens competitivas. A complexidade das cadeias exige uma grande capacidade de integração ou visão sistêmica, reunindo várias áreas de conhecimento em camadas que não se sobrepõem e sim se integram. Um dos desafios no ensino tecnológico é desenvolver nos jovens adultos esta habilidade. Uma das áreas de maior desafio logístico é o do agronegócio, o Brasil tem se apresentado ao longo das décadas um país capaz de atender as demandas mundiais por commodities agrícolas tanto em qualidade como em quantidade, situando-se como terceiro maior exportador do mundo (Estadão Conteúdo, Globo Rural, 2018), ficando atrás apenas dos Estados Unidos e União Europeia. Nas atividades primárias da logística (transportes, armazenamento e processamento de pedidos) a armazenagem de grãos é um item que exige grande atenção pois a produção obedece aos ciclos sazonais e os embarques no modal aquaviário necessitam de pré-armazenagem em vários pontos da cadeia de suprimentos para a consolidação e assim minimizar os custos de transportes. Estas armazenagens requerem controles ambientais apurados dos silos não só para questões sanitárias e microbiológicas como também para a segurança das operações referente aos riscos para os trabalhadores com relação ao desmoronamento de massa, fogo, explosão e demais precauções conforme as normas regulamentadoras NRs 6, 7, 9, 12, 18, 31, e 33 (Secretaria de Inspeção do Trabalho - SIT, 2018). Por essa razão justiça-se despertar a atenção dos alunos sobre o tema por meio de um instrumento didático com relativo desafio acadêmico que foi a proposta da construção de uma maquete de um silo com ênfase em aspectos de controle e segurança. Para tal atividade

envolveram-se alunos do Curso Superior de Tecnologia de Redes de Computadores e de Logística Aeroportuária, não integrando apenas a interdisciplinaridade de disciplinas como também de cursos distintos. Os requisitos de controle do silo foi estabelecer um sistema de monitoramento que simulasse controles de umidade, temperatura e gás metano (Araújo, 2014). Para tal solução além da construção de um modelo que imitasse o silo, engenhosamente feito com um cesto plástico de 60 litros, os alunos utilizaram uma placa Arduino, que opera em baixa tensão e possui várias entradas e saídas lógicas e permitiu a inserção dos sensores de monitoramento que foram integrados ao silo, um dispositivo sonoro alertava a presença de gás metano, cuja geração de gás era imitada por meio do acionamento de um isqueiro, para simular elevações da temperatura utilizou-se um soprador de ar quente doméstico concomitante com um processo de transmissão sem fio dos dados para monitoramento via software. Com esta atividade os alunos puderam explorar tanto as características e cuidados com a movimentação e armazenagem de grãos como também entenderam a importância da tecnologia aplicada. O uso de instrumentos simples de ensino não diminuiu o desafio acadêmico proporcionando a aproximação de alguns aspectos práticos e as disciplinas de ambos os cursos ocorrendo a colaboração entre alunos de logística e redes de computadores.

REFERÊNCIAS

Araújo, Giovani Moraes de. 2014. *Normas Regulamentadoras Comentadas.* [ed.] Gerenciamento Verde Consultoria. 8. Rio de Janeiro : Gerenciamento Verde Editora, 2014. Vol. 3. 9788599331064.

Estadão Conteúdo, Globo Rural. 2018. Brasil passa a ser 3º maior exportador agrícola, mas clima ameaça futuro. *Revista Globo Rural.* [Online] Editora Globo S/A, 17 de setembro de 2018. [Citado em: 29 de Julho de 2019.] https://revistagloborural.globo.com/Noticias/Agricultura/noticia/2018/09/globo-rural-brasil-passa-a-ser-3o-maior-exportador-agricola-mas-clima-ameaca-futuro.html. Agricultura, 17 de Setembro de 2018.

Secretaria de Inspeção do Trabalho - SIT. 2018. SEGURANÇA E SAÚDE NO TRABALHO (SST) , NORMATIZAÇÃO, NR (Normas Regulamentadoras). *Escola Nacional da Inspeção do Trabalho.* [Online] Escola Nacional da Inspeção do Trabalho, 18 de dezembro de 2018. [Citado em: 29 de julho de 2019.] https://enit.trabalho.gov.br/portal/index.php/seguranca-e-saude-no-trabalho/sst-menu/sst-normatizacao/sst-nr-portugues?view=default.

Soluções tecnológicas de baixo custo aplicadas em pequenas propriedades rurais

Kelinne de Oliveira Guimarães[1], *Moisés Laurence de Freitas Lima Junior*[1]

[1]Núcleo Docente de Informática, Instituto Federal de Educação do Tocantins
Araguatins, Tocantins, Brasil
kelinne.guimaraes@ifto.edu.br, moises.junior@ifto.edu.br

A agricultura está em constante processo de modernização tecnológica, no entanto, os programas públicos de modernização agrícola implantados no Brasil beneficiaram atividades em regiões específicas do país, favorecendo principalmente os grandes produtores do Centro-Sul e seus produtos, em geral destinados à exportação. Desta maneira, os pequenos agricultores, não tiveram acesso a tais programas de fomento a agricultura, consequentemente o segmento familiar, possui números negativos em termos de adoção tecnológica, produção e renda. Haja vista que estes indivíduos possuem limitações para acessar recursos, políticas públicas, tecnologias e mercado (SOUZA et al., 2018).

No que se refere à modernização agrícola estudos mostram que o Tocantins possui um baixo índice de modernização agrícola (BECKMANN; SANTANA, 2019). Haja vista que o acesso a tecnologias por parte dos agricultores familiares no Brasil, em especial no norte do Tocantins ainda é uma barreira.

A tecnologia possui um papel importante no desenvolvimento econômico do produtor rural, de que sua adoção eleva a produtividade e prover uma economia de mão-de-obra. De acordo com Meirelles (2019), muitas tentativas de promover a modernização e mudanças nos sistemas de produção dos agricultores têm fracassado por não reconhecerem as condições reais dos agricultores, como nível de qualificação, disponibilidade de recursos para investimentos, acesso a informação e etc.

Nesse panorama, nasceu a proposta deste projeto que visa concepção de produtos tecnológicos de baixo custo que podem ser aplicados com pouco investimento por parte do produtor rural, onde pretende-se otimizar e melhorar os processos de produção agrícola, principalmente dos pequenos produtores, possiblitando que o pequeno produtor tenha acesso a tecnologia para resolver problemas em sua propriedade. Os produtos desenvolvidos terão como foco a

solução de algum problema ou a otimização de processos por meio de dispositvos na produção agrícola por parte de pequenos produtores.

O domínio de técnicas de montagens elétricas e eletrônicas são um universo de possibilidades para projetos nas mais variadas áreas do conhecimento humano, utilizando componentes de baixo custo e uma tecnologia acessível pretende-se neste projeto apresentar diversos dispositivos tecnológicos à serem aplicados em propriedades rurais. Os produtos a serem desenvolvidos terão como foco a solução de algum problema ou a otimização de processos por meio de dispositvos na produção agrícola por parte de pequenos produtores.

Dentre os produtos que constaram no portfólio, destaca-se o eletrificador de cerca que consiste em um circuito isolado da rede de energia por meio de um transformador, produzindo pulsos de curta duração com alta tensão mas corrente limitada de modo a não causar queimaduras ou outros problemas graves (BRAGA, 2016). Este projeto se caracteriza pelo baixo consumo energético, podendo ser utilizado em pastos, eliminando assim a necessidade de uma cerca completa. Outro produto destacado será o espanta-bichos ultrassônico que consiste em um dispositivo emissor de ultra-sons em uma faixa controlada, afim de evitar que animais daninhos se aproximem de hortas, silos e etc. Ainda são produtos a serem desenvolvidos, o gerador de energia a partir de matrizes renováveis, irrigadores automáticos, espantalho eletrônico e outros.

Espera-se que este projeto tenha em seu portfólio de soluções, diversos produtos de baixo custo, onde o produtor rural consiga instalar tais soluções em sua propriedade com pouco investimento. De tal forma que seja possível levar tecnologia barata e funcional para pequenos produtores rurais de Araguatins, desenvolvendo produtos que visam o aumento da produção e da qualidade das atividades desenvolvidas nas propriedades.

REFERÊNCIAS

BECKMANN, E.; SANTANA, A. C. Modernizać ao da agricultura na nova fronteira agrícola do brasil: Matopiba e sudeste do pará. *Revista em Agronegocio e Meio Ambiente*, v. 12, p. 81–102, 2019.

BRAGA, N. C. *Projetos eletrônicos para o meio ambiente*. [S.l.]: Instituto Newton C. Braga, 2016.

SOUZA, P. de et al. Tecnologia na agricultura brasileira: uma análise das desigualdades regionais para os segmentos familiar e não familiar. *Revista Econômica do Nordeste*, v. 49, 2018.

Arduino da pesquisa ao campo: Um sistema de cultivo de plantas automatizado

João Batista Tolentino Junior[1], *Talia Marafon*[2], *Willian Macanan*[2]

[1] Docente titular da Universidade Federal de Santa Catarina – UFSC, Curitibanos-SC, Brasil.
[2] Discente do curso em Engenharia Agronômica em Universidade Federal de Santa Catarina – UFSC, Curitibanos-SC, Brasil.

O Arduino foi criado em 2005 na Itália por cientistas que possuíam como objetivo desenvolver um dispositivo com preço acessível, fácil programação e que poderia ser manipulado por projetistas amadores. A placa do Arduino é semelhante com um computador, contendo menor potencial de processamento e tamanho, sendo composta de um microcontrolador com circuito de entrada e saída e sua programação acontece no IDE (Ambiente de Desenvolvimento Integrado) e a linguagem usada é a C++, também denominada de linguagem de programação de alto nível. Existem diversos modelos de placas com característica distintas e a escolha dependerá do projeto que será efetuado (BANZI, 2011). Os projetos que podem ser realizados com o Arduino vão de coisas mais simples como funcionamento de LEDs ou até projetos mais elaborados como controlar um robô, isso só é possível, pois o Arduino tem inúmeros sensores (luminosidade, peso, temperatura), módulos (relé) que podem ser instalados separadamente ou associados e seu hardware é livre. Então, desenvolvemos um sistema de irrigação utilizando o Arduino, pois a água é um recurso de extrema importância nos sistemas agrícolas e o manejo adequado da água pode refletir em bons resultados na produção de alimentos, ao passo que desperdícios não são interessantes no sistema produtivo (PAZ et al., 2000). O setor agrícola é o maior consumidor de água, e no Brasil mais de cinquenta porcento dessa água consumida é destinada a irrigação, que é uma das principais formas de aumentar a produtividade (PAZ et al., 2000) e devido ao crescimento populacional a demanda por alimentos aumentará e o uso de sistemas de irrigação será mais frequente nas áreas agrícolas. Portanto, será essencial investir em sistemas de irrigação que possuem capacidade de otimizar os recursos hídricos, no entanto muitos modelos disponibilizados no mercado possuem preço elevado e os pequenos produtores não conseguem ter acesso, portanto a utilização do Arduino se torna interessante para agricultura familiar (CUNHA; ROCHA, 2015). Então, o objetivo do sistema de irrigação automatizado é usar de maneira mais eficiente os recursos hídricos, disponibilizando as plantas a quantidade de água essencial para completar o seu ciclo para pequena produção. Para elaboração do projeto foi necessário utilizar o sensor de umidade do solo higrômetro, o sensor de temperatura DS18B20, a válvula solenoide, o LCD, um módulo Bluetooth, onde

todos os componentes estão associados em uma placa do Arduino modelo Uno (CUNHA; ROCHA, 2015). Para o funcionamento do sistema de irrigação que será pôr gotejamento a placa do Arduino será programada de acordo com a exigência hídrica da cultura e levando em consideração outras variáveis como o estádio fenológico e o tipo de solo. Sendo que, para olericolas se preconiza umidade ideal igual a capacidade de campo. Então, o sensor de umidade do solo higrômetro através da condutividade elétrica do solo, mandará a informação da umidade do solo para a placa do Arduino, que se precisará abrirá a válvula solenoide, ou seja, ligará o sistema de irrigação e quando atingir a umidade desejada a partir da leitura do sensor de umidade a placa do Arduino fechará a válvula solenoide, ou seja, desligará o sistema de irrigação. O módulo Bluetooth permitirá que os dados registrados sejam acompanhados no celular através de um aplicativo chamado de SPP Bluetooth que está disponível no Google Play e o LCD permitirá também a visualização dos dados também (Figura 1). Após conhecer o Arduino e desenvolver o sistema de irrigação, compreendemos que essa tecnologia de baixo custo ainda pouco explorada e aplicada possui potencial de utilização em inúmeras áreas e pode ser uma solução de fácil acesso e baixo custo para pessoas que não possuem condições de efetuar altos investimento.

Figura 1: Estrutura demonstrativa do sistema de irrigação utilizado o Arduino



Fonte: Talia Marafon, 2019

REFERÊNCIAS

BANZI, M. **Primeiros passos com o Arduino**. São Paulo: Novatec Editora, 2011.

CUNHA, K. C.; B. ROCHA, R. V. **Automação no processo de irrigação na agricultura familiar com plataforma Arduino.** Revista Eletrônica Competências Digitais para Agricultura Familiar v. 01, n. 02. ISSN: 2448-0452

PAZ, V. P. S.; TEODORO, R. E. F.; MENDONÇA, F. C. **Recursos Hídricos, Agricultura Irrigada e Meio Ambiente. Revista Brasileira de Engenharia Agrícola e Ambiental.** v. 4. n.3. p.465-473, 2000. Disponível em:< http://www.scielo.br/pdf/%0D/rbeaa/v4n3/v4n3a25.pdf>. Acesso em: jun.2019.

Um sistema para determinação do consumo de água pelas plantas utilizando Arduíno e sensores de baixo custo

João Batista Tolentino Júnior[1], *Erick Santo Machado*[2]

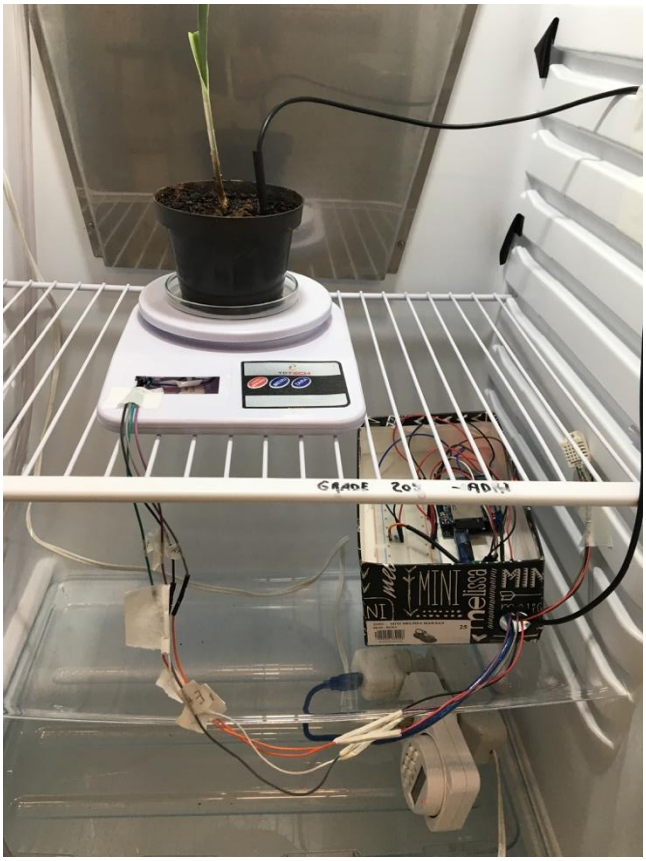
[1] Departamento de Agricultura, Biodiversidade e Florestas, Campus de Curitibanos, Universidade Federal de Santa Catarina, Curitibanos, Santa Catarina, Brasil, joao.tolentino@ufsc.br,

[2] Curso de Graduação em Agronomia, Campus de Curitibanos, Universidade Federal de Santa Catarina, Curitibanos, Santa Catarina, Brasil, ericksmach@gmail.com

Um dos modelos do avanço da modernização da produção de agrícola surgiu a partir da utilização de plataformas de prototipagem, sendo uma das mais utilizadas a plataforma Arduíno. De acordo com Melo (2012), esta plataforma consiste em um microcontrolador com circuitos de entrada e saída que permite ser controlado e programado via um Ambiente de Desenvolvimento Integrado que possibilita a interação com diferentes sensores de baixo custo na criação de projetos dinâmicos. O objetivo deste trabalho foi criar um protótipo que permite determinar a curva de consumo de água de uma planta, mensurando o balanço da água em um volume de solo ao longo do ciclo da planta, determinando a evapotranspiração nos diferentes estádios de desenvolvimento da planta. Para isso, utilizou-se uma balança feita com uma célula de carga de 10 kg (TAL220), montada em uma estrutura de plástico e instalada em uma câmara de crescimento com temperatura controlada. Inicialmente, procedeu-se a calibração da balança com pesos conhecidos, chegando a um fator de calibração de 268,3 para a massa em gramas. O registro dos dados foi feito por meio de um Shield DataLogger que conta com um módulo SD-card para a gravação dos dados e um RTC DS1307 (Relógio de Tempo Real) que permitiu o registro da data e hora das leituras. Além disso, utilizou-se um módulo relé para a ativação de uma lâmpada de LED que simulou um fotoperíodo de 12 horas. Ainda, foram adicionados um sensor de temperatura e umidade do ar (DHT22) e sensor de temperatura do solo (DS18B20), utilizados para monitorar as condições do ambiente. Os valores das variáveis obtidas pelos

sensores foram armazenados a cada 5 minutos em um cartão de memória em um arquivo de texto. Como teste, foi utilizado um vaso de 0,5 L com uma planta de alho. Diariamente, o vaso era irrigado com um volume de água igual à massa perdida pelo sistema desde a irrigação anterior.

Figura 1 – Protótipo montado em câmara de crescimento.



Fonte: Autores

O protótipo montado atendeu adequadamente aos objetivos propostos. A variação da massa do conjunto vaso+planta foi monitorado e o balanço diário foi utilizado como parâmetro para a reposição da água via irrigação ao conjunto. O registro frequente das variáveis pelo *datalogger* facilitou a análise destas e gerou uma grande quantidade de informações úteis para determinar o consumo de água pela planta com boa precisão.

REFERÊNCIAS

MELO, J. L. G. G; BARANIUK, J. **Mini curso: arduino**. Universidade Federal do Paraná. Curitiba, 2012. Acesso em 01 de agosto de 2019, disponível em <http://www.eletrica.ufpr.br/~james/Laboratorio%20V/arquivos/Mini%20Curso%20Arduino.pdf>

Identification of sugarcane straw availability through GIS tools

Rafaella Pironato Amaro[1], *Bruna Cristina Gama Campagnuci* [2], *Lauren Maine Santos Menandro*[1], *Sergio Gustavo Quassi de Castro*[1], *Thayse Ap. Dourado Hernandes*[1], *João Luis Nunes Carvalho*[1], *Ana Claudia dos Santos Luciano*[3]

[1] Laboratório Nacional de Biorenováveis (LNBR), Centro Nacional de Pesquisa em Energia e Materiais (CNPEM), CEP 13083-970, Campinas, São Paulo, Brasil, [rafaella.amaro, lauren.menandro, sergio.castro, thayse.hernandes e joao.carvalho]@lnbr.cnpem.br

[2] bruna0812@gmail.com [3] anaclauster@gmail.com

Geospatial data associated with the use of Geographic Information System (GIS) tools have assisted the monitoring of sugarcane production and the strategic planning of the sugar-energy sector. The objective of this work was to quantify the sugarcane straw available for bioenergy production, using geospatial and GIS data integrated with field information and climate data. Factors for straw removal in the field were determined according to the analysis of soil conservation and sugarcane yield data, considering the terrain slope (VALERIANO; ALBUQUERQUE, 2010), the soil texture (ROSSI, 2017), the crop management and the climate data. Field data, which consisted of information on yield, number of harvests, soil tillage and date of harvest at field level, were obtained from the partner mills for 2017/2018 crop season in region 1, which covers the municipalities of Miguelópolis; Ituverava; Ipuã; Guará; São Joaquim da Barra; and for the crop season 2016/2017 in region 2, covering the municipalities of Rancharia; João Ramalho; Quatá and Paraguaçu Paulista. Data were processed in the ArcGIS 10.4 GIS environment. Sugarcane straw yield was calculated assuming a 12% straw/stem yield ratio (MENANDRO, et al., 2017). Slope and solar radiation were divided in two classes each: less than or equal to 3% and greater than 3% and high (> 240 MJ) and low solar radiation (≤ 240 MJ) (HERNANDES et al., 2019), respectively. Field data referring to planting dates were classified according to the harvesting months during the harvest period: beginning (January - March); middle (May - August) and end of harvest (September - December). Regarding the sugarcane straw removal, areas were classified as Suitable, Restrict and Unsuitable. For Suitable areas, 2 tons of straw were maintained per hectare (dry mass) due to limitations of the operational conditions of straw collection in the field. For Restrict areas, we considered the

permanence of 7 tons of straw per hectare (CARVALHO et al., 2017), aiming at soil conservation regarding soil erosion. In region 1, 43,867 hectares were classified as Suitable and 20,493 hectares were classified as Restrict, representing 438 thousand tons of straw that could be removed and 66% of the total amount of straw generated by sugarcane. In region 2, 20,470 hectares were classified as Suitable and 10,389 hectares as Restrict, adding up to 111 thousand tons of straw available to be removed, which represents 32% of the total amount of straw generated by the crop. The amount of straw to be removed in region 1 is approximately twice the percentage of region 2. It can be explained by the predominant sandy soil texture in region 2, which limits the removal of straw due to their susceptibility to erosion (CARVALHO et.al, 2017). For region 1 clayey soils are predominant, which requires less straw to be left on the ground to prevent soil erosion. Despite the operational issues of sugarcane straw removal and the restriction to maintain 7 tons per hectare in areas classified as Restricted, which limits the potential for straw removal from the field, the total volume of straw that may be used for bioenergy purposes is significant. Geospatialized field information associated with the use of GIS tools has shown promising results in identifying the spatial distribution of Suitable, Restrict and Unsuitable areas, as well as the estimation of the sugarcane straw amount that can be removed. Thus, this is an important planning tool, as it enables appropriate strategies for straw removal that meet energy demands, contributing to the maintenance of the benefits of the straw mulching on soil conservation and in the yield performance of sugarcane areas.

REFERENCES

CARVALHO, J.L.N. et al. “Agronomic and environmental implications of sugarcane straw removal: a major review”. GCB Bioenergy. V 9, pp.1181–1195, 2017.

HERNANDES, T.A.D. et al. Agroclimatic zoning of straw removal and its impacts on sugarcane yield. 27th european biomass conference & exhibition, Lisbon, 2019.

MENANDRO, L. M. S. et al. Comprehensive assessment of sugarcane straw: implications for biomass and bioenergy production. Biofuels, Bioproducts and Biorefining, 11(3), 488-504, 2017.

ROSSI, M. “Mapa pedológico do Estado de São Paulo: revisado e ampliado”. Instituto Florestal, São Paulo, v 1, pp.118, 2017.

VALERIANO, M. M., ALBUQUERQUE, P. C. G. TOPODATA: Processamento dos dados SRTM. Projeto de Produtividade em Pesquisa ”Desenvolvimento de aplicações de dados topográficos SRTM. INPE, São José dos Campos, 81p., 2010.



Análise das Tendências de Pesquisa no Congresso Brasileiro de Agroinformática por Nuvem de Palavras

Bruno Ernesto Techera da Motta[1], *Alfredo Parteli Gomes*[1]

[1]Instituto Federal Sul-Rio-Grandense
Santana do Livramento, Rio Grande do Sul, Brasil
brunomotta@ifsul.edu.br, alfredogomes@ifsul.edu.br

A evolução da tecnologia da informação na última década trouxe novas oportunidades e desafios no contexto de dados. A redução de custos, possibilitou que a coleta, armazenamento, processamento, transmissão e compartilhamento da informação se tornassem cada vez mais acessíveis. Entretanto a quantidade de dados gerados pelos sistemas computacionais estão se tornando progressivamente mais volumosos. Dentro deste contexto, surge uma das tendências tecnológicas atuais mais significativas, chamada de Big Data, que possibilita a transformação, em tempo viável, de grandes quantidades de dados em conhecimento, através da descoberta de padrões ocultos ou correlações desconhecidas nos dados analisados. Uma das dificuldades encontradas é analisar os diferentes tipos de dados não estruturados que agrupam-se em dados textuais e dados não-textuais (AHMED, 2013).

Com a finalidade de avaliar as tendências de pesquisa em agroinformática no Brasil, foi delimitado o contexto do Congresso Brasileiro de Agroinformática (SBIAgro), considerando que é o principal evento científico nacional nesta área e que tem por objetivo fomentar a divulgação de resultados de pesquisas, trabalhos em andamento e debates sobre a inovação no setor. A partir de análises da literatura correlata (GOMES; CAMARGO; BELLINI, 2018), este trabalho tem como objetivo a identificação e apresentação dos tópicos de pesquisa mais proeminentes nas últimas oito edições do congresso, através do uso da técnica de nuvem de palavras para visualização dos termos mais utilizados nos títulos dos artigos publicados. Como objetivo secundário, serão identificadas as tendências de pesquisa dos assuntos mais frequentes nos quatro grupos de dois períodos. A partir desta análise, são apresentados os tópicos com tendência crescente e decrescente de relevância.

Como impactos esperados a partir dos resultados apresentados neste trabalho, pretende-se tornar claras as tendências de pesquisa na agroinformática do Brasil, a fim de contribuir com a

escolha de pesquisadores que tenham interesse em desenvolver trabalhos na área.

A metodologia utilizada foi a seguinte: 1) Obtenção dos dados com a organização do evento; 2) Criação de um arquivo em texto puro; 3) Agrupamento dos títulos em quatro grupos com duas edições (2003 e 2005; 2007 e 2009; 2011 e 2013; 2015 e 2017); 4) Remoção de *stop words*, tais como artigos, preposições, pronomes e conjunções; 5) Contagem do número de vezes que cada palavra aparece em cada grupo (contagem absoluta); 6) Criação de nuvens de palavras, onde a dimensão dos tópicos reflete a sua frequência de ocorrência conforme a Figura 1; 7) Cálculo da frequência relativa das palavras em cada grupo, através da divisão do valor da contagem absoluta pela quantidade de trabalhos em cada grupo; 8) Para cada palavra, foi criado um modelo de regressão linear, utilizando as frequências relativas relacionadas com cada grupo.

Este estudo permitiu identificar os termos de pesquisa mais frequentes, possibilitando a identificação de crescimento ou redução da importância destes termos ao longo das edições do congresso. Foi possível evidenciar os termos de maior e menor tendência respectivamente e o valor da inclinação da reta de regressão, por exemplo: Sensor (0,032), Vant (0,025), Solo (0,015), Dados (0,014) e Sistema (-0,076), Software (-0,031), Informação (-0,028), Desenvolvimento (-0,022).

Figura 1: Nuvens de Palavras dos termos mais frequentes



Fonte (Elaborada pelo autor)

REFERÊNCIAS

AHMED, Z. Data management and big data text analytics. *National Conference on Cloud Computing and Big data*, p. 140–144, 2013.

GOMES, A. P.; CAMARGO, S. D. S.; BELLINI, Y. N. S. Tendências de pesquisa em agroinformática na argentina: uma análise histórica. *X Congreso de AgroInformática (CAI)-JAIIO 47 (CABA, 2018)*, p. 446–454, 2018.

Eventos climáticos extremos: Uma análise do clima de Campinas

Lucas H. Ueda[1], *Paula D. P. Costa*[1], *Ana M. H. de Avila*[2], *Eliana C. de Faria*[3]

[1]Faculdade de Engenharia Elétrica e Computação (FEEC), [2]Centro de Pesquisas Meteorológicas e Climáticas Aplicadas à Agricultura (CEPAGRI), [3]Faculdade de Ciências Médicas (FCM), Universidade Estadual de Campinas (UNICAMP), Campinas, São Paulo, Brasil

[1]lucashueda@gmail.com e paulad@unicamp.br, [2]avila@cpa.unicamp.br, [3]cottadefaria@gmail.com

Diversos setores dependem de condições climáticas adequadas, como a agricultura, a pecuária e as usinas hidrelétricas, fazendo com que estudos sobre desvios de elementos do clima terrestre sejam cada vez mais relevantes e frequentes.

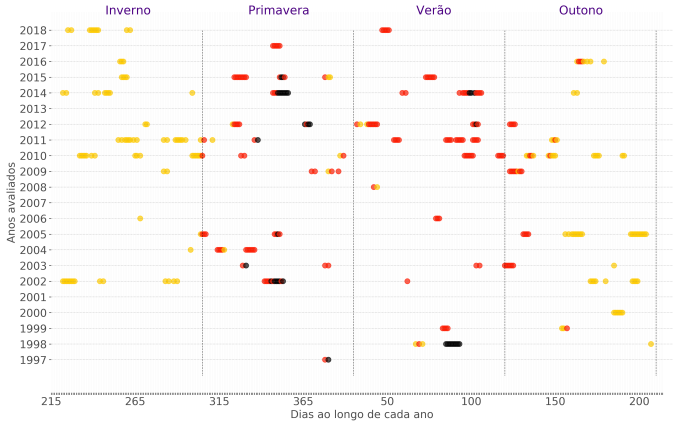
O presente trabalho buscou obter um panorama da ocorrência de eventos climáticos extremos na cidade de Campinas a partir da análise de medidas de umidade relativa do ar e ondas de calor no período de 1997 a 2018 (22 anos). O estudo utilizou dados coletados pela estação metereológica do Centro de Pesquisas Meteorológicas e Climáticas Aplicadas à Agricultura (CEPAGRI), localizada no campus "Zeferino Vaz" da Unicamp, no distrito de Barão Geraldo.

Na ausência de um consenso na definição de ondas de calor, o presente trabalho baseou-se nos estudos de Geirinhas et al. (2018) e Russo et al. (2014), que definem uma onda de calor como sendo um período de 2 ou mais dias consecutivos com a temperatura máxima do dia maior que o percentil 90 das temperaturas máximas, dentro de uma janela de 31 dias centrada no dia de avaliação. Extrapolamos esta metodologia para a definição de ondas de umidades mínimas, com a finalidade de observar períodos de umidade relativa do ar abaixo de um limiar (percentil 10), em um período prolongado de dias.

Seguindo esta abordagem, foi possível observar a ocorrência de ondas de calor também nas estações mais frias (outono e inverno). Em particular, observa-se que as ocorrências deste fenômeno se concentram nos últimos 10 anos de observação (Figura 1). Destaca-se que, uma vez que a definição de onda de calor adotada considera uma janela temporal centrada num dia específico do ano, essas ondas apresentam temperaturas máximas mais baixas (pontos em amarelo da Figura 1), como era esperado.

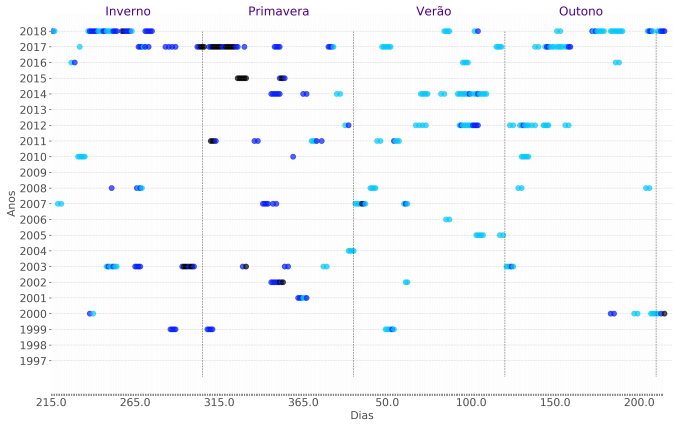
Nota-se também pela Figura 1 que as ondas de calor mais intensas (em vermelho e preto no gráfico) ocorrem apenas nas estações de verão e primavera (fenômeno esperado, por serem as estações mais quentes do ano).

Figura 1: Dias sob efeito de ondas de calor por ano em Campinas (Pontos em vermelho e preto indicam as ondas mais intensas no período).



Já para os fenômenos extremos de umidades relativas mínimas é possível notar como esse efeito está ligado à transição do inverno para a primavera, período no qual se concentram os eventos mais extremos do fenômeno (em preto e azul escuro no gráfico da Figura 2). Destaca-se também a grande concentração de dias sob esse efeito nos últimos dois anos (2017 e 2018).

Figura 2: Dias sob efeito extremo de umidade relativa mínima por ano em Campinas (Pontos em azul escuro e preto indicam os fenômenos de umidade mínima mais extremos no período).



Estes achados indicam que eventos climáticos extremos podem estar se tornado mais frequentes, representando assim uma potencial ameaça à agricultura regional e à segurança alimentar da população.

REFERÊNCIAS

GEIRINHAS, J. et al. Climatic characterization of heat waves in brazil. *Anuario do Instituto de Geociencias*, v. 41, 09 2018.

RUSSO, S. et al. Magnitude of extreme heat waves in present climate and their projection in a warming world. *Journal of Geophysical Research: Atmospheres*, v. 119, n. 22, p. 12,500–12,512, 2014.

Desenvolvimento de um sistema em rede estrela para administração de estufas inteligentes com *IOT*

Igor Mauro Silva De Almeida[1] *,Iraildo Barbosa*[2]*, Everlaine Leopoldino Dias*[3]*, Alberto E. Pereira de Araujo*[4]

[1]Graduando em Bacharelado em Ciência da Computação, Unidade Acadêmica de Garanhuns, Universidade Federal Rural de Pernambuco, Pernambuco, Brasil, igoralmeeida@hotmail.com.

[2]Graduando em Agronomia, Unidade Acadêmica de Garanhuns, Universidade Federal Rural de Pernambuco, Pernambuco, Brasil,iraildodasilva2016@gmail.com.

[3]Graduanda em Agronomia, Unidade Acadêmica de Garanhuns, Universidade Federal Rural de Pernambuco, Pernambuco, Brasil, everlaineldias@gmail.com

[4]Unidade Acadêmica de Garanhuns, Universidade Federal Rural de Pernambuco, Pernambuco, Brasil, alberto.araujo@ufrpe.br

RESUMO

O desenvolvimento da produção agrícola apresenta um rico histórico no enfrentamento de desafios, tais como a domesticação e melhoramento das plantas e sua adequação em diversas condições ambientais Regiões com condições hídricas ou climáticas desfavoráveis impõem ônus adicionais ao produtor rural. Neste contexto, as mudanças ambientais causadas pelo uso intensivo dos recursos minerais e a super-urbanização, demandam novas técnicas de cultivo. Tais técnicas cada vez mais, envolvem o uso em larga escala da tecnologia. Nesse cenário, a automação do cultivo com a administração ou manejo de dados remotos demanda a sincronização entre sensores e atuadores através de conexão com a internet. Porém, os locais onde acontece a produção agrícola podem apresentar uma baixa ou ineficiente cobertura de internet. Como solução, pode-se centralizar o acesso a rede externa em um servidor central e implementar uma rede local. Neste trabalho apresentamos o desenvolvimento de um sistema de automação sensor/atuador para uso em cultivo em ambientes fechados, com rede local. Para tanto, foram utilizados como componentes físicos: micro-controlador Arduino UNO, uma placa ATX-AsRock AD525PV3, roteador Smart Lan Aprio 150 e conectores de fios, estrutura metálica de suporte e como componentes lógicos: a linguagens de programação Python com o kit Anaconda, com destaque para as bibliotecas Pyserial e Socket, e a linguagem C++, implementada no Arduino. Com isso, utilizando da sua biblioteca Pyserial, foi-se conectado na placa ATX-AsRock o Arduino, através da conexão serial. Para teste do sistema foi usado um Módulo Sensor De Umidade Do Solo Higrômetro Yl-69, o qual realiza leitura da umidade do solo e retorna um número inteiro no intervalo [0,1024]. Com o uso da biblioteca PySerial, o Arduino foi

programado a se comunicar através da porta serial COM4 e retorna ao interpretador Python a umidade registrada no módulo Y1-69. Os dados foram registrados em um arquivo de texto no formato .txt para posteriormente gerar um gráfico da umidade presente no solo, através da biblioteca do mathplotlib do Python.

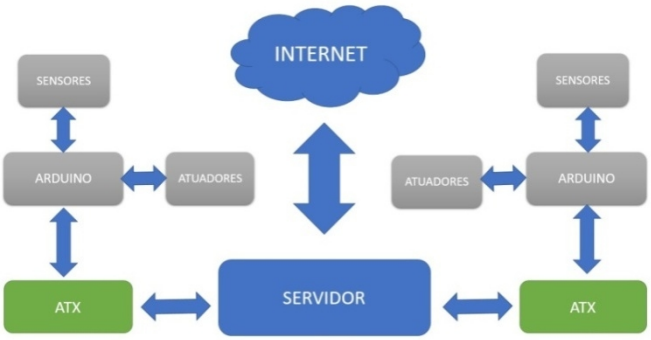


Figura 1 - Diagrama do sistema apresentando uma rede tipo estrela.

O experimento foi realizado na cidade de Catende-PE, utilizando como unidade experimental a planta babosa (*Aloe vera*) em um vaso com solo, sem a adição de água durante a análise. Os dados foram coletados durante 24h com início 21:30 do primeiro dia, sendo o monitoramento realizado via internet, utilizando uma rede VPN com uso da ferramenta TeamViewer 14.

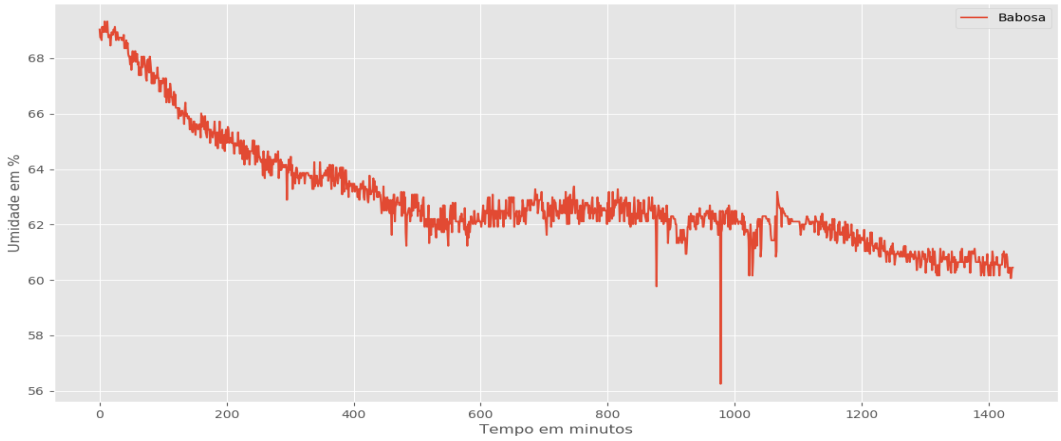


Figura 2 - Medida da umidade

Através do gráfico da figura 2, pode-se observar que o solo apresentou uma taxa de umidade acima de 68% tendo uma queda de 8% durante o experimento. Com essa medida demostramos a funcionalidade do sistema construído com aquisição de dados e controle a distância via VPN. Pretende-se adicionar outros sensores e atuadores para um total controle do ambiente interno a estufa, permitindo proteger a planta de agentes patógenos externos, bem como otimizar as condições ambientais, aumentando sua produtividade.

SARIPAN - Sistema de Alerta de Risco de Incêndio para o Pantanal

Marcelo Gonçalves Narciso[1], Balbina Maria Araújo Soriano[2]

[1] Embrapa Arroz e Feijão, Goiânia, Goiás, Brasil, marcelo.narciso@embrapa.br

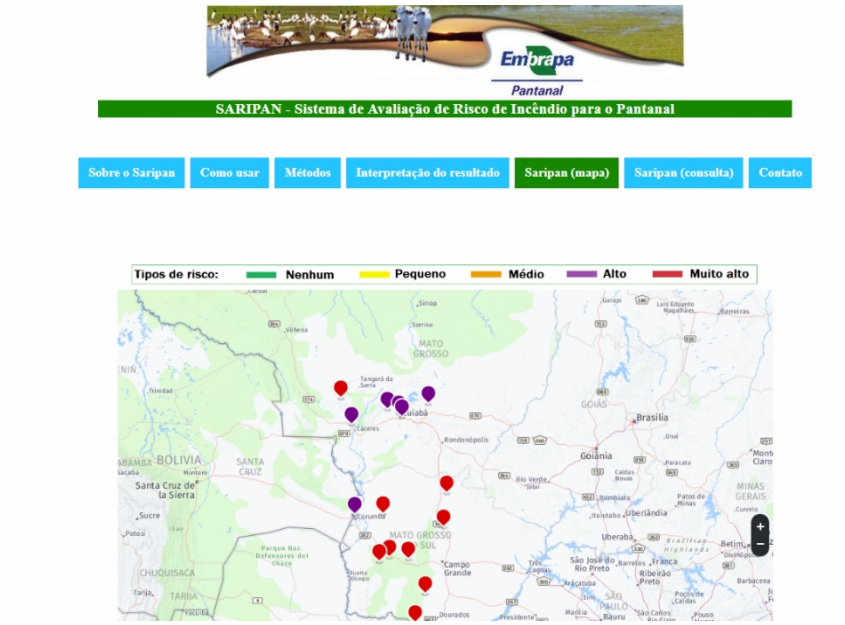
[2] Embrapa Pantanal, Corumbá, Mato Grosso do Sul, Brasil, balbina.soriano@embrapa.br

No Pantanal o manejo das pastagens naturais é complexo e dinâmico, em razão da grande variedade de fitofisionomias, que variam espacialmente e temporalmente, principalmente em função das condições climáticas. Muitas dessas fitofisionomias são propensas a incêndios que podem ocorrer acidentalmente ou provocados por práticas de manejo inadequadas de queimas em pastagem. Os incêndios podem levar a prejuízos irrecuperáveis, e mobilizar uma grande soma de esforços e recursos do setor público nas operações de prevenção e combate. Para ajudar na prevenção de incêndios no Pantanal, foi construído um sistema de previsão de risco de incêndio, através de software, conhecido por Saripan.

O software Saripan é acessado pela web e sua base de dados contém uma série de dados climáticos da região do Pantanal, de 2017 até a presente data, os quais são usados como entrada para fornecer resultados sobre índices de risco de incêndio na região desejada no Pantanal. O risco de incêndio recomendado para a região do Pantanal é a fórmula de Monte Alegre (SOARES, 1998), porém o usuário poderá usar outras fórmulas para o cálculo do risco de incêndio descritas na literatura (VOLPATO, 2002) que são a Fórmula de Monte Alegre Modificada (FMA+), índice Nesterov, índice logarítmico Telecyn e índice de Angstron. Caso o usuário tenha os dados climáticos de sua região, que pode estar fora da região do Pantanal, o sistema permite a inserção destes dados, que deverão estar contidos em arquivo texto, para o cálculo do índice de risco de incêndio.

Os dados climáticos do sistema são atualizados diariamente após às 14 horas (horário de Brasília) e são provenientes da base de dados do sistema de monitoramento Agrometeorológico Agritempo (AGRITEMPO, 2019) e, para o caso de algum dado faltante, são usados dados climáticos da base de dados climáticas de estações virtuais do sítio da Nasa (NASAPOWER, 2019). O acesso ao sítio do Saripan é através do link www.cnpaf.embrapa.br/saripan. A Figura 1 ilustra uma consulta ao sistema.

Figura 1 – Consulta a risco de incêndio em Campo Verde-MT.



Fonte: (sítio do Saripan, 2019)

A importância do Saripan, para a região do Pantanal, é quanto a existir um sistema de alerta que informe o risco de incêndio para cada sub-região do Pantanal, possibilitando às autoridades ou interessados tomarem medidas preventivas quanto a um possível incêndio que venha ocorrer.

REFERÊNCIAS

AGRITEMPO. Sistema de Monitoramento Agrometeorológico, Disponível em <https://www.agritempo.gov.br/agritempo/index.jsp>. Sítio visitado em 01/08/2019.

NASAPOWER. Sistema de Monitoramento Agrometeorológico, Disponível em <https://power.larc.nasa.gov/data-access-viewer/>. Sítio visitado em 01/08/2019.

SOARES, R. V. Desempenho da "Fórmula de Monte Alegre" índice brasileiro de perigo de incêndios florestais. Cerne, Lavras, v. 4, p. 87 - 99, 1998.

VOLPATO, M.M.L. Imagens avhrr-noaa para determinação do potencial de incêndios em pastagens. Tese (Doutorado em Engenharia Agrícola) – Faculdade de Engenharia Agrícola, Universidade Estadual de Campinas. Campinas-SP, p. 97. 2002.

Previsão de Indicadores Econômicos Individuais na Pecuária Leiteira

Helison Cavalheiro[1], Thales Maciel[1], Oldemar Weiller[2], Gustavo da Silva[3], Vinicius Lampert[3]

[1] Instituto Federal Sul-rio-grandense, campus Bagé, Brasil, thalesmaciel@ifsul.edu.br

[2] Empresa de Assistência Técnica e Extensão Rural, Bagé, Brasil, ohweiller@emater.tche.br

[3] Empresa Brasileira de Pesquisa Agropecuária, Bagé, Brasil, vinicius.lampert@embrapa.br

O agronegócio é uma das principais atividades de fonte de renda no Brasil, especialmente no Estado do Rio Grande do Sul, onde se destacam a prática de alimentação de rebanhos baseada em pastagens e as certificações dos processos de produção (VARELLA; CARDOSO, 2014).

A bovinocultura leiteira, ramo mais específico da pecuária brasileira, caracteriza uma das maiores produções do país, que figura como o 5º maior produtor de leite do mundo (MAPA, 2018). Este relatório ainda afirma que a produção leiteira tem passado por dois processos evolutivos, que são a incorporação da tecnologia para o melhor gerenciamento da produção e a eliminação, do mercado, dos produtores que não estejam dispostos a incorporar estas tecnologias em seus métodos de produção.

Um exemplo disto é o trabalho de Lopes et al. (2011), onde objetivou-se a identificação de características que exercem influência sobre os custos finais da produção leiteira. Utilizando um software estatístico próprio para o processamento de dados, os autores apresentaram os custos de alimentação como de maior influência no custo operacional, embora tenham deixado clara a necessidade de ampliação do conjunto de dados analisado e exploração de métodos diferenciados de análise.

Com um propósito similar, o presente trabalho objetivou a obtenção de um modelo preditivo e didático capaz de representar a influência dos indicadores de desempenho da produção leiteira nos aspectos econômicos da atividade.

Para tal, procedeu-se à descoberta de conhecimento em banco de dados através de tarefas de mineração de dados, especificamente a regressão (RENCHER; CHRISTENSEN, 2012). Os dados foram obtidos de unidades de produção na região Noroeste do estado do Rio Grande do Sul entre os anos de 2015 e 2016 e fazem de um estudo sobre o desempenho produtivo e econômico da atividade leiteira nessa região.

O conjunto de dados analisado foi composto por 42 instâncias e 13 atributos

numéricos, nominalmente a produção anual de leite em litros por hectare, contagem de células somáticas, custo do leite por litro (R$), preço do litro do leite (R$), margem de lucro anual do leite (R$/ha), margem de lucro anual de grãos (R$/ha), saldo de caixa mensal (R$), percentual de parcelamento, percentual de alimento oferecido em cocho, custo da silagem (R$/kg e R$/ha) .

Em pré-processamento, foi realizada a normalização dos atributos do conjunto de dados, visto que compreendiam escalas numéricas diferentes entre eles. Desta forma, foi possível formatar os dados de forma que, no modelo a ser descoberto no processamento, os pesos apresentados para cada atributo sejam proporcionais entre si. Em processamento, a parametrização da implementação do algoritmo de regressão linear utilizada ocorreu para previsão da margem de lucro anual do leite, sendo preditores os demais atributos.

Quadro 1 – Modelo descoberto para predição da margem de lucro anual do leite.

```
margem de lucro anual do leite =
          0.57 * produção anual de leite +
      -9785.20 * custo do leite por litro +
       9069.86 * preço do leite por litro +
          0.19 * saldo de caixa mensal +
      -1291.26 * porcentagem de alimento oferecido no cocho +
      15412.97 * custo da silagem +
      -5771.92
```

O modelo descoberto, conforme apresentado no Quadro 1, apresenta o custo da silagem como de maior influência na margem de lucro anual, o aspecto econômico escolhido como alvo da previsão no experimento. Também apresentaram alta relevância, no modelo, o custo do leite por litro, o preço do leite por litro e a porcentagem de alimento oferecido no cocho, este com proporcionalmente menor influência, em comparação aos demais. O coeficiente de correlação obtido para o modelo nos testes de predição foi 0,98.

Tendo cumprido o objetivo proposto, vislumbra-se a experimentação com outros algoritmos de previsão de valores numéricos e a aplicação dos modelos descobertos pela incorporação dos métodos utilizados em softwares especializados em gestão pecuária.

REFERÊNCIAS

VARELLA, A.; CARDOSO, F. Rumos da pecuária de corte do RS e desafios para a pesquisa científica na visão da Embrapa Pecuária Sul. Jornada do Núcleo de Estudos em Sistemas de Produção de Bovinos de Corte e Cadeia Produtiva. UFRGS, 2014.

Ministério da Agricultura, Pecuária e Abastecimento (MAPA). Projeções do Agronegócio: Brasil 2017/18 à 2027/28 - Projeções de Longo Prazo. Brasília, 2018.

LOPES, M.;, DOS SANTOS, G.; RESENDE, M.; CARVALHO, F.; CARDOSO, M. Estudo da rentabilidade de sistemas de produção de leite no município de Nazareno, MG. Ciência Animal Brasileira, p. 58-69. 2011.

RENCHER, A.; CHRISTENSEN, W. Multivariate Regression. Methods of Multivariate Analysis, capítulo 10. Wiley, 2012.

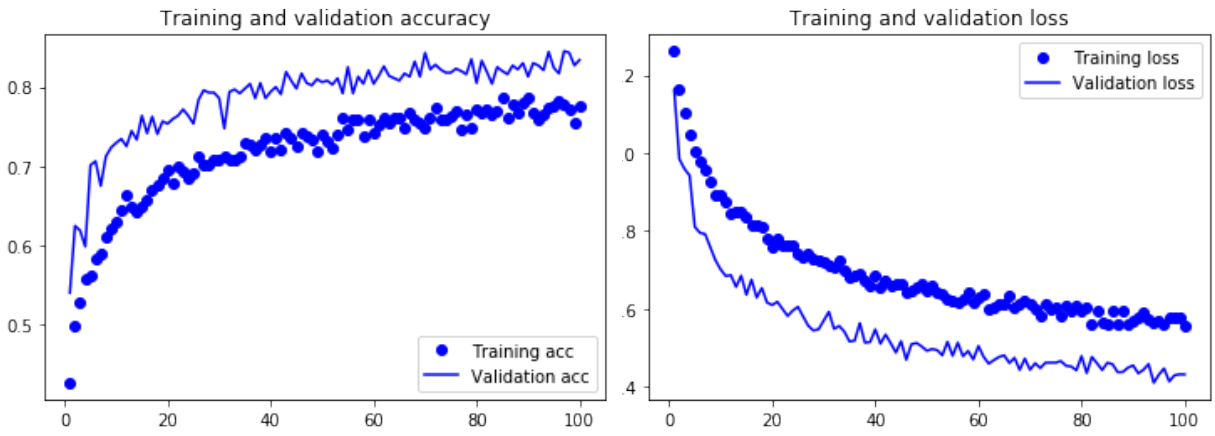
Deep Learning como ferramenta para classificação visual de folhas de soja

Eduardo Cintra, Alysson A. Naves Silva , Darlan Einstein do Livramento
Libertas - Faculdades Integradas Integradas, São Sebastião do Paraíso – MG

A utilização de técnicas de *Deep Learning* está cada vez mais frequente e tem redefinido o estado da arte em diversas áreas como: visão computacional, classificação de imagens, processamento da linguagem natural, aplicações industriais, medicina, agricultura, dentre outras (PONTI, 2017). Na agricultura há aplicações de monitoramento de culturas por imagens aéreas, sensores para captura de informações do solo, ar, água e planta, classificação de grãos, sementes e folhas são alguns exemplos de importantes tarefas para garantia da qualidade de um produto e otimização na aplicação de insumos. Com base nisto, surge a importância de desenvolver um sistema autônomo capaz de classificar folhas de soja sem interferência de um especialista humano durante o processo. A análise de folhas de sojas pode ser realizada com as seguintes finalidades: determinar os teores de nutrientes foliares da planta ou realizar o diagnóstico de doenças foliares através dos sintomas, esses observados no campo. A análise foliar da soja, com a finalidade de identificar e quantificar doenças é uma técnica onde se avalia visualmente o limbo foliar e para isso se utilizam de tabelas, ou escalas diagramáticas onde determinam quantidade de área foliar lesionada pela doença. Esse tipo de avaliação de campo pode ocasionar erros de estimativas de evolução de doenças, segundo Soares et al. (2009) o uso da escala diagramática é necessário para a avaliação de manchas na folha de soja, onde foi comprovado que os inspetores que a não utilizaram, obtiveram uma variância maior nos resultados. O objetivo deste trabalho foi desenvolver um classificador de folhas de soja com uso de técnicas de aprendizado de máquinas com o uso de imagens, sendo capaz de classificar a imagem da folha em folha doente (severidade baixa, média e alta) ou folha saudável. Para isto, foi utilizado um conjunto de imagens de folha de soja como entrada do modelo, estas imagens são de uma base de dados de um pesquisador na área agrícola e serviu como ponto de partida para o entendimento das técnicas de *Deep Learning*. O modelo foi desenvolvido através da biblioteca de código aberto Keras (CHOLLET, 2018) para

aprendizado de máquina e como *backend* o TensorFlow (ABADI, 2016). Foram construídos e testados 5 modelos distintos e o melhor resultado obtido pode ser verificados na Figura 1. A precisão do modelo foi de aproximadamente 90% mostrando-se que *Deep Learning* é uma técnica robusta quando aplicada na classificação de folhas de soja por imagens e com potencialidade de melhoria.

Figura 1- Treinamento de uma Rede Neural Convolucional para a classificação de folhas de soja (a) Acurária (b) Perda



(a) (b)

Fonte: Autores (2019)

REFERÊNCIAS

ABADI, M. et al. Tensorflow: A system for large-scale machine learning. In:12th Symposium on Operating Systems Design and Implementation (OSDI 16). 2016. p. 265-283.

CHOLLET, F. Deep Learning with Python. NY: Manning Publications Co, 2018.

PONTI, M. A.; COSTA, G. B. P. da. Como funciona o deep learning. Tópicos em gerencia-mento de dados e informações 2017.Uberlândia, p. 35-49, 2017.

SOARES, Rafael M.; GODOY, Cláudia V.; OLIVEIRA, MCN de. Escala diagramática para avaliação da severidade da mancha alvo da soja. Tropical Plant Pathology, v. 34, n. 5, p. 333-338, 2009.

Uso do 5G em Área Remota

Cledson Akio Sakurai[1], *Cecilia Emi Yamanaka Matsumura*[2],

[1] Instituto do Mar, Universidade Federal de São Paulo, São Paulo Brasil, akio.sakurai@unifesp.br

[2] Departamento de Engenharia da Computação e Sistemas Digitais, Escola Politécnica da Universidade de São Paulo , São Paulo, Brasil, cecilia.yamanaka@usp.br

A rede móvel de quinta geração (5G) traz novas possibilidades de serviços a serem oferecidos para os diversos usuários, além de conviver sem interferência com o sinal da televisão e das redes celulares existentes, como: 2G, 3G e 4G. O projeto em desenvolvimento com a comunidade europeia chamado 5G Range vem para superar os desafios que existem ao utilizar a tecnologia 5G em redes de longa distância para acesso a áreas remotas, de forma a permitir taxas de transmissão de dados a 100 Mbps em um raio de alcance de 50 quilômetros a partir da estação rádio base (ERB), neste caso atingindo cerca de 100 quilômetros por localidade. Apesar da taxa de transmissão ser menor que a prevista para os centros urbanos, que é estimado na ordem de 20Gbps, a taxa prevista no projeto 5G Range viabiliza diversos serviços e permite maior alcance a um custo mais apropriado para áreas remotas, e traz uma maior capacidade e cobertura que o padrão atual em 4G. Esta característica do projeto permite que boa parte da população brasileira que estão longe da infraestrutura de acesso a internet em locais remotos, possam ser atendidas por este sistema de 5G de longo alcance (EUROPEAN COMISSION, 2019).

O grande desafio do projeto consiste no desenvolvimento de uma camada cognitiva que permite o acesso sem comprometer os acessos existentes, através de um controle de acesso que usa uma abordagem de alocação dinâmica de espectro de frequência e uma camada física para reduzir emissões fora da banda, o que permite agilidade, baixa interferência e melhor aproveitamento do espectro de frequência disponível, e assim, trazer uma relação custo benefício otimizada.

Com esta característica de baixo custo de implementação, longo alcance e taxa de transmissão apropriada, a tecnologia proposta no projeto 5GRange permite ao agronegócio fazer um uso

mais intenso das tecnologias em locais onde a tecnologia atual não antedê a demanda, além de contar com baixa latência, robustez, sustentabilidade e baixo consumo de energia, viabilizando assim a agricultura de precisão e inteligente. (MA, 2017)

A tecnologia do 5GRange traz a possibilidade de aumentar o uso da tecnologia na produção agrícola, intensificando o uso de geoposicionamento, varredura do solo, gerenciamento de dados e de soluções que se utilizam de tecnologias da Internet das Coisas. Espera se que com o uso desta tecnologia será possível ter dados mais precisos sobre as condições co campo e assim aumentar de forma significativa a eficácia da operação na produção agrícola, como uso mais efetivo dos defensivos agrícolas e fertilizantes, ou ainda monitorar com maior precisão as necessidades de cada animal e ajustar o seu manejo de forma a reduzir doenças e custos produtivos. (BACCO, 2018) (KHANNA, 2018)

O 5GRange traz características de transmissão de dados que juntamente com outras técnicas computacionais, como a computação em nuvem e a inteligência artificial, permite o uso de equipamentos agrícolas com maior precisão e eficiência, além de compreender melhor as condições operacionais e ambientais do campo, e assim permite ações mais efetivas para cada situação, sempre preocupado com o aumento da produtivade mas com redução dos custos envolvidos.

REFERÊNCIAS

BACCO M. et Al. Smart farming: Opportunities, challenges and technology enablers. IEEE IoT Vertical and Topical Summit on Agriculture. 2018.

EUROPEAN COMISSION. Remote area Access Network for 5th Generation. Disponível em: https://cordis.europa.eu/project/rcn/212224/factsheet/en. Acessado em 20/09/2019.

KHANNA A. Evolution of Internet of Things (IoT) and its significant impact in the field of Precision Agriculture. Elsevier B.V. 2018.

MA R. et Al. A GaN PA for 4G LTE-Advanced and 5G: Meeting the Telecommunication Needs of Various Vertical Sectors Including Automobiles, Robotics, Health Care, Factory Automation, Agriculture, Education, and More. IEEE Microwave Magazine. Volume 18. Issue 7. Nov-Dec 2017.

Monitoramento Digital de Doenças em Tomateiros Baseado em Aplicativos Datacêntricos

Nicolas T. Maia[1], Guilherme Andrade[1]; Sérgio Manuel Serra da Cruz[1],

[1] Departamento de Computação, Universidade Federal Rural do Rio de Janeiro, Seropédica, Rio de Janeiro, Brasil, serra@pet-si.ufrrj.br

RESUMO

A agricultura e a pecuária brasileiras estão passando por uma revolução silenciosa, trata-se da Agricultura 4.0 (MASSRUHÁ e LEITE, 2017). Big Data, inteligência artificial, robôs, drones, nuvens de computadores e aplicativos móveis já podem fazer parte do trabalho no campo. Esses artefatos estão trazendo melhor monitoramento, gestão e controle da produção agrícola, com redução de custo e menos desperdícios. Mas há também limitações, por exemplo, a cobertura de telefonia móvel no campo é parcial e não há uma grande oferta de ferramentas de baixo custo voltadas para os pequenos produtores, em especial os produtores de tomate (*Solanum lycopersicon*).

A cadeia produtiva do tomate é provavelmente a mais importante entre todas as hortaliças cultivadas no estado do Rio de Janeiro. O estado produz aproximadamente 75 milhões de quilos. De acordo com a Empresa de Assistência Técnica e Extensão Rural do Estado do Rio de Janeiro, só no Ceasa, do Rio, sua comercialização movimentou R$ 73 milhões em 2010. Além disso, a demanda pelo tomate com reduzida carga de agroquímicos vem crescendo, portanto, a adoção de novas práticas da agricultura digital na gestão de propriedades e lavouras apoiadas por aplicativos se faz cada vez mais necessária.

Este resumo apresenta os avanços de arquitetura datacêntrica (centrada em tratamento de grandes volumes de dados) denominada *MyFarm++* proposta por (LOPES e CRUZ, 2018), ela é baseada em computação distribuída que armazena e compartilha dados georeferenciados sobre as culturas, safras, aplicações de agroquímicos e ocorrências de doenças e pragas em lavouras de tomateiros. A arquitetura suporta a geração de gráficos, mapas e relatórios acerca do planejamento e da produção e dos gastos com insumos de uma

propriedade. A arquitetura pode auxiliar agricultores e agrônomos para que gerenciem a produção, registrando e racionalizando o uso de insumos, possivelmente contribuindo com aumentando da produtividade, sustentabilidade e diminuindo a degradação do meio ambiente.

A metodologia adotada neste trabalho foi a da Fábrica de Software baseada em Métodos Ágeis (FSMA) (CRUZ et al. 2013) que alinha com a escola de pensamento proposta por Vygostky (2007) onde a educação é considerada como fonte de desenvolvimento. A ênfase da metodologia é capacitar os alunos-desenvolvedores para trabalharem em equipes multidisciplinares e despertar o interesse a adoção de boas práticas em TI e desenvolvimento ágil em atividades voltadas para a inovação tecnológica oriundas em ambientes de programas de educação tutorial (PET).

Dentre os primeiros resultados alcançados, destacamos a concepção da arquitetura ilustrada na Figura 1. As setas indicam os fluxos de dados semiestruturados entre o backend da ferramenta em ambientes na nuvem e o aplicativo *MyFarm*++. O aplicativo é apenas uma interface, sendo parte da ferramenta que é classificada como um sistema agrícola de próxima geração (ANTLE et al., 2017); é multiusuária, multiplataforma e de fácil utilização, utiliza os *frameworks* React-Native, *Google Maps, AWS Lambda Serverless Computing* e NodeJS (CouchDB), comporta microsserviços e opera sob paradigma dos *offline-first* permitindo operar em ambientes rurais com conectividade intermitente.

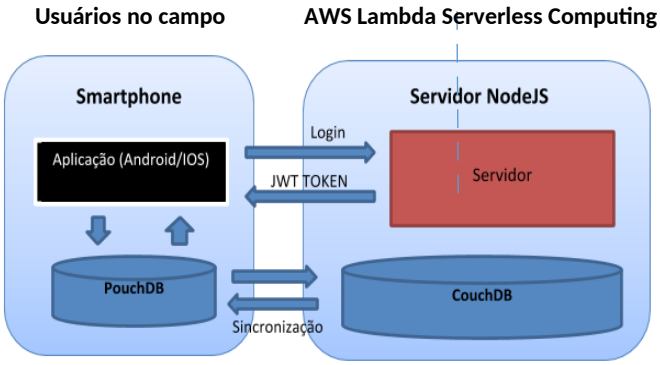
Através do aplicativo os usuários têm acesso a um menu (Figura 2) e *dashboard* onde é capaz de registrar de modo georreferenciado suas propriedades, talhões e plantas, registrar os acontecimentos referentes aos cultivos e safras. *MyFarm*++ permite armazenar/organizar as imagens de uma propriedade, talhões ou de ocorrências de pragas e doenças, podendo ser compartilhadas com serviços agronômicos credenciados e previamente registrados.

MyFarm++ computa dados históricos sobre perdas, sementes, insumos, equipamentos e índices de produtividade por safra. Oferece uma calculadora que calcula o quantitativo de sementes/há, pulverizações e regulagens para cada aplicação visando melhor o aproveitamento dos insumos na lavoura. A ferramenta produz relatórios e mapas georreferenciados com médias de pragas por área e um mapa de calor que exibe o nível e local exato da infestação, facilitando a percepção da rastreabilidade das doenças no decorrer das safras além das pesquisas por talhões onde foram registradas ocorrências e o acesso a informações históricas detalhadas.

MyFarm++ pode se conectar aos serviços de API de parceiros para obter informações regionalizadas, por exemplo, oferecendo notícias, mapas e previsão do tempo para intervalos

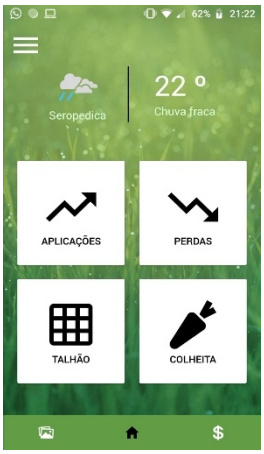
de tempo ou se conectando ao *backend* da plataforma *OpenSoils*[1] (CRUZ et al. 2018) para recuperar informações sobre tipos de solos, estudos de fertilidade da propriedade ou da região do entorno da propriedade.

Figura 1 - Diagrama do fluxo de dados da arquitetura.



Fonte: OS AUTORES (2019)

Figura 2 – Dashboard MyFarm++.



Atualmente a arquitetura está sendo desenvolvida e avaliada experimentalmente junto a pequenos produtores de tomates e serviços agronômicos do estado do RJ. Como trabalhos futuros estimasse desenvolver experimentos controlados da ferramenta.

AGRADECIMENTOS

O presente trabalho foi parcialmente realizado com apoio da Programa de Educação Tutorial (PET-SI), da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001 e do Conselho Nacional de Pesquisas (CNPq) – Código de Financiamento produtividade DT-II (315399/2018-0) e pela bolsa de PIBIC/UFRRJ.

REFERÊNCIAS

ANTLE J.M., et al. Next generation agricultural system models and knowledge products: synthesis and strategy. Agricultural Systems. V.155, July, p. 179-185, 2017.

CRUZ, S. M. S, et al. Relato de um experimento piloto de uma fábrica de software baseada em métodos ágeis. In: XVIII Encontro Nacional de Grupos PET, 2013, Recife. Anais do ENEPAET 2013, p. 1-5.

[1] www.opensoils.org

CRUZ, S. M. S, et al. Towards an e-infrastructure for Open Science in Soils Security. In: Brazilian E-Science Workshop (BRESCI), 12. , 2018, Natal. Anais do XII Brazilian e-Science Workshop. Porto Alegre: Sociedade Brasileira de Computação.

LOPES, B. C. do P.; CRUZ, S. M. S. MyFarm++: Monitoramento Online De Doenças Em Tomateiros Baseado Em Smartphones. In: VI Reunião Anual de Iniciação Científica da UFRRJ, 2018, Seropédica. Anais da RAIC 2018, p. 1-2.

MASSRUHÁ, S. M. F. S.; LEITE, M. A. de A. AGRO 4.0 no meio rural. Futuro das TICs na AGRO 4.0. 2017.

VYGOSTKY, L. S. Formação social da mente. Ed.Martins Fontes, 2ª Edição. 2007.



Utilizando processamento em cascata e agrupamento em imagens para otimizar modelos de classificação de solos

Gabriel Teston Vasconcelos[1], Kleber Xavier Sampaio de Souza[2], João Camargo Neto[3], Stanley Robson de Medeiros Oliveira[4]

[1] Instituto de Computação, Universidade Estadual de Campinas, Campinas, SP, Brasil, gtestonvasconcelos@gmail.com

[2,3,4] Embrapa Informática Agropecuária, Campinas, SP, {kleber.sampaio, joao.camargo, stanley.oliveira}@embrapa.br

RESUMO

O presente artigo propõe duas novas abordagens para a classificação automática de solos, com o objetivo de aperfeiçoar o processo classificatório. As abordagens consistem no processamento em cascata para cada nível categórico e a organização dos horizontes do um perfil em forma de imagens possibilitando a classificação de perfis de solo como um todo. Constatou-se uma leve melhora, contudo a mesma carrega um grande custo computacional.

PALAVRAS-CHAVE: Árvores de decisão, Floresta aleatória, SVM, Mineração de dados, Atributos de solo.

ABSTRACT

This article proposes two new approaches to the automatic soil classification, with the objective to improve the classification process. Those approaches consist in the cascate processing for each classification level and the organization of the soil horizons into an image being able to classify the soil profile as a whole. It was found some improvement but it comes with a huge computational cost.

KEYWORDS: Decision Tree, Random Forest, SVM, Data Mining, Soil Attributes.

INTRODUÇÃO

Com o intuito de aperfeiçoar a classificação automática de perfis de solos em aplicações práticas (Vasconcelos et al., 2018) propõe-se neste trabalho o processamento em cascata levando em consideração a hipótese que classificadores de níveis diferentes podem não concordar durante o processo de classificação.

Também se propõe uma maneira de obter a classificação de perfis de solos levando em consideração seus *n* horizontes simultaneamente através da representação dos atributos de forma contínua.

O objetivo final desta publicação é avaliar a viabilidade do uso das técnicas propostas em aplicações práticas de classificação automática de solos.

MATERIAL E MÉTODOS

Origem dos dados

Para o treinamento dos modelos analisados foram utilizados o dados providos pelo Mapeamento de Recursos Naturais do Brasil, disponíveis no Instituto Brasileiro de Geografia e Estatística (IBGE, 2018). Dos dados obtidos, foram considerados os referentes aos atributos físicos e químicos dentro do contexto da pedologia. Resultando em 23.534 horizontes com 95 atributos cada, referentes a 5013 perfis de solo.

Tratamento dos dados

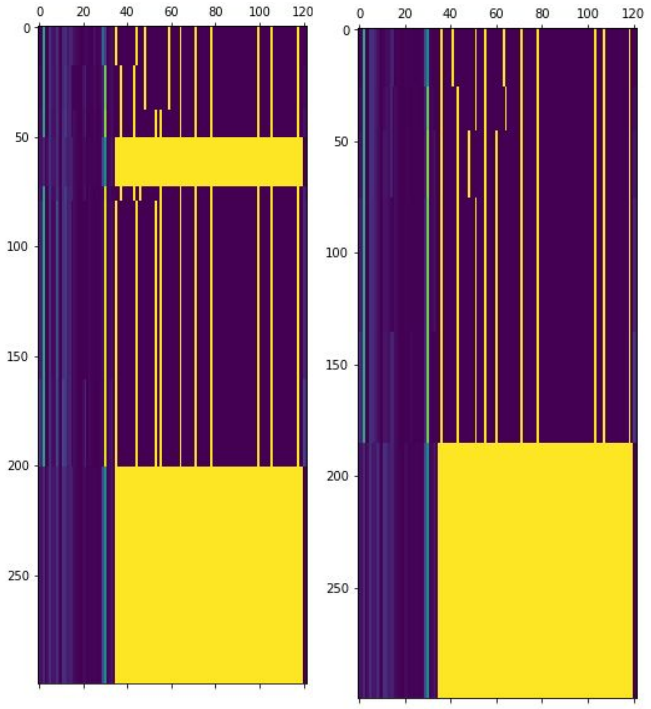
Dos 95 atributos originais foram removidos os atributos de identificação, que não são utilizados para classificação de solos e os atributos com mais de 50% de valores faltantes.

A base de dados é composta de atributos numéricos, assim como atributos categóricos. Para os atributos numéricos foi aplicada uma transformação comprimindo os valores entre 0 e 1 e valores faltantes substituídos pela média (Geron, 2017), para os valores categóricos foi aplicada a técnica de one-hot-bit encoding (Geron, 2017), transformando cada atributo categórico em n atributos binários, sendo n o número de categorias presentes no atributo original e valores faltantes representando uma das categorias. Também foram removidos os atributos categóricos com mais de 30 categorias, pois a sua presença apenas piorava a classificação final.

Por fim foram considerados os horizontes até os 3 metros de profundidade, dada que a classificação feita por pedólogos é feita considerando esse intervalo, para a criação de

“imagens” representando os perfis de maneira “contínua”. Onde cada faixa de valores descrita pelos atributos de profundidade de cada horizonte foi populada com os valores do horizonte a qual pertencia e em casos de valores faltantes completados com os valores centrais correspondentes. Cada imagem é composta de um canal, com dimensões de (300, 122) pixels como na Figura 1.

Figura 1 – Imagem representando perfis de solo de diferentes classes.



Cambissolo Latossolo

Fonte: (Autor, 2019)

Algoritmos de aprendizado de máquina

Foram utilizados três algoritmos de Aprendizado de Máquina (AM) todos contidos na biblioteca Scikit-learn (Pedregosa *et al, 2011*).

Árvore de Decisão (sklearn.tree.DecisionTreeClassifier)

Consiste na criação de um grafo direcional em formato de árvore começando da raiz, onde cada nó contém uma pergunta que leva ao próximo dependendo da resposta da mesma.

Floresta Aleatória (sklearn.ensemble.RandomForestClassifier)

Consiste na criação de um comitê de árvores de decisão, onde cada uma tem acesso a uma combinação aleatória dos atributos.

Máquinas de Vetor de Suporte (SVM) (sklearn.svm.SVC)

Consiste na divisão do espaço euclidiano gerado pelos atributos utilizando hiperplanos de separação ótima no intuito de dividir esse espaço nas classes desejadas.

Processamento dos dados em cascata

O processamento em cascata foi definido de maneira que a classificação do nível *n* depende dos atributos do perfil mais a classificação dos níveis anteriores. portanto a classificação nível 4 (Ordem, Subordem, Grande Grupo, Subgrupo), depende dos atributos do perfil mais a classificação dos níveis 1,2 e 3.

Diferentemente da abordagem anterior (Vasconcelos et al., 2018) onde se classificavam horizontes, cada modelo foi treinado para classificar um perfil, tornando o processo de classificação mais próximo ao realizado por pedólogos. O processamento em cascata possibilita que durante o processo de classificação o modelo que classifica o nível n+1 aperfeiçoe a classificação do modelo do nível anterior, uma vez que o de nível n+1 é mais especializado e capaz de verificar outras relações entre os atributos.

Validação dos modelos

A validação dos modelos foi feita em duas etapas, **Relativa** onde assumiu-se o acerto das classificações anteriores e **Absoluta** onde a classificação foi realizada a partir do processamento em cascata usando os modelos gerados, tendo como métrica de avaliação a acurácia.

Para a criação dos modelos foi usada uma divisão entre treino e teste de 70% dos perfis para treino e 30% para teste.

RESULTADOS E DISCUSSÃO

Com os modelos em casca treinados foi possível comparar o resultado na classificação e de cada algoritmo para cada nível de classificação com a classificação de um modelo treinado nos mesmos dados, porém que classifica cada nível individualmente, como pode ser visto nas Tabelas 1 a 3.

O ganho médio na acurácia utilizando o processamento em cascata foi de 4,3%, sendo possível notar que o modelo baseado em SVM foi que tirou maior proveito do processamento em cascata com um ganho médio na acurácia de 7,5%.

Tabela 1 – Acurácia obtida com Decision Tree para os diferentes métodos de classificação.

Nível	Tipo de classificação (Decision Tree)		
	Sem cascata	Com cascata (Relativa)	Com Cascata (Absoluta)
Ordem	0.532	0.532	0.532
Subordem	0.327	0.6613	0.360
Grande grupo	0.249	0.7184	0.303
Subgrupo	0.209	0.650	0.254

Tabela 2 – Acurácia obtida com Random Forest para os diferentes métodos de classificação.

Nível	Tipo de classificação (Random Forest)		
	Sem cascata	Com cascata (Relativa)	Com Cascata (Absoluta)
Ordem	0.674	0.674	0.674
Subordem	0.482	0.553	0.521
Grande grupo	0.402	0.473	0.438
Subgrupo	0.340	0.396	0.3512

Tabela 3–Acurácia obtida com SVM os diferentes métodos de classificação.

Nível	Tipo de classificação (SVM)		
	Sem cascata	Com cascata (Relativa)	Com Cascata (Absoluta)
Ordem	0.527	0.527	0.527
Subordem	0.330	0.355	0.354
Grande grupo	0.120	0.260	0.259
Subgrupo	0.094	0.231	0.231

Tabela 4 – Ganho de acurácia (Absoluta) com processamento em cascata.

Nível	Classificador		
	Decision Tree	Random Forest	SVM
Ordem	0.000	0.000	0.000
Subordem	0.033	0.039	0.024
Grande grupo	0.054	0.036	0.140
Subgrupo	0.045	0.012	0.137

Média	0.033	0.022	0.075

CONCLUSÕES

Mesmo com um possível ganho na acurácia com o processamento em cascatas não é possível constatar uma melhora significativa dessa técnica, tornando pouco viável sua utilização em na aplicação proposta em (Vasconcelos et al., 2018) uma vez que cada classificação em cascata chega a ser 4 vezes mais complexa computacionalmente.

Porém, foi possível estabelecer um método para a classificação de perfis, ao invés de horizontes, tornando a classificação automática de solos mais próxima à realizada por pedólogos.

AGRADECIMENTOS

Os autores agradecem ao programa CNPq/PIBIC pela concessão da bolsa de Iniciação Científica, processo Nº125043/2018-0 para o aluno Gabriel Teston Vasconcelos e à equipe do projeto SmartSolos da Embrapa pelo apoio oferecido durante o desenvolvimento.

REFERÊNCIAS

GERON, A. Hands-On Machine Learning with Scikit-Learn & Tensor Flow - Concepts, Tools and Techniques to Bild Intelligent Systems. O'Reilly Media Inc, Sebastopol, USA. 2017.

PEDREGOSA, F.; VAROQUAUX, G.; GRAMFORT, A.; MICHEL, V.; THIRION, B.; GRISEL, O.; BLONDEL, M.; PRETTENHOFER, P.; WEISS, R.; DUBOURG, V.; VANDERPLAS, J.; PASSOS, A.; COURNAPEAU, D.; BRUCHER, M.; PERROT, M.; DUCHESNAY, E., Scikit-learn: Machine Learning in Python, Pedregosa et al., JMLR 12, pp. 2825-2830, 2011.

VASCONCELOS, G. T.; SOUZA, K. X. S. de; OLIVEIRA, S. R. de M.; CAMARGO NETO, J. Montagem de ambiente para classificação de solos usando ScikitLearn. In: MOSTRA DE ESTAGIÁRIOS E BOLSISTAS DA EMBRAPA INFORMÁTICA AGROPECUÁRIA, 14., 2018, Campinas. Resumos expandidos... Brasília, DF: Embrapa, 2018. p. 104-110. (Embrapa Informática Agropecuária. Eventos técnicos & científicos, 1). Editores técnicos: Carla

Geovana do Nascimento Macário, Carla Cristiane Osawa, Flávia Bussaglia Fiorini, Maria Fernanda Moura, Poliana Fernanda Giachetto.

Estudo de ferramentas para design e modelagem de sistemas de serviço aplicado ao Agronegócio

Maria de Jesus de Souza[1], Amanda Roberta Camargo.[1], Valter Castelhano de Oliveira[1], Edson Luiz Pereira[1]

[1] Faculdade de Tecnologia Dr. Achimedes Lamoglia – Indaiatuba, São Paulo, Brasil,

mariadejesusdesouza599@gmail.com, amanda.camargo6@fatec.sp.gov.br,

voliveira@fatecindaiatuba.edu.br, edson.luiz@fatec.sp.gov.br

RESUMO

O objetivo deste trabalho é estudar modelagem e ferramentas de suporte ao design de sistema de serviço, buscando elaboração de material instrucional de apoio ao estudo e a utilização do ambiente de modelagem Enterprise Architect (EA) e do *Service Oriented Modeling Framework* (SOMF). A metodologia apresenta pesquisa bibliográfica sobre a mudança de paradigma da orientação a bens tangíveis para a orientação a serviços, juntamente com as consequências para o processo de modelagem de sistemas de serviço, e pesquisa de campo na empresa Agritech, fabricante de equipamentos agrícolas, visando aplicar os conceitos de modelagem de sistemas de serviço, recursos do EA e SOMF, visando compreender e modelar o relacionamento entre fábrica, concessionário e produtor agrícola no tocante ao pós-venda de equipamentos agrícolas. Como resultados foram obtidos o manual de recursos básicos da ferramenta EA, o manual do SOMF e amodelagem do sistema de serviço associado ao relacionamento das concessionárias e a área de pós-vendas da Agritech.

PALAVRAS-CHAVE: sistemas de serviço, modelagem de sistemas, enterprise architect, SOMF, equipamentos agrícolas.

ABSTRACT

The objective of this work is to study modeling and tools to support service system design, seeking elaboration of instructional material to support the study and utilization of Enterprise Architect (EA) modeling environment and the *Service Oriented Modeling Framework* (SOMF). The methodology brings literature search about the goods oriented paradigm change

to a services oriented paradigm, with the consequences to the service modeling systems process, and a field research at Agritech, agricultural equipments manufacturer, aiming to apply the services systems modeling concepts, EA resources and SOMF and, to understand and model the relationship between factory, dealers and producers related to post sales of agricultural equipments. As results, were developed the EA tool Basic Resources Manual, the SOMF Manual and the modeling of the service system associated to the relationship between dealers and the post sales area at Agritech.

KEYWORDS: service systems, systems modeling, enterprise architect, SOMF, agricultural equipment.

INTRODUÇÃO

Desde o início deste século ocorre a mudança de paradigma de uma abordagem dominada por bens materiais (lógica G-D) para uma abordagem dominada por serviços (*service-dominant logic* ou lógica S-D) (LUSCH; VARGO; WESSELS, 2008; MOUSSA; TOUZANI, 2010).

Nesse contexto os sistemas de serviço vêm se tornando foco de atenção pelo seu aspecto integrador de inovação tecnológica, automação e controle de forma ubíqua (OSTROM et al., 2010, 2015). Entretanto, a automação e abstração que estes sistemas provêm dependem tanto da inovação tecnológica quanto das soluções adotadas para o gerenciamento de recursos, da informação, e das restrições inerentes à aplicação. Portanto, as técnicas clássicas aplicadas ao desenvolvimento de sistemas devem ser revistas para se adequar aos novos desafios propostos para o desenvolvimento destes novos sistemas, os quais passam a ser o meio de sustentação da lógica S-D, sendo caracterizados pela configuração dinâmica de pessoas, de tecnologias, de organizações e, também, das informações compartilhadas entre estes elementos, que permite aos fornecedores criar e entregar valor aos seus clientes.

O conceito chave associado a sistemas de serviço é que estes interagem para a co-criação de valor envolvendo fornecedores e clientes em um mesmo processo (MAGLIO et al., 2009; SPOHRER et al., 2007).

Entretanto é comum deparar-se com a questão de como viabilizar a automação envolvendo máquinas, pessoas e processos como recursos, onde o produto final agora é o serviço (intangível, inseparável, variável e perecível) (GRÖNROOS, 2004; LOVELOCK; WRIGHT, 2002; VARGO; LUSCH, 2006). Todo esse quadro potencializa a importância de tratar o design dos sistemas de serviço de forma sistemática (BARDHAN et al., 2010;

BARILE; POLESE, 2010; KIM; NAM, 2009; MAGLIO et al., 2009; OSTROM et al., 2010, 2015; SPOHRER et al., 2007; STANICEK; WINKLER, 2010).

O uso de modelos no processo de design de sistemas de serviço segue a tendência atual proposta pela OMG[1] e pelo INCOSE[2]. A aderência do design por modelos (*model driven design*) (KENT, 2002) a sistemas de serviço se justifica pela premissa de que o processo de automação aplicado aos sistemas de serviço seja implementado por meio dos sistemas de informação, como uma forma interativa e que admite: co-criação, flexibilidade, aderência a requisitos voláteis, aderência a sistemas com um alto número de *viewpoints* e facilidade de reutilização.

Neste trabalho parte-se do pressuposto que a pesquisa permita relacionar os conceitos associados a sistemas de serviço e necessidade de modelagem e design destes sistemas aos recursos oferecidos pelas ferramentas de modelagem por meio da análise da documentação existente, prioritariamente em língua inglesa, e resultar uma nova documentação em língua portuguesa aplicável a um processo de disseminação da utilização das ferramentas de modelagem de sistemas de serviço, tomando sempre como referência o pós-venda de equipamentos agrícolas e os relacionamentos entre a fabricante Agritech, seus concessionários e os clientes, representados pelos produtores agrícolas.

MATERIAL E MÉTODOS

Nesta pesquisa foram aplicadas basicamente duas técnicas de pesquisa: bibliográfica e análise de conteúdo (MARCONI; LAKATOS, 2011). A pesquisa bibliográfica visa abordar os conceitos associados a sistemas em geral e, especificamente, sistemas de serviço, no intuito de identificar as necessidades de modernização dos processos de modelagem e design destes sistemas em busca de métodos de design e engenharia de sistemas de serviços.

A análise de conteúdo foi aplicada para tratar de forma objetiva e sistemática o conteúdo de livros, publicações, manuais e documentos técnicos associados aos ambientes e ferramentas de modelagem de sistemas a fim de nortear a elaboração da primeira versão do manual de recursos básicos.

O projeto também contou com a pesquisa de campo descritiva exploratória para conhecimento e documentação do sistema, gerando insumo para a eliciação e análise dos requisitos do negócio, nesse caso o serviço pós-venda da Agritech, empresa de máquinas agrícolas localizada na Região Metropolitana de Campinas (RMC). Foram realizadas entrevistas desestruturadas com os colaboradores do setor de pós-venda da empresa.

[1] OMG - *Object Management Group* (http://www.omg.org/).
[2] INCOSE - The International Council on Systems Engineering (http://www.incose.org).

As entrevistas tiveram como foco identificar os envolvidos no sistema de serviço pós-venda e os processos pertencentes a ele, buscando uma visão de como o sistema é (*as-is*) (CRUZ, 2015).

Inicialmente estudaram-se os conceitos associados a sistemas de serviço e os métodos e técnicas empregados na modelagem de sistemas, por intermédio de pesquisa bibliográfica.

A próxima etapa inicio-se com a análise do conteúdo de comunicação associado ao Enterprise Architect – EA (SPARX, 2018), selecionando e classificando a documentação disponível (majoritariamente em língua inglesa), criando padrões para análise e o consequente registro dos resultados obtidos por meio de captura de tela, visando a realização de práticas que permitam compreender o funcionamento da ferramenta EA e os recursos de modelagem de sistemas, bem como, elaborar documentação em língua portuguesa - Manual de Recurso Básicos EA[3] (SOUSA, 2017),

Na etapa seguinte, o foco foi realizar a análise de conteúdo de comunicação associado ao *Service Oriented Modeling Framework* – SOMF (BELL, 2008), também majoritariamente em língua inglesa, visando a realização de práticas que sustentem a compreensão e utilização do SOMF no desenvolvimento de sistemas de serviço por meio de padrões de análise e registro de resultados e, a coleta e análise de conteúdos para a elaboração de material instrucional visando o aprendizado de tais ferramentas em formato de curso EAD. Uma versão completa do Manual de Recursos Básicos do SOMF[4] (SOUSA, 2018) também foi desenvolvida e concluída, mostrando sua aplicação na modelagem de um sistema de serviço simplificado do agronegócio, fundamentada pelos princípios de modelagem de sistemas de serviços de Michel Bell (2008) e tendo como referência a análise de diversos documentos disponibilizados em língua inglesa pela Sparx Systems[5] sobre o módulo SOMF do EA.

Como padrão de análise e registro de resultados foi selecionada a metodologia DOMP (CRUZ, 2015). Todos os processos relacionados ao sistema de serviços pós-venda foram identificados e desenhados na ferramenta Bizagi *Modeler*[6] com a notação BPMN[7] (*Business Processes Model and Notation)*.

[3] Manual de Recursos Básicos EA – Disponibilizado no ambiente de apoio educacional da Fatec Indaiatuba e fornecido à SPARX pelo acordo de licenciamento educacional.

[4] Manual de Recursos Básicos SOMF - Disponibilizado no ambiente de apoio educacional da Fatec Indaiatuba e fornecido à SPARX pelo acordo de licenciamento educacional.

[5] *Sparx System - SOMF in Enterprise Architect* (http://www.sparxsystems.com/somf)

[6] Bizagi (http://www.bizagi.com/pt)

[7] BPMN - *Business Processes Model and Notation* (http://www.bpmn.org/)

O projeto de iniciação tecnológica foi realizado na Fatec Indaiatuba e envolveu os seguintes recursos: computadores do laboratório de informática da Fatec Indaiatuba; 30 licenças do EA, obtidas com acordo entre a Fatec Indaiatuba e a Sparx (SPARX, 2018), que viabiliza instalação gratuita nos computadores dos laboratórios da Fatec e disponibiliza vasta documentação (manuais, artigos, tutoriais e relatórios técnicos), originalmente na língua inglesa; dois alunos do curso de Tecnologia em Gestão de Serviços da Fatec Indaiatuba, sendo um aluno com bolsa PIBIT-CNPq; e Docentes da Fatec Indaiatuba atuando como orientador e coorientador.

A comunicação no projeto ocorre via e-mail institucional da FATEC-ID, reuniões de acompanhamento, seminários e a ferramenta de colaboração acadêmica Mendeley[8]

RESULTADOS E DISCUSSÃO

A elaboração do material de apoio ao estudo e utilização do ambiente de modelagem Enterprise Architect – EA foi realizado por meio do registro e documentação das atividades desenvolvidas na prática de tais ferramentas resultando no manual do usuário do EA (SOUSA, 2017), utilizado regularmente como referência em disciplinas do curso de Gestão de Serviços.

O manual concentra-se no uso de tutoriais disponíveis no site da Sparx Systems (SPARX, 2018) e informações disponibilizadas junto com a ferramenta EA, visando mostrar sua prática e não apenas uma tradução dos conteúdos existentes, salientando-se o uso de termos em inglês como coletados de sua origem e, em alguns momentos, intercalados com suas respectivas versões em português visando um melhor aproveitamento do material pelo usuário final.

O manual apresenta um roteiro prático de utilização da ferramenta com explicações detalhadas de como criar um projeto, diagramas, pacotes e utilizar os elementos e recursos de construção de cada modelo.

A modelagem orientada a serviços SOMF foi proposta por Michel Bell (2008) como uma linguagem de modelagem holística e antropomórfica para o desenvolvimento de software que emprega a disciplina e uma linguagem universal para fornecer soluções táticas para problemas empresariais.

Os estudos sobre o SOMF tem uma visão aplicada à modelagem do sistema de serviço de pós venda da Agritec e primeiramente foram estudados os conceitos do SOMF presentes em três documentos, respectivamente *Service-Oriented Modeling Framework* (SOMF™)[9], SOMF - 2011 - *Service-Oriented Conceptualization Model Language Specifications* e SOMF

[8] Mendeley (https://www.mendeley.com/)

- 2011 - *Constructing a Service-Oriented Attribution Model*. Tais documentos possibilitaram aplicar os conceitos de forma prática utilizando o EA (SPARX, 2018), pois este implementa integralmente as especificações e recursos do SOMF.

O resultado desta etapa do trabalho foi o Manual de Modelagem SOMF (SOUSA, 2018), utilizado regularmente em disciplinas do curso de Gestão de Serviços da Fatec Indaiatuba.

Modelagem sistema de serviço pós-venda

A pesquisa teve como objeto, a empresa Agritech (Indaiatuba/SP), especificamente o departamento de Pós-vendas, e uma concessionária. A Agritech desenvolve e produz tratores e implementos agrícolas e, sua produção é voltada para a agricultura familiar.

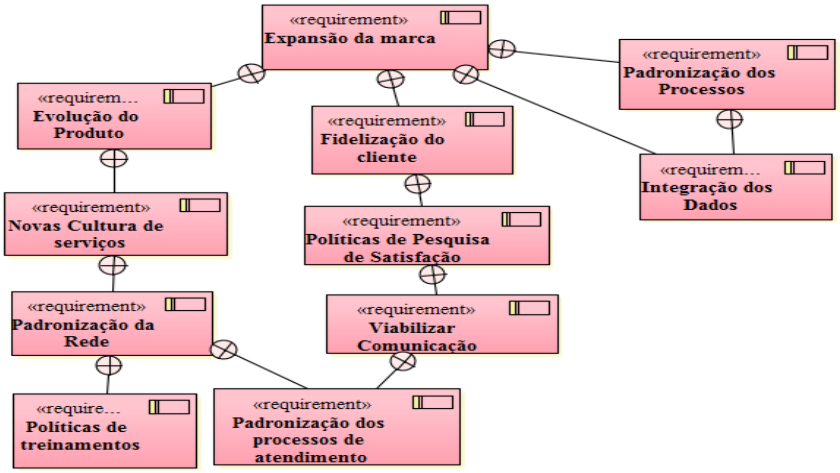
As unidades de revenda são heterogenias, contando com concessionárias e revendedores autorizados, dificultando a padronização. As concessionárias comercializam desde pequenos a grandes tratores numa determinada região de atendimento, enquanto, o revendedor vende apenas cultivadores motorizados e, não há nenhuma limitação quanto a região de atendimento. A fábrica não tem contato direto como agricultor, apenas por meio das concessionárias.

A primeira etapa da modelagem de um sistema de serviço é compreender os processos necessários para que ele funcione, os agentes envolvidos e a dinâmica adotada pelos prestadores do serviço, ou seja, os componentes do sistema, visando a eliciação e analise de seus requisitos (OLIVEIRA, 2013).

A Figura 1 apresenta o diagrama de requisitos SysML (*System Modeling Language*) (FRIEDENTHAL; MOORE; STEINER, 2009), resultado da eliciação de requisitos junto a equipe de pós-vendas de equipamentos agrícolas da Agritech e representa o ponto de vista da fábrica.

[9] *Service-Oriented Modeling Framework* (SOMF™) (https://www.sparxsystems.com.au/downloads/whitepapers/SOMF-2.1-Architecture-Model-Language-Specifications.pdf/)

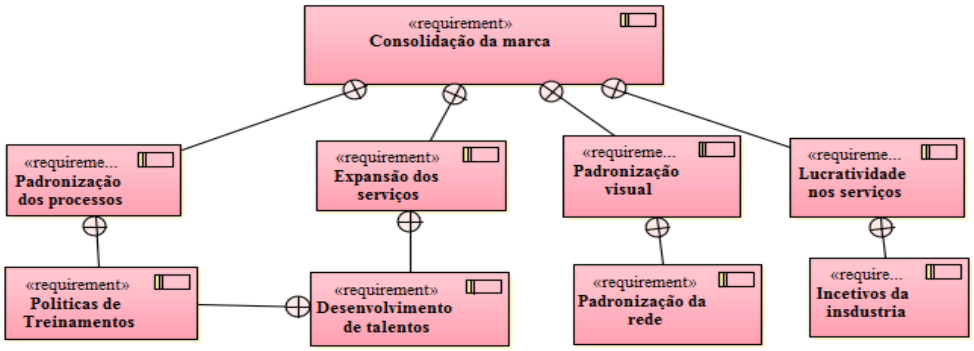
Figura 1 - Requisitos do Pós-vendas da empresa Agritech



Fonte: autoria própria

A fábrica da Agritech é dependente da concessionária, para que possa se manter dentro do mercado competitivo, sendo a concessionária o meio de comunicação com o agricultor, que é o cliente final. A Figura 2 apresenta o diagrama de requisitos sob o ponto de vista da concessionária, extraídos das entrevistas com o dono da concessionária e vendedor.

Figura 2 - Requisitos da concessionária

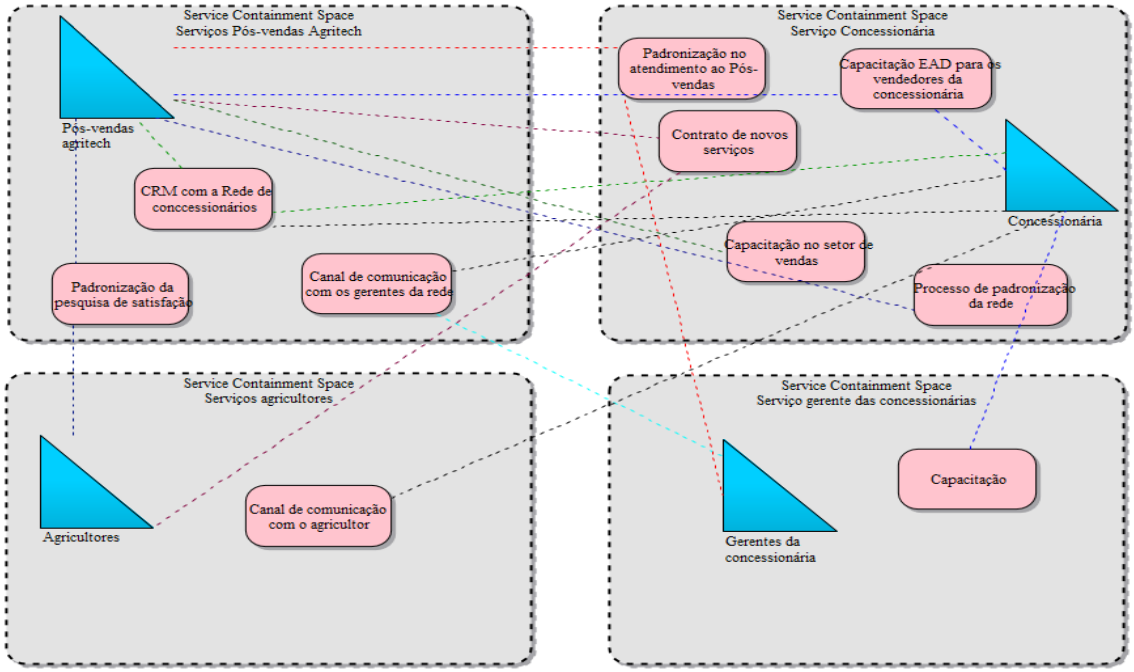


Fonte: autoria própria

O modelo de atributos SOMF é um diagrama, no qual se define os atributos principais em termos de características, qualidade e propriedades (OLIVEIRA, 2013), ou seja, são as descrições propostas que o serviço visa atender. Considerando-se os requisitos e a documentação obtida, para o pós venda da Agritech foram selecionado os atributos padronização, comunicação, talentos e rentabilidade, os quais auxiliam nas decisões e priorizações no decorrer da modelagem SOMF, principalmente, como base para a identificação dos serviços.

A Figura 3 apresenta o modelo de associação conceitual, contendo os serviços (retângulos na notação SOMF) relacionados com os consumidores (triângulos) e os espaçoes de contenção associados aos consumidores (retângulos em cinza).

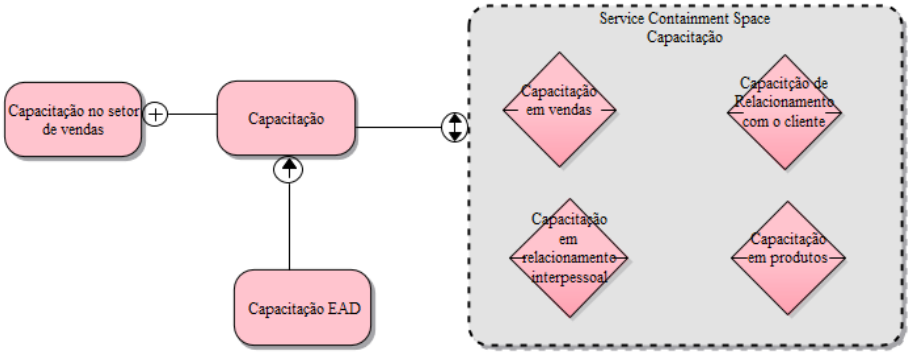
Figura 3 - Diagrama de associação conceitual



Fonte: autoria própria

A Figura 4 apresenta o diagrama de análise conceitual SOMF que alinha os serviços mostrando apenas a identidade do serviço Capacitação, representando inclusive os serviços primitivos que o compõem (losangos na notação SOMF). Este modelo apresenta a estrutura interna do serviço Capacitação, fornecendo sua identidade e o contexto, possibilitando alinhar os novos serviços com as necessidades de negócios da fábrica Agritech e das concessionárias.

Figura 4 - Diagrama conceitual do serviço

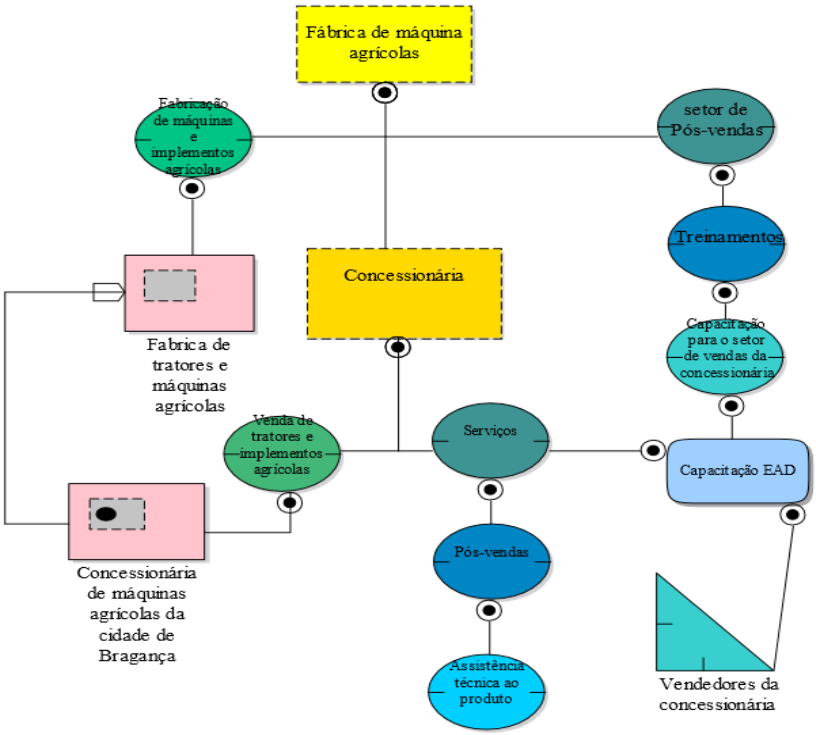


Fonte: autoria própria

A Figura 5 apresenta o diagrama de integração geográfica de negócio da fábrica Agritech e das concessionárias, juntamente aos impactos do novo serviço modelado dentro do contexto empresarial da fábrica. Este modelo oferece dois tipos de modelagens: contextual e estrutural. O contextual refere-se aos imperativos de negócios da empresa tais como, o que ela vende, para quem ela vende e como ela gera renda. O estrutural trabalha com a arquitetura de

negócio, representando os vários segmentos da empresa, por meio de seus domínios de negócios, eles representam as linhas de negócios da empresa como ocupação, produtos, departamentos divisões locais e geográficas (BELL, 2008). A modelagem desse diagrama trabalha em camadas lógicas e físicas, a camada lógica refere-se a hierarquia dos domínios de negócios da empresa, a camada física está relacionada ao todo da empresa, ou seja, é considerada a entidade organizacional.

Figura 5 - Diagrama de integração regional de negócio



Fonte: autoria própria

CONCLUSÕES

A pesquisa concentrou-se no estudo sobre os conceitos associados a sistema de serviço e os métodos e técnicas empregados na modelagem de sistemas, por meio da qual se fez a análise dos conteúdos de comunicação associados ao Enterprise Architect, no decorrer do projeto foi possível perceber que o conhecimento prévio dessa ferramenta contribuiu de forma significativa para a prática dos conceitos.

Os manuais e tutorias sobre o EA disponíveis no site da Sparx foram de grande ajuda, uma vez que forneceram insumos para as práticas que permitiram compreender o funcionamento da ferramenta e os recursos de modelagem de sistemas, bem como, elaborar a documentação em língua portuguesa, ou seja, o manual do usuário EA.

As visitas e entrevistas realizadas no ambiente da empresa Agritech foram de fundamental importância para entendimento e aplicação dos recursos associados às ferramentas de modelagem de sistemas estudados no presente projeto.

Durante a pesquisa foram percebidas necessidades de estudos além do esperado, como conteúdos voltados a modelagem de processos, que geraram subsídios para o início da modelagem do sistema de serviço pós-venda voltado ao Agronegócio. Os estudos voltados a elaboração do material instrucional ocorreu de maneira simultânea a elaboração do manual, visando que o mesmo pudesse ser aproveitado para a produção do curso EAD.

O projeto está em continuidade com a participação de novos alunos pesquisadores, os quais estão reutilizando os materiais e conhecimentos já produzidos para reformular o manual de recursos básicos do EA com a nova versão da ferramenta e dar continuidade ao conteúdo produzido acerca do framework SOMF visando a finalização do manual de usuário.

REFERÊNCIAS

BARDHAN, I. R.; DEMIRKAN, H.; KANNAN, P. K.; KAUFFMAN, R. J.; SOUGSTAD, R. An Interdisciplinary Perspective on IT Services Management and Service Science. **Journal of Management Information Systems**, [s. l.], v. 26, n. 4, p. 13–64, 2010. Disponível em: <http://mesharpe.metapress.com/openurl.asp?genre=article&id=doi:10.2753/MIS0742-1222260402>

BARILE, S.; POLESE, F. Smart Service Systems and Viable Service Systems: Applying Systems Theory to Service Science. **Service Science Journal**, [s. l.], v. 2, n. 1/2, p. 21–40, 2010. Disponível em: <http://www.sersci.com/ServiceScience/upload/12704925720.pdf>. Acesso em: 29 jul. 2010.

BELL, M. **Service-Oriented Modeling: Service Analysis, Design, and Architecture**. New Jersey: Wiley, 2008.

CRUZ, T. **Sistemas, Métodos & Processos: Administrando organizações por meio de processos de negócios**. 3o edição ed. São Paulo, SP, Brasil. NULL.

FRIEDENTHAL, S.; MOORE, A.; STEINER, R. **A Practical Guide to SysML: The Systems Modeling Language**. First edit ed. San Francisco, CA: Morgan Kaufmann, 2009.

GRÖNROOS, C. **Marketing : gerenciamento e serviços**. Rio de Janeiro: Elsevier/Campus, 2004.

KENT, S. Model driven engineering. In: INTEGRATED FORMAL METHODS 2002, **Anais**... : Springer, 2002. Disponível em: <http://www.springerlink.com/index/9VUQB4HP8FYG2ADV.pdf>. Acesso em: 1 ago. 2011.

KIM, Y. J.; NAM, K. Service Systems and Service Innovation: Toward the Theory of Service Systems. In: AMCIS 2009 PROCEEDINGS 2009, **Anais**... [s.l: s.n.]

LOVELOCK, C.; WRIGHT, L. **Principles of Service Marketing and Management**. Upper Saddle River, NJ: Prentice Hall, 2002.

LUSCH, R.; VARGO, S.; WESSELS, G. Towards a conceptual foundation for service science: Contributions from service-dominant logic. **IBM Systems Journal**, [s. l.], p. 1–9, 2008.

MAGLIO, P. P.; VARGO, S. L.; CASWELL, N.; SPOHRER, J. The service system is the basic abstraction of service science. **Information Systems and e-Business Management**, [s. l.], v. 7, n. 4, p. 395–406, 2009. Disponível em: <http://www.springerlink.com/index/10.1007/s10257-008-0105-1>

MARCONI, M. de A.; LAKATOS, E. M. **Técnicas de pesquisa**. 7a edição ed. São Paulo, SP, Brasil: Editora Atlas, 2011.

MOUSSA, S.; TOUZANI, M. A literature review of service research since 1993. **Journal of Service Science**, [s. l.], v. 2, n. 2, p. 173–212, 2010. Disponível em: <http://www.springerlink.com/index/10.1007/s12927-010-0008-9>. Acesso em: 16 nov. 2012.

OLIVEIRA, V. C. **Modelagem e design de sistemas de serviço para automação**. 2013. Escola Politécnica da Universidade de São Paulo, [s. l.], 2013.

OSTROM, A. L.; BITNER, M. J.; BROWN, S. W.; BURKHARD, K. a.; GOUL, M.; SMITH-DANIELS, V.; DEMIRKAN, H.; RABINOVICH, E. Moving Forward and Making a Difference: Research Priorities for the Science of Service. **Journal of Service Research**, [s. l.], v. 13, n. 1, p. 4–36, 2010.

OSTROM, A. L.; PARASURAMAN, A.; BOWEN, D. E.; PATRICIO, L.; VOSS, C. a. Service Research Priorities in a Rapidly Changing Context. **Journal of Service Research**, [s. l.], v. 18, n. 2, p. 127–159, 2015. Disponível em: <http://jsr.sagepub.com/cgi/doi/10.1177/1094670515576315>

SOUSA, M. de J. De. **Manual de Recursos Básicos do Enterprise Architect 12.1**, Fatec Indaiatuba - Curso de Gestão de Serviços, 2017. Disponível em: <http://moodle.fatecid.com.br/mod/resource/view.php?id=5613>

SOUSA, M. de J. De. **Manual - Modelagem SOMF**, Fatec Indaiatuba - Curso de Gestão de Serviços, 2018. Disponível em: <http://moodle.fatecid.com.br/mod/resource/view.php?id=5919>

SPARX. **Sparx Systems Pty Ltd.** 2018. Disponível em: <http://www.sparxsystems.com.au/>. Acesso em: 20 jun. 2018.

SPOHRER, J.; MAGLIO, P. P.; BAILEY, J.; GRUHL, D. Steps Toward a Service Systems. **IEEE Computer Society**, [s. l.], n. January, p. 71–77, 2007.

STANICEK, Z.; WINKLER, M. Service Systems Through The Prism of Conceptual Modeling. **Service Science Journal**, [s. l.], v. 2, p. 112–125, 2010.

VARGO, S. L.; LUSCH, R. F. Service-dominant Logic: What It Is, What It Is Not, What It Might Be. In: VARGO, S. L.; LUSCH, R. F. (Eds.). **The Service-dominant Logic of Marketing: Dialog, Debate, and Directions**. Armonk: M.E. Sharpe, 2006. p. 43–56.

Development and analysis of a labeled dataset of Brazilian agricultural market news titles for sentiment analysis

Tiago Santa Maria Rodrigues Marto[1], Roberto Fray da Silva[1], Angel Felipe Magnossão de Paula[1], Anna Helena Reali Costa[1], Carlos Eduardo Cugnasca[1]

[1] Departamento de Engenharia de Computação e Sistemas Digitais, Escola Politécnica da Universidade de São Paulo, São Paulo, São Paulo, Brasil, tiago.marto@usp.br, roberto.fray.silva@gmail.com, angel.magnossao@gmail.com, anna.reali@usp.br, carlos.cugnasca@usp.br

RESUMO

A análise de sentimentos pode ser definido como o uso de técnicas de processamento de linguagem natural para identificar, extrair e quantificar informação subjetiva e estados emocionais a partir de dados textuais de diferentes origens. No entanto, não existem *datasets* no domínio agrícola de alta qualidade que possuam valores de sentimentos. O principal objetivo deste artigo é criar e analisar, dos pontos de vista estatístico e linguístico, um *dataset* de títulos de notícias dos mercados agrícolas em português. Foram utilizadas três fontes de notícias de mercado confiáveis e relevantes. O *dataset* contém 335 notícias, de 01-03-2019 a 31-05-2019, rotuladas em duas dimensões: valor de sentimento e urgência. A primeira é relacionada ao impacto esperado das notícias no mercado agrícola e contém cinco categorias: muito negativo, negativo, neutro, positivo e muito positivo. A segunda é relacionada à probabilidade de que a notícia em questão impacte no mercado agrícola no dia de sua publicação e contém duas categorias: urgente e não urgente. A maioria das notícias é positiva e não urgente, refletindo a situação do mercado no período em questão. O *dataset* pode ser utilizado para estimar o sentimento do mercado e melhorar tanto previsões de mercado quanto a tomada de decisões.

PALAVRAS-CHAVE: Agricultura, Análise de sentimentos, Inteligência artificial

ABSTRACT

Sentiment analysis can be defined as the use of natural language processing techniques to identify, extract and quantify subjective information and emotional states from textual data from different sources. Nevertheless, there are no high-quality datasets with sentiment values for the agricultural domain. The main objective of this paper is to build and analyze, from the statistical and linguistic points of view, a dataset of agricultural market news titles in Brazilian portuguese. Three trustworthy and relevant market news sources were used. The dataset contains 335 news from 01-03-2019 to 31-05-2019, which were labeled in two dimensions: sentiment value and urgency. The first is related to the expected impact of the news on the agricultural market, and contains five categories: very negative, negative, neutral, positive, and very positive. The second is related to the probability of impacting the agricultural market at the same day that the news was released, and contains two categories: urgent and non-urgent. The majority of the news is positive and non-urgent, reflecting the market situation at the period. This dataset can be used to estimate market sentiment and help on improving market prediction and decision making.

KEYWORDS: Agriculture, Sentiment analysis, Artificial intelligence

INTRODUCTION

The agroindustrial domain comprises all companies involved in activities related to food production, from production at the farm to processing, transportation, warehousing and retailing. The companies, also referred to as agents, and their relations, are the field of study of supply chain management. This considers all companies that are directly or indirectly involved in the fulfillment of a customer's group demands (CHOPRA, MEINDL, 2013).

One important aspect of this domain, that directly influences decision making, is the market sentiment. Sentiment analysis, also referred to as opinion analysis, is the use of natural language processing techniques to identify, extract and quantify subjective information and emotional states from textual data from different sources, such as news, message boards, social media, among others (LIU, 2012; AGGARWAL, ZHAI, 2012). Several domains explores these techniques, such as: finance, using sentiment analysis to evaluate the market sentiment (REN, WU, LIU, 2018; NASSIRTOUSSI et. al, 2014); product reviews, to evaluate how customers rank different products (MUKHERJEE, BHATTACARYYA, 2012); movie reviews, to help to predict the box office (PANG, LEE, 2004); among others.

Sentiment analysis applied on agricultural market news could help to better understand what is the current sentiment on the agroindustrial domain, providing useful information for decision making. Nevertheless, there are almost no papers that explore sentiment analysis on the agroindustrial domain. One such example is Froehlich et. al (2017), which evaluates the worldwide public perception in the aquaculture sector. This lack of papers is even more present on the agroindustrial domain in Brazil.

One of the reasons that explain this lack of papers is the fact that are, to the best of our knowledge, no open datasets for agricultural market news in Brazil. Without a labeled, high-quality dataset, it is not possible to use sentiment analysis to estimate market sentiment. Our objective in this paper is to develop and analyze a labeled dataset with agricultural market news, which can be used for sentiment analysis on this domain.

The news were gathered from three relevant and trustworthy news sources (Valor Econômico, UOL Economia and InfoMoney), and were labeled independently by three researchers on two dimensions: sentiment value and urgency. The first dimension is related to five sentiment classes (very negative, negative, neutral, positive and very positive), as several sentiment analysis papers show that only three sentiment classes (negative, neutral and positive), as used in most of the literature, may not be enough to capture all the sentiments in a text (BOLLEN, MAO, ZENG, 2011; NASSIRTOUSSI et. al, 2014).

The second dimension is related to the subjective probability that the specific news will impact on the agricultural market on the same day of its release. Even though our dataset can be used for both regression and classification problems, because it provides the individual sentiment scores as well as the aggregated ones, we believe it will provide better results on classification tasks. Finally, we conduct an exploratory analysis of the dataset, considering both statistical and natural language processing aspects, and provide suggestions for its use.

METHODOLOGY

The methodology used in this paper contained five main steps:

1. **Identification of the main market news sources that provide publicly available news on the agricultural domain.** Three sources were identified: Valor Econômico (https://www.valor.com.br/), InfoMoney (https://www.infomoney.com.br/) and UOL Economia (https://economia.uol.com.br/). These are considered trustworthy and among the most relevant news sources for financial and market information;

2. **Data collection.** First, we developed multiple web crawlers with the scrapy library (https://scrapy.org/), considering the different websites and their sections. Then, we used them to gather agricultural market news titles from the sources identified in Step 1, from 01-03-2019 to 31-05-2019. It is important to observe that only news that were publicly available were gathered;
3. **Data preprocessing.** In this step, we evaluated the dataset and removed incorrect, irrelevant and redundant news. The main criteria used for identifying irrelevant news was: spams and news that had not even a remote connection with agriculture and the agricultural markets. The main library used in this Step was Pandas (https://pandas.pydata.org/);
4. **Data labeling,** using Pandas and Google Sheets (https://docs.google.com/spreadsheets/u/0/). In this Step, we labeled the data on two main dimensions: (i) urgency, related to the probability that the specific news would impact on the market at the day of its release; and (ii) sentiment value, considering the most probable impact on the market of the specific news. The first one consisted of two classes (high and low probability). The second one consisted of five categories (very negative, negative, neutral, positive, and very positive). Every data point was independently labeled on the two dimensions by three researchers. Then, the resulting labels for each data point in each dimension were calculated. For urgency, the final label was the majority of the labels for that data point. For sentiment value, a rounded up simple average for the labels of the data point was used**;**
5. **Exploratory data analysis.** This step consisted of an analysis of the distribution of the labels on the whole dataset and aggregated by day, using a simple average of the data points of that day. We also analyzed the content of the news titles, in terms of: average number of words, most important words, and words that are specific for the agricultural domain, which may impact on the sentiment analysis task. The main libraries used on this Step were: Pandas, Matplotlib (https://matplotlib.org/), Spacy (https://spacy.io/) and Seaborn (https://seaborn.pydata.org/). The results of this step will be described in the Results Section.

RESULTS

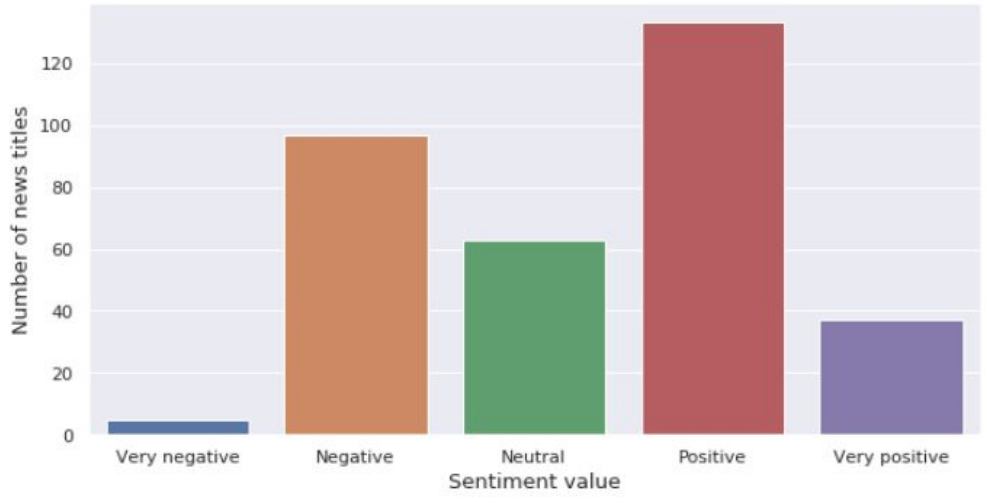
Exploratory data analysis

The raw data collected using the crawlers resulted in a dataset of 2.134 news titles. These were composed of news titles from different markets, such as the stock, commodities markets (such as agricultural commodities, iron, and oil), and electronics markets. Around 10% of these were spam. After carefully analyzing and cleaning the data, it resulted in a dataset of 335 agricultural markets news titles that are relevant for the domain.

After the data cleaning process, we labeled the data in the dimensions of sentiment value and urgency. Figure 1 and Table 2 illustrate the aggregated results of the sentiment value dimension. It is important to observe that the dataset is considerably positive (around 50% of the total dataset), while only around 1.5% is very negative. This reflects the current market situation, in which the exchange rates are favorable for exporters, impacting positively the agricultural domain. Table 2 also shows an example of a sentence for each category.

As for the urgency dimension, only around 15% of the news titles were considered urgent, meaning that they would have a high probability of impacting the market sentiment on the day of its release. Table 3 shows an example of a sentence for each category in this dimension.

Figure 1. Distribution of the sentiment classes on the dataset



As news sentiment values may vary widely within the same day (for example, on some days there were very positive and very negative news), it is essential to aggregate multiple sentiment values in a given period to correctly interpret the market sentiment. The daily sentiment was calculated considering a simple average of the sentiment values for each

day in the dataset. The same procedure was used for calculating the weekly sentiment. Figure 2 illustrates these two variables. It is possible to observe that, even though the daily sentiment varies considerably from one day to the next, the weekly sentiment shows a stable rising trend with some positive peaks at the end of March and beginning of May.

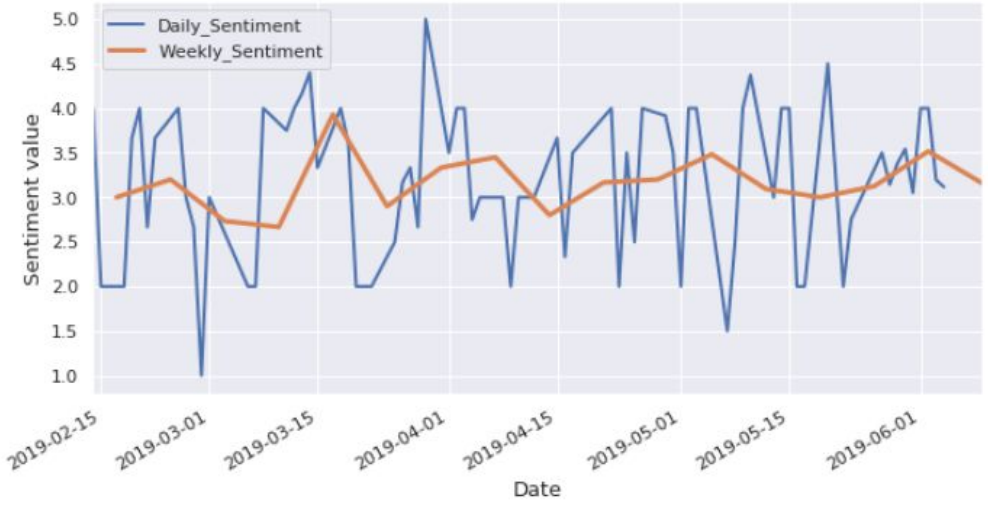
Table 2. Number of data points, percentage and example of sentences for the sentiment dimension

Sentiment class	Number of data points	Percentage of the total dataset	Example of a sentence
Very negative	5	1.49%	BRF loss almost triples in the fourth quarter and totals R$ 2.1 billion
Negative	97	28.97%	Soybeans still under pressure in Chicago and closing prices show a downtrend
Neutral	63	18.83%	CRAs emissions totaled R$ 4 billion in the first quarter
Good	133	39.72%	Conab and IBGE raise projections for grain harvests
Very good	37	10.99%	Brazilian coffee exports increase by 36% in February
Total	335	100%	

Table 3. Number of data points, percentage and example of sentences for the urgency dimension

Sentiment class	Number of data points	Percentage of the total dataset	Example of a sentence
Urgent	52	15.52%	Brazilian coffee exports grow 123% in May, Secex reveals
Non-urgent	283	84.48%	Paraná could lose second place in soybean production to the Rio Grande do Sul
Total	335	100%	

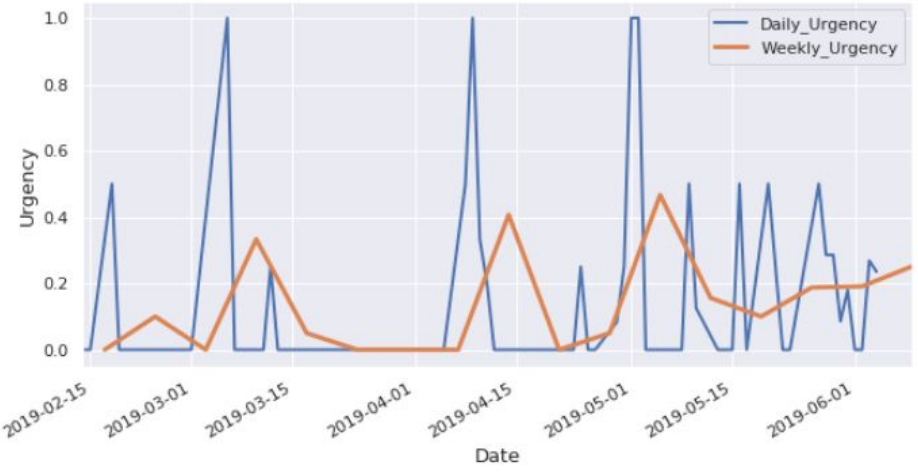
Figure 2. Daily and weekly sentiments of the news on the dataset.



These results reflect the reality: the market was indeed increasing its expectations from the beginning of 2019 up to the end of the dataset. The differing volatility between daily and weekly sentiments is also observed in the financial market, and we believe that our results reflect the reality of the agricultural markets. With a larger dataset, the weekly and monthly sentiment could be used to point the overall trend in sentiment value, and the daily sentiment could be used for speculation purposes.

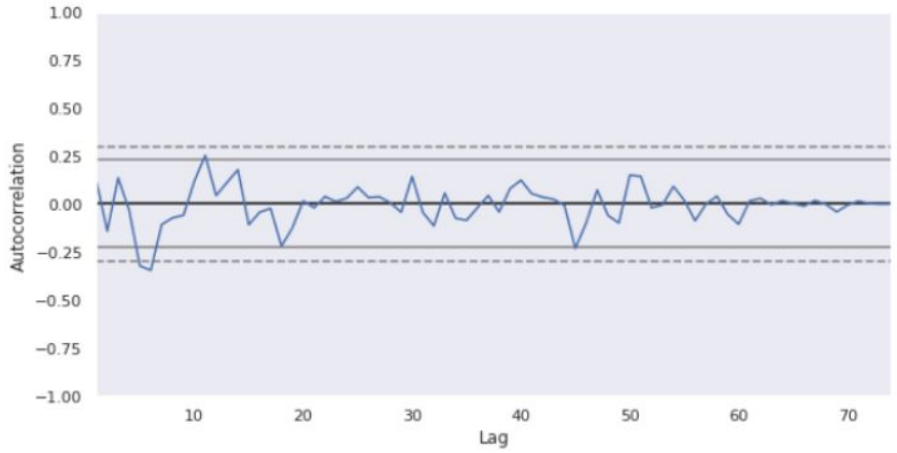
Figure 3 illustrates the daily and weekly urgencies of the news on the dataset. Some partial conclusions that can be drawn from this chart (as more data is needed in this dimension to clearly evaluate the frequency of urgent news related to non-urgent ones) are that: (i) it does present some correlation with the peaks and valleys of sentiment value; and (ii) that it may be influenced by the sentiment volatility in the market (or itself influence this volatility). These results are interesting because this dimension could also be an input for decision making, as well as the sentiment values, especially if the agent wants to better identify the periods of higher volatility in the market sentiment.

Figure 3. Daily and weekly urgency of the news on the dataset.



An analysis of autocorrelation of these two dimensions (sentiment value and dimension) on both daily and weekly frequencies showed that there is autocorrelation on daily sentiment. This is illustrated in Figure 4. Nevertheless, more data is needed to make a thorough evaluation of autocorrelation. It is not possible to discard, at the moment, if weekly sentiment or the daily and weekly urgencies are not autocorrelated variables.

Figure 4. Autocorrelation plot for Daily Sentiment.



A brief analysis of the linguistic properties of the dataset showed that it contains 773 unique words, from which 95 are considered stopwords for the Portuguese language. This results in 678 relevant words that can be used for sentiment analysis. Nevertheless, a considerable portion of these is not relevant for the domain, such as: dirigente (leader), days of the week, month names, and location identifications (such as south, center-west, etc), among others.

An analysis of the morphology of the dataset shows that the most important classes, in terms of frequency of words, are: verbs, propositions, numerals, and nouns. An analysis of the entities on the sentences, using Spacy's named entities recognition module (https://spacy.io/usage/linguistic-features#named-entities), showed some interesting results: (i) only very large multinational companies have their entities on their dictionary (the main examples are Marfrig and BRF, very important food companies); and (ii) countries are correctly identified (such as the USA and China).

DISCUSSIONS

One of the biggest barriers for conducting sentiment analysis on the agricultural market, as pointed before, is the lack of datasets with sentiment values. In this paper, we developed a labeled dataset, considering two dimensions: sentiment values and urgency. We also conducted statistical and linguistic analysis.

There were two main difficulties in conducting this research: (i) creating the dataset itself, which involved choosing trustworthy sources and gathering data from them; and (ii) a lack of a well-established methodology for labeling the dataset. We believe that the methodology used can be used for generating datasets for other domains, as the results were

satisfactory. One important aspect to consider is the necessity to label the dataset using additional dimensions, what depends on domain-specific characteristics.

Although more data is needed to create robust sentiment prediction models, we are currently conducting experiments with several machine learning models, such as artificial neural networks, long short-term memory networks, and support vector machines. Future work is needed on improving the dataset, including more data points, further refining the sentiment value and urgency dimensions, and also implementing sentiment analysis and prediction models. We believe that the results of these researches could improve financial prediction models applied to the agricultural markets, improving decision making.

It is very important to observe that, as a considerable portion of the unique identified words in the dataset are domain-specific (related to agricultural processes, products, or simply the relation between market and environmental variables and the quantity and quality of products on a given season), we can infer that a general lexicon with sentiment values would not be able to correctly identify and predict sentiments in this domain.

CONCLUSIONS

Sentiment analysis techniques could help improving decision making in the agricultural sector. Nevertheless, there are no few papers that conduct this type of research on the Brazilian agricultural market, and there are no datasets that contain sentiment values available for research. The main objective of this paper was to develop and analyze a dataset of agricultural market news titles that could be used for sentiment analysis on this domain, as well as analyze it from the statistical and linguistic points of view.

The dataset developed contains 335 news titles, labeled in two dimensions: (i) sentiment value, which contains five categories (very negative, negative, neutral, positive, and very positive); and (ii) urgency, which contains two categories (urgent and non-urgent). On the first dimension, it contains mostly positive news. On the second dimension, it contains mainly non-urgent news.

Through the analyses that were conducted, we conclude that the dataset is suitable for sentiment analysis tasks, which will be developed on further researches. The methodology used can be applied to develop labeled datasets for other domains, improving the sentiment analysis and prediction in those domains as well.

ACKNOWLEDGEMENTS

This work was supported by Itaú Unibanco S.A. through the Itaú Scholarship Program, at the Centro de Ciência de Dados (C2D), Universidade de São Paulo, Brazil.

REFERENCES

AGGARWAL, C.C.; ZHAI, C. Mining text data. Springer Science & Business Media, 2012, 527 pp.

BOLLEN, J.; MAO, H; ZENG, X. Twitter mood predicts the stock market. Journal of Computational Science, v. 2, n. 1, p.1-8, 2011.

CHOPRA, S.; MEINDL, P. Supply chain management: Strategy, planning, and operation, 5th ed., New Jersey, USA: Pearson Education, 2013, 528pp.

FROEHLICH, H.E.; GENTRY, R.R.; RUST, M.B.; GRIMM, D.; HALPERN, B.S. Public perceptions of aquaculture: evaluating spatiotemporal patterns of sentiment around the world. PloS one, v. 12, n. 1, p.e0169281, 2017.

HUTTO, C.J; GILBERT, E.E. VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. In: Proceedings of the Eighth International Conference on Weblogs and Social Media (ICWSM-14). Ann Arbor, MI, 2014.

LIU, B. Sentiment analysis and opinion mining. Synthesis lectures on human language technologies, v. 5, n. 1, p.1-167, 2012.

MUKHERJEE, S.; BHATTACHARYYA, P. Feature specific sentiment analysis for product reviews. In: International Conference on Intelligent Text Processing and Computational Linguistics. Springer, Berlin, Heidelberg, p. 475-487, 2012.

NASSIRTOUSSI, A.K.; AGHABOZORGI, S.; WAH, T.Y.; NGO, D.C.L. Text mining for market prediction: A systematic review. Expert Systems with Applications, v. 41, n. 16, p.7653-7670, 2014.

PANG, B.; LEE, L. A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts. In: Proceedings of the 42nd annual meeting on Association for Computational Linguistics. Association for Computational Linguistics, 2004.

REN, R.; WU, D.D.; LIU, T. Forecasting stock market movement direction using sentiment analysis and support vector machine. IEEE Systems Journal, v. 99, p.1-11, 2018.



SVM com DenseNet para classificação de mudas do guaraná a partir da nervura central do folíolo

Brenda Talyne Costa Martins[1], *Marcos Filipe Alves Salame*[2]

[1]Bolsista de Iniciação Científica FAPEAM, Embrapa Amazônia Ocidental
Manaus, Amazonas, Brasil, brendatalyne@gmail.com

[2]Analista de Tecnologia da Informação, Embrapa Amazônia Ocidental
Manaus, Amazonas, Brasil, marcos.salame@embrapa.br

RESUMO

O guaranazeiro *(Paullinia cupana var. sorbilis)* é um importante e tradicional cultivo do Estado do Amazonas. Com o aumento da demanda por guaraná, a Embrapa Amazônia Ocidental desenvolveu e registrou 20 cultivares e lançou 18 até o momento, com o objetivo de aumentar a produtividade, reduzir o tempo de formação, fornecer alto teor de cafeína e mais resistência às principais pragas. Entretanto, o procedimento de identificação desses cultivares é difícil e feito somente por especialistas. Diante deste cenário, foi criado um *dataset* de mudas de guaraná, composto por 10 cultivares, através de um protótipo desenvolvido para dispositivos móveis, de forma a evitar a necessidade de qualquer pré-processamento manual e visando a proximidade com o ambiente real do produtor. Foram realizados experimentos com 5 dos cultivares coletados, utilizando aprendizado de máquina com SVM e DenseNet para classificação a partir de imagens da nervura central dos folíolos e obteve-se como resultado uma acurácia de 88%.

PALAVRAS-CHAVE: Aprendizado de Máquina, guaranazeiro, cultivares.

ABSTRACT

The guarana plant (Paullinia cupana var. Sorbilis) is an important and traditional crop of state of Amazonas. Due to increase of demand for guarana, Embrapa Amazônia Ocidental developed and registered 20 cultivars and launched 18 so far, aiming to increase productivity, reduce formation time, provide high caffeine content and more resistance to major pests. However, the identification procedure of these cultivars is difficult and it is made only by specialists. Based on this scenario, we created a dataset of guarana seedlings, composed by 10 cultivars, through a prototype developed for mobile devices, in order to avoid the need for any manual preprocessing and aiming at proximity to the environment of the rural producer. Experiments

were carried out on 5 of the cultivars collected, using machine learning with SVM and DenseNet for classification from images of the central vein of the leaflets and we obtained, as a result, an accuracy of 88%.

KEYWORDS: Machine Learning, guarana, cultivars.

INTRODUÇÃO

O guaranazeiro *(Paullinia cupana var. sorbilis)* dentre muitas espécies amazônicas, é reconhecido pelo seu potencial econômico (GONÇALVES et al., 2006). Com o aumento da demanda e importância, a Embrapa Amazônia Ocidental desenvolveu e registrou 20 cultivares ao longo dos anos através do programa de melhoramento genético, que visa selecionar e recomendar cultivares com grande potencial de produtividade, alto teor de cafeína e com mais resistência às principais pragas. As informações dos cultivares podem ser encontradas em (MAPA, 2019).

Desses cultivares, 19 foram obtidos pelo método de estaquia (reprodução assexuada). Uma das grandes diferenças desses cultivares para os que são produzidos por sementes (reprodução sexuada) é que segundo (TRICAUD; PINTON; PEREIRA, 2016), o tempo de formação das mudas é de 7 meses, enquanto que a da muda tradicional por semente, demora 12 meses até ficar pronta para ir ao campo.

De acordo com (CORRÊA, 1982), é recomendável que as mudas, antes do plantio, possuam no mínimo de quatro a seis folhas completas, com folíolos bem desenvolvidos, maduros e de coloração normal. Até a formação das primeiras folhas, as mudas no viveiro recebem tratamento através da irrigação intermitente, ou seja, o fornecimento controlado de água para as plantas em quantidade suficiente e no momento que precisam de irrigação, assegurando a produtividade e a sobrevivência da plantação. Segundo (PEREIRA, 2005), após esse período, a planta é levada para um segundo viveiro que lhe proporcionará melhores resultados na propagação vegetativa com a nebulização, que lhe proporciona umidade necessária para o bom desempenho do enraizamento das estacas, promovendo uma proteção eficiente da superfície até a formação das mudas para que possam ser levadas para o seu local definitivo de plantio.

No entanto, a distinção desses cultivares, principalmente em mudas, que é quando ocorre o período de transporte dessas plantas para o local definitivo, ainda é feito por mecanismos manuais (PEREIRA, 2005). Isso resulta em um processo árduo e muitas vezes impreciso, podendo resultar numa aquisição incorreta por parte do agricultor. Esse cenário estimula a pesquisa em novas tecnologias para identificação dessas variedades genéticas, uma vez que as folhas apresentam muita semelhança entre si e geram resultados bem distintos.

Nos experimentos realizados por (SOUSA et al., 2017), foram utilizadas imagens de 2 cultivares que foram obtidas de 2 ângulos diferentes, sendo 20 imagens por cultivar de folhas inteiras, uma de cada amostra. O *dataset* utilizado pode ser considerado pequeno quando comparado a outros trabalhos relacionados de classificação por imagens, portanto, eles expandiram através de *data augmentation*, gerando também mais variações. Com a utilização de Redes

Neurais Convolucionais foi possível obter o resultado de 97% de acurácia neste experimento.

Entretanto, houve um pré-processamento realizado manualmente com redimensionamento, retirada de fundo, e centralização da folha na imagem, inviabilizando de poder ser usado de forma mais prática.

Em (SOUSA; SALAME, 2019) é possível identificar resultados para análises com 3 cultivares considerados de alta produtividade. Os autores utilizaram 434 imagens de plantas adultas e aplicaram técnicas de aprendizado de máquina supervisionado para a classificação realizada pelo formato dos folíolos e pelos padrões de venação, que são mais visíveis na superfície abaxial das folhas. A precisão de reconhecimento para este experimento foi de 89%. Ou seja, reduziu a acurácia, mas o processo foi facilitado, devido a ser apenas o folíolo, reduzindo os trabalhos de pré-processamento manual.

Dessa forma, este trabalho propõe uma abordagem computacional para identificação dos cultivares de guaraná a partir de imagens dos folíolos das mudas de guaranazeiros utilizando aprendizado de máquina. As imagens foram coletadas de forma prática, o mais próximo do ambiente real de produção, sem necessidade de qualquer tipo de pré-processamento manual e utilizando apenas a nervura central dos folíolos, com o intuito de que a solução, futuramente, possa ser usada pelos produtores rurais através do celular.

MATERIAL E MÉTODOS

Pelo fato de não haver um *dataset* público dos cultivares de guaranazeiro criados pela Embrapa e também pelo intuito de transformar a solução em algo que possa ser usado por produtores rurais no dia-a-dia, foi desenvolvido um protótipo em Ionic para dispositivos móveis com sistema operacional Android, com o propósito de coletar as imagens dos folíolos das mudas para compôr um *dataset*, de forma a evitar qualquer tipo de pré-processamento manual. Para a criação do mesmo, foram utilizados os *frameworks Ionic* e nodeJs com as linguagens *javaScript* e *typescript*

O protótipo também pode ser usado para validar a classificação em ambiente de produção, podendo vir, futuramente, a se tornar um produto. Ele contém a funcionalidade de câmera com filtro e corte e também possui informações referentes aos cultivares, obtidas a partir de publicações científicas e entrevistas com especialistas da área.

Criação do dataset

A coleta teve enfoque na nervura central dos folíolos e foi utilizado um smartphone com uma câmera de 16 *megapixels* e Android 7.0. Ao todo foram coletados 10 cultivares, onde em cada cultivar foram capturados 2 folíolos de cada folha, sendo 10 fotos por folíolo e 10 folhas, totalizando 200 imagens por cultivar. A captura das imagens foi feita dos folíolos 2 e 3 de cada folha da parte abaxial (parte inferior), conforme Figura 1. Foram fotografadas várias áreas, com o cuidado para que a captura não excedesse o perímetro dos folíolos. Na coleta foi observada a

distância entre o *smartphone* e o folíolo, além de variações de tamanhos dos folíolos conforme Figura 2.

Os cultivares coletados foram desenvolvidos pelo método de estaquia (reprodução assexuada), sendo eles: BRS-Luzéia, BRS-Maués, BRS CG612, BRS Andirá, BRS Cereçaporanga, BRS-Amazonas, BRS CG 372, BRS CG608, BRS CG610 e BRS CG850.

Figura 1: Nervura central dos folíolos capturadas com uso do protótipo para criação do dataset



Figura 2: Tamanhos dos folíolos e aproximação do *smartphone* na captura das imagens.

Tam. (Altura)	Aprox. da Câmera	Período
10 à 12 cm	06 à 08 cm	Jan à Fev
13 à 16 cm	08 à 09 cm	Jan à Fev
17 à 20 cm	08 à 10 cm	Maio
21 à 22 cm	08 à 11 cm	Maio

Os cultivares BRS Mundurucânia, BRS CG611, BRS Saterê e BRS Onhiamuaçabê, não foram coletados pois não haviam mudas no viveiro. Já os cultivares BRS Marabitana, BRS CG505, BRS CG648, BRS CG189, BRS CG882, ainda serão coletados para serem acrescentados ao *dataset*. Segundo (KARASAWA, 2009) há diferenças entre os sistemas de reprodução de plantas. O BRS Noçoquém (ATROCH; FILHO; PEREIRA, 2015) foi o primeiro cultivar lançado de reprodução via semente. Este cultivar não foi capturado por não fazer parte dos experimentos de reprodução pelo método de estaquia, como é o caso dos outros 19 cultivares registrados.

Cultivares selecionadados

Para os experimentos deste trabalho, devido ao tempo de processamento e à importancia, foram escolhidos 5 cultivares que apresentam alta produtividade: BRS-Luzéia, BRS-Maués, BRS-

Andirá, BRS-Cereçaporanga e BRS-CG612, com idades variando entre 8 meses e 2 anos, mas ainda consideradas mudas, plantadas dentro de sacos.

De acordo com (JAIN; DUIN; MAO, 2000), o reconhecimento de padrões é o estudo de como máquinas podem observar o ambiente, aprender a distinguir padrões e tomar decisões sobre as categorias dos padrões. Sendo assim, podemos concluir que a importância do reconhecimento desses padrões está ligada a encontrar métodos que automatizem atividades, que possam ser trabalhosas e também sujeitas a falhas.

Técnicas de classificação

Foi utilizada uma CNN (*Convolutional neural network*) como extrator de características. Conforme (ABDEL-HAMID et al., 2012), a CNN consiste em um ou mais pares de camadas de convolução. As camadas convolucionais aplicam filtros que processam pequenos locais de uma imagem, e são replicados por toda a imagem.

Para a implementação dos algoritmos, foi utilizada a linguagem de programação Python e bibliotecas como NumPy, *OpenCv*, *Scikit-learn*. Para o aprendizado de máquina, foram utilizadas diversas técnicas, desde técnicas clássicas até *deep learning*. O melhor resultado obtido foi com SVM (*Suport Vector Machine*) e DenseNet (Dense Convolutional Network). O melhor conjunto de parâmetros encontrado no conjunto de desenvolvimento nesta clasificação foi: C:1, gamma: 0.1 e Kernel: linear.

RESULTADOS E DISCUSSÃO

As técnicas abordadas no experimento apresentaram resultados satisfatórios na classificação dos cultivares de guaranazeiros. O resultado de algumas métricas utilizadas pode ser visto na Figura 3. Através de SVM com DenseNet para classificação a partir de imagens da nervura central dos folíolos do guaraná, obteve-se como resultado uma acurácia de 88%.

Figura 3: Resultado da classificação

Cultivares	Precision	Recall	F1-score
BRS Andirá	0.85	1.00	0.92
BRS Luzéia	0.83	0.96	0.89
BRS Cereçaporanga	0.84	0.76	0.79
BRS Maués	0.95	0.78	0.86
BRS CG612	0.96	0.92	0.94

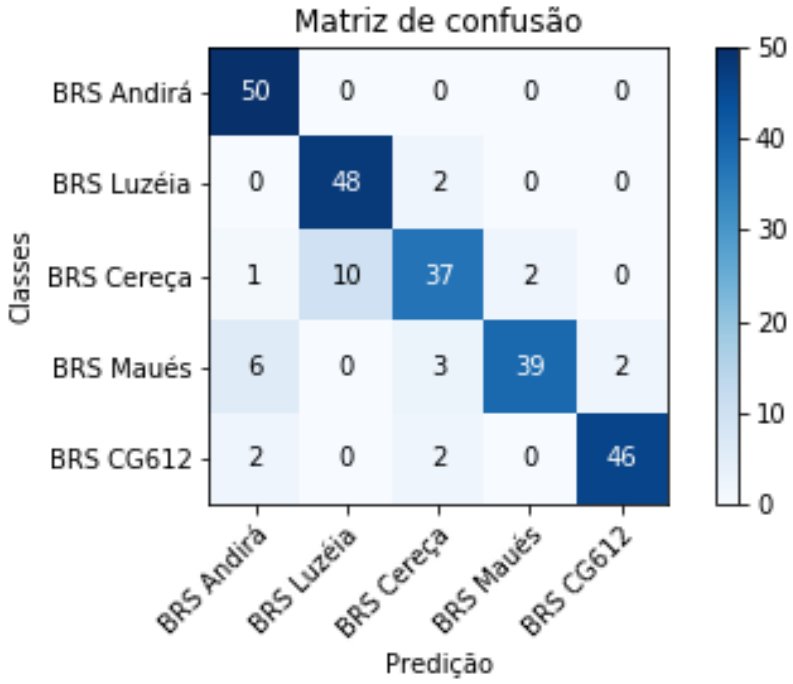
Segundo (FACELI et al., 2011), com princípios apoiados na teoria de aprendizado estatistico, o SVM destaca-se por apresentar uma boa capacidade de generalização. Uma outra ca-

racterística atrativa é a convexidade do problema da otimização formulado por seu treinamento, que implica a existência de um único mínimo global.

Conforme (JÉGOU et al., 2017), a DenseNet tem mostrado excelentes resultados em tarefas de classificação de imagens. A ideia da DenseNet baseia-se na observação de que, se cada camada estiver diretamente conectada para todas as outras camadas em um feed-forward, a rede será mais precisa e mais fácil de treinar. A DenseNet requer menos parâmetros do que uma CNN tradicional equivalente, já que não há necessidade de aprender mapas de recursos.

Na Figura 4, é apresentada a matriz de confusão do experimento. Nela, pode ser observado que o BRS Andirá apresentou 100% de verdadeiro positivo, sendo o único com tal feito. No entanto, falsos positivos ocorreram. Dentre os outros cultivares, podemos observar também que o BRS Luzéia obteve um bom resultado e o BRS Cereçaporanga foi o que obteve menos acertos.

Figura 4: Matriz de confusão da arquitetura de classificação



CONCLUSÕES

Com os resultados do experimento, é possível verificar que há viabilidade na implementação de uma solução para reconhecer cultivares de guaraná a partir na nervura central dos folíolos, que pode ser considerado o dna da planta. As técnicas utlizadas, SVM com DenseNet, apresentaram um resultado promissor na classificação.

Ao todo, foram coletadas 2 mil imagens. Todavia, no experimento, foram utilizadas mil imagens de 5 cultivares e obteve-se como resultado uma acurácia de 88%. Nenhum pré-

processamento manual foi necessário para o *dataset*, tornando real a possibilidade da solução vir a se tornar um produto.

Como trabalhos futuros, pretende-se coletar imagens dos cultivares restantes e realizar vários experimentos com todos os 19, desenvolvidos por método de estaquia, até conseguir um resultado satisfatório. Para dessa forma, posteriormente, ajustar o algoritmo no servidor, realizar validação com o protótipo em ambiente de produção e, se tudo der certo, disponibilizar para o público, podendo oferecer relevante complemento aos métodos atuais utilizados na identificação de cultivares.

AGRADECIMENTOS

Aos pesquisadores em guaranaicultura Firmino José do Nascimento Filho e André Luiz Atroch, cuja disponibilidade e conhecimentos sobre o guaranazeiro transmitidos foram fundamentais; aos especialistas Orbélio Mota Campos Silva e Luciano Simões Malcher pela ajuda no viveiro; aos colegas Allex de Lima Sousa e Richardson Allan Ferreira de Souza pelo auxílio nos estudos e à Fundação de Amparo à Pesquisa do Estado do Amazonas pela disponibilização da bolsa de iniciação científica.

REFERÊNCIAS

ABDEL-HAMID, O. et al. Applying convolutional neural networks concepts to hybrid nn-hmm model for speech recognition. In: IEEE. *2012 IEEE international conference on Acoustics, speech and signal processing (ICASSP)*. [S.l.], 2012. p. 4277–4280.

ATROCH, A.; FILHO, F. do N.; PEREIRA, J. Brs noçoquém: primeira cultivar de guaranazeiro de reprodução via sementes para cultivo no estado do amazonas. *Embrapa Amazônia Ocidental-Comunicado Técnico*, Manaus: Embrapa Amazônia Ocidental, 2015., 2015.

CORRÊA, M. *Formação de mudas de guaraná*. [S.l.]: EMBRAPA-UEPAE de Manaus, 1982.

FACELI, K. et al. Inteligência artificial: Uma abordagem de aprendizado de máquina. 2011.

GONÇALVES, J. F. de C. et al. Características fisiológicas e anatômicas de folhas de dois clones de guaraná. *Pesquisa Agropecuária Brasileira*, v. 41, n. 3, p. 393–398, 2006.

JAIN, A. K.; DUIN, R. P. W.; MAO, J. Statistical pattern recognition: A review. *IEEE Transactions on pattern analysis and machine intelligence*, Ieee, v. 22, n. 1, p. 4–37, 2000.

JÉGOU, S. et al. The one hundred layers tiramisu: Fully convolutional densenets for semantic segmentation. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*. [S.l.: s.n.], 2017. p. 11–19.

KARASAWA, M. Diversidade reprodutiva de plantas, uma perspectiva evolutiva e bases genéticas. *Editora SBG, Sociedade Brasileira de Genética, Ribeirão Preto, SP: p. 113*, 2009.

MAPA. *Registro Nacional de Cultivares - RNC*. 2019. Disponível em: <http://www.agricultura.gov.br/guia-de-servicos/registro-nacional-de-cultivares-rnc.>. Acesso em: 1 jul. 2019.

PEREIRA, J. C. Cultura do guaranazeiro no amazonas. *Embrapa Amazônia Ocidental-Sistema de Produção*, Manaus: Embrapa Amazônia Ocidental, 2005., 2005.

SOUSA, A. d. L. et al. Redes neurais convolucionais aplicadas ao processo de classificação de cultivares de guaranazeiros. In: IN: ENCONTRO NACIONAL DE INTELIGÊNCIA ARTIFICIAL E COMPUTACIONAL, 14. [S.l.], 2017.

SOUSA, A. de L.; SALAME, M. F. A. Classificação de guaranazeiros através do padrão de venação e formato de folíolos utilizando aprendizagem de máquina. *Anais do Computer on the Beach*, p. 676–685, 2019.

TRICAUD, S.; PINTON, F.; PEREIRA, H. dos S. Saberes e práticas locais dos produtores de guaraná (paullinia cupana kunth var. sorbilis) do médio amazonas: duas organizações locais frente à inovação. *Boletim do Museu Paraense Emílio Goeldi. Ciências Humanas*, Museu Paraense Emílio Goeldi, v. 11, n. 1, p. 33–53, 2016.



Detecção de Folíolos do Guaranazeiro Atacados pelo Tripes Utilizando CNNs Pré-Treinadas

Richardson Allan Ferreira de Souza[1], *Marcos Filipe Alves Salame*[2]

[1]Bolsista de Iniciação Científica FAPEAM, Embrapa Amazônia Ocidental
Manaus, Amazonas, Brasil, rafds.snf@uea.edu.br

[2]Analista de Tecnologia da Informação, Embrapa Amazônia Ocidental
Manaus, Amazonas, Brasil, marcos.salame@embrapa.br

RESUMO

O cultivo do guaranazeiro tem grande importância socioeconômica na Região Amazônica, sendo fonte de renda para pequenos agricultores rurais. O ataque do Tripes (*Pseudophilotrips adisi*) tem interferido no sistema de agricultura familiar prejudicando a produção de grãos e limitando a exploração comercial do guaraná. Com isso, abordagens computacionais para identificar automaticamente os danos causados pelo Tripes podem auxiliar no seu controle, entre outras aplicações. Este trabalho utilizou Redes Neurais Convolucionais pré-treinadas como extratores de características, através da técnica de *Transfer Learning*, para avaliar os classificadores *Support Vector Machine* e *Multilayer Perceptron* na tarefa de reconhecer danos em folíolos jovens do guaranazeiro atacados pelo Tripes com base em imagens obtidas de *smartphones*. Embora a base de imagens seja pequena e desbalanceada, nenhuma avaliação obteve um índice kappa abaixo de 0.8. No melhor cenário foi possível atingir um índice kappa de 0,9436 e uma acurácia balanceada de 0,9609 com o classificador *Support Vector Machine*.

PALAVRAS-CHAVE: *Transfer learning*, Tripes, Guaraná.

ABSTRACT

Guarana crops has a great socioeconomic importance in the Amazon Region, being a source of income for small rural farmers. The attack of the Thrips (*Pseudophilotrips adisi*) interferes in the family farming system, damaging the production of grains and limiting the commercial exploration of guarana. Based on this, computational approaches to automatically identify the damage caused by the thrips can help in its control, among other applications. This work uses features extracted from pre-trained Convolutional Neural Networks through the Transfer Learning technique to evaluate the Support Vector Machine and Multilayer Perceptron classifiers for the recognition of damage in young guarana leaflets caused by the Thrips attack based on

images obtained from smartphones. Although the images base be small and unbalanced, no assessment had a Cohen's Kappa rating below 0,8. In the best scenario it was possible to achieve a Cohen's Kappa of 0,9436 and a balance accuracy of 0,9609 with the Support Vector Machine classifier.

KEYWORDS: Transfer learning, Thrips, Guarana.

INTRODUÇÃO

A cultura do guaraná vem assumindo importância cada vez maior no cenário econômico e social do país, especialmente na Região Amazônica. A importância econômica é evidenciada pelo crescimento da demanda de sementes pelas indústrias de bebidas para atender o promissor mercado de refrigerantes, em nível nacional e internacional. Considerando também as outras formas de utilização, o guaranazeiro representa fonte de renda para inúmeros produtores rurais da Região Amazônica, sendo indiscutível sua importância socioeconômica, contando com cerca de 2.500 produtores, com área cultivada média de 2 hectares/produtor (SANTOS, 2018; SENRA; GARCIA; TAVARES, 2008).

Proposto por (SAHA; EKBAL, 2013)

A ocorrência da praga Tripes, *Pseudophilothrips adisi* (zur Strassen), tem interferido de forma direta e indireta na produtividade, limitando a exploração comercial do guaraná em sistemas de agricultura familiar. Em relação aos danos diretos, ao se alimentarem do conteúdo celular das estruturas jovens das plantas de guaranazeiro, pequenas pontuações são formadas de onde sugam a seiva causando alterações morfológicas e de desenvolvimento. Alternativas de controle mais seguras para o homem e para o ambiente têm sido o principal alvo das pesquisas desenvolvidas na Embrapa (FONTES; TAVARES, 2018; TAVARES; GARCIA, 2009; TAVARES; GARCIA; NASCIMENTO-FILHO, 2007).

A literatura apresenta diversos trabalhos que buscam aplicar técnicas de Aprendizado de Máquina (*Machine Learning* ou ML), Visão Computacional (*Computer Vision* ou CV) e Aprendizagem Profunda (*Deep Learning* ou DL) no reconhecimento de injúrias e doenças que acometem diversos tipos de cultivos. O reconhecimento de imagem oferece uma tecnologia econômica e escalonável para utilização na agricultura e a DL está rapidamente se tornando a técnica padrão para classificação de imagens. Contudo, abordagens tradicionais necessitam de uma etapa orientada a extrair características mais representativas de cada imagem (BARBEDO, 2019; MOHANTY; HUGHES; SALATHÉ, 2016).

A técnica de *transfer learning* pode ser aplicada ao se utilizar uma Rede Neural Convolucional (*Convolutional Neural Networks* ou CNNs) pré-treinada como extrator de características para um classificador qualquer. Em razão de seus bons resultados, as CNNs estão liderando o estado da arte em tarefas de visão computacional (RAMCHARAN et al., 2017; NETO et al., 2018).

Com base nesse cenário, foram utilizadas arquiteturas de redes convolucionais por meio de *transfer learning*, para extrair características de imagens digitais obtidas de *smartphones* e

avaliar os classificadores *Support Vector Machine* (SVM) e *Multilayer Perceptron* (MLP) na identificação dos danos em folíolos jovens do guaranazeiro atacados pelo Tripes.

A estrutura do artigo está organizada como segue. A seção Material e Métodos descreve a coleta das imagens para compor a base de dados, a estratégia usada para elaborar o *dataset*, o uso de CNNs pré-treinadas como extratores de características e demais técnicas empregadas. A seção Resultados e Discussão expõem os resultados alcançados com um comparativo entre os classificadores. Por fim, a seção Conclusões resume os resultados obtidos e fornece perspectiva de trabalhos futuros(SENRA; GARCIA; TAVARES, 2008).

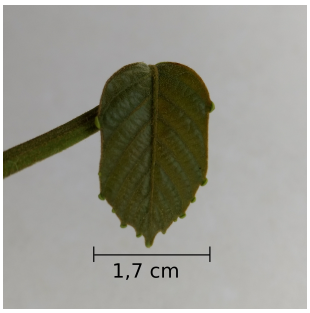
MATERIAL E MÉTODOS

Os experimentos foram realizados na Embrapa Amazônia Ocidental, situada à Rodovia AM-010, Km 29, Manaus, AM. Os indivíduos de *P. adisi* adultos e folhas jovens foram obtidos de plantas de guaranazeiro cultivadas nas áreas do Banco Ativo de Germoplasma (BAG) da Embrapa Amazônia Ocidental.

Aquisição das imagens

Para isolar os danos causados pelo Tripes de outros eventos, folíolos de até 1,7 centímetro, do cultivar BRS-Maués, foram isolados em placas de Petri com até dez exemplares de *P. adisi* adultos e monitorados por um período de 27 horas como mostra a Figura 1a e 1b. Na Figura 1c pode-se observar as injúrias provocadas pelo ataque do Tripes na parte adaxial do folíolo após esse período.

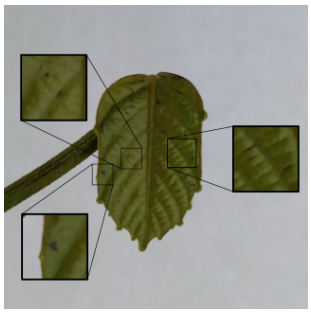
Figura 1: Processo para estimular o aparecimento de injúrias provocadas pelo Tripes.



(a) Folíolo saudável.



(b) Folíolo com Tripes.



(c) Folíolo com injúrias.

As imagens foram obtidas pela câmera digital de um *smatphone* da Motorola Moto G5 XT1672 com as funções de HDR e Flash desativadas tendo apenas a luz natural durante a coleta das imagens. Um papel cartão de cor branco e, posteriormente, outro de cor preta foram colocados a uma distância de 10 a 20 centímetro do folíolo, que ajudou na redução dos ruídos de fundo. A Figura 2 mostra uma aplicação que foi desenvolvida com o framework Ionic v3, o plugin de câmera nativo cordova-plugin-camera e uma biblioteca chamada angular-cropperjs

v0.1.5. Essa aplicação permitiu reduzir a área de interesse, redimensionar as imagens para 1872 x 1872 pixeis e gerar variações através da rotação e inversão da mesma em tempo de captura.

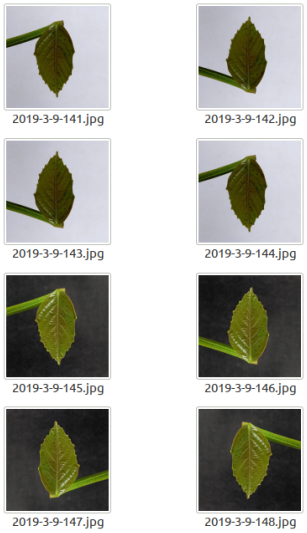
Figura 2: Aplicação para auxiliar na captura e padronização das imagens.



(a) Tela de confirmação.



(b) Imagem editada.



(c) Imagens padronizadas.

Conjunto de dados

Para compor o conjunto de dados (*dataset*), 26 imagens foram manualmente selecionadas, dimensionadas para 512 x 512 e subdivididas em 6240 regiões de interesse de 32 x 32 pixels. Como pode ser observado na Figura 3, as regiões que apresentavam pequenas pontuações, evidenciando o ataque, foram separadas para compor a base positiva e as que não apresentavam nenhum tipo de injúria, ver Figura 4, foram separadas para base negativa chegando ao número de 520 amostras da classe positiva e 3.123 amostras da classe negativa.

Figura 3: Exemplos de amostras positivas.

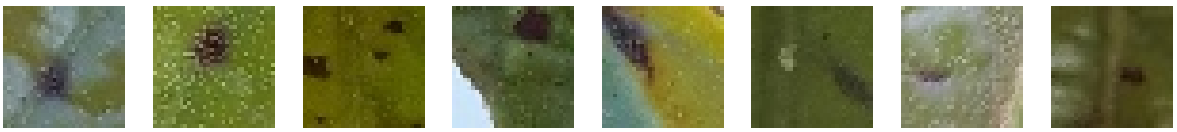
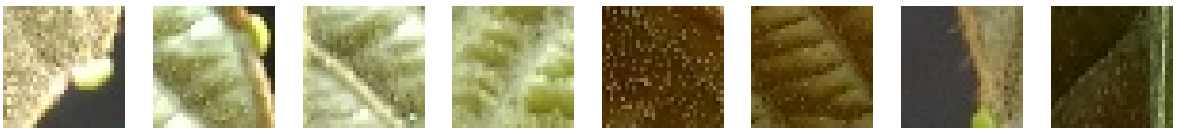


Figura 4: Exemplos de amostras negativas.



Por fim, como mostra a Tabela 1, as imagem foram aleatoriamente separadas uma vez em um conjunto de treino com 70% e de teste com 30% do total.

Tabela 1: Separação das amostras para treino e teste.

Conjunto	Amostras Positivas	Amostras Negativas	Total
Treino	373	2177	2550
Teste	147	946	1093

Extrator de Características

Uma CNN pré-treinada pode ser usada como extrator de características removendo a última camada totalmente conectada da rede e utilizando a nova saída para obter o vetor de característica que representa a imagem de entrada. Essa técnica conhecida como *transfer learning* também pode ser aplicada reaproveitando redes já treinadas com uma grande base de dados para resolver problemas em bases menores ajustando seus pesos.

Nesse artigo, foram utilizadas as arquiteturas canônicas VGG16, Densenet201, Resnet50 e Xception presentes no framework Keras com os pesos treinados pelo desafio da ImageNet (DENG et al., 2009).

Figura 5: Extração de característica com a arquitetura VGG16.

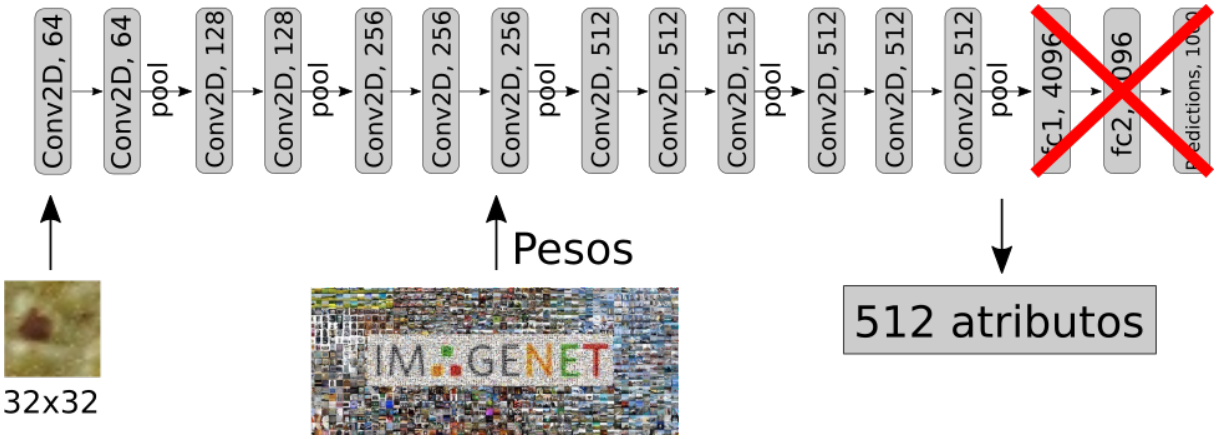


Tabela 2: Quantidade de atributos obtidos e o tamanho da entrada utilizado em cada arquitetura.

Extratores	Entrada	Quantidade de Atributos
VGG16	32 x 32 x 3	512
Resnet50	32 x 32 x 3	2048
Densenet201	32 x 32 x 3	1920
Xception	71 x 71 x 3	2048

Classificadores

Para implementar os classificadores SVM e MLP foi utilizada a linguagem de programação e definido um conjunto de hiper-parâmetros, ver Tabela 3, para método *GridSearchCV* disponível na biblioteca *scikit-learn*.

Tabela 3: Conjunto de hiper-parâmetros utilizados para os algoritmos SVM e MLP.

Algoritmo	**Hiper-parâmetro**	**Valor**
SVM	kernel	linear, rbf
	gamma	0.1 , 0.01 , 0.001
	C	1, 10, 100, 1000
MLP	hidden_layer_sizes	*Geometric Pyramid Rule*
	learning_rate_init	0.1 , 0.01 , 0.001
	activation	relu
	solver	sgd, adam

O hiper-parâmetro *hidden_layer_sizes* encontrado no classificador MLP indica o número de neurônios na camada oculta e para estimar essa quantidade de neurônios foi utilizada a Regra da Pirâmide Geométrica (*Geometric Pyramid Rule*) proposta por Masters (1993). Onde, para uma camada oculta com n entradas e m saídas, a camada oculta terá $\sqrt{n \times m}$ neurônios. Também foram estimadas a quantidade de neurônios para duas e três camadas ocultas sendo obtidas, respectivamente, por:

- $r = \sqrt[3]{\frac{n}{m}}$, $m \times r^2$ neurônios na primeira camada e $m \times r$ neurônios na segunda camada.
- $r = \sqrt[4]{\frac{n}{m}}$, $m \times r^3$ neurônios na primeira camada, $m \times r^2$ neurônios na segunda camada e $m \times r$ neurônios na terceira camada.

Os classificadores foram treinados com cada vetor de características descritos na subseção anterior utilizando o serviço em nuvem *Google Colaboratory* (*Colab*), totalizando oito treinamentos.

Métricas de avaliação

Para avaliar a capacidade de generalização dos classificadores, as medidas de acurácia balanceada, precisão, sensitividade, *f1 score* e índice Kappa de Cohen foram utilizadas, que de acordo com (PEDREGOSA et al., 2011):

- Acurácia Balanceada (*Balanced Accuracy* ou BA): Indicada em problemas classificação binária e multiclasse para lidar com conjuntos de dados desbalanceados. Baseada em uma Matriz de Confusão a acurácia balanceada é dada por $\frac{1}{2}\left(\frac{TP}{P} + \frac{TN}{N}\right)$;
- Precisão (*Precision*): É a relação $tp/(tp + fp)$, em que tp é o número de positivos verdadeiros e o número de falsos positivos. A precisão é intuitivamente a capacidade do classificador de não rotular como positiva uma amostra que é negativa;

- Sensitividade (*Recall*): É a relação $tp/(tp + fn)$ onde tp é o número de positivos verdadeiros e fn o número de falsos negativos. O *Recall* é intuitivamente a capacidade do classificador de encontrar todas as amostras positivas;
- F1-Score: Pode ser interpretado como uma média ponderada da Precisão e da Sensitividade dada pela Equação 1:

$$F1 = 2 \times \frac{(Precision \times Recall)}{(Precision + Recall)} \tag{1}$$

- Índice Kappa de Cohen (*Cohen's Kappa*): É um método estatístico para avaliar o nível de concordância ou reprodutibilidade entre dois conjuntos de dados e é definido pela Equação 2:

$$\frac{(p_o - p_e)}{(1 - p_e)} \tag{2}$$

RESULTADOS E DISCUSSÃO

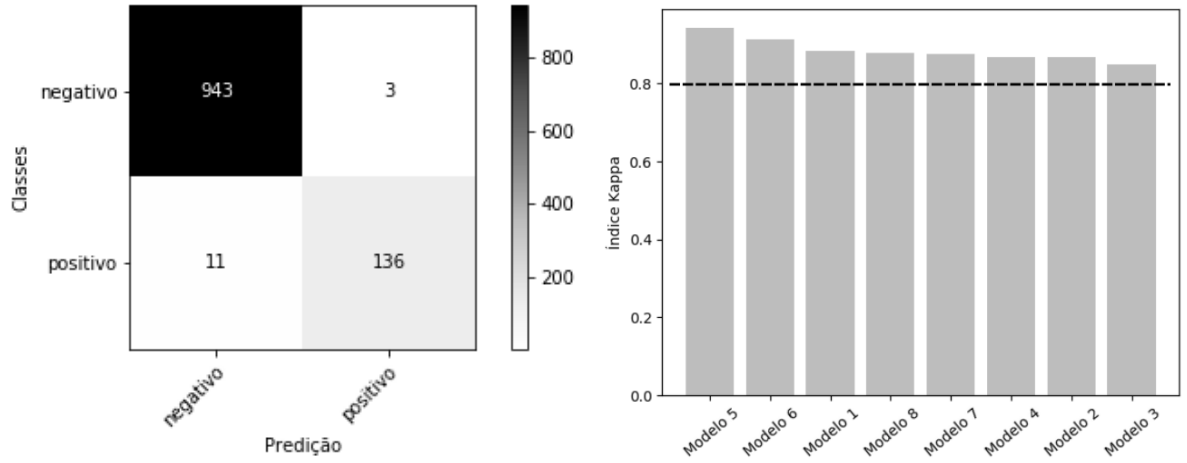
Após encontrar a melhor configuração durante o treinamento, cada modelo foi avaliado com as amostras de teste descritas na seção Conjunto de dados. A Tabela 4 mostra os resultados alcançados na avaliação dos oito modelos, comparando os valores de precisão, sensitividade, *f1 score* e acurácia balanceada de cada um, tanto para classe negativo quanto para a classe positivo.

Tabela 4: Resultado da Precisão, Sensitividade, F1 Score e Acurácia Balanceada de cada modelo.

Modelo	Classe	Precisão	Sensitividade	F1 Score	BA
1 (SVM/VGG16)	Negativo	0,98	0,99	0,98	0,9358
	Positivo	0,92	0,88	0,90	
2 (MLP/VGG16)	Negativo	0,98	0,99	0,98	0,9251
	Positivo	0,91	0,86	0,89	
3 (SVM/Resnet50)	Negativo	0,98	0,99	0,98	0,9143
	Positivo	0,90	0,84	0,87	
4 (MLP/Resnet50)	Negativo	0,98	0,99	0,98	0,9141
	Positivo	0,94	0,84	0,88	
5 (SVM/Densenet201)	Negativo	0,99	1,00	0,99	0,9609
	Positivo	0,98	0,93	0,95	
6 (MLP/Densenet201)	Negativo	0,98	0.99	0,99	0,9429
	Positivo	0,96	0,89	0,93	
7 (SVM/Xception)	Negativo	0,98	0,99	0,98	0,9319
	Positivo	0,91	0,88	0,89	
8 (MLP/Xception)	Negativo	0,98	0,98	0,98	0,9462
	Positivo	0,88	0,91	0,90	

Como pode ser visto na Tabela 4, o modelo 5 não só teve uma acurácia balanceada de 0,9609 como também foi o modelo que mais identificou amostras da classe positivo, alcançando uma sensitividade de 0,93. A Figura 6a é apresenta a Matriz de Confusão com o valores resultantes da classificação do modelo 5, onde as linhas equivalem a classe verdadeira e as colunas as classes preditas pelo modelo.

Figura 6: .



(a) Matriz de Confusão do Modelo 5.

(b) Índice Kappa de todos os modelos.

É possível observar na Figura 6a a alta densidade de valores nas células mais escuras. Demonstrando a classificação correta de muitas amostras do conjunto de teste, que mesmo desbalanceado apresentou bons resultados ao identificar a classe positivo com poucos elementos.

Na Figura 6b é possível observar que todos os modelos atingiram um desempenho de classificação pelo índice kappa acima de 0,8 que em Landis e Koch (1977) é rotulado como excelente. O modelo 5 destaca-se dos demais por ter alcançado um índice kappa de 0,9436.

CONCLUSÕES

Por meio deste trabalho, verificou-se que o emprego de CNNs pré-treinadas como extratores de características pode contribuir na tarefa de construir vetores mais representativos para o problema. Da mesma forma, analisar regiões de interesse ao invés de considerar todo o folíolo, ajudou a elaborar um *dataset* com maior quantidade de variações e mais inerente ao problema de identificação dos danos causados pelo Tripes. Como trabalho futuro pretende-se aumentar a base de amostras positivas fazendo uma base de dados mais equilibrada e gerar modelos que possam ser avaliados em um cenário de uso real. Uma base de imagens mais robusta permitirá avaliar outras técnicas de *Deep Learning* como fine-tuning.

AGRADECIMENTOS

Ao pesquisador Adauto Mauricio Tavares pelas informações técnicas referentes ao tripes e guaraná; à Julliane da Silva Fontes pelo auxílio nos experimentos e à Fundação de Amparo à Pesquisa do Estado do Amazonas pela disponibilização da bolsa de iniciação científica.

REFERÊNCIAS

BARBEDO, J. G. A. Plant disease identification from individual lesions and spots using deep learning. *Biosystems Engineering*, Elsevier, v. 180, p. 96–107, 2019.

DENG, J. et al. ImageNet: A Large-Scale Hierarchical Image Database. In: *CVPR09*. [S.l.: s.n.], 2009.

FONTES, J. d. S.; TAVARES, A. Aspectos da biologia da linhagem sexuada do tripes-do-guaranazeiro (pseudophilothrips adisi, zur strassen)(thysanoptera: Phlaeothripidae). In: IN: JORNADA DE INICIAÇÃO CIENTÍFICA DA EMBRAPA AMAZÔNIA OCIDENTAL, 14., 2017. *Embrapa Amazônia Ocidental-Artigo em anais de congresso*. [S.l.], 2018.

LANDIS, J. R.; KOCH, G. G. The measurement of observer agreement for categorical data. *biometrics*, JSTOR, p. 159–174, 1977.

MASTERS, T. *Practical Neural Network Recipes in C++*. [S.l.]: Acad. Press, 1993.

MOHANTY, S. P.; HUGHES, D. P.; SALATHÉ, M. Using deep learning for image-based plant disease detection. *Frontiers in plant science*, Frontiers, v. 7, p. 1419, 2016.

NETO, H. da S. et al. Transfer learning for facial emotion recognition. 2018.

PEDREGOSA, F. et al. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, v. 12, p. 2825–2830, 2011.

RAMCHARAN, A. et al. Deep learning for image-based cassava disease detection. *Frontiers in plant science*, Frontiers, v. 8, p. 1852, 2017.

SAHA, S.; EKBAL, A. Combining multiple classifiers using vote based classifier ensemble technique for named entity recognition. *Data & Knowledge Engineering*, Elsevier, v. 85, p. 15–39, 2013.

SANTOS, L. Sistema mecanizado de processamento pós-colheita de guaraná: nova tecnologia para o agronegócio e a agricultura familiar. *Embrapa Amazônia Ocidental-Capítulo em livro científico*, In: MACHADO, MWK (Org.). A engenharia de produção na contemporaneidade, 2018.

SENRA, T.; GARCIA, M.; TAVARES, A. Comportamento e controle de tripes do guaranazeiro, liothrips adisi (thysanoptera: Phlaeothripidae). In: IN: SEMINÁRIO DE PÓS-GRADUAÇÃO NA EMBRAPA AMAZÔNIA OCIDENTAL, 2008, MANAUS. *Embrapa Amazônia Ocidental-Artigo em anais de congresso*. [S.l.], 2008.

TAVARES, A.; GARCIA, M. Tripes do guaranazeiro: Liothrips adisi zur strassen, 1977 (thysanoptera: Phlaeothripidae, phlaeothripinae). *Embrapa Amazônia Ocidental-Documentos*, Manaus: Embrapa Amazônia Ocidental, 2009., 2009.

TAVARES, A. M. T.; GARCIA, M. V. B.; NASCIMENTO-FILHO, F. *Tripes do guaranazeiro: estado atual e perspectivas*. [S.l.]: Embrapa Amazônia Ocidental, 2007.