



# Tecnologias da Informação e Comunicação

e suas relações com a agricultura

*Silvia Maria Fonseca Silveira Massruhá  
Maria Angelica de Andrade Leite  
Ariovaldo Luchiari Junior  
Luciana Alvim Santos Romani*  
Editores Técnicos

**Embrapa**



# **Tecnologias da Informação e Comunicação**

**e suas relações com a agricultura**

*Empresa Brasileira de Pesquisa Agropecuária  
Embrapa Informática Agropecuária  
Ministério da Agricultura, Pecuária e Abastecimento*

# **Tecnologias da Informação e Comunicação e suas relações com a agricultura**

**Silvia Maria Fonseca Silveira Massruhá  
Maria Angelica de Andrade Leite  
Ariovaldo Luchiari Junior  
Luciana Alvim Santos Romani**  
Editores Técnicos

*Embrapa  
Brasília, DF  
2014*

# Editores Técnicos

Exemplares desta publicação podem ser solicitados à:

**Embrapa Informática Agropecuária**  
Av. André Tosello, 209 - Barão Geraldo  
Caixa Postal 6041 - 13083-886 - Campinas, SP  
Telefone: (19) 3211-5700  
www.embrapa.br  
www.embrapa.br/fale-conosco/sac

## Unidade responsável pelo conteúdo e edição

Embrapa Informática Agropecuária

## Comitê de Publicações da Embrapa Informática Agropecuária

Presidente: *Silvia Maria Fonseca Silveira Massruhá*

Secretária: *Carla Cristiane Osawa*

Membros: *Adhemar Zerlotini Neto, Stanley Robson de Medeiros Oliveira, Thiago Teixeira Santos, Maria Goretti Gurgel Praxedes, Adriana Farah Gonzalez, Neide Makiko Furukawa, Carla Cristiane Osawa*

Membros suplentes: *Felipe Rodrigues da Silva, José Ruy Porto de Carvalho, Eduardo Delgado Assad, Fábio César da Silva*

Supervisão editorial: *Stanley Robson de Medeiros Oliveira, Neide Makiko Furukawa*

Revisão de texto: *Adriana Farah Gonzalez, Daniela dos Santos, Nadir Rodrigues Pereira*

Normalização bibliográfica: *Maria Goretti Gurgel Praxedes, Márcia Izabel Fugisawa Souza*

Projeto gráfico/edição eletrônica/arte final: *Neide Makiko Furukawa*

Foto/arte capa: *Neide Makiko Furukawa*

## 1ª edição

On line (2014)

1ª impressão (2015): 1.000 exemplares

Tradução dos capítulos 18, 19 e 20 do espanhol para português: *Pangea Idiomas, São Paulo, SP*

Revisão da tradução:

*Daniela dos Santos*

Colaboração:

*André Fachini Minitti, Gustavo Vasques,*

*João Vila da Silva, Júlia Felício Capuano,*

*Paula Regina Kuser Falcão,*

*Rafaela Barbosa dos Santos, Thiago Teixeira Santos*

## Todos os direitos reservados.

A reprodução não autorizada desta publicação, no todo ou em parte, constitui violação dos direitos autorais (Lei nº 9.610).

## Dados Internacionais de Catalogação na Publicação (CIP)

Embrapa Informática Agropecuária

Tecnologia da informação e comunicação e suas relações com a agricultura / Silvia Maria Fonseca Silveira Massruhá ... [et al.], editores técnicos.- Brasília, DF : Embrapa, 2014.

411 p.: il. color. ; 17 cm x 25 cm

ISBN 978-85-7035-414-3 (On line)

ISBN 978-85-7035-487-7

1. Tecnologia da informação. 2. Automação das cadeias produtivas. 3. Biotecnologia. 4. Recursos naturais. 5. Mudanças climáticas. 6. Inovação tecnológica. I. Massruhá, Silvia Maria Fonseca Silveira. II. Leite, Maria Angélica de Andrade. III. Luchiari Júnior, Ariovaldo. IV. Romani, Luciana Alvim Santos. V. Embrapa Informática Agropecuária. VI. Programa Cooperativo para el Desarrollo Tecnológico Agroalimentario y Agroindustrial del Cono Sur - Proccisur.

CDD (21.ed.) 004

© Embrapa 2014

# Autores

**Adauto Luiz Mancini**

Cientista da computação, mestre em Ciências da Computação e Matemática Computacional pesquisador da Embrapa Informática Agropecuária, Campinas, SP

**Alan Massaru Nakai**

Cientista da computação, doutor em Ciência da Computação analista da Embrapa Informática Agropecuária, Campinas, SP

**Aldemir Chaim**

Engenheiro-agrônomo, mestre em Agronomia pesquisador da Embrapa Meio Ambiente, Jaguariúna, SP

**Alexandre Camargo Coutinho**

Biólogo, doutor em Ciências Ambientais, pesquisador da Embrapa Informática Agropecuária, Campinas, SP

**Antônio Márcio Buainain**

Economista, doutor em Economia, pesquisador do Instituto Nacional de Ciência e Tecnologia em Políticas Públicas, Estratégia e Desenvolvimento (INCT-PPED), professor livre docente do Instituto de Economia da Universidade Estadual de Campinas (Unicamp), Campinas, SP

**Ariovaldo Luchiari Junior**

Engenheiro-agrônomo, Ph.D. em Agronomia e Solos pesquisador da Embrapa Informática Agropecuária, Campinas, SP

**Aryeverton Fortes de Oliveira**

Economista, doutor em Economia Aplicada, pesquisador da Embrapa Informática Agropecuária, Campinas, SP

**Carla Geovana do Nascimento Macário**

Tecnóloga em Processamento de dados, doutora em Ciência da Computação, pesquisadora da Embrapa Informática Agropecuária, Campinas, SP

**Carlos Alberto Alves Meira**

Matemático, doutor em Engenharia Agrícola, pesquisador da Embrapa Informática Agropecuária, Campinas, SP

**Cássia Isabel Costa Mendes**

Advogada, doutoranda em Desenvolvimento Econômico pelo Instituto de Economia da Universidade Estadual de Campinas (Unicamp), analista da Embrapa Informática Agropecuária e membro do Instituto Nacional de Ciência e Tecnologia em Políticas Públicas, Estratégia e Desenvolvimento (INCT-PPED), Campinas, SP

**Fábio Danilo Vieira**

Tecnólogo em Processamento de dados, mestre em Engenharia Agrícola, analista da Embrapa Informática Agropecuária, Campinas, SP

**Fábio Rogério de Moraes**

Físico, doutor em Genética e Biologia Molecular, assistente de pesquisa da Universidade Estadual Paulista Júlio de Mesquita Filho, São José do Rio Preto, SP

**Glauber José Vaz**

Cientista da computação, mestre em Ciência da Computação, analista da Embrapa Informática Agropecuária, Campinas, SP

**Goran Neshich**

Físico químico, Ph.D. em Biofísica Molecular, pesquisador da Embrapa Informática Agropecuária, Campinas, SP

**Gustavo Alejandro Chacón Cruz**

Engenheiro de computação, pesquisador do Instituto de Investigaciones Agropecuarias, Inia La Platina, Santiago, Chile

**Inácio Henrique Yano**

Tecnólogo em Processamento de dados e Economista, mestre em Gestão de Redes de Telecomunicações, doutorando em Engenharia Agrícola, analista da Embrapa Informática Agropecuária, Campinas, SP

**Isabel Rodrigues Gerhardt**

Engenheira-agrônoma, doutora em Biologia Molecular, pesquisadora da Embrapa Informática Agropecuária e da Unidade Mista de Pesquisa em Genômica Aplicada a Mudanças Climáticas (UMiP GenClima), Campinas, SP

**Ivan Mazoni**

Tecnólogo em Processamento de dados, doutorando em Genética e Biologia Molecular, analista da Embrapa Informática Agropecuária, Campinas, SP

**Ivo Pierozzi Júnior**

Biólogo, doutor em Ecologia, pesquisador da Embrapa Informática Agropecuária, Campinas, SP

**Izabella Agostinho Pena Neshich**

Bióloga, doutoranda em Genética e Biologia Molecular do Instituto de Biologia da Universidade Estadual de Campinas, SP

**Jayme Garcia Arnal Barbedo**

Engenheiro-eletricista, doutor em Engenharia Elétrica, pesquisador da Embrapa Informática Agropecuária, Campinas, SP

**João Camargo Neto**

Engenheiro-eletricista, doutor em Engenharia de Biosistemas, analista da Embrapa Informática Agropecuária, Campinas, SP

**João dos Santos Vila da Silva**

Matemático, doutor em Engenharia Agrícola, pesquisador da Embrapa Informática Agropecuária, Campinas, SP

**João Francisco Gonçalves Antunes**

Matemático, doutor em Engenharia Agrícola, pesquisador da Embrapa Informática Agropecuária, Campinas, SP

**Jorge Luiz Corrêa**

Cientista da computação, mestre em Ciência da computação, analista da Embrapa Informática Agropecuária, Campinas, SP

**José Augusto Salim**

Engenheiro de computação, mestrando em Engenharia da Computação da Faculdade de Engenharia Elétrica e Computação da Universidade Estadual de Campinas (FEEC-Unicamp), Campinas, SP

**José Eduardo Boffino de Almeida Monteiro**

Engenheiro-agrônomo, doutor em Agrometeorologia, pesquisador da Embrapa Informática Agropecuária, Campinas, SP

**José Gilberto Jardine**

Engenheiro de alimentos, doutor em Biotecnologia, pesquisador da Embrapa Informática Agropecuária, Campinas, SP

**Juan Manuel Soares de Lima Lapetina**

Engenheiro-agrônomo, doutor em Estatística e Otimização, pesquisador do Programa Nacional de Carne e Lã do Instituto Nacional de Investigación Agropecuaria (Inia), Tacuarembó, Uruguay

**Juliana Erika de Carvalho Teixeira Yassitepe**

Engenheira-agrônoma, doutora em Agronomia, pesquisadora da Embrapa Informática Agropecuária e da Unidade Mista de Pesquisa em Genômica Aplicada a Mudanças Climáticas (UMiP GenClima), Campinas, SP

**Júlio César Dalla Mora Esquerdo**

Engenheiro-agrônomo, doutor em Engenharia Agrícola, pesquisador da Embrapa Informática Agropecuária, Campinas, SP

**Kleber Xavier Sampaio de Souza**

Engenheiro-eletricista, doutor em Engenharia Elétrica, pesquisador da Embrapa Informática Agropecuária, Campinas, SP

**Leandro Carrijo Cintra**

Cientista da computação, doutor em Bioinformática, analista da Embrapa Informática Agropecuária, Campinas, SP

**Leandro Henrique Mendonça de Oliveira**

Cientista da computação, doutor em Ciências da Computação e Matemática Computacional, analista da Embrapa Informática Agropecuária, Campinas, SP

**Leonardo Ribeiro Queiros**

Cientista da computação, doutor em Engenharia Agrícola, analista da Embrapa Informática Agropecuária, Campinas, SP

**Letícia Sayuri Nishimura**

Farmacêutica, bolsista do Hospital de Clínicas da Universidade Estadual de Campinas, Campinas, SP

**Luciana Alvim Santos Romani**

Cientista da computação, doutora em Ciência da Computação, pesquisadora da Embrapa Informática Agropecuária, Campinas, SP

**Luís Gustavo Barioni**

Engenheiro-agrônomo, Ph.D. em Ciência Animal e Pastagens, pesquisador da Embrapa Informática Agropecuária, Campinas, SP

**Luiz César Borro**

Bacharel em Informática, mestre em Ciências de Computação e Matemática Computacional, doutorando em Genética e Biologia Molecular na Universidade Estadual de Campinas (Unicamp), Campinas, SP

**Marcia Izabel Fugisawa Souza**

Bibliotecária, doutora em Educação, analista da Embrapa Informática Agropecuária, Campinas, SP

**Maria Angélica de Andrade Leite**

Engenheira-civil, doutora em Engenharia de Computação, pesquisadora da Embrapa Informática Agropecuária, Campinas, SP

**Maria do Carmo Ramos Fasiaben**

Engenheira-agrônoma, doutora em Desenvolvimento Econômico, pesquisadora da Embrapa Informática Agropecuária, Campinas, SP

**Maria Fernanda Moura**

Estatística, doutora em Ciências Matemáticas e da Computação, pesquisadora da Embrapa Informática Agropecuária, Campinas, SP

**Martha Delphino Bambini**

Engenheira-química, mestre em Política Científica e Tecnológica, analista da Embrapa Informática Agropecuária, Campinas, SP

**Miguel Angel Luengo**

Licenciado em Informática, gerente de Gestão da Informação do Instituto Nacional de Tecnología Agropecuaria (Inta), Buenos Aires, Argentina

**Poliana Fernanda Giachetto**

Zootecnista, doutora em Produção Animal, pesquisadora da Embrapa Informática Agropecuária, Campinas, SP

**Rafael Vieira de Sousa**

Engenheiro eletricista, doutor em Engenharia Mecânica, professor da Faculdade de Zootecnia e Engenharia de Alimentos da Universidade de São Paulo (FZEA-USP), Pirassununga, SP

**Ricardo Augusto Dante**

Biólogo, Ph.D. em Plant Sciences, pesquisador da Embrapa Informática Agropecuária e da Unidade Mista de Pesquisa em Genômica Aplicada a Mudanças Climáticas (UMiP GenClima), Campinas, SP

**Ricardo Yassushi Inamasu**

Engenheiro-mecânico, doutor em Engenharia Mecânica, pesquisador da Embrapa Instrumentação, São Carlos, SP

**Roberto Hiroshi Higa**

Engenheiro-eletricista, doutor em Engenharia Elétrica, pesquisador da Embrapa Informática Agropecuária, Campinas, SP

**Sergio Aparecido Braga da Cruz**

Engenheiro-eletricista, doutor em Computação Aplicada, pesquisador da Embrapa Informática Agropecuária, Campinas, SP

**Silvia Maria Fonseca Silveira Massruhá**

Analista de sistemas, doutora em Computação Aplicada, pesquisadora da Embrapa Informática Agropecuária, Campinas, SP

**Sílvio Roberto Medeiros Evangelista**

Estatístico, doutor em Engenharia Elétrica, analista da Embrapa Informática Agropecuária, Campinas, SP

**Stanley Robson de Medeiros Oliveira**

Cientista da computação, Ph.D. em Ciência da Computação, pesquisador da Embrapa Informática Agropecuária, Campinas, SP

**Tércia Zavaglia Torres**

Administradora, doutora em Educação, analista da Embrapa Informática Agropecuária, Campinas, SP

**Thiago Teixeira Santos**

Cientista da computação, doutor em Ciências da Computação,  
pesquisador da Embrapa Informática Agropecuária, Campinas, SP

**Wellington Carlos Lopes**

Tecnólogo em Processamento de dados, mestre em Engenharia Mecânica,  
pesquisador da Escola de Engenharia de São Carlos da Universidade de São Paulo (EESC-USP), São Carlos, SP

# Agradecimentos

Como organizadores desta iniciativa multi-institucional de colaboração técnico e científica, desejamos registrar nossos sinceros agradecimentos aos colegas e às instituições, os quais, por meio do trabalho colaborativo e compartilhamento de ideias, tornaram possível a realização deste documento sobre desafios e oportunidades de desenvolvimento do tema Tecnologias de Informação e Comunicação (TIC) na área de pesquisa agropecuária.

Inicialmente agradecemos ao Programa Cooperativo para el Desarrollo Tecnológico Agroalimentario y Agroindustrial del Cono Sur (Procisur), por estabelecer o grupo de Tecnologia da Informação aplicada à Agricultura, como constituinte da Plataforma de Tecnologias Emergentes, no âmbito do qual surgiu a ideia da elaboração deste documento. Dentro deste grupo, agradecemos as contribuições e o apoio dos colegas do Instituto Nacional de Tecnología Agropecuaria (Inta Argentina), Instituto Nacional de Investigaciones Agropecuaria (Inia Chile), Instituto Nacional de Investigación Agropecuaria (Inia Uruguay) por elaborarem os capítulos referentes aos trabalhos que vêm sendo desenvolvidos em seus respectivos países. Ao Instituto Interamericano de Cooperación para Agricultura (IICA-Procisur) agradecemos pelo suporte financeiro.

De forma decisiva, foi a participação dos empregados da Embrapa Informática Agropecuária. O corpo técnico se dedicou, com afinco, a elaborar capítulos deste livro não medindo esforços para divulgar as pesquisas e os trabalhos que estão sendo realizados neste centro de pesquisa sob sua responsabilidade. Com igual empenho agradecemos aos participantes do Comitê de Publicações (CP), pela revisão técnica; ao Núcleo de Comunicação Organizacional (NCO), pelas atividades de editoração e revisão gramatical; às bibliotecárias, pela revisão bibliográfica; e à equipe administrativa, por todo o suporte recebido.

Também agradecemos aos colegas do Instituto de Economia da Universidade Estadual de Campinas (IE-Unicamp), da Faculdade de Zootecnia e Engenharia de Alimentos (FZEA-USP), da Escola de Engenharia de São Carlos (EESC-USP), e da Embrapa Instrumentação que se uniram a nós trazendo sua experiência técnica e científica.

À Diretoria Executiva da Empresa Brasileira de Pesquisa Agropecuária (Embrapa) agradecemos imensamente pelo seu apoio constante e por acreditar em nosso trabalho. Às Unidades Centrais e Descentralizadas da Embrapa agradecemos por sua colaboração. Por fim, agradecemos a todos os nossos parceiros e clientes que nos desafiam continuamente a desenvolver soluções de TIC que suportem o tratamento dos dados, da informação e do conhecimento referentes às pesquisas da agricultura brasileira.

# Apresentação

A Empresa Brasileira de Pesquisa Agropecuária (Embrapa) participa do Programa Cooperativo para el Desarrollo Tecnológico Agroalimentario y Agroindustrial del Cono Sur (Procisur) que se constitui em uma iniciativa conjunta dos Institutos Nacionais de Pesquisa Agrícola do Cone Sul, visando promover sua cooperação, junto com outros atores globais envolvidos em ciência, tecnologia e inovação, para ajudar a melhorar a produtividade, a competitividade e a sustentabilidade dos recursos naturais, segurança alimentar, desenvolvimento territorial rural e equidade social da agricultura regional. Os países representados no Procisur são: Argentina, Bolívia, Brasil, Chile, Paraguai e Uruguai.

O Procisur se organiza em Plataformas Regionais, com o objetivo de articular e integrar os diferentes atores públicos e privados, de um setor específico, em empreendimentos cooperativos para promover o desenvolvimento tecnológico e a inovação em uma área de interesse comum. A Embrapa Informática Agropecuária coordena o grupo de Tecnologia da Informação aplicada à agricultura no escopo da Plataforma Regional de Uso das Tecnologias Emergentes.

Visando promover o intercâmbio de conhecimento e estabelecer a participação das novas tecnologias sobre a competitividade e a sustentabilidade da agricultura regional, a Embrapa Informática Agropecuária tomou a iniciativa de organizar a elaboração deste livro que trata do uso das tecnologias da informação e comunicação e suas relações com a agricultura. Este é o fruto de um trabalho conjunto com o Instituto Nacional de Tecnología Agropecuaria (Inta Argentina), Instituto Nacional de Investigaciones Agropecuaria (Inia Chile), o Instituto Nacional de Investigación Agropecuaria (Inia Uruguay) e o Instituto Interamericano de Cooperación para Agricultura (IICA-Procisur).

O leitor terá a oportunidade de acompanhar os principais resultados nas áreas de biologia computacional e bioinformática, automação e agricultura de precisão, gestão da informação e do conhecimento, modelagem e simulação, processamento de alto desempenho, mercado, transferência de tecnologia e perspectivas futuras das TIC sob a ótica de seu uso, para aplicação em temas como recursos naturais, mudanças climáticas, segurança fitossanitária e soluções envolvendo dados espaciais. A expectativa inicial é difundir a situação atual do uso das TIC na agricultura dos países do cone sul. Além disto, é o ponto de partida para o estabelecimento de uma base para futuras discussões, que permitam o intercâmbio de experiências, o fortalecimento de parcerias e a promoção do desenvolvimento sustentável das cadeias produtivas agrícolas do cone sul. Como resultado, espera-se o estabelecimento de incentivos que possibilitem que as TIC exerçam cada vez mais seu papel de impulsionar o crescimento do setor agrícola e a melhoria das condições sociais, ambientais e econômicas do meio rural.

*Maurício Antônio Lopes*  
Diretor-Presidente  
da Embrapa

*Waldyr Stumpf Junior*  
Diretor-Executivo de Transferência de Tecnologia  
da Embrapa

# Prefácio

Este livro é o resultado de um esforço que vem sendo realizado pelo grupo de Tecnologia da Informação aplicada à Agricultura no âmbito da Plataforma de Tecnologias Emergentes do Programa Cooperativo para el Desarrollo Tecnológico Agroalimentario y Agroindustrial del Cono Sur (Procisur). Na última reunião deste grupo do Procisur, ocorrida em julho de 2013, em Montevideo/Uruguay, a Embrapa Informática Agropecuária juntamente com os representantes do Instituto Nacional de Tecnología Agropecuaria (Inta Argentina), Instituto Nacional de Investigaciones Agropecuaria (Inia Chile), Instituto Nacional de Investigación Agropecuaria (Inia Uruguay), Instituto Paraguayo de Tecnología Agraria (IPTA Paraguay) e Instituto Interamericano de Cooperación para Agricultura (IICA-Procisur) identificaram a oportunidade de relatar e consolidar em um documento o que vem sendo discutido com as instituições participantes sobre os desafios e oportunidades de desenvolvimento do tema Tecnologias de Informação e Comunicação (TIC) na área de pesquisa agropecuária.

Globalmente, as TIC estão sendo consideradas como o terceiro pilar da pesquisa científica, junto com a teoria e a experimentação. Avanços em TIC têm um caráter estratégico e político em todas as cadeias produtivas da economia, incluindo a agricultura.

Os novos desafios da agricultura onde é necessário maior produtividade e eficiência, por meio da otimização do uso dos recursos naturais e ambientais, demandam cada vez mais a necessidade do uso de diferentes tecnologias de informação. A integração das TIC permite desenvolver soluções para a mecanização, a fitossanidade, o bem-estar animal, a rastreabilidade e segurança alimentar, o desenvolvimento de biocombustíveis, a agricultura de precisão, o gerenciamento de risco agrícola, a avaliação de impacto e mitigação das mudanças climáticas e a reprodução e biotecnologia.

Novos desenvolvimentos como a simulação, a modelagem, a gestão do conhecimento (ontologia, web semântica), os sistemas especialistas, a inteligência artificial, as redes sem fio, a mobilidade, o processamento de imagem em conjunto com padrões abertos para facilitar a troca de informações, permitem o progresso cada vez mais rápido em pesquisa e inovação, oferecendo soluções em benefício da agricultura.

A sociedade industrial encontra-se em transição para uma sociedade baseada na informação. Neste conceito, as TIC são um conjunto de ferramentas que oferece um grande potencial para a disseminação e a transferência de tecnologia. Os novos canais e as formas de comunicação de massa estão começando uma nova era da revolução da informação. É essencial que os países latino-americanos priorizem o desenvolvimento em ciência e tecnologia, inovação e particularmente em infraestrutura e recursos humanos capacitados na área das TIC.

Neste livro é apresentado como os países do Procisur estão contornando as questões que envolvem os sistemas produtivos agrícolas por meio do uso das TIC. Em particular, com relação ao Brasil, é apresentado um mapeamento do que está sendo desenvolvido na área de TIC, nos últimos cinco anos, considerando os projetos de pesquisa que a Embrapa Informática Agropecuária vem desenvolvendo em parceria com os 46 Unidades de Pesquisa da Embrapa. Seu objetivo é proporcionar subsídios para se fazer uma reflexão das principais perspectivas tecnológicas das TIC para os próximos anos, junto com os demais países participantes do Procisur.

# Sumário

Com este propósito, o livro foi dividido em sete partes, onde é descrita a situação atual das TIC em diversos temas. Na parte I, é apresentada uma visão geral de possibilidades de inovação tecnológica a partir das ferramentas de TIC e de suas aplicações (AgroTIC) na agricultura bem como o mapeamento da inserção de TIC no meio rural.

A Parte II tem um foco mais voltado para o papel das TIC na Biotecnologia distribuídos em quatro capítulos sobre genômica, bioinformática, biologia computacional e o papel da visão computacional na fenotipagem de plantas.

Na parte III é apresentado o papel das TIC no escopo dos recursos naturais e mudanças climáticas. Aqui, os capítulos abordam desde a questão de TIC na agrometeorologia, zoneamento de risco climático, mudanças climáticas bem como aplicações que envolvem dados geoespaciais.

Na parte IV é apresentado o papel das TIC no âmbito das cadeias produtivas. O primeiro capítulo aborda o uso de TIC na segurança fitossanitária. Os demais capítulos abordam o tema de automação e agricultura de precisão. Nestes capítulos, contou-se com a experiência e “expertise” dos pesquisadores da Faculdade de Zootecnia e Engenharia de Alimentos (FZEA-USP), da Escola de Engenharia de São Carlos (EESC-USP) e da Embrapa Instrumentação.

Na Parte V, o objetivo é discutir sobre algumas ferramentas de TIC e o desenvolvimento tecnológico na área de TIC para construção de aplicações na agricultura (AgroTIC). Com este objetivo, essa parte foi dividida em cinco capítulos que descrevem desde a infraestrutura para armazenamento e processamento de alto desempenho e de grande volume de dados, passando pela gestão de dados, informação e conhecimento e culminando em uma discussão sobre modelagem e simulação e mineração de dados.

Na parte VI, é abordada a questão de geração de tecnologias, aplicações AgroTIC, bem como sua transferência. Primeiramente, são descritas as aplicações de AgroTIC e seus mercados, e também alguns estudos de casos de modelos de transferência de tecnologia para produtos AgroTIC. Finalmente, algumas tecnologias emergentes assim como o futuro e as perspectivas de evolução tecnológica de AgroTIC são apresentadas.

A parte VII, refere-se às experiências que vêm sendo desenvolvidas, nas áreas citadas nos capítulos anteriores, em outros países participantes do Procisur como Chile, Argentina, e Uruguai.

*Silvia Maria Fonseca Silveira Massruhá*  
Empresa Brasileira de Pesquisa Agropecuária (Embrapa, Brasil)

*Gustavo Chacón*  
Instituto de Investigaciones Agropecuarias (Inia, Chile)

*Miguel Luengo*  
Instituto Nacional de Tecnología Agropecuaria (Inta, Argentina)

*Juan Manuel Soares de Lima*  
Instituto Nacional de Investigación Agropecuaria (Inia, Uruguay)

<b>Parte I</b>	
<b>O papel das TIC na agricultura</b> .....	21
<b>Capítulo 1</b> - Os novos desafios e oportunidades das tecnologias da informação e da comunicação na agricultura (AgroTIC) .....	23
<b>Capítulo 2</b> - Uso de computador e Internet nos estabelecimentos agropecuários brasileiros. ....	39
<b>Parte II</b>	
<b>TIC na Biotecnologia</b> .....	53
<b>Capítulo 3</b> - Genômica e biotecnologia aplicadas a adaptação a mudanças climáticas.....	55
<b>Capítulo 4</b> - Bioinformática aplicada à agricultura.....	67
<b>Capítulo 5</b> - Fenotipagem de plantas em larga escala: um novo campo de aplicação para a visão computacional na agricultura .....	85
<b>Capítulo 6</b> - Biologia computacional molecular e suas aplicações na agricultura.....	101
<b>Parte III</b>	
<b>TIC nos recursos naturais e mudanças climáticas</b> .....	119
<b>Capítulo 7</b> - TIC em agrometeorologia e mudanças climáticas.....	121
<b>Capítulo 8</b> - Tecnologias da informação aplicadas aos dados geoespaciais .....	139
<b>Parte IV</b>	
<b>TIC na automação das cadeias produtivas</b> .....	157
<b>Capítulo 9</b> - TIC na segurança fitossanitária das cadeias produtivas.....	159
<b>Capítulo 10</b> - AgroTIC em agricultura de precisão e automação agrícola.....	191
<b>Capítulo 11</b> - Automação de máquinas e implementos agrícolas: eletrônica embarcada, robótica e sistema de gestão de informação.....	215
<b>Parte V</b>	
<b>TIC na construção das aplicações para a agricultura</b> .....	233
<b>Capítulo 12</b> - Gestão da informação e do conhecimento .....	235
<b>Capítulo 13</b> - Desenvolvimento de simuladores na agropecuária.....	259
<b>Capítulo 14</b> - Mineração de dados: conceitos e um estudo de caso sobre certificação racial de ovinos .....	273
<b>Capítulo 15</b> - Métodos, procedimentos e técnicas utilizadas na construção de AgroTIC .....	293

## **Parte VI**

<b>Transferência de Tecnologia das AgroTIC e Futuro</b> .....	303
---	-----

<b>Capítulo 16</b> - Mercado de AgroTIC e transferência de tecnologia .....	305
---	-----

<b>Capítulo 17</b> - Tecnologias emergentes - futuro e evolução tecnológica das AgroTIC .....	331
---	-----

## **Parte VII**

<b>Experiência de TIC na América do Sul</b> .....	351
---	-----

<b>Capítulo 18</b> - Tecnologias da Informação e da Comunicação e sua relação com a agricultura - Chile .....	353
---	-----

<b>Capítulo 19</b> - Tecnologias de Informação e Comunicação e sua relação com a agricultura - Argentina .....	365
--	-----

<b>Capítulo 20</b> - Tecnologias de Informação e Comunicação e sua relação com a agricultura - Uruguai.....	377
---	-----

**Parte |**

**O papel das TIC na agricultura**



## Os novos desafios e oportunidades das tecnologias da informação e da comunicação na agricultura (AgroTIC)

Silvia Maria Fonseca Silveira Massruhá  
Maria Angélica de Andrade Leite  
Maria Fernanda Moura

### 1 Introdução

Nos dias atuais, a humanidade está passando por um momento de transformação em que se vive em uma grande dicotomia. Por um lado, as pessoas vivem em tempos extraordinários de grande prosperidade, vida longa saudável, tecnologias disponíveis associadas ao acesso à informação e ao conhecimento e crescimento do nível de educação. Por outro lado, deparam-se com vários riscos ao planeta, níveis altos de pobreza, enfermidades e necessidade de melhoria da qualidade da educação.

Este mundo contemporâneo e globalizado remete todos a uma busca por uma economia mais sustentável e mais justa, em que a bioeconomia ganha força e visibilidade porque a sustentabilidade entrou definitivamente como uma das prioridades da sociedade. No escopo deste trabalho, a bioeconomia é considerada um ramo da atividade humana que promete reunir todos os setores da economia que utilizam recursos biológicos (seres vivos) para oferecer soluções coerentes, eficazes e concretas para grandes desafios - como as mudanças climáticas, substituição de insumos de origem fóssil, segurança alimentar e saúde da população (EMBRAPA, 2014).

Neste contexto, em que o foco é a saúde, a qualidade de vida e o bem-estar, cada vez mais os avanços em tecnologia de informação terão um caráter estratégico e político para o Brasil e para o mundo. No relatório elaborado pela National Science Foundation dos Estados Unidos da América, *Converging Technologies for Improving Human Performance Nanotechnology, Biotechnology, Information Technology and Cognitive Science*, foi apontada a sinergia entre a nanotecnologia, a tecnologia da informação, a biotecnologia e a ciência cognitiva como o maior potencial de futuro para a humanidade nos próximos 20 anos (ROCO; BAINBRIDGE, 2002).

As tecnologias da informação e da comunicação (TIC) têm contribuído, há várias décadas, de forma impactante, para as diversas áreas de conhecimento, permitindo o armazenamento e processamento de grandes volumes de dados, automatização de processos e o intercâmbio de informações e de conhecimento. Seu grande potencial reside na sua transversalidade podendo agregar valor e benefício para as diversas áreas de negócios, mercado, agricultura e meio ambiente.

Algumas das inovações mais recentes em TIC prometem alavancar as pesquisas na agricultura gerando novas AgroTIC. Neste trabalho, AgroTIC é definido como um conjunto de aplicações específicas para agricultura que utilizam ferramentas baseadas em TIC, tais como sistemas de in-

formação geográfica (SIG), sistemas baseados em conhecimento, sistemas de suporte à decisão e modelos que são incorporados em novas tecnologias empregadas no campo. Dentre as aplicações no campo pode-se destacar: sistema de irrigação inteligente, agricultura de precisão envolvendo a aplicação de inteligência embarcada, automação e rede de sensores locais para mapeamento de solos, monitoramento de doenças e de variáveis meteorológicas. Além dessas aplicações tem-se atividades de sensoriamento remoto visando obter mais dados sobre a produção e aspectos ambientais e climáticos.

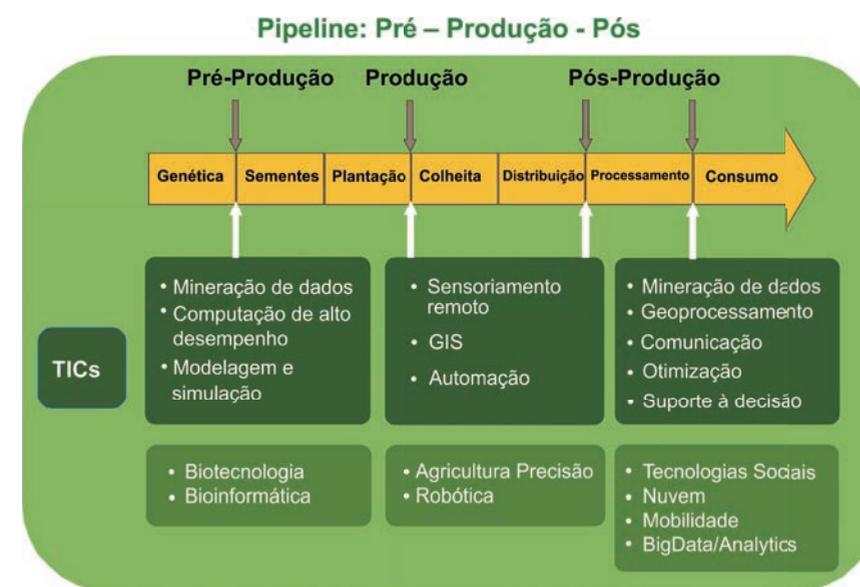
Em todas estas aplicações é produzido um grande volume de dados, também denominado *big data*, aos quais se pode aplicar técnicas de mineração de dados visando identificar padrões, de modo a gerar informações e conhecimentos para uso do setor agrícola. Adicionalmente às técnicas de mineração de dados pode-se utilizar modelos de inteligência computacional e simulação para emissão de alertas e suporte à decisão agropecuária. Pode-se destacar também o uso de SIG, bem como sistemas de informação e gestão do conhecimento implementados em aplicações web e dispositivos móveis.

No relatório elaborado pelo governo norte-americano Computational Science: Ensuring America's Competitiveness (ESTADOS UNIDOS, 2005), o Comitê Assessor de TI para o Presidente (PITAC) apontou que a tecnologia da informação (TI) constitui o terceiro pilar da investigação científica, com a teoria e a experimentação, permitindo aos cientistas construir e simular modelos de fenômenos complexos – tais como mudanças climáticas, testes de estresse estrutural em avião e explosões estelares – que não poderiam ser replicados em laboratório. Outro exemplo que retrata o papel estratégico da TI nos dias de hoje é o trabalho dos vencedores do prêmio Nobel de Química de 2013, que envolveu a criação de modelos de computador que simulam reações químicas. A academia sueca comparou o estudo dos vencedores como “levar a química do tubo de ensaio ao ciberespaço”.

Por outro lado, na agricultura são encontrados alguns desafios que devem ser superados para atingir o desejado aumento de produtividade: heterogeneidades inerentes ao ecossistema de produção agrícola (sistemas biológicos, químicos e físicos); eventos extremos da atmosfera; grande dispersão geoespacial; requisitos de segurança alimentar e alimentos seguros, além das limitações da agricultura (TING et al., 2011). A agricultura trabalha com sistemas biológicos que são inerentemente heterogêneos refletindo no ecossistema de produção agrícola. Os campos podem variar em tipo de solo e teor de umidade para a resolução de um metro quadrado. Padrões climáticos podem variar espacialmente e temporalmente em termos da luz solar e da chuva. As matérias-primas podem ter variações genéticas básicas de planta para planta e de animal para animal. De fato, a variação genética é muitas vezes biologicamente útil para aumentar a resistência a doenças e pragas.

Processos agrícolas são muito mais vulneráveis às perturbações inesperadas do que muitos outros processos industriais. Variações climáticas podem causar inundações ou trazer tempestades de granizo que, por sua vez, podem devastar plantações. Pestes ou infestações de doenças podem afetar rapidamente, se não acabar com grandes quantidades de matéria-prima. Quando se compara esse ambiente natural com o ambiente cuidadosamente controlado de uma sala limpa de indústria de fabricação de semicondutores, imediatamente pode-se entender que, por causa de forças externas, os níveis de precisão na cultura ou o rendimento do rebanho são muito mais baixos do que em outras indústrias.

Outro grande desafio da agricultura é a grande dispersão geoespacial. Vários pontos em uma cadeia de suprimentos agrícolas estão muito dispersos e o sistema agrícola global pode ser dividido em subprocessos interligados nas três principais etapas de uma cadeia produtiva, a saber: pré-produção, produção e pós-produção (Figura 1). Um desafio relacionado à dispersão geoespacial refere-se ao tratamento de algumas matérias-primas, como a pecuária e culturas perecíveis. Assim, na etapa de pós-produção, a logística e as longas distâncias, entre os pontos de processamento da cadeia de abastecimento, podem apresentar riscos para a viabilidade global de todo o processo.



**Figura 1.** Pipeline da cadeia de produção agrícola.

Fonte: Adaptado de Ting (2011).

Outras questões estão relacionadas com os requisitos de segurança alimentar e alimentos seguros que são fundamentais para os sistemas agrícolas em duas escalas de tempo separadas. Em uma escala de tempo curto de dias ou semanas, a segurança dos alimentos é crítica porque muitos produtos são eventualmente ingeridos por seres humanos. Proteger a saúde humana requer um processo de gestão da cadeia de oferta de produtos agrícolas a uma escala global. Em uma escala de anos ou décadas de tempo mais longo, a sustentabilidade do ambiente natural é fundamental para a saúde de longo prazo da sociedade. Melhorias agrícolas, como pesticidas ou fertilizantes, devem ser usadas de modo a aumentar a produtividade sem afetar negativamente a qualidade de vida global. Da mesma forma, os recursos, como a terra e a água, devem ser utilizados de uma forma que possam ser mantidos indefinidamente.

Todos esses desafios devem ser atendidos dentro das limitações inerentes ao processo agrícola. Por exemplo, a quantidade de terra arável é relativamente fixa no mundo, especialmente nos países mais desenvolvidos. O tempo também impõe uma restrição, particularmente em sistemas sensíveis a prazo, como gado e produtos perecíveis. Há uma janela de tempo finito durante o qual estes produtos agrícolas são viáveis durante o processamento. Adicionalmente, estes desafios devem ser atendidos dentro dos limites de viabilidade econômica (ou seja, custo).

Diante destes grandes desafios da agricultura, surgem novas oportunidades para a utilização de inovações na área de tecnologias da informação e comunicação (AgroTIC) em todas as etapas da cadeia produtiva. Conforme apresentado na Figura 1, na pré-produção tem-se a oportunidade de utilizar técnicas de modelagem e simulação, mineração de dados e computação de alto desempenho para, por exemplo, tratar o grande volume de dados moleculares gerados no melhoramento genético com apoio da Biotecnologia e da Bioinformática. Na etapa de produção, pode-se destacar as técnicas de sensoriamento remoto, SIG para automação das etapas de plantação e colheita, avançando para uma agricultura de precisão e robótica. Finalmente, em relação à pós-produção, as TIC já estão amplamente utilizadas nas etapas de distribuição, processamento e consumo, conforme apresentados na Figura 1 (computação em nuvem para armazenar grandes volumes de dados, análises de dados para orientação de mercado e logística, além de dispositivos móveis e tecnologias sociais para monitorar o mercado).

Na próxima seção, o papel das TIC e suas aplicações na pesquisa agropecuária são apresentados. Na seção 3 é apresentado o mapeamento de tendências das publicações em TI na agricultura, bem como o centro de pesquisa da Empresa Brasileira de Pesquisa Agropecuária (Embrapa) que tem como missão viabilizar soluções de Pesquisa, Desenvolvimento e Inovação (PD&I) em TI para a agricultura. A seção 4 trata das perspectivas de TI na agricultura. Finalmente, na seção 5, são apresentadas as considerações finais deste capítulo.

## 2 O papel das TIC na pesquisa agropecuária

A evolução recente da PD&I para o setor agrícola brasileiro pode ser discutida em função do macroambiente (mundo e Brasil) e sua relação com o desenvolvimento sustentável. O enorme avanço da agricultura brasileira, nos últimos 40 anos, confirmou a convicção existente na década de 1970 de que era necessária a criação de tecnologias adaptadas ao ambiente tropical, fundamentadas em pesquisa científica contínua e bem planejada. O avanço tecnológico evidenciado por essa crescente produtividade foi conseguido graças ao fortalecimento do Sistema Nacional de Pesquisa Agropecuária, que inclui a Embrapa, as Organizações Estaduais de Pesquisa Agropecuária (Oepa) e as universidades. Em relação à inserção do tema de tecnologia da informação no agronegócio brasileiro, pode-se afirmar que ela ocorreu no período de 1984-1985 com a criação do Centro de Informática (Ciagri) da Escola Superior de Agricultura “Luiz de Queiroz” da Universidade de São Paulo (Esalq/USP) em Piracicaba, SP, da Embrapa Instrumentação, em São Carlos, SP e da Embrapa Informática Agropecuária, em Campinas, SP, conforme descrito por Zambalde et al. (2011).

Em 1985, por meio de uma ação visionária, a Diretoria da Embrapa criou o Núcleo Tecnológico de Informática Agropecuária (NTIA) que, a partir de 1993, passou a ser chamado Centro Nacional de Pesquisa Tecnológica em Informática para Agricultura (CNPTIA), como um centro de pesquisa voltado à excelência na pesquisa e na geração de conhecimento e tecnologia em TI para a agricultura brasileira. Atualmente, este centro de pesquisa é conhecido pelo nome-síntese Embrapa Informática Agropecuária. Em 1989, foi criado o Laboratório de Automação Agrícola (LAA) da Escola Politécnica da USP, com o objetivo de desenvolver pesquisas na área de eletrônica embarcada, agricultura de precisão, zootecnia de precisão, controle de ambientes e processamento pós-colheita. No período entre 1990 e 2000, o tema TIC no Brasil caminhou a passos

largos. Em 1996, foi criada a Sociedade Brasileira de Informática Aplicada à Agropecuária e Agroindústria (SBIAGro) que, posteriormente, passou a ser denominada Associação Brasileira de Agroinformática. A partir de 1999, o LAA passou a trabalhar também na área de tecnologia de informação aplicada ao ambiente, em particular à biodiversidade, a chamada Informática na Biodiversidade (UNIVERSIDADE DE SÃO PAULO, 2014). Paralelamente, em 2000, a Embrapa Informática Agropecuária ampliou sua atuação, iniciando suas atividades de pesquisa em Bioinformática.

Desde então, a Embrapa Informática Agropecuária<sup>1</sup> tem comprovado a transversalidade da TI por meio da execução de seus projetos de pesquisa, aplicando métodos, técnicas e ferramentas de modelagem e simulação, inteligência artificial, reconhecimento de padrões e geoprocessamento, apoiados na gestão da informação e do conhecimento e no uso de tecnologias emergentes e padrões abertos. A atuação da área de pesquisa e desenvolvimento pauta-se pela visão estratégica, focada no desenvolvimento de soluções de TI, especialmente nas áreas de agroinformática e bioinformática.

A Embrapa, por sua vez, de modo a otimizar sua programação de pesquisa, tem instituído novas figuras programáticas como Portfólios e Arranjos no Sistema Embrapa de Gestão (SEG)<sup>2</sup>. Portfólios são instrumentos de apoio gerencial para a organização de projetos afins, segundo visão temática, com o objetivo de direcionar, promover e acompanhar a obtenção dos resultados finalísticos a serem alcançados naquele tema, considerando-se os objetivos estratégicos da Empresa. Dentre os 17 portfólios instituídos na Empresa, a Embrapa Informática Agropecuária participa mais ativamente do comitê gestor de três deles que envolvem vários centros de pesquisa da Embrapa, a saber: o Portfólio de Mudanças Climáticas, o Portfólio de Dinâmica e Uso da Cobertura da Terra (Geotecnologias) e o Portfólio de Automação e Agricultura de Precisão.

Enquanto os portfólios têm uma abordagem top-down, os arranjos são bottom-up. Arranjos são conjuntos de projetos convergentes, complementares e sinérgicos organizados para fazer frente a desafios prioritários em determinado tema, preferencialmente a partir da visão conjunta de mais de uma Unidade da Embrapa. Atualmente são 63 arranjos aprovados, trabalhando de forma sinérgica em temas como: melhoramento genético, sustentabilidade e sistemas de produção vegetal e animal; Huanglongbing (HLB) dos citros e a mosca-das-frutas, pragas e toxinas de grãos armazenados, entre outros. A Embrapa Informática Agropecuária participa de 35 arranjos, o que comprova a complementaridade das TIC neste novo cenário científico.

Em relação à bioinformática, diversas áreas do conhecimento, especialmente a biologia molecular, experimentaram nas três últimas décadas um crescimento exponencial na capacidade de gerar dados e, conseqüentemente, do volume de dados disponível. A bioinformática, embora originalmente lidasse com sequências proteicas na década de 1960, ganhou importância e foi reconhecida como área distinta a partir do papel decisivo nos primeiros projetos genoma, no final da década de 1980. Desde então, passou a atuar também nas áreas de expressão gênica, marcadores moleculares, evolução, regulação da expressão, modelagem de sistemas biológicos, predição de estrutura proteica e interação molecular, entre outras.

<sup>1</sup> Disponível em: <<https://www.embrapa.br/informatica-agropecuaria>>.

<sup>2</sup> Disponível em: <[www.embrapa.br](http://www.embrapa.br)>.

A Embrapa ainda tem um longo caminho a percorrer na incorporação da bioinformática em seus programas de melhoramento genético, fazendo com que somente sejam levados a campo os experimentos com alto potencial de sucesso. Estão em execução projetos na área de prospecção de novas tecnologias para obtenção de dados genômicos; bases de conhecimento em nível molecular, desvendando as principais forças que regem a comunicação e a interação entre macromoléculas biológicas; caracterização estrutural e funcional das proteínas identificadas através dos proteomas brasileiros com impacto no agronegócio; e pipeline de utilização de softwares de bioinformática na Embrapa visando a sua interoperabilidade; entre outros.

A Embrapa tem estimulado a criação de laboratórios multiusuários para atender às demandas de alta complexidade científica, envolvendo equipamentos modernos e equipes multidisciplinares altamente qualificadas. Desde 2011, o Laboratório Multiusuário de Bioinformática (LMB), que visa viabilizar soluções de bioinformática para projetos de pesquisa, desenvolvimento e inovação em um ambiente colaborativo, está sediado na Embrapa Informática Agropecuária<sup>3</sup>.

Outra iniciativa da Embrapa são as Unidades Mistas de Pesquisa (Umip). A Embrapa Informática Agropecuária participa deste novo desafio por meio da Unidade Mista de Pesquisa em Genômica Aplicada a Mudanças Climáticas (Umip GenClima), que é um laboratório conjunto entre a Embrapa e a Universidade Estadual de Campinas (Unicamp) que visa à união de esforços técnicos, científicos, materiais, operacionais e de recursos humanos voltados à geração de tecnologias genéticas e biotecnológicas, a partir da prospecção genômica, que serão utilizadas para o desenvolvimento de plantas melhor adaptadas às mudanças climáticas.

A Umip GenClima<sup>4</sup> foi implantada para identificar e validar novos genes de alto valor biotecnológico e desenvolver construções genéticas que possuam valores científicos e comerciais e que contenham novos genes que possam ser transferidos, por meio de transformação genética, para variedades comerciais de plantas desenvolvidas pela Embrapa.

As experiências com o Laboratório Multiusuário de Bioinformática e com a Umip GenClima estão sendo estendidas para outras áreas que a Embrapa visa atender, por meio de seus centros de pesquisa e instituições parceiras públicas e privadas.

Neste contexto em que a pesquisa agrícola gera grandes volumes de dados e informações, é importante um ambiente virtual para pesquisa científica, o que envolve um trabalho na área de organização da informação. Um projeto previsto na Embrapa, denominado Agropedia, prevê a construção semiautomática de mapas conceituais por meio de geração semiautomática de ontologias baseadas em mineração de textos e processamento de linguagem natural (RODRIGUES et al., 2013). A execução desse projeto dará uma importante contribuição para a inserção da Embrapa no panorama global de integração de informações previsto na Web Semântica, inclusive vinculando três das quatro áreas do relatório NBIC (ROCO; BAINBRIDGE, 2002): TI, biotecnologia e ciência cognitiva.

Adicionalmente, para contemplar os novos desafios da agricultura brasileira, além de projetos na área de organização e estruturação da informação agropecuária, estão sendo desenvolvidos

projetos no âmbito do monitoramento de fronteira agrícola e dos biomas; uso de ferramentas inteligentes na gestão de recursos hídricos, no diagnóstico de doenças e no licenciamento ambiental; sistemas de suporte à decisão para análise de impactos ambientais; sistemas de modelagem e simulação de cenários agrícolas futuros frente às mudanças climáticas; modelagem e simulação de sistemas de produção agrícola; e sistemas de rastreabilidade para produtos de origem animal e vegetal, entre outros.

Avanços na área de modelagem e simulação de crescimento de plantas são outro grande desafio na pesquisa agropecuária. Aplicações como análise funcional-estrutural de plantas, desenvolvimento de modelos de crescimento de plantas, análise de fenótipos para genômica animal e vegetal e realidade aumentada para instrumentação e controle envolvem a construção automática de modelos tridimensionais a partir de imagens digitais, de modo que possam ser produzidas e utilizadas em larga escala.

A construção automática de modelos tridimensionais para objetos simples, provenientes de atividades humanas, tais como prédios e móveis, recebeu muita atenção da comunidade de visão computacional nos anos 1990. Entretanto, a construção de modelos tridimensionais (3D) de plantas a partir de imagens digitais é muito mais complexa, dadas as estruturas orgânicas das plantas e as condições de ambiente externo que influenciam seu crescimento, tais como luz e sombra. Trabalhos de pesquisa nessa área envolvem estudos e investigações na área de computação gráfica, processamento de imagens e reconhecimento de padrões. Algoritmos desenvolvidos no âmbito desta linha de trabalho visam auxiliar as pesquisas em botânica, ambiente, genômica e proteômica. Em um horizonte mais longo, outros campos que se beneficiariam dos resultados seriam sistemas de realidade aumentada aplicada à instrumentação em agricultura e o uso de robôs em agricultura de precisão.

Conforme apresentado acima, diante dessas considerações, é possível ver que o cenário científico atual passa por um processo de mudança de paradigma, no qual cada vez mais é comum projetos de pesquisa utilizarem tecnologias capazes de adquirir e/ou gerar *terabytes/petabytes* de dados de alta qualidade e a um custo razoavelmente baixo. Além disso, esses grandes volumes de dados exigem a utilização de equipamentos de alto poder computacional para serem analisados de forma integrada, visando obter o máximo de informação e conhecimento.

Para fazer frente a esse novo cenário, a Embrapa Informática Agropecuária aprovou o arranjo denominado “Armazenamento e Processamento de Dados Experimentais da Embrapa – DataExp”, que tem por objetivo estruturar e organizar uma infraestrutura computacional de hardware e software para suporte a projetos de pesquisa da Empresa, que estejam enfrentando a questão de armazenamento, processamento e análise integrada de grandes volumes de dados. Por meio deste arranjo será criado um centro de dados da pesquisa agropecuária brasileira. Para tanto está sendo construído um novo laboratório que irá abrigar a infraestrutura computacional a ser adquirida para a sua instalação.

A Embrapa Informática Agropecuária também vem participando de outra proposta de arranjo, denominado “Métodos Quantitativos Avançados e Computação Científica na Pesquisa Agropecuária - AgroMQCC”, envolvendo 23 centros de pesquisa da Embrapa. O foco deste arranjo está no desenvolvimento e aplicação de modelos e métodos quantitativos e de computação científica para análise de dados da agropecuária em estudos prospectivos e descritivos. O arranjo AgroMQCC poderá atuar junto ao arranjo DataExp para disponibilizar e validar métodos,

<sup>3</sup> Disponível em: <<https://www.embrapa.br/informatica-agropecuaria/infraestrutura/laboratorio-multiusuario-de-bioinformatica>>.

<sup>4</sup> Disponível em: <<https://www.embrapa.br/informatica-agropecuaria/infraestrutura/unidade-mista-de-pesquisa-em-genomica-aplicada-a-mudancas-climaticas-umip-genclima>>.

técnicas, modelos e recursos de software e, ainda, fazer uso dos dados que serão armazenados no DataExp.

### 3 Mapeamento de tendências de produção científica de TI aplicada à agricultura

Para auxiliar a prospecção das atuais contribuições e tendências futuras de vários grupos de interesse que se ocupam da aplicação da TI na pesquisa agropecuária, a Embrapa Informática Agropecuária vem realizando um estudo de avaliação das publicações científicas brasileiras e internacionais nessa área. Considerando as duas vertentes de atuação: computação e agricultura, tem-se considerado representativa a comparação entre as publicações dos congressos internacionais de agroinformática, tais como o European Federation for Information Technologies in Agriculture, Food and the Environment (Efita), os congressos da American Society of Agricultural and Biological Engineers (Asabe) e os congressos da Associação Brasileira de Agroinformática (SBIAgro). Soma-se a essa comparação a produção científica da Embrapa Informática Agropecuária, para que se tenha um retrato da sua atuação frente aos tópicos cobertos por esses congressos. Em Massruhá et al. (2011) foi apresentada uma análise dessa comparação entre 1997 a 2009. Como as publicações dos Congressos da Asabe não estão publicamente disponíveis em formato digital, para se ter uma comparação estatisticamente mais justa, tem-se considerado apenas as publicações dos demais congressos a partir de 2003, que é quando o congresso da SBIAgro passou a disponibilizar publicamente seus anais em formato digital. Acredita-se que este fato não prejudica uma análise global porque Efita, SBIAgro e Asabe participam da rede Internacional Network for Information Technology (Infita) e, desta forma, os avanços norte-americanos estão também contemplados nos anais do Efita.

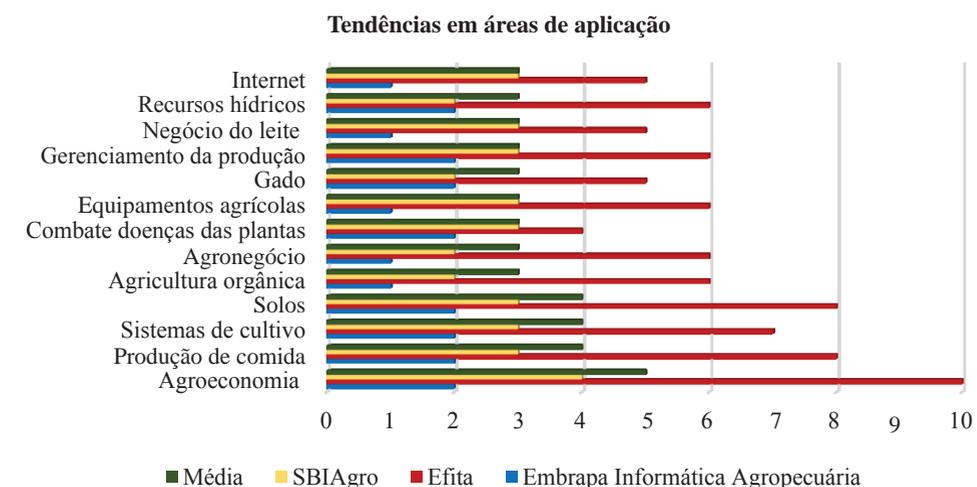
Desta forma, nas últimas análises conduzidas, tem-se os resultados dos últimos dez anos (2003 a 2013), considerando os anais do Efita, do SBIAgro e da produção científica da Embrapa Informática Agropecuária. Nessas análises, tem-se utilizado dicionários de vocabulário controlado para solucionar o problema de diferenças entre tópicos e subtópicos entre os anais dos congressos, bem como a publicação técnico-científica da Embrapa. Para encontrar esses tópicos e subtópicos comuns a todos, utilizou-se os vocábulos do Thesagro, tesouro do Ministério da Agricultura, Pecuária e Abastecimento (Mapa) e do Agrovoc da Food and Agriculture Organization of the United Nations (FAO), para cobrir os termos da área agrícola, consideradas como “áreas de aplicação”. Na área de modelos computacionais presentes nas publicações, optou-se por utilizar a taxonomia da Association for Computing Machinery (ACM) e uma tradução elaborada pela Embrapa Informática Agropecuária.

Para identificação de uma taxonomia de classes de modelos e aplicação foram utilizados métodos e técnicas de mineração de textos, de acordo com a metodologia TopTax (MOURA, 2009), na qual a informação automaticamente extraída das publicações e comparada aos vocábulos dos diversos tesouros é agrupada, descrita e, posteriormente, avaliada por um especialista do domínio de conhecimento, em um processo retroalimentativo. Desta forma, a categorização de assuntos, aqui apresentada, é resultado de um processo semiautomático, no qual o julgamento semântico subjetivo é realizado por especialistas em agroinformática.

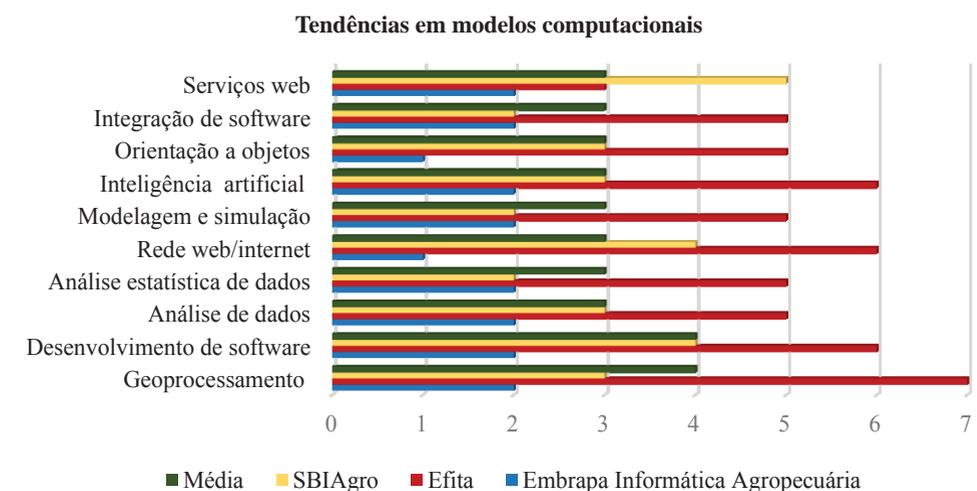
Como resultado final desses dez últimos anos, foram observadas 147 áreas de aplicações e 49 diferentes modelos computacionais. Os principais resultados obtidos podem ser observados nas Figuras a seguir, considerando-se os grupos:

- Embrapa Informática como os dados da produção científica da Embrapa Informática Agropecuária.
- Efita como as publicações dos congressos da European Federation for Information Technologies in Agriculture, Food and the Environment.
- SBIAgro como as publicações dos congressos da Associação Brasileira de Agroinformática.

Na Figura 2 são apresentadas as áreas de aplicação em agricultura 10% mais citadas nas publicações em relação à média geral dos grupos. Na Figura 3 são apresentados os modelos computacionais aplicados à agricultura também 10% mais citados nas publicações em relação à média geral dos grupos.



**Figura 2.** Principais áreas de aplicação das publicações entre 2003 e 2013.



**Figura 3.** Principais modelos computacionais utilizados nas publicações entre 2003 e 2013.

Tanto na Figura 2 quanto na Figura 3, podem-se observar que as publicações da SBIAgro e da Embrapa apresentam tendências semelhantes, embora a SBIAgro seja mais representativa (engloba todas as instituições e universidades do Brasil que participam do congresso desta associação), se comparados aos resultados apresentados pelo Efit. Na Figura 3, observa-se que os modelos computacionais apresentam um comportamento semelhante à análise apresentada em Massruhá et al. (2011), pois a importância de geoprocessamento, análise de dados, modelos para internet e modelos de inteligência computacional continuam entre os primeiros mais frequentes. Aparecem, como novidade, entre os modelos mais frequentes a integração de software e *web services* nesses últimos dez anos. Na Figura 2, pode-se observar que as áreas de aplicação tiveram um maior foco em agroecologia, produção de alimentos, sistemas de cultivo, agricultura orgânica, combate às doenças e recursos naturais (aparecem em destaque solos e recursos hídricos). Na análise anteriormente publicada em Massruhá et al. (2011), análise de mercado, agrometeorologia, gestão de propriedades e recursos naturais eram os mais frequentes. Ou seja, parece que esta nova análise reflete a mudança de foco para a produção de alimentos integrada aos sistemas de cultivo em relação às questões de mercado.

Em relação aos modelos computacionais, pode-se observar um destaque na produção científica da Embrapa Informática Agropecuária em geoprocessamento, modelos de inteligência artificial, software para análise de dados, análise estatística de dados, nos modelos de simulação, integração de base de dados e serviços web. Deve-se observar que os trabalhos na área de geoprocessamento costumam ser publicados em congressos específicos da área; assim, por vezes, não foram apresentados nos congressos considerados. Esse tipo de tecnologia é amplamente utilizada em sistemas de suporte à decisão para zoneamento, monitoramento territorial e recuperação de áreas degradadas (integração lavoura, pecuária e floresta; energia). É importante ressaltar também que o domínio desses modelos computacionais é essencial para incorporação de tecnologias avançadas no agronegócio, tais como automação, nanotecnologia, biotecnologia, sistemas de suporte à decisão, que serão citados nos próximos capítulos.

## 4 Perspectivas das TIC na agricultura

A Embrapa, ciente dos novos desafios neste mundo dinâmico e moderno, tem procurado inovar nas suas áreas de pesquisa e desenvolvimento e transferência de tecnologia. No nível estratégico, a Empresa criou o Sistema de Inteligência Estratégica, denominado Agropensa. No documento de visão 2014-2034: O Futuro do Desenvolvimento Tecnológico da Agricultura Brasileira, gerado no âmbito do Agropensa, foi proposta a criação de “Observatórios de Estudos e Tendências”, que visam capturar as principais tendências sobre o setor agropecuário no Brasil e no exterior, envolvendo suas Unidades Centrais e Descentralizadas, bem como seus laboratórios virtuais no exterior (Labex) (EMBRAPA, 2014).

Este monitoramento de tendências e perspectivas ocorre em sintonia com as cadeias produtivas agropecuárias e, para isso, foram definidos oito macrotemas que emulam o fluxo de inovação nas cadeias, a saber: recursos naturais e mudanças climáticas; novas ciências (biotecnologia, nanotecnologia, geotecnologias); automação, agricultura de precisão e tecnologias de informação e comunicação (TIC); segurança zoofitossanitária na cadeia produtiva; sistemas de produção; tecnologia agroindustrial da biomassa e química verde; segurança dos alimentos, nutrição e saúde;

além dos temas transversais: mercado, políticas e desenvolvimento rural; agricultura familiar, produção orgânica e agroecológica; inovações gerenciais nas cadeias produtivas agropecuárias; comunicação e a busca de um novo olhar sobre a agricultura.

No contexto do Sistema Agropensa, a Embrapa Informática Agropecuária, que tem como missão viabilizar soluções de pesquisa, desenvolvimento e inovação em tecnologia da informação para agricultura, está fazendo uma reflexão do que está sendo gerado nesta área e das principais tendências nos próximos 20 anos.

Nesta reflexão é importante pensar sobre o papel da Embrapa Informática Agropecuária não apenas no Brasil, mas também na América do Sul e no mundo. A Embrapa participa do Programa Cooperativo para o Desenvolvimento Tecnológico Agroalimentar e Agroindustrial do Cone Sul (Procisur)<sup>5</sup> com o Instituto Nacional de Tecnologia Agropecuária (Inta, Argentina), Instituto Nacional de Investigación Agropecuária (Inia, Chile e Uruguay), Instituto Paraguayo de Tecnología Agraria (IPTA, Paraguai) e Instituto Interamericano de Cooperación para Agricultura (IICA). A Embrapa Informática Agropecuária participa do grupo de tecnologia da informação aplicada à agricultura, no âmbito da Plataforma de Tecnologias Emergentes do Procisur, e vem discutindo com as instituições participantes sobre os desafios e oportunidades de desenvolvimento deste tema na área de pesquisa agropecuária dos países envolvidos na plataforma.

Nos últimos anos, na era da globalização e da comunicação, uma rápida transformação tem acontecido na área de TIC e uma terceira plataforma define uma nova TI. Enquanto a primeira plataforma era baseada no *mainframe* para atender milhões de usuários e a segunda plataforma era baseada na internet e em redes locais, numa arquitetura cliente/servidor para atender a centenas de milhões de usuários, a terceira plataforma é motivada por quatro importantes tendências para atender a bilhões de usuários: computação em nuvem, mobilidade, *big data*/análise preditiva e plataformas sociais.

A tecnologia de *big data* inclui o processamento de alto desempenho e armazenamento distribuído. Em função desta tecnologia é possível armazenar e processar imensos volumes de dados resultantes, por exemplo, das varreduras de satélites, produzindo mapas de alta resolução e alta frequência de imageamento, em que podem ser analisados dados de recursos naturais, uso da terra e mudanças climáticas no ambiente.

Uma vez que se tenham armazenadas as características dos solos, dos recursos hídricos, dos microclimas, dos ecossistemas, dos organismos, e seus genomas e proteomas, pode-se entender os processos globais que envolvem a natureza e a agricultura e suas influências na biosfera, incluindo os efeitos antrópicos. Através do uso de técnicas de inteligência artificial, modelagem e simulação e otimização de sistemas complexos, será possível agregar o conhecimento de todos os elos das cadeias produtivas para permitir entender o seu comportamento mediante a modelagem das variáveis biofísicas, econômicas, sociais e ambientais envolvidas em sua logística. Aplicações nesta área vão desde a simulação de crescimento de plantas, simulação de experimentos, predição da produção até o armazenamento e a distribuição otimizada dos produtos e a logística reversa envolvendo o monitoramento e o descarte controlado de resíduos.

Atualmente, fala-se em Internet das Coisas (Internet of Things) (LEE et al., 2013) considerando o aumento da oferta de dispositivos conectados à internet podendo ser tanto móveis como fixos,

<sup>5</sup> Disponível em: <www.procisur.org.uy>.

como, por exemplo, refrigeradores, equipamentos de transporte, controladores de estoque de silos e armazéns. Aliando a conectividade dos equipamentos à internet com a tecnologia de identificação por radiofrequência (RFID, na sigla em inglês), em que cada produto vegetal ou animal pode ser etiquetado, vislumbra-se aplicações de controle de estoque e distribuição controlada de produtos. Além disto, será possível acompanhar os produtos nas diversas etapas da cadeia de distribuição e, no caso de algum tipo de contaminação, eles poderão ser rastreados para verificar sua origem, contribuindo desta forma para a segurança alimentar, nutrição e saúde.

Na área da gestão da informação e do conhecimento um tópico, importante é garantir a disponibilidade, o acesso aberto e a interoperabilidade dos dados relacionados com a agricultura, bem como sua geoespacialização.

Em um futuro próximo, espera-se poder incorporar na agricultura algumas aplicações de realidade aumentada, como os aplicativos para smartphones existentes na Austrália, que permitem ao cliente fotografar uma imagem na embalagem de um lanche e disparar uma aplicação de realidade aumentada. O aplicativo transforma dados do sistema de gestão da cadeia de suprimentos do lanche, como fazendas, fornecedores, ingredientes, horário, data, meteorologia, localização geográfica e outras variáveis, em uma animação que envolve rostos e vozes dos fazendeiros reais.

Plataformas sociais são outra tendência para tornar os softwares mais colaborativos. Assim, as pessoas trabalharão em softwares que vão suportar relacionamentos profissionais, interpessoais e transacionais em um mesmo lugar como se fosse uma rede social misturada, gerando novos modelos de produção e financiamento para as empresas, como, por exemplo, *crowdsourcing* e *crowdfunding*.

Além das áreas de *big data*, análise preditiva e plataformas sociais, a mobilidade e a computação em nuvem são alguns dos pilares para inovação nas empresas. O novo perfil econômico da sociedade brasileira e a sua relação com a tecnologia, com destaque para os dispositivos móveis, estimula um modelo de data center melhor preparado para atender à tendência de crescimento de fluxo de informações, buscando eficiência, dinamismo, alta disponibilidade e baixo custo. Com a popularização dos smartphones, tablets e, mais recentemente, com a chegada do conceito BYOD (Bring Your Own Device), a mobilidade se consolida como um dos pilares fundamentais sobre os quais esta reestruturação está baseada. Outra tendência é a computação em nuvem, caminhando para uma segunda geração de nuvem, mais avançada, em que é possível ter uma nuvem privada além da nuvem pública ou um modelo híbrido.

No nível molecular, a biotecnologia também possui linhas de pesquisa que muito se beneficiam dos avanços em TIC. A genômica, e também as demais ciências ômicas (proteômica, metabólômica e transcriptômica), penetram em diferentes dimensões nos organismos e populações gerando massas de informações inimagináveis. Com a ajuda da bioinformática e da biologia computacional, é possível a análise, o processamento e o desenvolvimento de aplicações relacionadas à prospecção de dados genômicos e da estrutura das proteínas e seus efeitos colaterais em drogas farmacêuticas. Hoje já se fala em biossimulação e biopredição, isto é, a simulação e a predição no mundo biológico, no mundo vivo, na medicina e na agricultura, similar ao que acontece na indústria automobilística e aeronáutica, em que os engenheiros desenvolvem grande parte de seus produtos por computador, simulando as partes mecânicas, evitando protótipos defeituosos e minimizando tempo e custo. Nos dias atuais, já se tem exemplos da integração de nanotecnologia, jogos, *crowdsourcing* e dispositivos conectados, emergindo como componentes

importantes na área de bioinformática. Por exemplo, o jogo chamado Dizeez visa solucionar questões relacionadas à prospecção de dados genômicos e da estrutura das proteínas, conforme publicado pela *Communications of ACM*, de maio de 2014.

No âmbito das cadeias produtivas do agronegócio, o conhecimento nas suas diversas etapas, aliado à nanotecnologia, à robótica e à agricultura e pecuária de precisão representadas por nanossensores, nanorobôs, veículos aéreos não tripulados (vant) e máquinas agrícolas, tornará possível o desenvolvimento de aplicações inteligentes capazes de capturar dados dos nanossensores e vant e criar aplicações de mapeamento sofisticadas para a tomada de decisão. Estas aplicações poderão ser adaptadas às máquinas agrícolas, vant e aos nanorobôs, para que atuem de acordo com a necessidade da cultura e em função do entendimento das suas necessidades, em ações de irrigação, aplicação de fertilizantes e pesticidas e realização de colheita de forma inteligente.

Dada a complexidade e heterogeneidade das tecnologias emergentes como as TIC, a nanotecnologia, a biotecnologia, a robótica e a agricultura de precisão, e suas aplicações na agroindústria, é natural que ainda não se tenha uma perspectiva sistemática, integrada e interdisciplinar entre elas, como apontado no relatório NBIC (ROCO; BAINBRIDGE, 2002).

Em um exemplo mais direto apontado pelo relatório NBIC (ROCO; BAINBRIDGE, 2002), a agricultura poderia aumentar grandemente sua produtividade e reduzir o desperdício com o uso de redes de sensores baratos que monitorariam constantemente as condições e necessidades das plantas, animais e insumos de uma fazenda. O relatório vai além, apontando que os desenvolvimentos recentes em abordagens sistêmicas, matemática e computação permitirão, pela primeira vez, entender o mundo natural, a sociedade humana e a pesquisa científica como sistemas complexos, hierárquicos e fortemente acoplados. Estão previstos impactos na eficiência do trabalho e aprendizado, melhoria da capacidade cognitiva e sensorial individual, mudanças drásticas na medicina, melhora na criatividade individual e coletiva, formas de comunicação altamente eficientes, incluindo comunicação cérebro-cérebro e interface homem-máquina, entre outros.

Em artigos apresentados em fevereiro de 2008 na edição Especial Robótica da Scientific American Brasil, especialistas afirmam que por volta de 2025 existirão computadores custando US\$ 1 mil com poder de processamento de 100 milhões de instruções por segundo (mip), o equivalente a um cérebro humano, capazes de imitar o raciocínio humano para diversas aplicações práticas. Os mais otimistas afirmam que, em 2055, o computador pessoal terá o poder de processamento de todos os cérebros humanos juntos.

Também na Europa e no Japão, há um crescente interesse na utilização da computação em modelagem e simulação. Dentro do Programa Quadro 7 (UNIÃO EUROPÉIA, 2006), financiado pela Comunidade Europeia, no âmbito do tema tecnologias da informação e comunicação, existem projetos destinados ao progresso conjunto das TIC e ciências biológicas e de estudo do funcionamento do cérebro. No projeto Sistemas Computacionais Auto-Construídos (Self-constructed Computing Systems - SECO Project), a partir de partes mais simples, cujo comportamento se pode determinar, busca-se explicar o surgimento de funções mais complexas; por exemplo, como a mente surge a partir de alguns poucos tipos de neurônios inibidores e excitadores interconectados existentes no neocórtex. A última fronteira - construção de um robô humanóide pensante - está ainda distante. Entretanto, importantes passos já foram dados: os robôs humanóides japoneses já apresentam um alto grau de avanço, mas ainda não apresentam consciência.

O Programa Quadro 7 (UNIÃO EUROPÉIA, 2006) estabeleceu também como desafio a construção de bibliotecas digitais e de sistemas gestores de conhecimento que incorporem algum grau de inteligência e tratamento semântico. A Web Semântica tem sido usada com sucesso nas áreas médica e farmacêutica: a indústria Eli Lilly a tem usado para organizar dados heterogêneos de diferentes fontes, tais como registros de pacientes, estruturas químicas, sequências de ácido desoxirribonucleico (DNA), imagens, processos biológicos e artigos científicos, com o objetivo de priorizar alvos biológicos para descoberta de novas drogas.

O documento *Visões para o Futuro da Nanotecnologia* (SCHMIDT, 2007), organizado pela National Science Foundation e National Institute of Health dos Estados Unidos, apontou o papel crucial da TI tanto na organização da informação de uma biblioteca sobre o mundo nanométrico (Nano Library), quanto no tratamento da informação gerada por nanosensores. A manipulação da matéria no nível de átomos e moléculas, realizada em escala nanométrica, está beneficiando a chamada biologia sintética: o projeto e construção de novas partes, dispositivos e sistemas que não existem no mundo natural, bem como o reprojeto dos sistemas biológicos existentes para executar tarefas específicas.

Em um artigo elaborado pelo ETC Group para o governo canadense (EXTREME GENETIC ENGINEERING, 2009), comenta-se que não existe barreira técnica à síntese de plantas e animais e que isto ocorrerá logo que alguém se disponha a financiar tais projetos. Neste mesmo artigo, um pesquisador do Massachusetts Institute of Technology (MIT) prevê que os engenheiros biológicos do futuro começarão o trabalho em seus laptops, não nos laboratórios.

Para contemplar estes desafios em escalas antes inimagináveis, um conjunto de ferramentas e tecnologias se tornam necessárias. A síntese da tecnologia da informação e ciência para lidar com grandes volumes de dados de forma colaborativa e multidisciplinar é o que está se denominando *e-science* (BELL et al., 2009). Outra nova área de estudo é a *Data Science*, que se refere ao estudo sistemático, a partir de conhecimento extraído de grandes volumes de dados, para gerar explicações e predições nas várias áreas do conhecimento, conforme artigo publicado na revista *Communications of the ACM* (Association of Computer Machinery) de dezembro de 2013.

As principais empresas de tecnologia como a Google e a Microsoft vêm concentrando seus esforços na área de *deep-learning*, um campo relativamente novo de pesquisa em inteligência artificial baseado no estudo de redes neurais voltado à criação de produtos que podem entender e aprender a partir de imagens, textos e vídeos. Atividades como o reconhecimento facial em vídeo ou palavras na fala humana, com tradução de voz em tempo real, são alguns dos resultados já alcançados nesta área, conforme descrito em artigo publicado em *Communications of the ACM* (Association of Computer Machinery), em outubro de 2013.

Mirando todos estes novos desafios e para manter sua competitividade no cenário internacional e conquistar novos mercados, os países-membros do ProciSur deverão aumentar a eficiência de seus sistemas produtivos em termos de uso de insumos agrícolas, incluindo o provimento de alternativas orgânicas, biológicas ou naturais, além do uso otimizado de água e energia. A essas restrições somam-se questões como a necessidade de preservação dos biomas, os mecanismos de sequestro de carbono, certificação de qualidade dos produtos e rastreabilidade dos alimentos, garantia de bem-estar animal, equilíbrio social, as mudanças climáticas e a intensificação da agricultura na matriz energética mundial por meio dos biocombustíveis. Por ser uma área transversal,

a tecnologia da informação tem o potencial de aplicação em todas essas questões. Quando se fala em ciência, em qualquer que seja a área de conhecimento, depende-se da computação para processar gigantescas massas de dados ou simular novos e complexos fenômenos. Na agricultura brasileira, não seria possível antecipar as mudanças climáticas, realizar previsões meteorológicas, monitorar o desmatamento da floresta amazônica e realizar as pesquisas genéticas se não fossem os avanços alcançados na área de TI.

## 5 Considerações finais

No âmbito do ProciSur, a Embrapa Informática Agropecuária representa a Embrapa no grupo de tecnologia da informação aplicada à agricultura, da Plataforma de Tecnologias Emergentes. Neste capítulo, tendo como perspectiva os novos cenários agrícolas e o desenvolvimento rural sustentável, foi apresentada uma visão geral de como as TIC têm contribuído para as diversas áreas de conhecimento, focando especialmente em suas aplicações na agropecuária.

Também foi mostrada uma análise das tendências de publicações dos principais congressos de TI na agricultura, considerando os congressos do Efitá, da Asabe e da SBIAgro, comparando com a produção científica da Embrapa Informática Agropecuária, para construir um retrato da atuação desta frente aos tópicos cobertos por esses congressos. Ainda no que se refere à Embrapa Informática Agropecuária, outro ponto de destaque foi o mapeamento da sua atuação na área de TIC, nos últimos cinco anos, tendo como base os projetos de pesquisa por ela liderados e executados em parceria com os demais 46 centros de pesquisa da Embrapa. Com relação à Embrapa, foi descrito seu sistema de gestão da pesquisa, por meio das figuras programáticas, e de iniciativas corporativas como os laboratórios multiusuários e unidades mistas de pesquisas, bem como seu Sistema de Inteligência Estratégica Agropensa.

No âmbito da Plataforma de Tecnologias Emergentes, as instituições participantes vêm discutindo sobre os desafios e oportunidades de desenvolvimento das TIC na área de pesquisa agropecuária. Tecnologias como *big data*, análise preditiva e plataformas sociais, a mobilidade e a computação em nuvem são alguns dos pilares para inovação nas empresas, além de modelos computacionais como técnicas de geoprocessamento, modelos de inteligência artificial, software para análise de dados, análise estatística de dados, modelos de simulação, integração de base de dados e serviços web que permitem aos cientistas construir e simular modelos de fenômenos complexos na agricultura.

Neste livro estão sendo abordadas as perspectivas de TIC para incorporação destas tecnologias avançadas no agronegócio, em áreas como biotecnologia, recursos naturais e mudanças climáticas, automação e agricultura de precisão, assim como uma visão de transferência destas tecnologias. Desta forma, nos capítulos que se seguem, espera-se provocar uma reflexão dos principais desafios e oportunidades do uso das TIC na pesquisa agropecuária, em todos os elos da cadeia produtiva e sua inserção no meio rural visando à geração de inovação na agricultura nos países do cone Sul.

## 6 Referências

- BELL, G.; HEY, T.; SZALAY, A. Beyond the data deluge. **Science**, Washington, D.C, v. 323. n. 5919, p. 1297-1298. Mar. 2009. Disponível em: <<http://www.sciencemag.org/content/323/5919/1297>>. Acesso em: 2 out. 2014. DOI: 10.1126/science.1170411
- EMBRAPA. **Visão 2014-2034: o futuro do desenvolvimento tecnológico da agricultura brasileira: síntese**. Brasília, DF, 2014. 53 p.
- ESTADOS UNIDOS. President's Information Technology Advisory Committee. **Computational Science: Ensuring America's Competitiveness**. Arlington: National Coordination Office for Information Technology Research Development, 2005. 104 p. Report to the President.
- EXTREME genetic engineering: an introduction to synthetic Biology. [S.l.]: ETC Group, 2007. 72 p.
- LEE, M.; HWANG, J.; YOE, H. Agricultural Production System Based on IoT. In: IEEE INTERNATIONAL CONFERENCE ON COMPUTATIONAL SCIENCE AND ENGINEERING, 16., 2013, Sydney. **Proceedings...** Los Alamitos: IEEE, 2013, p. 833-837. DOI: 10.1109/CSE.2013.126. Conference Publishing Series.
- MASSRUHÁ, S. M. F. S.; SOUZA, K. X. S. de; LEITE, M. A. de A.; MOURA, M. F.; SARAIVA, A. M. Tendências e perspectivas da tecnologia da informação aplicada à agricultura. In: MENDES, C. I. C.; OLIVEIRA, D. R. M. dos S.; SANTOS, A. R. dos. (Ed.). **Estudo do mercado brasileiro de software para o agronegócio**. Campinas: Embrapa Informática Agropecuária, 2011. Cap. 6. p. 147-168.
- MOURA, M. F. **Contribuições para a construção de taxonomias de tópicos em domínios restritos utilizando aprendizado estatístico, 2009**. 137 f. Tese (Doutorado em Ciências de Computação e Matemática Computacional) - Instituto de Ciências Matemáticas e de Computação, USP, São Carlos, SP.
- ROCO, M. C.; BAINBRIDGE, W. S. (Ed). **Coverging technologies for improving human performance: nanotechnology, biotechnology, information technology and cognitive science**. Dordrecht; Boston: Kluwer Academic, 2002. 482 p. il.
- RODRIGUES, L. A.; LACERDA, M. G. de; VAZ, G. J.; PIEROZZI JÚNIOR, I. Arquitetura da informação dos sites da Agropedia brasilis. In: MOSTRA DE ESTAGIÁRIOS E BOLSISTAS DA EMBRAPA INFORMÁTICA AGROPECUÁRIA, 9., 2013. **Resumos...** Brasília, DF: Embrapa, 2013. p. 51-54.
- SCHMIDT, K. **Nanofrontiers: visions for the future of nanotechnology**. Washington, D. C.: Woodrow Wilson International Center for Scholars, 2007. 51 p. il. Project on Emerging Technologies.
- TING, K.C.; ABDELZAHER, T.; ALLEYNE, A.; RODRIGUEZ L. Information technology and agriculture: global challenges and opportunities. **The Bridge**, Washington, D.C., v. 41, n. 3, p. 6-13, 2011.
- UNIÃO EUROPEIA. Comissão Europeia. **FP7: As respostas do amanhã começam hoje**. Bruxelas: Comunidades Europeias, 2006. 32 p. (Investigação comunitária). Disponível em: <[http://ec.europa.eu/research/fp7/pdf/fp7-factsheets\\_pt.pdf](http://ec.europa.eu/research/fp7/pdf/fp7-factsheets_pt.pdf)>. Acesso em: 22 out. 2014.
- UNIVERSIDADE DE SÃO PAULO. Escola Politécnica. **Laboratório de Automação Agrícola: apresentação**. São Paulo, [2014]. Disponível em: <http://www.pcs.usp.br/~laa/html/pagina.php?p=apresentacao>. Acesso em: 22 set. 2014.
- ZAMBALDE, A. L.; SCHNEIDER, H.; LOPES, M. A.; PAGLIS, C. M.; BAMBINI, M. D. **Tecnologia da informação no agronegócio**. In: MENDES, C. I. C.; OLIVEIRA, D. R. M. dos S.; SANTOS, A. R. dos. (Ed.). **Estudo do mercado brasileiro de software para o agronegócio**. Campinas: Embrapa Informática Agropecuária, 2011. Cap. 2. p. 39-72.

## Uso de computador e Internet nos estabelecimentos agropecuários brasileiros

Cássia Isabel Costa Mendes  
Antônio Márcio Buainain  
Maria do Carmo Ramos Fasiaben

### 1 Introdução

As tecnologias da informação e comunicação (TIC) são responsáveis coadjuvantes pelas profundas transformações nos modelos de produção e acumulação até então vigentes e configuram o surgimento da denominada Sociedade da Informação (IBGE, 2009)<sup>1</sup>.

Os impactos revolucionários das TIC sobre a sociedade contemporânea são por demais visíveis e já assimilados no cotidiano das famílias e das instituições. No entanto, não é tão visível quando se trata da agropecuária, ainda que isto não signifique que não seja importante. Ao contrário, a aplicação das TIC na agropecuária é uma realidade, como na agricultura de precisão, nas máquinas que interagem com sinais de satélites, em sistemas de monitoramento das condições de solo.

O uso de TIC na agropecuária pode trazer como benefícios agrícolas e econômicos a melhoria da gestão da produção e da propriedade rural, a disseminação de importantes informações do setor, melhoria no planejamento, monitoramento e acompanhamento de produção integrada e o acesso aos mais recentes resultados de pesquisa na área (GELB; VOET, 2009).

Goyal e Gonzáles-Velosa (2012) complementam o rol de benefícios esperados com o uso de TIC no espaço rural indicando o papel que podem ter na:

- 1) Redução dos custos de comunicação entre os agentes (como os relacionados ao uso de celulares e ferramentas de internet).
- 2) Redução dos custos de acesso a serviços (como seguro e crédito) e informação (de mercado ou tecnológica).
- 3) Redução dos riscos relacionados a eventos climáticos, pragas e desastres naturais, viabilizados por sistemas de monitoramento e de informação acessíveis.
- 4) Ampliação de renda, qualidade e atendimento às demandas dos consumidores de produtos agrícolas.

---

<sup>1</sup> O intercâmbio entre agricultura e TIC fez surgir uma área denominada agroinformática, que estuda as aplicações de TIC para o ambiente rural, em níveis mundial e nacional. Para informações sobre o surgimento e evolução da agroinformática no Brasil, ver Zambalde et al. (2011).

A despeito dos benefícios esperados com o uso de TIC como um dos elementos motrizes da agricultura, o produtor rural brasileiro ainda tem acesso restrito e usa pouco tais tecnologias em seu empreendimento agrícola.

O último Censo Agropecuário brasileiro inseriu na categoria de eletrodomésticos utilizados na propriedade rural o acesso a computador e à internet. Embora se admita que alguns dados referentes a uso de computador e internet, relativos a 2006, estejam ultrapassados - dada a rapidez na propagação desses instrumentos -, considera-se que questões de fundo subsistem, e ainda são responsáveis pela desigualdade no acesso à tecnologia da informação ao longo do território nacional. Entre as questões, podem ser mencionadas as atinentes ao nível de instrução dos produtores, a condição do produtor em relação às terras, o grau de dinamismo das diferentes atividades econômicas, entre outras variáveis.

Há imperativa necessidade de se conhecer e entender estes dados - tanto por instituições públicas e privadas de pesquisa, extensão, ensino e fomento que atuam com TIC rural, como por usuários atuais e potenciais - para que se tenha uma noção do caminho já percorrido na busca de tentar prover as condições necessárias para acesso a computador e a internet na área rural, e estimar a distância que falta para alcance desta meta.

Este capítulo analisa os dados de acesso ao computador e à internet pelos agricultores brasileiros, segundo o Censo Agropecuário 2006, em relação ao nível de instrução, condição do produtor, direção do estabelecimento, grupos de atividade econômica e distribuição por região. O capítulo estrutura-se em quatro seções, incluindo esta introdução. Para contextualizar a relevância do tema na esfera internacional, a próxima seção relata alguns estudos sobre a adoção de TIC no campo em países da Europa e nos Estados Unidos. Em seguida, analisam-se os dados de acesso, pelos agricultores brasileiros, ao computador e à internet em seus empreendimentos agrícolas. Por último, seguem as considerações finais.

## 2 TIC em países da Europa e nos Estados Unidos

A adoção de TIC no campo nos países da Europa e nos Estados Unidos tem sido objeto dos estudos de Gelb e Voet (2009) e Gelb (2012, 2013), que abrangem o período de 1999 a 2013. Estes autores efetuaram levantamento de dados junto a especialistas em agroinformática, participantes<sup>2</sup> dos congressos bienais da *European Federation Information Technologies in Agriculture, Food and Environment* (EFITA).

Para Gelb (2012), há uma clara tendência de que a adoção da TIC na agricultura é reconhecida, pela maioria dos participantes da conferência EFITA, como um problema permanente. De 2001 até 2013, em sete eventos consecutivos, os especialistas em agroinformática reforçaram essa realidade, como demonstra a Tabela 1.

Os problemas para adoção de TIC no campo apresentam como fatores limitantes os listados na Tabela 2.

As interpretações das tendências apresentadas na Tabela 2 são: desde 1999, o custo e os aspectos de infraestrutura não foram observados como uma restrição/limitação dominante. Segundo Gelb

**Tabela 1.** Há problemas com a adoção de TIC na agricultura? (% de “SIM” nas respostas).

Montpellier (França) 2001	Debrecen (Hungria) 2003	Villa Real (Portugal) 2005	Glasgow (Escócia) 2007	Wageningen (Países Baixos) 2009	Praga (República Checa) 2011	Torino (Itália) 2013
72.0	72.5	96.7	94.4	90.3	90.0	78.9

Fonte: Gelb (2012, 2013).

**Tabela 2.** Fatores limitantes do uso de TIC pelos agricultores (% de “SIM” nas respostas).

Conferências EFITA	Inabilidade em usar TI	Aspectos de infraestrutura	Custo	Falta de treinamento
1999 - Bonn (Alemanha)	22,4	18,9	17,6	8,6
2001 - Montpellier (França)	29,3	0,0	32,3	16,9
2003 - Debrecen (Hungria)	5,9	19,6	39,2	35,3
2005 - Villa Real (Portugal)	45,0	35,0	23,0	16,7
2007 - Glasgow (Escócia)	12,5	28,6	42,2	17,9
2009 - Wageningen (Países Baixos)	45,2	23,8	29,0	58,0
2011 - Praga (República Checa)	<45,0	0,0	<25,0	>65,0
2013 - Torino (Itália)	56,1	21,0	19,3	54,3

Fonte: Gelb (2012, 2013).

(2012, 2013), as questões de infraestrutura parecem ter sido resolvidas em grande parte no contexto mais amplo de serviços de comunicação, tanto regionais como nacionais para as realidades europeia e norte-americana.

Nos Estados Unidos, no período de 2003 a 2013, verificou-se um crescimento no uso de computadores, que chegou a 70%, em 2013, e a 65% o acesso à Internet, no mesmo ano. A utilização nos negócios agrícolas também tem aumentado ao longo deste período, que saltou de 30% para 40% (ESTADOS UNIDOS, 2013).

No entanto, como se verá mais adiante neste capítulo, no Brasil<sup>3</sup> a agricultura vive uma situação bem diversa dos países europeus e americano.

De acordo com os dados da Tabela 2, os custos representaram um baixo fator limitante. Talvez uma tendência futura seja supor que os custos das inovações em TIC serão compensados por menores custos de produção e maior eficiência de equipamentos, sistemas, comunicações e da capacidade das pessoas envolvidas com o desenvolvimento de soluções em TIC (por exemplo, a automação).

No entanto, a falta de melhoria na capacidade do agricultor em utilizar a TI é um fator limitante bem maior do que os fatores custo e infraestrutura. Complementarmente, a falta de treinamento é uma restrição crescente e influente. Gelb (2012) pondera que, dentro dos limites de precisão das respostas, estes fatores podem indicar que uma das causas refere-se ao aumento das complexidades da TIC desde 1999.

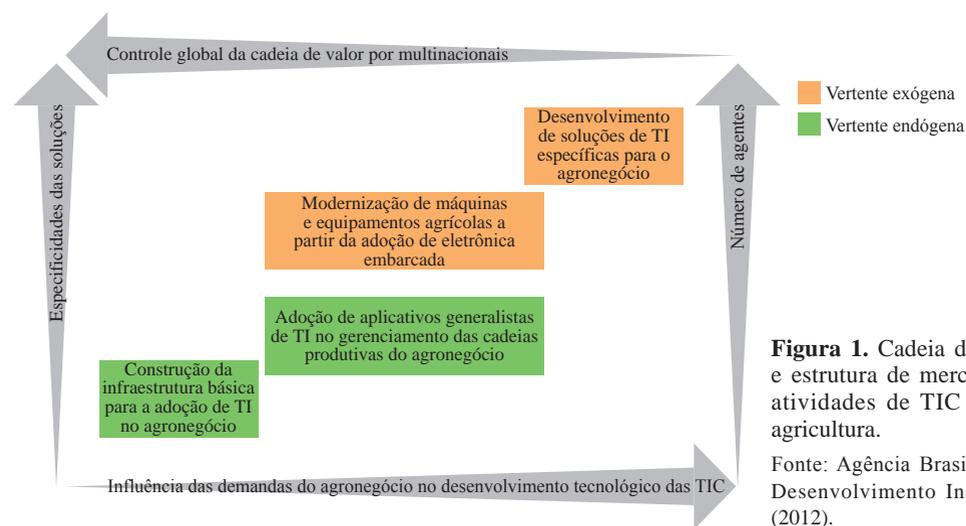
<sup>3</sup> São reconhecidas as enormes diferenças - social, econômica, estrutural e cultural - entre Brasil e Estados Unidos. A apresentação dos dados norte-americanos serve para evidenciar como o Brasil está distante de proporcionar condições básicas de infraestrutura mínima aos agricultores brasileiros, iniciando pelo acesso à internet e computador, como se discute na seção seguinte deste capítulo.

<sup>2</sup> A lista e o perfil dos especialistas em agroinformática que responderam os questionários estão em Gelb (2012, p. 10).

Para Gelb e Voet (2009), a identificação de problemas específicos de adoção em TIC pode levar a benefícios significativos para os formuladores de políticas públicas, instituições de pesquisa, extensão rural e ensino que atuam com tecnologia da informação rural.

No Brasil, a adoção de TIC na agricultura ocorre a partir de duas grandes vertentes. A primeira, denominada exógena, relaciona-se à disseminação de soluções de TIC de outros setores econômicos para o setor agrícola. Esta vertente divide-se em dois blocos: a disseminação da infraestrutura básica de TIC e a posterior adoção de soluções em TIC de gerenciamento empresarial. A segunda vertente, de caráter endógeno, diz respeito ao desenvolvimento de aplicações/soluções em TIC específicas para a agricultura. Isso se deve ao esforço de empresas em desenvolver soluções de TIC e à sofisticação de máquinas e equipamentos que requerem o uso de eletrônica embarcada.

A Figura 1 ilustra estas duas grandes forças simultâneas de expansão da TIC na agricultura nacional que, ao atuarem, constroem a cadeia de valor de TIC agrícola, de acordo com estudo da Agência Brasileira de Desenvolvimento Industrial (2012).



**Figura 1.** Cadeia de valor e estrutura de mercado de atividades de TIC para a agricultura. Fonte: Agência Brasileira de Desenvolvimento Industrial (2012).

Segundo a Agência Brasileira de Desenvolvimento Industrial (2012), a expansão de infraestrutura em TIC pode configurar um vetor inicial de disseminação no espaço rural. Essa disseminação pode se tornar um catalisador da adoção inicial da agroinformática, mesmo por parte de propriedades de pequeno e médio porte.

Como mostra a Figura 1, a vertente endógena refere-se ao desenvolvimento de aplicações em TIC específicas para a agricultura, que pode contribuir para a sua expansão no campo. Neste sentido, a Embrapa Informática Agropecuária busca desenvolver soluções em TIC aplicadas à agricultura, como evidenciado em outros capítulos deste livro.

Os estudos corroboram a relevância de se avançar no conhecimento sobre o tema, principalmente atinente aos fatores condicionantes para uso da TIC. A infraestrutura de acesso a computadores e a internet é um dos pré-requisitos essenciais para que essa adoção ocorra. Como apresentado no início do capítulo, nos estudos de Gelb (2012, 2013) a infraestrutura não foi considerada uma restrição/limitação dominante para a realidade americana e europeia. A realidade brasileira, po-

rém, é muito diferente e reflete a indigência digital dos produtores rurais no que tange ao acesso a computador e internet, como retrata a seção seguinte.

### 3 Acesso a computador e internet na agricultura brasileira

Esta seção apresenta o retrato do acesso a instrumentos que lembram o uso de TIC - computador e internet - com base no Censo Agropecuário de 2006.

O Censo Agropecuário de 2006 (IBGE, 2006) inovou sua pesquisa ao incluir, no rol de eletrodomésticos utilizados nos estabelecimentos agropecuários, o uso de computador e de internet. No Brasil, algumas pesquisas, em áreas de menor abrangência geográfica, também buscaram levantar estes dados, como é o caso do Levantamento Censitário de Unidades de Produção Agropecuária (LUPA)<sup>4</sup> e da TIC Domicílios do Comitê Gestor da Internet no Brasil (CGI.br).

Note-se que a posse do computador não pode ser interpretada, de maneira automática, como uso da TI, uma vez que o mesmo tende a ser utilizado para muitas outras funções no âmbito da família. Mas ainda que aceitássemos os computadores como uma aproximação de uso da TI, o Censo Agropecuário de 2006 confirmou a indigência mencionada anteriormente: do total de estabelecimentos agropecuários pesquisados, apenas 4,54% tinham computador e 1,87% tinham acesso à internet no Brasil, conforme mostra a Tabela 3. Ou seja, era mesmo muito baixo o índice de acesso a estes instrumentos que remetem ao uso de tecnologias da informação.

**Tabela 3.** Número de estabelecimentos agropecuários com computador e acesso à internet no Brasil, segundo dados do Censo Agropecuário 2006.

Tipo de eletrodoméstico utilizado	Número de estabelecimentos agropecuários	
	Unidades	Percentual
Computador	183.623	4,54
Acesso à internet	75.407	1,87

Fonte: Mendes et al. (2013).

Os dados do Censo Agropecuário 2006 (IBGE, 2006) apontam, para o estado de São Paulo, que 16,87% dos estabelecimentos tinham computador e que 9,48% deles tinham acesso à internet.

Outra fonte de informação sobre o uso de computador e internet no Brasil é a TIC Domicílios, realizada desde 2005 pelo Comitê Gestor da Internet no Brasil (2012). Trata-se de uma pesquisa amostral, a qual em sua edição de 2012, analisou 25 mil domicílios. Na pesquisa são considerados como equipamentos de TIC: televisão, antena parabólica, televisão por assinatura, rádio, telefone fixo, telefone celular, console de jogo (videogame), computador de mesa (desktop), computador portátil (notebook), computador de mão (palm top). No presente trabalho serão tomados da TIC Domicílios os dados referentes a computador e ao uso da internet, de modo a tecer comparações de suas ocorrências nas áreas rurais e urbanas do Brasil.

<sup>4</sup> Trata-se do Censo das Unidades de Produção Agropecuária paulistas, realizado pelo Instituto de Economia Agrícola (IEA) e pela Coordenadoria de Assistência Técnica Integrada (CATI), órgãos da Secretaria da Agricultura e Abastecimento do Estado de São Paulo.

A TIC Domicílios analisou, a partir de 2009, separadamente, os dados quanto ao uso de TIC obtidos em municípios urbanos e rurais. Cabe esclarecer que o domicílio rural nem sempre implica em uma propriedade agropecuária. A pesquisa de 2012 evidencia a disparidade envolvendo a posse de computadores em áreas urbanas e rurais, descrita na Tabela 4.

**Tabela 4.** Percentual de domicílios com computador e com internet no Brasil, segundo CGI.br para o ano de 2012.

Tipo de eletrodoméstico utilizado	Perímetro	Sim	Não
Computador	Urbano	51	49
	Rural	16	84
Acesso à internet	Urbano	43	57
	Rural	10	90

Fonte: Mendes et al. (2013).

Nas duas pesquisas fica evidente a necessidade de aumentar o acesso ao computador e à internet na área rural. Este quadro é apenas uma reprodução da conhecida marginalização pela qual foi historicamente submetido o espaço rural no Brasil, carente de investimentos em infraestrutura de desenvolvimento e na provisão de serviços sociais básicos.

### 3.1 Regiões do país

Para caracterizar o produtor rural que tem acesso a estes instrumentos de TI foram utilizados os dados desagregados do Censo Agropecuário (IBGE, 2006). A Tabela 5 ilustra o número de estabelecimentos agropecuários com computador e acesso à internet segundo as regiões do País.

**Tabela 5.** Número de estabelecimentos agropecuários com computador e acesso à internet, distribuídos por região do Brasil, segundo dados do Censo Agropecuário 2006.

Regiões	Tipo de eletrodoméstico utilizado	Número de estabelecimentos agropecuários	
		Unidades	Percentual
Sul	Computador	83.330	9,59
	Acesso à internet	29.795	3,43
Sudeste	Computador	61.859	8,11
	Acesso à internet	30.144	3,95
Centro-Oeste	Computador	12.939	5,19
	Acesso à internet	5.244	2,1
Norte	Computador	4.284	1,31
	Acesso à internet	1.295	0,4
Nordeste	Computador	21.211	1,16
	Acesso à internet	8.929	0,49

Fonte: Mendes et al. (2013).

As regiões Sul e Sudeste são as que se destacavam, respondendo, respectivamente, por 9,59% e 8,11% com acesso a computador, e por 3,43% e 3,95% à internet, respectivamente. São também

estas duas regiões que apresentam a maior participação no Produto Interno Bruto (PIB) brasileiro. Isso confirma que as desigualdades regionais também se traduzem em desigualdade de acesso às tecnologias da informação.

Outros indicadores de modernização da agricultura também mostram esta concentração no Sul e Sudeste do país, em detrimento das demais regiões (DELGADO, 2005).

Corroboram com esta afirmação Souza Filho et al. (2011) a partir da análise de dados do Censo Agropecuário 2006. Segundo estes autores, 36% dos estabelecimentos agropecuários nacionais usavam adubação e 16% faziam aplicação de calcário e/ou outro corretivo de pH do solo. Das propriedades rurais, 69% tinham acesso à energia elétrica, facilitando o uso de técnicas mais avançadas de manejo, processamento da produção e de gestão da propriedade.

Quanto à orientação técnica, um dos meios de acesso à informação e às novas tecnologias, 22% dos estabelecimentos agropecuários mencionaram ter recebido assistência técnica ocasionalmente ou regularmente. Todavia, as diferenças entre as regiões brasileiras eram evidentes. Por exemplo, para os indicadores uso de adubação, aplicação de corretivos do pH do solo e de uso de serviços de orientação técnica, as regiões Sul (72%, 39% e 48% dos estabelecimentos rurais daquela região, respectivamente) e Sudeste (53%, 31% e 31%, respectivamente) apresentavam o maior número de estabelecimentos agropecuários com acesso a esses itens. Por outro lado, na região Nordeste, 20% dos estabelecimentos rurais utilizavam adubação, apenas 3% usavam algum corretivo do pH do solo e 8% recebiam orientação técnica (IBGE, 2006).

### 3.2 Grupo de atividade econômica

No que tange ao grupo de atividade econômica, os dados da Tabela 6 apresentam que predominava o acesso a estes instrumentos pelo produtor que trabalhava com pecuária e criação de outros animais.

Surpreende o destaque dos estabelecimentos cuja atividade principal era a pecuária no que se refere à presença de computador e acesso à internet. Isto porque na pecuária brasileira, apesar dos notáveis progressos nos indicadores de produtividade, ainda predominam os sistemas extensivos de produção.

Entretanto, na análise dos dados nacionais do Censo Agropecuário de 2006, se considerar conjuntamente as atividades ligadas à produção vegetal, e mais especificamente, aquelas passíveis de serem enquadradas como um conjunto de lavouras (somatório de lavoura temporária; horticultura e floricultura; lavoura permanente; sementes, mudas e outras formas de propagação vegetal), os percentuais de uso de computador e internet se aproximam aos da pecuária, com vantagem para o conjunto de lavouras.

Este fato ocorre no caso do Brasil, das regiões sudeste e sul e do estado de São Paulo (Tabela 7). Já nos casos das regiões Norte, Nordeste e Centro Oeste predominam os estabelecimentos que tinham computador e que utilizavam a internet no grupo de atividade econômica pecuária e criação de outros animais.

Interessante ressaltar que a importância relativa de acesso ao computador e à internet na pecuária (Tabela 7) coincide com a maior disponibilidade de software agronegócio para a área de manejo animal, o que permitiria inferir que a indústria e estrutura de prestação de serviços estão atentas a essa demanda (MENDES et al., 2011).

**Tabela 6.** Número de estabelecimentos agropecuários com computador e acesso à internet em relação aos grupos de atividade econômica, no Brasil, segundo dados do Censo Agropecuário 2006.

Tipo de eletrodoméstico utilizado	Grupos de atividade econômica	Número de estabelecimentos agropecuários	
		Unidades	Percentual <sup>1</sup>
Computador	<b>Total</b>	<b>183.623</b>	<b>4,54</b>
	Lavoura temporária	53.493	1,32
	Horticultura e floricultura	12.011	0,3
	Lavoura permanente	26.265	0,65
	Sementes, mudas e outras formas de propagação vegetal	305	0,01
	Pecuária e criação de outros animais	85.861	2,12
	Produção florestal - florestas plantadas	3.324	0,08
	Produção florestal - florestas nativas	873	0,02
	Pesca	78	0
	Aquicultura	1.413	0,03
	Acesso à internet	<b>Total</b>	<b>75.407</b>
Lavoura temporária		21.303	0,53
Horticultura e floricultura		5.320	0,13
Lavoura permanente		11.881	0,29
Sementes, mudas e outras formas de propagação vegetal		165	0
Pecuária e criação de outros animais		33.967	0,84
Produção florestal - florestas plantadas		1.647	0,04
Produção florestal - florestas nativas		385	0,01
Pesca		25	0
Aquicultura		714	0,02

<sup>1</sup> Percentual em relação ao total de estabelecimentos do Brasil.  
Fonte: Mendes et al. (2013).

**Tabela 7.** Percentual de estabelecimentos agropecuários com computador e internet para o Brasil, nas macrorregiões, segundo dados do Censo Agropecuário 2006.

Tipo de eletrodoméstico utilizado	Grandes grupos de atividade econômica <sup>1</sup>	Brasil	Norte	Nordeste	Sudeste	Sul	Centro-Oeste
Computador	Total	4,54	1,31	1,16	8,11	9,59	5,19
	Lavouras	2,28	0,34	0,52	4,16	5,22	1,76
	Pecuária e criação de outros animais	2,12	0,91	0,60	3,76	4,02	3,34
	Florestas	0,10	0,04	0,02	0,13	0,29	0,05
	Pesca	0	0	0	0	0	0
	Aquicultura	0,03	0,02	0,02	0,05	0,06	0,04
Acesso à internet	Total	1,87	0,40	0,49	3,95	3,43	2,10
	Lavouras	0,95	0,09	0,21	2,07	1,92	0,81
	Pecuária e criação de outros animais	0,84	0,29	0,25	1,77	1,35	1,24
	Florestas	0,05	0	0,01	0,07	0,14	0,03
	Pesca	0	0	0	0	0	0
	Aquicultura	0,02	0,01	0,01	0,03	0,03	0,02

<sup>1</sup> Adaptado pelos autores, onde: a) “Lavouras” correspondem à soma de lavoura temporária; horticultura e floricultura; lavoura permanente; sementes, mudas e outras formas de propagação vegetal e b) “Florestas” correspondem à soma de Produção florestal - florestas plantadas e Produção florestal - florestas nativas.

Fonte: Mendes et al. (2013).

Em pesquisa realizada com 162 empresas desenvolvedoras, prevalece a oferta de software rural para a área animal (Tabela 8).

A análise da categoria “manejo animal” (Tabela 8), desagregada por áreas de aplicação, confirma oferta de software dedicado à produção animal (Tabela 9), com destaque para bovinos de corte e de leite. Não é clara a razão da relativa concentração da TIC na pecuária, uma vez que a produção agrícola apresenta indicadores de modernização mais elevados que os da pecuária. Mendes et al. (2011) esclarecem que os produtos de software voltados para a pecuária auxiliam em algumas etapas da produção animal, tais como para controle de custos de produção, balanceamento de rações, controle da reprodução por meio da genética e controle sanitário, incluindo a rastreabilidade.

**Tabela 8.** Percentual de software rural ofertado, por categorias.

Categorias	Total de respostas por categoria (somando as áreas de aplicação da categoria) <sup>1</sup>	% do total de respostas em relação às quatro categorias
Administração/Gerenciamento	467	40,9
Manejo animal	235	20,6
Cultivo vegetal	155	13,6
Controle de processo e/ou de atividades rurais	286	25,0
<b>Total</b>	<b>1143</b>	<b>100,0</b>

<sup>1</sup> Permite respostas múltiplas.

Dados referentes ao total de empresas privadas ofertantes de software para o agronegócio: 162  
Fonte: Mendes et al. (2011).

**Tabela 9.** Softwares ofertados para o agronegócio, por áreas de aplicação na categoria Manejo Animal.

Área de aplicação	Nº de software <sup>1</sup>	% da área de aplicação em relação ao total de software da categoria Manejo Animal
Bovinos de corte	53	22,6
Bovinos de leite	45	19,2
Suínos	31	13,2
Aves	19	8,1
Ovinos (ovelhas)	19	8,1
Bubalinos (criação de búfalos)	18	7,7
Caprinos (cabras)	18	7,7
Equídeos (cavalo, burro, mula, jumento)	15	6,4
Peixes	9	3,8
Frutos do mar (camarão, ostra, etc.)	6	2,6
Abelhas	2	0,9
<b>Total</b>	<b>235</b>	<b>100</b>

<sup>1</sup> Permite respostas múltiplas.

Dados referentes ao total de empresas privadas ofertantes de software para o agronegócio: 162

Fonte: Mendes et al. (2011).

Por outro lado, tendo como base os dados estaduais do Levantamento de Unidades de Produção Agropecuária na agricultura paulista, Francisco e Caser (2007) mostraram que no período de 2000 a 2006 a maior concentração de uso de computador e acesso à internet na agricultura paulista ocorreu em unidades produtoras de grãos, cana-de-açúcar, citros, café, ou seja, em atividades nas quais São Paulo se destaca na produção nacional.

### 3.3 Escolaridade

Algumas variáveis condicionam a incorporação da TIC e de inovações nesta área, entre as quais a disponibilidade de energia elétrica, o acesso à rede pública de comunicações e o nível de instrução do responsável pelo estabelecimento.

A Tabela 10 apresenta o nível de instrução do responsável pelo estabelecimento.

Em relação à variável nível de instrução do produtor os dados do Censo do IBGE mostraram, como era de se esperar, uma concentração no uso de computador e internet nos estabelecimentos onde as pessoas que os dirigem têm maior grau de instrução (segundo grau completo e ensino superior). Este fato se observa nas diferentes dimensões territoriais analisadas (Brasil e grandes regiões). Em relação do nível de escolaridade, ficam evidenciadas as grandes diferenças regionais quanto ao acesso a computador e internet, com Sul e Sudeste apresentando as maiores médias para todos os níveis de escolaridade, com destaque para o Estado de São Paulo, cujas médias ultrapassam as dessas duas regiões. A Região Norte foi a que apresentou a pior situação nacional, seguida pelo Nordeste. O acesso à internet, em 2006, representava uma proporção bem menor que a presença do computador no estabelecimento, chegando próximo à metade, ou mesmo menos, nas regiões Norte e Centro Oeste.

O nível de instrução integra as principais variáveis determinantes da adoção de tecnologia reportadas na literatura, dentre outras, como: tamanho da propriedade; área destinada à atividade;

**Tabela 10.** Percentual de estabelecimentos agropecuários com computador e internet segundo o nível de instrução do produtor, para o Brasil, nas macrorregiões, segundo dados do Censo Agropecuário 2006.

Tipo de eletrodoméstico utilizado	Nível de instrução da pessoa que dirige o estabelecimento	Brasil	Norte	Nordeste	Sudeste	Sul	Centro-Oeste
Computador	<b>Total</b>	<b>4,54</b>	<b>1,31</b>	<b>1,16</b>	<b>8,11</b>	<b>9,59</b>	<b>5,19</b>
	Alfabetização de adultos	1,26	0,31	0,38	2,18	4,41	1,28
	Ensino fundamental incompleto (1º grau)	3,6	0,78	0,83	4,53	6,7	2,66
	Ensino fundamental completo (1º grau)	8,29	2,19	2,36	10,08	14,77	5,67
	Ensino médio ou 2º grau completo	16,7	7,26	8,5	20,74	25,32	12,19
	Ensino Superior	32,32	20,16	26,58	35,92	39,49	21,07
	Nenhum, mas sabe ler e escrever	0,95	0,48	0,49	1,74	3	1,46
Não sabe ler e escrever	0,24	0,2	0,17	0,55	0,89	0,66	
Acesso à internet	<b>Total</b>	<b>1,87</b>	<b>0,4</b>	<b>0,49</b>	<b>3,95</b>	<b>3,43</b>	<b>2,1</b>
	Alfabetização de adultos	0,31	0,07	0,09	0,62	1,03	0,27
	Ensino fundamental incompleto (1º grau)	1,03	0,18	0,24	1,59	1,74	0,82
	Ensino fundamental completo (1º grau)	3,05	0,58	0,91	4,19	5,04	2
	Ensino médio ou 2º grau completo	8,02	2,54	4,03	10,95	11,56	5,27
	Ensino Superior	19,78	8,48	15,73	23,2	24,29	10,88
	Nenhum, mas sabe ler e escrever	0,25	0,08	0,12	0,54	0,75	0,38
Não sabe ler e escrever	0,06	0,02	0,04	0,16	0,19	0,16	

Fonte: Mendes et al. (2013).

mercado; idade; sexo (gênero); renda anual; atividade; acesso à informação; comportamento de adoção no passado; experiência; comportamento de risco financeiro; infraestrutura e regulamentação governamental (MACHADO, 2008).

A condição do produtor em relação às terras também nos ajuda a caracterizar os produtores com acesso a estes instrumentos de informática. Conforme demonstra a Tabela 11, predominava o acesso entre os proprietários da terra.

Tanto o nível de instrução como a condição do produtor em relação à terra são fatores que influenciam o acesso a computador e à internet no estabelecimento agropecuário. Estes aspectos estão inseridos num grupo mais abrangente de fatores que impactam na decisão de agricultores em adotar, não adotar ou retardar a adoção de uma tecnologia. Souza Filho et al. (2011) agrupam tais fatores segundo a natureza das variáveis envolvidas:

- 1) Condições socioeconômicas e características do produtor.
- 2) Características da produção e da propriedade rural.
- 3) Características da tecnologia.

**Tabela 11.** Número de estabelecimentos agropecuários com computador e acesso à internet no Brasil quanto à condição do produtor em relação à terra, de acordo com dados do Censo Agropecuário 2006.

Tipo de eletrodoméstico utilizado	Condição do produtor	Número de estabelecimentos agropecuários	
		Unidades	Percentual
<b>Computador</b>	<b>Total</b>	<b>183.623</b>	<b>4,54</b>
	Proprietário	166.559	4,12
	Assentado sem titulação definitiva	1.400	0,03
	Arrendatário	8.687	0,21
	Parceiro	1.775	0,04
	Ocupante	3.695	0,09
	Produtor sem área	1.507	0,04
<b>Acesso à internet</b>	<b>Total</b>	<b>75.407</b>	<b>1,87</b>
	Proprietário	68.224	1,69
	Assentado sem titulação definitiva	332	0,01
	Arrendatário	4.537	0,11
	Parceiro	717	0,02
	Ocupante	1.116	0,03
	Produtor sem área	481	0,01

Fonte: Mendes et al. (2013).

#### 4) fatores sistêmicos.

Souza Filho et al. (2011) detalham a análise de tais fatores. As condições socioeconômicas do produtor e de sua família referem-se às características que podem ter papel de destaque na trajetória da unidade de produção, tais como a experiência e a capacidade de obter e processar informações, a habilidade no uso de técnicas agrícolas e de métodos de gerenciamento mais sofisticados que podem contribuir para o sucesso do empreendimento.

No que concerne às características da produção, verifica-se qual é o papel que a tecnologia exerce na determinação do desempenho econômico-financeiro do estabelecimento, pois ela pode permitir elevar a produtividade do trabalho e criar elos a montante e a jusante. Quanto à característica da tecnologia, interessa apontar se ela possibilita ter como os efeitos esperados a elevação da produtividade e a economia de mão-de-obra, que correspondem às principais necessidades dos agricultores familiares.

Quanto aos fatores sistêmicos, analisam-se as condições dos segmentos da cadeia produtiva em que a exploração agrícola está inserida, bem como as instituições e organizações que lhe provêm suporte tecnológico, de informações e financeiro. A infraestrutura física (energia, telecomunicação, armazenamento), a infraestrutura de ciência e tecnologia (institutos de pesquisa, universidades) e serviços de educação básica são de fundamental importância no sentido de gerar externalidades positivas para ações de adoção de tecnologia.

## 4 Considerações finais

Embora se admita que alguns dados referentes a uso de computador e internet, relativos a 2006, estejam obsoletos – dada a velocidade de propagação desses instrumentos –, sabe-se que questões de fundo subsistem, e ainda são responsáveis pela desigualdade no acesso à tecnologia da informação ao longo do território nacional. Entre elas estão o nível de instrução dos produtores, sua condição em relação às terras, o grau de dinamismo das diferentes atividades econômicas, entre outras variáveis.

Os resultados evidenciaram a concentração da posse de computador e do uso de internet nas regiões Sul e Sudeste do Brasil, bem como entre os produtores que possuíam maior nível de escolaridade. Também ressaltaram essa concentração entre os proprietários, em relação às outras condições do produtor em relação à terra.

Estes dados nos alertam que ainda há um longo caminho a percorrer para conferir condições isonômicas de acesso a computador e internet para o produtor rural brasileiro. A concentração de uso destes instrumentos nas regiões sudeste e sul do país reforçam as desigualdades regionais e a heterogeneidade da agricultura do País.

A concentração do acesso aos recursos de TIC entre os produtores com maior nível de escolaridade, principalmente no que se relaciona ao acesso à internet, reflete uma tendência apontada em alguns estudos de que, quanto maior o nível de instrução do produtor, maior sua capacidade de apropriar conhecimentos e tecnologias, o que pode contribuir para a exclusão digital das pessoas com baixo nível de instrução.

Diante deste quadro, é imprescindível ampliar o acesso do produtor rural ao computador e à internet como uma das condições para que ele possa usufruir dos benefícios do uso da tecnologia da informação aplicada ao campo, sob pena da ampliação das assimetrias entre as denominadas agricultura moderna e atrasada.

Para tanto, são relevantes ações coordenadas e integradas de fomento à adoção de tecnologias da informação pelos agricultores por diversos agentes públicos e privados – como organizações de pesquisa, ensino, extensão, bem como políticas de inclusão digital que facilitem o acesso à tecnologia da informação pelas populações mais excluídas.

O quadro é apenas uma reprodução da conhecida marginalização pela qual foi historicamente submetido o espaço rural no Brasil, carente de investimentos em infraestrutura de desenvolvimento e na provisão de serviços sociais básicos.

Dada a importância do setor agrícola, se a agricultura brasileira não tiver condições para se capacitar, absorver e utilizar inovações em geral, e tecnologias da informação, mais especificamente, isso poderá comprometer a competitividade dinâmica do setor.

## 5 Referências

AGÊNCIA BRASILEIRA DE DESENVOLVIMENTO INDUSTRIAL. **TIC agronegócio**: oportunidades de desenvolvimento tecnológico em tecnologias de informação e comunicação para o agronegócio. [Campinas], 2012. 40 p. (Relatório de acompanhamento setorial).

COMITÊ GESTOR DA INTERNET NO BRASIL. **Pesquisa sobre o uso das tecnologias da informação e da comunicação no Brasil 2011**. São Paulo, 2012.

DELGADO, G. C. A questão agrária no Brasil: 1950-2003. In: JACCOUD, L. (Org.). **Questão social e políticas sociais no Brasil contemporâneo**. Brasília, DF: IPEA, 2005. p. 51-90. Disponível em: <[http://www.ipea.gov.br/portal/images/stories/PDFs/livros/Cap\\_2-10.pdf](http://www.ipea.gov.br/portal/images/stories/PDFs/livros/Cap_2-10.pdf)>. Acesso em: 10 out. 2014.

ESTADOS UNIDOS. Department of Agriculture. **Farm computer usage and ownership**. Washington, D. C., 2013. 30 p. Disponível em: <<http://usda.mannlib.cornell.edu/usda/current/FarmComp/FarmComp-08-20-2013.pdf>> Acesso em: 10 abr. 2014.

FRANCISCO, V. L. F. dos S.; CASER, D. V. Adoção da internet em Fazendas no Estado de São Paulo. In: CONGRESSO BRASILEIRO DE AGROINFORMÁTICA, 6., 2007, São Pedro, SP. **Anais...** Campinas, 2007. p. 316-320.

GELB, E. **The EFITA Bonn – Turino Conferences (1999-2013): ICT adoption questionnaire update**. [2013]. Disponível em: <[http://www.informatique-agricole.org/gazette/Efita\\_Documents/ICT%201999%20-%202013%20Questionnaire%20summary%20with%20Ossiach%20-%20updated%20\(3\).pdf](http://www.informatique-agricole.org/gazette/Efita_Documents/ICT%201999%20-%202013%20Questionnaire%20summary%20with%20Ossiach%20-%20updated%20(3).pdf)>. Acesso em: 10 abr. 2014.

GELB, E. **The EFITA ICT adoption questionnaire: 1999-2011 priority indicators for the future**. [2012]. 11 p. Disponível em: <<http://departments.agri.huji.ac.il/economics/gelb-efita-2012.pdf>>. Acesso em: 10 abr. 2014.

GELB, E.; VOET, H. **ICT adoption trends in agriculture: a summary of the EFITA ICT adoption questionnaires (1999-2009)**. [2009]. Disponível em: <<http://departments.agri.huji.ac.il/economics/voet-gelb.pdf>>. Acesso em: 6 fev. 2013.

GOYAL, A.; GONZÁLES-VELOSA, C. **Improving agricultural productivity and market efficiency in Latin America and the Caribbean: how ICTs can make a difference?** Washington, D. C.: World Bank, 2012. 20 p. (LCSSD Occasional Paper Series on Food Prices, 68255). Disponível em: <<https://openknowledge.worldbank.org/bitstream/handle/10986/18017/682550WP0P1247018B0Final0ICT0report.pdf?sequence=1>>. Acesso em: 10 out. 2014.

IBGE. **Censo agropecuário 2006**. Rio de Janeiro, 2006. Disponível em: <<http://www.sidra.ibge.gov.br/bda/acervo/acervo2.asp?e=v&p=CA&z=t&o=11>>. Acesso em: 3 abr. 2013.

IBGE. **O setor de tecnologia da informação e comunicação no Brasil 2003-2006**. Rio de Janeiro, 2009. 82 p. (Estudos e pesquisas: informação econômica, nº 11). Disponível em: <<http://www.ibge.gov.br/home/estatistica/economia/stic/publicacao.pdf>>. Acesso em: 10 out. 2014.

MACHADO, J. G. de C. F. **Adoção da tecnologia da informação na pecuária de corte**. 2008. 219 f. Tese (Doutorado em Engenharia de Produção) - Universidade Federal de São Carlos, São Carlos.

MENDES, C. I. C.; BUAINAIN, A. M.; FASIABEN, M. do C. R. Acesso ao computador e à internet na agricultura brasileira: uma análise a partir do Censo Agropecuário. In: CONGRESSO DA SOCIEDADE BRASILEIRA DE ECONOMIA, ADMINISTRAÇÃO E SOCIOLOGIA RURAL, 51., Belém, PA, 2013. **Novas fronteiras da agropecuária no Brasil e na Amazônia: desafios da sustentabilidade: anais**. Belém, PA: SOBER, 2013.

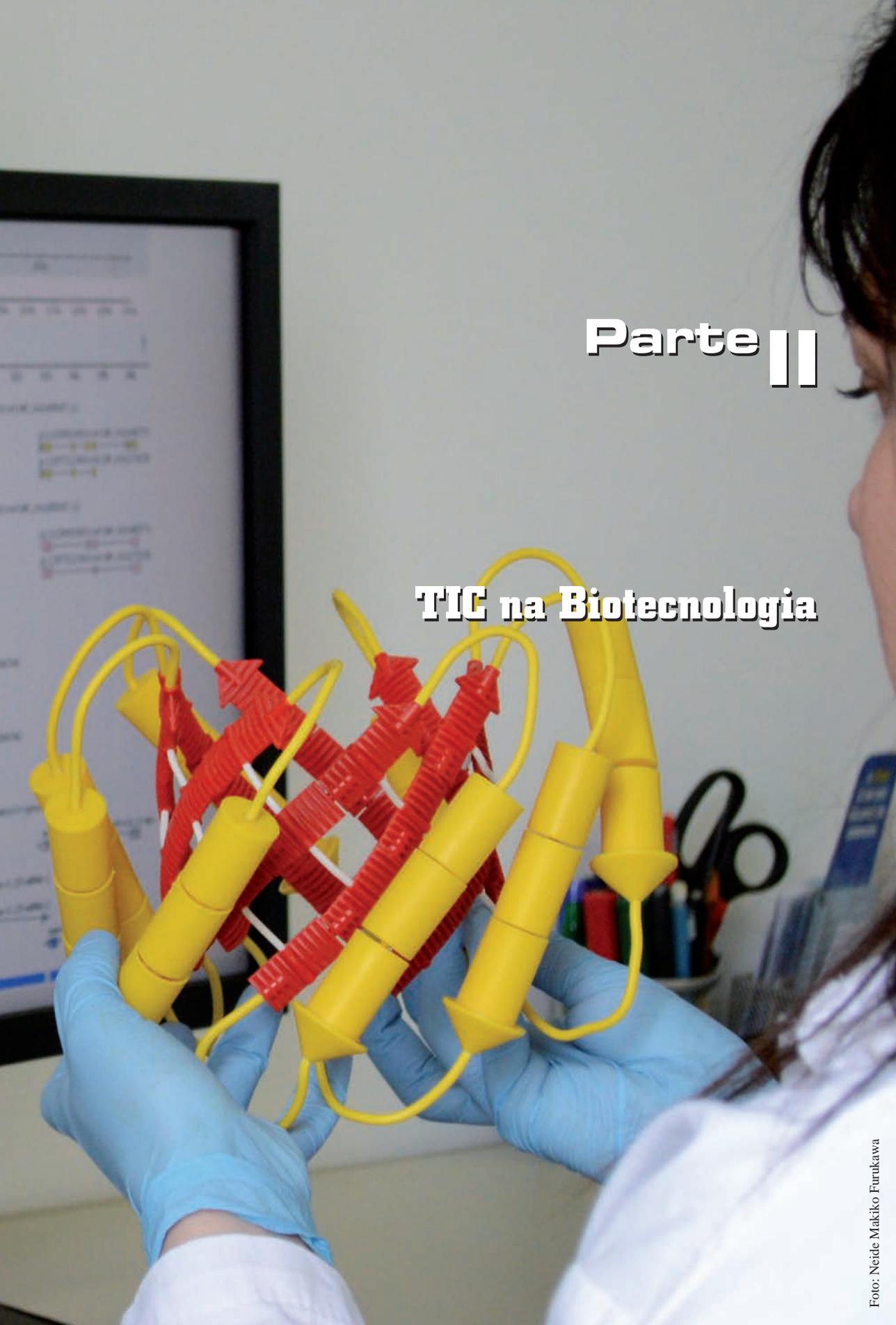
MENDES, C. I. C.; OLIVEIRA, D. R. M. dos S.; SANTOS, A. R. (Ed.). **Estudo do mercado brasileiro de software para o agronegócio**. Campinas: Embrapa Informática Agropecuária: Campinas, 2011. 187 p.

SOUZA FILHO, H. M. de; BUAINAIN, A. M.; SILVEIRA, J. M. F. J. da; VINHOLIS, M. de M. B. Condicionantes da adoção de inovações tecnológicas na agricultura. **Cadernos de Ciência & Tecnologia**, Brasília, DF, v. 28, n. 1, p. 223-255, jan./abr. 2011.

ZAMBALDE, A. L.; SCHNEIDER, H.; LOPES, M. A.; PAGLIS, C. M. BAMBINI, M. D. Tecnologia da informação no agronegócio. In: MENDES, C. I. C.; OLIVEIRA, D. R. M. dos S.; SANTOS, A. R. (Ed.). **Estudo do mercado brasileiro de software para o agronegócio**. Campinas: Embrapa Informática Agropecuária: Campinas, 2011. cap. 2, p. 38-72.

# Parte II

## TIC na Biotecnologia





## Genômica e biotecnologia aplicadas a adaptação a mudanças climáticas

Isabel Rodrigues Gerhardt  
Ricardo Augusto Dante

### 1 Introdução

#### 1.1 Mudanças climáticas globais e seus impactos na agricultura

Países em desenvolvimento, muitos dos quais situados em regiões de clima tropical, subtropical e semiárido, têm sua produção agrícola e florestal especialmente vulnerável às mudanças climáticas globais (SOLOMON et al., 2007). Além dos efeitos diretos decorrentes das elevações tanto da concentração atmosférica de dióxido de carbono como da temperatura global média, esses setores deverão sofrer grande impacto em consequência do aumento da frequência de eventos extremos, como ondas de calor, estiagens, secas, alagamentos e inundações (SOLOMON et al., 2007). No Brasil, mudanças climáticas podem levar a perdas de produção agrícola da ordem de R\$ 7,4 bilhões em 2020 (DECONTO, 2008). Além das perdas econômicas diretas, profundos efeitos sociais associados a menor disponibilidade de alimentos e energia para a população deverão ocorrer consequentemente.

Nas últimas três décadas, milho e trigo tiveram sua produtividade global afetada negativamente por alterações de temperatura e de precipitação, enquanto que culturas que não sofreram impacto global, como arroz e soja, foram afetadas negativamente por esses estresses nos níveis nacionais e regionais (LOBELL; GOURDJI, 2012; LOBELL et al., 2011). Também nesse período, as produtividades de milho, soja e trigo no Brasil sofreram impactos climáticos negativos (LOBELL et al., 2011; SAKURAI et al., 2011). Milho e soja sofreram impactos das alterações de temperatura e precipitação que variaram fenológica, temporal e geograficamente (BERGAMASCHI et al., 2007; FERREIRA; RAO, 2011; SAKURAI et al., 2011). Na Região Sul, a qual responde por aproximadamente 30% da produção brasileira de milho e soja, diversas safras agrícolas apresentaram desde o final dos anos 1990 graves perdas por secas associadas a ocorrências da fase La Niña do fenômeno El Niño-Oscilação Sul (CONAB, 2013; SLEIMAN; SILVA, 2010). Notavelmente, a perda de produção agrícola nessa região foi de aproximadamente 25% na safra 2004/05 em comparação a safras imediatamente anteriores e posteriores, as quais foram menos afetadas por redução de precipitação.

Exemplos de eventos climáticos extremos com consequências drásticas na produtividade agrícola foram observados recentemente nos EUA e no Brasil. A redução da produção de milho na safra 2012 norte-americana, devido à seca e ao calor recordes, foi estimada em cerca de 40 milhões de

toneladas (Mt), valor equivalente à média da safra anual brasileira no período 2005-10 (49 Mt) (ESTADOS UNIDOS, 2013; FAO, 2013). No Nordeste brasileiro, a seca mais severa das últimas cinco décadas provocou ao menos R\$ 3,6 bilhões em perdas agrícolas diretas e graves consequências sociais (BARROS-NETO; MOURA, 2013; NAÇÕES UNIDAS, 2013).

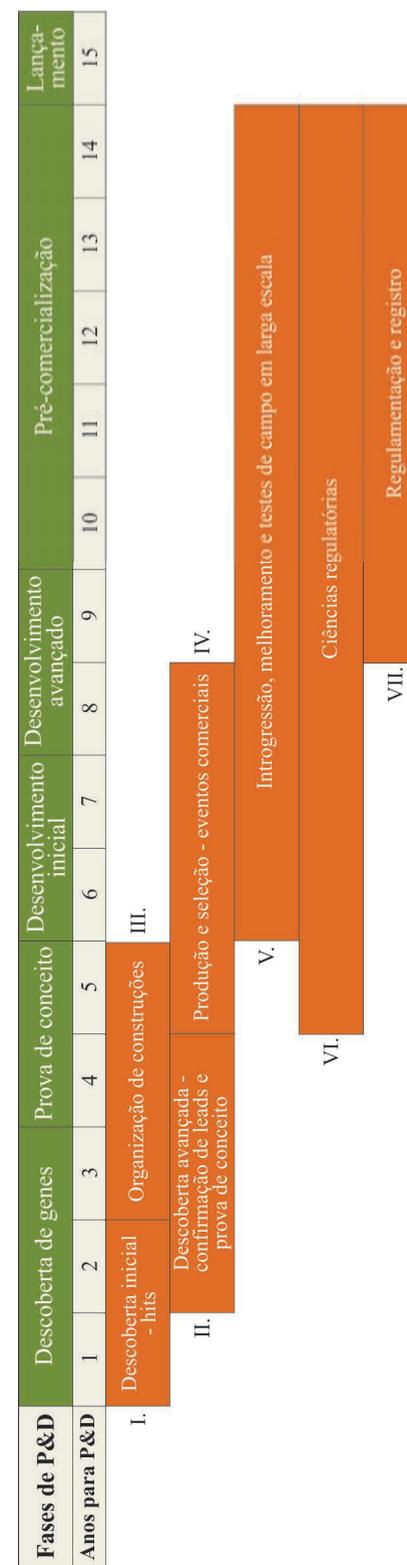
A produção mundial de grãos deverá crescer 70% (FAO 2009) para satisfazer a demanda por alimentos da população global em 2050, estimada em mais de 9 bilhões de pessoas. Agravadas por investimentos insuficientes em pesquisa e desenvolvimento agrícola, as baixas taxas de crescimento da produtividade agrícola observadas nas últimas duas décadas podem ter consequências profundas na produção e nos preços futuros (ALSTON et al., 2009). Para que os efeitos das mudanças climáticas globais sobre a agricultura possam ser minimizados, é fundamental que o melhoramento genético vegetal desenvolva continuamente variedades mais adaptadas aos estresses decorrentes que possibilitem maior estabilidade à produção agrícola de forma ambiental, econômica e socialmente viáveis. Para que o melhoramento genético faça contribuições efetivas, são necessários esforços que visam a ampliação da diversidade genética disponível em bancos de germoplasma e a geração de novas variedades com desempenho agrícola superior. Nesse sentido, a biotecnologia moderna pode contribuir significativamente, pela introdução de novos genes que confirmam capacidade adicional de adaptação à restrição hídrica e ao aumento da temperatura às variedades desenvolvidas pelo melhoramento genético convencional. Para tanto, a descoberta e a validação de genes com esse potencial devem ser realizadas por meio de pipelines de pesquisa e desenvolvimento (P&D) de grande capacidade (Figura 1), que reúnam ferramentas avançadas de biologia molecular e genômica, como a caracterização em larga escala de genomas, de expressão gênica e de fenótipos, marcadores moleculares e transgenia (PARK et al., 2011; PELLEGRINO et al., 2007; TESTER; LANGRIDGE, 2010).

A genômica funcional é caracterizada pela análise de função gênica em larga escala (portanto, almejando centenas ou milhares de genes-alvo) e envolve, entre outros recursos e abordagens experimentais, a utilização de grandes coleções de genótipos contendo variantes alélicas (incluindo mutações pontuais e inserções de T-DNAs e elementos de transposição que modificam a expressão ou função do locus-alvo) e de eventos geneticamente modificados com expressão alterada de genes-alvo. No âmbito da genômica funcional e da biotecnologia agrícola moderna, duas possibilidades distintas, mas não mutuamente exclusivas, de geração de variedades geneticamente melhoradas são:

- a) Desenvolvimento de marcadores moleculares (mais comumente genéticos) para seleção assistida de genes ou alelos de interesse em programas de melhoramento genético.
- b) Identificação e validação, em sistemas ou plantas-modelo, de novos genes e alelos de diferentes fontes, seguidas de introdução em genótipos-elite.

**1.2 Pesquisa e desenvolvimento de tecnologias transgênicas**

Ao longo da última década, a transgenia foi a tecnologia mais rapidamente adotada na história da agricultura (CHASSY, 2007). A redução da aplicação de cerca de 224 milhões de quilos de pesticidas juntamente com os ganhos de produtividade e a diminuição nos custos de produção devido ao cultivo de plantas transgênicas contribuíram, no período de 1996 a 2004, com a alocação de valores entre US\$ 20 a US\$ 30 bilhões na renda dos agricultores em todo o mundo (HERDT, 2006).



**Figura 1.** Cronograma, fases e estágios de P&D do pipeline de geração de tecnologias transgênicas das grandes empresas transnacionais de biotecnologia agrícola. I. Descoberta Inicial: Identificação de genes com o objetivo de obter o caráter de interesse (hits). II. Descoberta Avançada: Introdução dos hits em plantas-modelo. Se observado o caráter de interesse na planta transformada, tem-se a sequência lead. III. Otimização de Construções: Otimização temporal, espacial e de nível de expressão do transgene. A cultura-alvo é transformada e tem fenótipo avaliado em casa de vegetação e/ou campo. IV. Produção e Seleção: Construções gênicas com os leads são usadas para produzir eventos de qualidade comercial, eliminando-se eventos com múltiplas inserções. Avaliação fenotípica ainda é realizada. V. Introgessão, Melhoramento e Testes de Campo: Eventos de qualidade comercial são introgredidos em germoplasma-elite para produção de híbridos ou variedades para avaliação da performance em testes de campo. VI. Ciência Regulatória: Condução de estudos regulatórios: caracterização completa do evento e confirmação de segurança alimentar e ambiental. VII. Regulamentação e Registro: Preparação e submissão de documentação para plantio comercial em diferentes países.

Fonte: adaptado de Prado et al. (2014).

Uma das principais características das grandes empresas de biotecnologia agrícola atuantes no mercado de sementes é a capacidade de inovação no desenvolvimento de novas cultivares que incorporam biotecnologia avançada. Para isso, contam com pipelines robustos de pesquisa e desenvolvimento (P&D) que combinam melhoramento genético (convencional e assistido por marcadores moleculares) e a modificação genética (transgenia) para a geração de novas cultivares mais produtivas que incorporem resistência a herbicidas, resistência a pragas e doenças e, mais recentemente, cultivares tolerantes à seca, entre outros caracteres. Esses pipelines avaliam anualmente inúmeras construções gênicas introduzidas diretamente em plantas economicamente importantes, como milho e soja.

Estima-se que o desenvolvimento de uma tecnologia transgênica, da descoberta inicial de muitos genes até a pré-comercialização de um único evento transformado, leve de 10 a 15 anos e exija investimentos da ordem de US\$ 136 milhões ao longo desse período (PRADO et al., 2014). Esse valor compreende os custos com recursos humanos especializados, infraestruturas laboratorial e operacional, propriedade intelectual e aspectos regulatórios. Os altos custos (cerca de US\$ 35 milhões) e a complexidade de operações da etapa de regulamentação, envolvendo segurança alimentar e ambiental, restringem a participação de empresas públicas nesse processo, além de limitar o desenvolvimento da tecnologia preferencialmente para culturas commodities e não de subsistência, principalmente em países do Terceiro Mundo.

A geração e a seleção em larga escala de plantas transgênicas com fenótipos de valor comercial normalmente compreende seis fases distintas de P&D (Figura 1), caracterizadas por diferentes estágios e atividades. A fase de descoberta de genes compreende a avaliação de milhares de genes e tem como objetivo identificar aqueles capazes de conferir as características desejadas para a cultura alvo. A genômica funcional apresenta-se como uma poderosa ferramenta para auxiliar esse processo.

## 2 Genômica funcional para descoberta de genes de valor agrônomico

A genômica funcional e demais ciências genômicas relacionadas (entre elas o sequenciamento de genomas, a quantificação em larga escala de expressão nos níveis transcricional, protéico, metabólico e fenotípico), auxiliadas pela Biologia de Sistemas (integração, via bioinformática e modelagem computacional, das observações feitas por meio das Ciências Genômicas), têm emergido como abordagens eficientes para a caracterização integrada de fenótipos e genótipos, levando à compreensão da estrutura, da dinâmica e do controle de respostas adaptativas, majoritariamente multigênicas e complexas. Diversas espécies vegetais de importância econômica, como arroz, soja, milho, sorgo e eucalipto, foram alvo de programas de sequenciamento nos últimos anos (BOLGER et al., 2014). O rápido avanço nas tecnologias de sequenciamento de nova geração abre novas possibilidades de geração de informação, capazes de contribuir com caracteres envolvidos na adaptação a estresses abióticos, bióticos, aumento de produtividade e qualidade nutricional de cultivos agrícolas. É possível expandir os recursos genômicos para culturas agrícolas e florestais por meio de:

a) Prospecção de espécies da biodiversidade.

- b) Sequenciamento do genoma de genótipos contrastantes presentes nos bancos de germoplasma de culturas agrícolas e florestais de interesse.
- c) Desenvolvimento de ferramentas de bioinformática para comparação de múltiplos genomas, de modo a identificar variabilidade nas sequências gênicas ou regulatórias que possa estar relacionada a características de interesse.

Abaixo citamos alguns dos recursos e estratégias que, com o suporte da bioinformática, são a base para descoberta de genes pelas abordagens de Genômica Funcional.

### 2.1 Bancos de dados genômicos

Bancos de dados baseados na *World Wide Web* (WWW) de vários genomas de plantas têm se revelado um recurso inestimável para os biólogos de plantas em todo o mundo. Pesquisadores podem facilmente adquirir informação genética diversa tanto de sequências codificantes como não codificantes, elementos regulatórios, famílias de genes, polimorfismos que potencialmente originam marcadores moleculares e variabilidade genética entre diferentes espécies de plantas. Os recursos dos bancos de dados de genoma podem servir como uma fonte substancial para a identificação de genes responsivos a estresses bióticos e abióticos, associados a aumento de produtividade, qualidade nutricional e outras características agrônômicas de interesse. Entre os vários bancos de dados existentes, o Phytozome<sup>1</sup> é um exemplo de banco de dados e portal que permite acesso à informação de conjuntos completos de genomas de plantas, genes e sequências homólogas, informação funcional e de famílias gênicas, além de alinhamento de sequências (GOODSTEIN et al., 2012).

### 2.2 Genômica comparativa

A disponibilidade de sequências de genomas de plantas, juntamente com o acúmulo de dados de expressão gênica relacionados às mais variadas características agrônômicas, representam valiosos recursos para descoberta de genes e novas vias metabólicas envolvidas em processos biológicos de interesse. Uma das principais vantagens da genômica comparativa é a transferência da informação de anotações de genes de plantas modelo para culturas agrícolas recém sequenciadas, onde estudos funcionais ainda são rudimentares (MA et al., 2012). Um requisito importante, no entanto, na utilização da genômica comparativa é a disponibilidade de conjuntos de dados de genes ortólogos (genes homólogos que evoluíram a partir de um gene ancestral comum e estão separados por um evento de especiação), uma vez que esses genes frequentemente apresentam funções similares, que são mantidas entre as espécies que apresentam um ancestral comum (MA et al., 2012). Um exemplo de uso bem sucedido da genômica comparativa foi a identificação de fatores de transcrição responsivos a estresses em soja e diversos cereais, a partir da comparação com fatores de transcrição previamente caracterizados em *Arabidopsis* e arroz (MOCHIDA et al., 2011; TRAN; MOCHIDA, 2010).

Com o rápido avanço das tecnologias de sequenciamento e a geração de grandes volumes de informação sobre a constituição genômica de milhares de organismos eucariotos e procariotos, a genômica comparativa entre espécies e mesmo entre táxons distantemente relacionados, tais como reinos, abre a possibilidade para a identificação de respostas a diferentes condições

<sup>1</sup> Disponível em: <[www.phytozome.net](http://www.phytozome.net)>.

ambientais tanto conservadas evolutivamente como espécie-específicas (DEYHOLOS, 2010; PITZSCHKE; HIRT, 2010). A regulação da expressão de genes associados a respostas celulares basais, incluindo metabolismo de carbono e indução de proteínas de choque térmico, é geralmente conservada entre os organismos. No entanto, genes que codificam proteínas de sinalização celular e fatores de transcrição em resposta a estresses são muitas vezes espécie-específicos (MUSTROPH et al., 2010). Por exemplo, a análise comparativa do transcriptoma de *Arabidopsis* e *Thellungiella halophila* (espécie da mesma família da *Arabidopsis*, tolerante à alta salinidade, à seca e ao frio) destacou três importantes mecanismos de adaptação de plantas a condições extremas:

- a) Especificidade - *Thellungiella* regula um conjunto de genes específicos em resposta a situações de estresse.
- b) Antecipação - *Thellungiella* é constitutivamente preparada para responder ao estresse.
- c) Sensibilidade - *Thellungiella* necessita de estresses mais intensos para a indução de respostas transcricionais (AMTMANN, 2009; GONG et al., 2005).

### 2.3 Análise filogenética de famílias gênicas

Famílias gênicas são grupos de genes homólogos descendentes de um ancestral comum, cujo evento de separação é uma duplicação genômica que retém similaridade de sequência e, em muitos casos, similaridade de função. As informações genômicas disponíveis em bases de dados públicas fez com que fosse possível o estudo da origem e expansão das famílias de genes. Ao longo da evolução, os genomas de plantas foram submetidos a duplicações e rearranjos, de modo que alguns dos genes duplicados ganharam novas funções, alguns mantiveram uma parte ou a totalidade suas funções ancestrais, enquanto outros perderam completamente a funcionalidade original (DEMUTH; HAHN, 2009). Avanços em genômica e bioinformática têm ajudado pesquisadores na elucidação da evolução das famílias de genes de plantas. Essa informação pode contribuir para uma melhor compreensão da diversidade funcional de genes individuais que pertencem a uma mesma família.

A análise de famílias de genes em diferentes genótipos de uma mesma espécie tem sido realizada em várias plantas para identificar genes relacionados a mecanismos de controle de estresse e outras características. Por exemplo, a análise filogenética da família de fatores de transcrição NAC de soja previu que 58 genes estão envolvidos na resposta à desidratação (LE et al., 2011). Da mesma forma, estudos de filogenia associados a padrões de expressão gênica tanto de outras famílias de fatores de transcrição, como AP2/ERF, HD-Zip e MYB, como de microRNAs também foram capazes de identificar membros responsivos a estresses abióticos em diferentes espécies vegetais (BARIK et al., 2014; JAIN et al., 2007; LIU et al., 2013; WANG et al., 2014).

### 2.4 Aplicações de tecnologias de sequenciamento de próxima geração

Abordagens para gerar sequências genômicas necessitam de grandes investimentos financeiros e de recursos humanos. O desenvolvimento e a aplicação das tecnologias de sequenciamento de próxima geração (*Next Generation Sequencing* (NGS), na sigla em inglês), no entanto, têm facilitado enormemente a geração de sequências genômicas e de transcritos de um número cada vez maior de espécies de plantas de interesse agropecuário, bem como de indivíduos fenotípicamente distintos de uma mesma espécie. Como consequência, um ampla gama de oportunidades para

a descoberta de genes e rotas metabólicas relacionadas a um grande número de características que podem servir como base para o melhoramento das culturas agrícolas de interesse tem sido criada. Por exemplo, a tecnologia de NGS pode ser associada com a determinação massal do perfil transcricional para investigar alterações de expressão gênica em resposta a variados tipos de estresses como salinidade, submersão, seca e altas temperaturas, ou para compreender as alterações genotípicas responsáveis pelas diferenças fenotípicas entre indivíduos de uma mesma espécie (EGAN et al., 2012; GARG et al., 2013; MOLINA et al., 2011; OH et al., 2012).

Embora a tecnologia de NGS ainda esteja na fase inicial de sua aplicação, ela tem provado ser uma ferramenta robusta para a identificação da variabilidade genética intraespecífica existente no genoma. A adoção da tecnologia de NGS, combinada com a seleção genômica ampla, tem sido utilizada para identificar potenciais marcadores moleculares, tais como Single Nucleotide Polymorphism (SNPs), inserções e deleções, que estão associados com múltiplas características, como crescimento e desenvolvimento e/ou respostas a estresses. O sequenciamento extensivo de populações naturais de álamo, milho e arroz contribuiu, por exemplo, para a identificação e melhor compreensão de mecanismos de biologia evolutiva, incluindo variação funcional e bases moleculares de adaptação, domesticação e produtividade (EVANS et al., 2014; HUANG et al., 2012; HUFFORD et al., 2012; MCKOWN et al., 2014).

### 2.1 Seleção em larga escala de eventos transgênicos

No âmbito das grandes empresas de biotecnologia agrícola e suas parceiras tecnológicas, destacam-se abordagens de genômica funcional que almejam a geração de tecnologias baseadas em plantas geneticamente modificadas visando ganhos de produtividade intrínseca e tolerância a estresses. O complemento de fatores de transcrição da espécie-modelo *Arabidopsis thaliana* tem sido caracterizado funcionalmente pela empresa Mendel Biotechnology<sup>2</sup>, enquanto que a análise de milhares de genes de arroz e outras espécies é foco da empresa CropDesign<sup>3</sup>. A caracterização *in silico* (RIECHMANN et al., 2000) e avaliação funcional por meio de superexpressão dos aproximadamente 1.500 genes que codificam fatores de transcrição em *Arabidopsis thaliana* possibilitou a identificação daqueles que conferem tolerância a estresses e aumento de produtividade em culturas agrícolas em condições de campo, gerando tecnologias de possível alcance comercial. Notavelmente, a superexpressão de fatores de transcrição do tipo NF-YB aumenta a capacidade fotossintética e minimiza a redução da produtividade do milho em condições hídricas limitantes, aparentemente através de vias de sinalização independentes de ácido abscísico (NELSON et al., 2007). A superexpressão do fator de transcrição da classe *B-box domain* BBX32 aumenta a produtividade de soja pela alteração de respostas reprodutivas à luz (PREUSS et al., 2012), enquanto que a superexpressão de um fator de transcrição do tipo II da classe HD-Zip em milho leva ao aumento da massa da inflorescência feminina, um componente de produtividade (RICE et al., 2014). Essas tecnologias ilustram o potencial de geração de tecnologias proporcionado por abordagens genômicas funcionais em espécies vegetais.

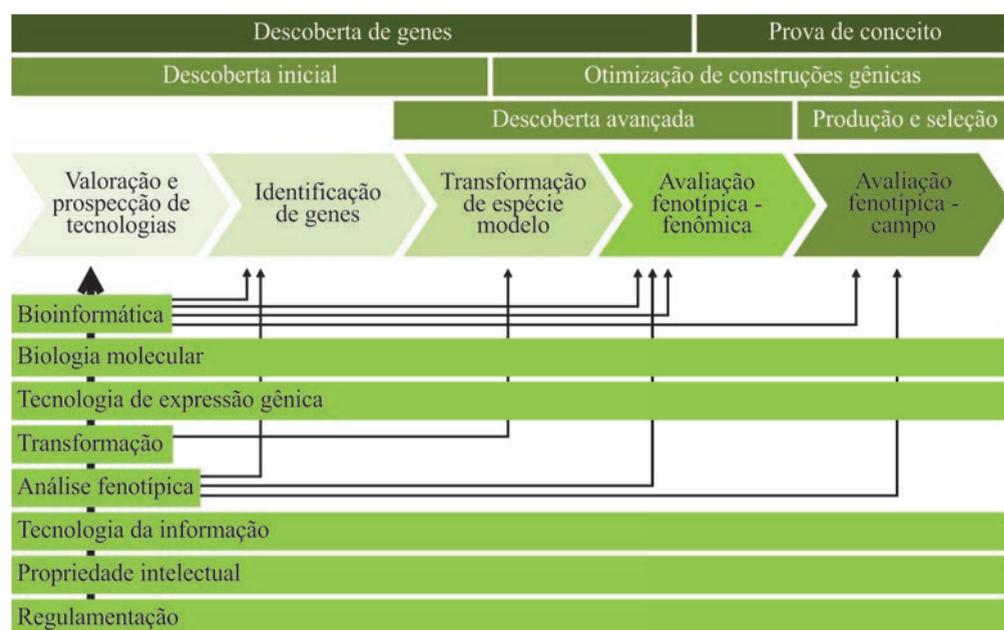
<sup>2</sup> Disponível em: <www.mendel.com>.

<sup>3</sup> Disponível em: <www.cropdesign.com>.

### 3 A Unidade Mista de Pesquisa em Genômica Aplicada a Mudanças Climáticas (UMiP GenClima)

Segundo um estudo realizado pelas empresas de consultoria Deloitte e Economist Intelligence Unit (DELOITTE, 2006), as principais estratégias para assegurar o sucesso de empresas de biotecnologia (agrícolas e farmacêuticas) são a existência de um pipeline robusto de P&D e a formação de parcerias institucionais. Nesse sentido, a Empresa Brasileira de Pesquisa Agropecuária (Embrapa) e a Universidade Estadual de Campinas (Unicamp) deram um importante passo para enfrentar o desafio do desenvolvimento de plantas transgênicas tolerantes a condições ambientais adversas ao firmarem, em dezembro de 2012, um acordo para a criação da Unidade Mista de Pesquisa em Genômica Aplicada a Mudanças Climáticas (UMiP GenClima).

A UMiP GenClima é uma iniciativa conjunta da Embrapa e da Unicamp para unir suas competências em ciência genômicas e biologia molecular de plantas. Essa parceria tem como objetivo a criação de um *pipeline* de genômica funcional (Figura 2) focado na descoberta e na validação de genes por meio de transgenia, visando à produção de variedades mais adaptadas a condições ambientais exacerbadas por mudanças climáticas (tais como episódios de seca e de calor e concentração aumentada de dióxido de carbono atmosférico). As instalações físicas próprias da UMiP GenClima serão instaladas no Parque Científico e Tecnológico da Unicamp e contará com laboratórios de bioinformática, biologia molecular, transformação genética e fenotipagem em larga escala em condições controladas de cultivo para a identificação de eventos transgênicos que confirmam tolerância a estresses abióticos. A operação da UMiP GenClima é conduzida por pesquisadores da Embrapa lotados na Embrapa Informática Agropecuária e por professores e pesquisadores da Unicamp. A gestão da UMiP GenClima é conduzida por um coordenador geral



**Figura 2.** Pipeline operacional e atuação das equipes da UMiP GenClima. Fases e estágios de P&D correspondentes, descritos na Figura 1, são mostrados na parte superior.

indicado pela Unicamp e por um conselho constituído por pesquisadores da Embrapa e professores da Unicamp. A UMiP GenClima tem como meta o desenvolvimento, em médio prazo, de tecnologia genética de adaptação a condições de seca, calor e alta concentração atmosférica de dióxido de carbono, usando como planta-modelo o milho, e potencialmente aplicável para outras culturas de grande impacto socioeconômico no Brasil. Consequentemente, a UMiP GenClima pretende contribuir, juntamente a outras unidades da Embrapa, para o fortalecimento da posição dessa empresa como um ator importante no cenário da biotecnologia agrícola mundial.

### 4 Considerações finais

A utilização de estratégias de seleção assistida por marcadores, juntamente com a incorporação de novos caracteres por meio da engenharia genética, permitirá, no futuro próximo, a geração de genótipos de culturas agrícolas mais produtivas e adaptadas aos novos cenários climáticos decorrentes de mudanças globais. Com a evolução das técnicas de NGS e fenotipagem de alto desempenho (Capítulo 8), será possível, respectivamente, a rápida caracterização molecular e fenotípica de uma grande diversidade de genótipos com propriedades contrastantes de crescimento, produtividade e adaptação a estresses abióticos e bióticos. Essa abordagem permitirá a identificação e a localização de genes e regiões do genoma responsáveis por diferenças fenotípicas, além da compreensão dos mecanismos que controlam as redes de expressão gênica. A Embrapa Informática Agropecuária, por seus Grupos de Pesquisa de Modelagem Agroambiental, de Bioinformática Aplicada, de Novas Tecnologias e da UMiP GenClima, vem atuando na criação e disponibilização de crescente infraestrutura computacional e nucleação de recursos humanos multidisciplinares necessários para a utilização efetiva dessas tecnologias.

### 5 Referências

- ALSTON, J. M.; BEDDOW, J. M.; PARDEY, P. G. Agricultural research, productivity, and food prices in the long run. *Science*, Washington, D. C., v. 325, n. 5945, p. 1209-1210, Sept. 2009. DOI: 10.1126/science.1170451.
- AMTMANN, A. Learning from evolution: thellungiella generates new knowledge on essential and critical components of abiotic stress tolerance in plants. *Molecular Plant*, Oxford, v. 2, n. 1, p. 3-12, Jan. 2009. DOI: 10.1093/mp/ssn094.
- BAKIK, S.; SARKARDAS, S.; SINGH, A.; GAUTAM, V.; KUMAR, P.; MAJEE, M.; SARKAR, A., K. Phylogenetic analysis reveals conservation and diversification of micro RNA166 genes among diverse plant species. *Genomics*, San Diego, v. 103, n. 1, p. 114-121, 2014. DOI: 10.1016/j.ygeno.2013.11.004.
- BARROS-NETO, N.; MOURA, R. Pior seca em 50 anos fecha empregos e arruína lucros no Nordeste. *Folha de S. Paulo*, São Paulo, 5 maio 2013.
- BERGAMASCHI, H.; WHEELER, T. R.; CHALLINOR, A. J.; COMIRAN, F.; HECKLER, B. M. M. Maize yield and rainfall on different spatial and temporal scales in Southern Brazil. *Pesquisa Agropecuária Brasileira*, Brasília, DF, v. 42, n. 5, p. 603-613, maio 2007.
- BOLGER, M. E.; WEISSHAAR, B.; SCHOLZ, U.; STEIN, N.; USADEL, B.; MAYER, K. F. Plant genome sequencing - applications for crop improvement. *Current Opinion in Biotechnology*, London, v. 26, p. 31-37, Apr. 2014. DOI: 10.1016/j.copbio.2013.08.019.

CHASSY, B. M. The history and future of GMOs in food and agriculture. **Cereal Foods World**, Minneapolis, v. 52, n. 4, p. 169-172, 2007. DOI: 10.1094/CFW-52-4-0169.

CONAB (Brasil). **Séries históricas relativas às safras 1976/77 a 2012/2013**. Disponível em: <<http://www.conab.gov.br/conteudos.php?a=1252&t=2>>. Acesso em: 3 maio 2013.

DELOITTE. **The future of the life sciences industries: strategies for success in 2015**. 2006. Disponível em: <http://www.deloitte.com/assets/Dcom-Global/Local20Assets/Documents/2006FutLifScnRptFINAL.pdf>>. Acesso em: 22 set. 2014.

DEMUTH, J. P.; HAHN, M. W. The life and death of gene families. **BioEssays**, Cambridge, v. 31, n. 1, p. 29-39, Jan. 2009. DOI: 10.1002/bies.080085.

DECONTO, J. G. (Coord.). **Aquecimento global e a nova geografia da produção agrícola no Brasil**. 2008. [Brasília, DF]: Embrapa; [Campinas]: Unicamp, 2008. 82 p. Disponível em: <<http://ainfo.cnptia.embrapa.br/digital/bitstream/item/102629/1/Aquecimento-Global-e-a-Nova-Geografia-na-Producao-Agricola-1.pdf>>. Acesso em 23 out. 2014.

DEYHOLOS, M. K. Making the most of drought and salinity transcriptomics. **Plant, Cell & Environment**, Oxford, v. 33, n. 4, p. 648-654, Apr. 2010. DOI: 10.1111/j.1365-3040.2009.02092.x.

EGAN, A. N.; SCHLUETER, J.; SPOONER, D. M. Applications of next-generation sequencing in plant biology. **American Journal of Botany**, Columbus, v. 99, n. 2, p. 175-185, Feb. 2012. DOI: 10.3732/ajb.1200020.

ESTADOS UNIDOS. Department of Agriculture. **Foreign Agricultural Service Circular Series WAP 1-13**. 2013.

EVANS, L. M.; SLAVOV, G. T.; RODGERS-MELNICK, E.; MARTIN, J.; RANJAN, P.; MUCHERO, W.; BRUNNER, A. M.; WENDY SCHACKWITZ, W.; GUNTER, L.; CHEN, J.-C.; TUSKAN, G. A.; DIFAZIO, S. P. Population genomics of *Populus trichocarpa* identifies signatures of selection and adaptive trait associations. **Nature Genetics**, New York, v. 46, n. 10, p. 1089-1096, Aug. 2014. DOI: 10.1038/ng.3075.

FAO. **Declaration of the world summit on food security**. Rome, 2009. Disponível em: <[http://www.fao.org/fileadmin/templates/wsfs/Summit/Docs/Final\\_Declaration/WSFS09\\_Declaration.pdf](http://www.fao.org/fileadmin/templates/wsfs/Summit/Docs/Final_Declaration/WSFS09_Declaration.pdf)>. 2009. Acesso em: 22 set. 2014.

FAO. **FAO Statistical Databases**. 2013. Disponível em: <<http://faostat.fao.org>>. Acesso em: 22 set. 2014.

FERREIRA, D. B.; RAO, V. B. Recent climate variability and its impacts on soybean yields in Southern Brazil. **Theoretical and Applied Climatology**, Wien, v. 105, p. 83-97, Dec. 2011. DOI: 10.1007/s00704-010-0358-8.

GARG, R.; VERMA, M.; AGRAWAL, S.; SHANKAR, R.; MAJEE, M.; JAIN, M. Deep transcriptome sequencing of wild halophyte rice, *Porteresia coarctata*, provides novel insights into the salinity and submergence tolerance factors. **DNA Research**, Oxford, p. 1-16, Oct. 2013. DOI: 10.1093/dnares/dst042.

GONG, Q.; LI, P.; MA, S.; RUPASSARA, I.; BOHNERT, H. J. Salinity stress adaptation competence in the extremophile *Thellungiella halophila* in comparison with its relative *Arabidopsis thaliana*. **The Plant Journal for Cell and Molecular Biology**, Oxford, v. 44, n. 5, p. 826-839, Dec. 2005. DOI: 10.1111/j.1365-313X.2005.02587.x.

GOODSTEIN, D. M.; SHU, S.; HOWSON, R.; NEUPANE, R.; HAYES, R. D.; FAZO, J.; MITROS, T.; DIRKS, W.; HELLSTEN, U.; PUTNAM, N.; ROKHSAR, D. S. Phytozome: a comparative platform for green plant genomics. **Nucleic Acids Research**, Oxford, v. 40, n. v.1, p. D1178-D1186, Jan. 2012. DOI: 10.1093/nar/gkr944.

HERDT, R. W. Biotechnology in Agriculture. **Annual Review of Environment and Resources**, Palo Alto, v. 31, n. 1, p. 265-295, Nov. 2006. DOI: 10.1146/annurev.energy.31.031405.091314.

HUANG, X. et al. Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. **Nature Genetics**, New York, v. 44, n. 1, p. 32-39, Dec. 2011. DOI: 10.1038/ng.1018.

HUFFORD, M. B. et al. Comparative population genomics of maize domestication and improvement. **Nature Genetics**, New York, v. 44, n. 7, p. 808-811, June 2012. DOI: 10.1038/ng.2309.

JAIN, M.; NIJHAWAN, A.; ARORA, R.; AGARWAL, P.; RAY, S.; SHARMA, P.; KAPOOR, S.; TYAGI, A. K.; KHURANA, J. P. F-Box proteins in rice. Genome-wide analysis, classification, temporal and spatial gene expression during panicle and seed development, and regulation by light and abiotic stress. **Plant Physiology**, Bethesda, v. 143, n. 4, p. 1467-1483, Apr. 2007. DOI: 10.1104/pp.106.091900.

LE, D. T. et al. Genome-Wide Expression Profiling of Soybean Two-Component System Genes in Soybean Root and Shoot Tissues under Dehydration Stress. **DNA Research**, Oxford, v. 18, n. 1, p. 17-29, Feb. 2011. DOI: 10.1093/dnares/dsq032.

LIU, Z. et al. Genome-wide identification, phylogeny, evolution and expression patterns of ap2/erf genes and cytokinin response factors in *Brassica rapa* ssp. **PLoS ONE**, v. 8, n. 12, p. e83444, 2013. DOI: 10.1371/journal.pone.0083444.

LOBELL, D. B.; GOURDJI, S. M. The influence of climate change on global crop productivity. **Plant Physiology**, Bethesda, v. 160, n. 4, p. 1686-1697, Dec. 2012. DOI: 10.1104/pp.112.208298.

LOBELL, D. B.; SCHLENKER, W.; COSTA-ROBERTS, J. Climate trends and global crop production since 1980. **Science**, Washington, D. C., v. 333, n. 6042, p. 616-620, 2011. DOI: 10.1126/science.1204531.

MA, Y.; QIN, F.; PHAN TRAN, L.-S. Contribution of genomics to gene discovery in plant abiotic stress responses. **Molecular Plant**, 2012. DOI: 10.1093/mp/sss085.

MCKOWN, A. D. et al. Genome-wide association implicates numerous genes underlying ecological trait variation in natural populations of *Populus trichocarpa*. **New Phytologist**, v. 203, n. 2, p. 535-553, 2014. DOI: 10.1111/nph.12815.

MOCHIDA, K. et al. In silico analysis of transcription factor repertoires and prediction of stress-responsive transcription factors from six major gramineae plants. **DNA Research**, v. 18, n. 5, p. 321-332, 2011. DOI: 10.1093/dnares/dsr019.

MOLINA, C. et al. The salt-responsive transcriptome of chickpea roots and nodules via deepSuperSAGE. **BMC Plant Biology**, London, v. 11, n. 1, p. 31, 2011. DOI: 10.1186/1471-2229-11-31.

MUSTROPH, A. et al. Cross-kingdom comparison of transcriptomic adjustments to low-oxygen stress highlights conserved and plant-specific responses. **Plant Physiology**, Bethesda, v. 152, n. 3, p. 1484-1500, 2010. DOI: 10.1104/pp.109.151845.

NAÇÕES UNIDAS. **Pior seca dos últimos 50 anos no nordeste brasileiro confirma estatísticas da ONU sobre escassez**. 2013. Disponível em: <<http://www.onu.org.br/pior-seca-dos-ultimos-50-anos-no-nordeste-brasileiro-confirma-estatisticas-da-onu-sobre-escassez>>. Acesso em: 22 de set. 2014.

NELSON, D. E. et al. Plant nuclear factor Y (NF-Y) B subunits confer drought tolerance and lead to improved corn yields on water-limited acres. **Proceedings of the National Academy of Sciences**, Washington, D. C., v. 104, n. 42, p. 16450-16455, 2007. DOI: 10.1073/pnas.0707193104.

OH, D.-H.; DASSANAYAKE, M.; BOHNERT, H. J.; CHEESEMAN, J. M. Life at the extreme: lessons from the genome. **Genome Biology**, v. 13, p. 241, 2012. DOI: 10.1186/gb-2012-13-3-241.

PARK, J. R. et al. The role of transgenic crops in sustainable development. **Plant Biotechnology Journal**, v. 9, n. 1, p. 2-21, 2011. DOI:10.1111/j.1467-7652.2010.00565.x.

PELLEGRINO, G. Q.; ASSAD, E. D.; MARIN, F. R. Mudanças climáticas globais e a agricultura no Brasil. **Revista Multiciência**, v. 8, p. 139-162, 2007.

PITZSCHKE, A.; HIRT, H. Bioinformatic and systems biology tools to generate testable models of signaling pathways and their targets. **Plant Physiology**, Bethesda, v. 152, n. 2, p. 460-469, 2010. DOI: 10.1104/pp.109.149583.

PRADO, J. R. et al. Genetically engineered crops: from idea to product. **Annu Rev Plant Biol**, v. 65, p. 769-90, 2014. DOI: 10.1146/annurev-arplant-050213-040039.

- PREUSS, S. B. et al. Expression of the Arabidopsis thaliana BBX32 gene in soybean increases grain yield. **PLoS One**, v. 7, n. 2, p. e30717, 2012. DOI: 10.1371/journal.pone.0030717.
- RICE, E. A. et al. Expression of a truncated ATHB17 protein in maize increases ear weight at silking. **PLoS One**, v. 9, n. 4, p. e94238, 2014. DOI: 10.1371/journal.pone.0094238.
- RIECHMANN, J. L. et al. Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes. **Science**, Washington, D. C., v. 290, n. 5499, p. 2105-2110, 2000. DOI: 10.1126/science.290.5499.2105.
- SAKURAI, G.; IIZUMI, T.; YOKOZAWA, M. Varying temporal and spatial effects of climate on maize and soybean affect yield prediction. **Climate Research**, Oldendorf, v. 49, p. 143-154, 2011. DOI: 10.3354/cr01027.
- SLEIMAN, J.; SILVA, M. E. S. Padrões atmosféricos associados a eventos extremos de precipitação sobre a região Sul do Brasil. **Revista do Departamento de Geografia**, São Paulo, v. 20, p. 93-109, 2010.
- SOLOMON, S.; QUIN, D.; MANNING, M.; MARQUIS, M.; AVERYT, K.; TIGNOR, M. M. B.; MILLER JUNIOR, H. L.; CHEN, Z. **Climate Change 2007: the physical science basis**. Cambridge: Cambridge University Press, 2007. 142 p. il. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change
- TESTER, M.; LANGRIDGE, P. Breeding technologies to increase crop production in a changing world. **Science**, Washington, D. C., v. 327, n. 5967, p. 818-822, 2010. DOI: 10.1126/science.1183700.
- TRAN, L.-S. P.; MOCHIDA, K. Identification and prediction of abiotic stress responsive transcription factors involved in abiotic stress signaling in soybean. **Plant Signaling & Behavior**, v. 5, n. 3, p. 255-257, 2010. DOI: 10.4161/psb.5.3.10550.
- WANG, H. et al. Genome-wide identification, evolution and expression analysis of the grape (*Vitis vinifera* L.) Zinc Finger-Homeodomain Gene Family. **International Journal of Molecular Sciences**, Basel, v. 15, n. 4, p. 5730-5748, 2014. DOI: 10.3390/ijms15045730.

## Bioinformática aplicada à agricultura

Poliana Fernanda Giachetto  
Roberto Hiroshi Higa

### 1 Introdução

Os avanços nas áreas de tecnologia de informação e das novas tecnologias de sequenciamento têm provocado uma necessidade cada vez maior do uso da bioinformática na agricultura, principalmente com relação ao melhoramento genético vegetal e animal. Uma vasta quantidade de dados genômicos tem sido gerada a partir de diversas espécies de plantas, animais e micro-organismos, trazendo desafios no sentido de se desenvolver novas ferramentas de análise e de integração dos dados, além de soluções para se armazenar e tratar esse grande volume de dados.

Dada a importância da Empresa Brasileira de Pesquisa Agropecuária (Embrapa) no cenário brasileiro da pesquisa agrícola, ações no sentido de prover à comunidade científica, condições para que informações biológicas possam ser obtidas a partir dos dados gerados pelas tecnologias genômicas e adequadamente utilizadas em programas de melhoramento genético, na caracterização de recursos genéticos e na biotecnologia, têm sido realizadas.

Este capítulo tem por objetivo colocar o leitor a par do estado da arte das principais aplicações da bioinformática na agricultura, particularmente àquelas relacionadas ao melhoramento genético animal e vegetal, executadas no âmbito da Embrapa.

### 2 A bioinformática e as novas tecnologias de sequenciamento

#### 2.1 Sequenciamento e montagem de genomas

Um grande desafio veio à tona com o advento das novas tecnologias de sequenciamento, ou sequenciamento de nova geração, do inglês *Next Generation Sequencing* (NGS), em termos de capacidade de armazenamento e processamento de dados, assim como a necessidade do desenvolvimento de novas ferramentas de análise.

Caracterizada por um dramático aumento na quantidade dos dados gerados, acompanhado de uma substancial redução nos custos para produzi-los, as novas tecnologias de sequenciamento geram sequências bastante curtas, comparadas ao sequenciamento utilizando a tecnologia de Sanger, ou sequenciamento tradicional (METZKER, 2010; POP; SALZBERG, 2008; STREANGER; SALZBERG, 2012). Essa característica representa um grande desafio à bioinformática, princi-

palmente na montagem de genomas: o curto tamanho das sequências resulta em dificuldades na desambiguação de regiões repetitivas, resultando em montagens fragmentadas e demandando ferramentas otimizadas para resolver essa questão, diferentes daquelas até então utilizadas com dados gerados pelo sequenciamento Sanger.

Dificuldades à parte, sem dúvidas o surgimento do sequenciamento de nova geração tem transformado várias áreas da pesquisa biológica, incluindo a agricultura. Os estudos genéticos foram largamente beneficiados com o avanço na obtenção de informações genômicas que podem ser aplicadas no pré-melhoramento e melhoramento genético de espécies animais e vegetais de interesse econômico, na caracterização de recursos genéticos com vistas à prospecção gênica e na descoberta de ativos biotecnológicos. As ferramentas desenvolvidas para a análise dos dados gerados pelos sequenciadores de nova geração incluem aquelas que permitem o alinhamento das sequências produzidas contra um genoma referência ou a montagem *de novo* das sequências geradas, resultando em um genoma “montado”, a detecção de polimorfismos do tipo SNP (*Single Nucleotide Polymorphism*) e ainda outras variações estruturais como deleções, inserções, rearranjos e variações no número de cópias de trechos do genoma - as CNVs (*Copy Number Variation*), a análise de transcriptomas em grande profundidade e a análise de metagenomas, a partir de comunidades microbianas.

Apesar da importância econômica e biológica das plantas, e do custo reduzido do sequenciamento nos dias de hoje comparados à tecnologia tradicional, poucas espécies tiveram seu genoma sequenciado. Isso é devido, em grande parte, à natureza complexa de muitos genomas vegetais, como a existência de elementos repetitivos, transposons, duplicações gênicas e variações nos níveis de ploidia, que dificultam sua adequada montagem e anotação. Ainda, a presença de grandes famílias gênicas e um número abundante de pseudogenes, derivados de eventos recentes de duplicações do genoma e da atividade de transposons (SCHNABLE et al., 2009), fazem com que as montagens obtidas sejam bastante fragmentadas. Mesmo com todas essas dificuldades, esforços têm sido empregados na obtenção da sequência de genomas de espécies vegetais, incluindo plantas modelo e aquelas de interesse comercial. A Embrapa está participando, sozinha ou em parceria com outras instituições de pesquisa, do sequenciamento e da montagem de 2 genomas de plantas de interesse econômico: genoma da palma de óleo (*Elaeis guineensis* Jacq.) e genoma do feijão comum, *Phaseolus vulgaris*, em projetos liderados pela Embrapa Agroenergia e pela Embrapa Arroz e Feijão, respectivamente.

Na área animal, a Embrapa participa de um grande projeto em rede, a Rede Genômica Animal II (RGAI) (CAETANO et al., 2012), que prevê o sequenciamento e a montagem do genoma do Nelore (*Bos indicus*), raça bovina que compõe cerca de 80% do rebanho brasileiro de gado de corte, animais Gir leiteiro, Guzerá e Girolanda, representantes de raças de gado leiteiro, e das espécies de peixe nativas do Brasil, Cachara (*Pseudoplatystoma reticulatum*) e Tambaqui (*Colossoma macropomum*). Ao contrário dos genomas vegetais, a montagem de genomas animais não apresenta alguns desafios inerentes à constituição genética das plantas, como um grande número de famílias multigênicas e elevada frequência de poliploidia, mas a presença de sequências repetitivas torna essa tarefa não trivial também no caso dessas espécies. Os principais genomas de animais de produção publicados até o momento são: suíno (ARCHIBALD et al., 2010a), caprino (DONG et al., 2013), genoma da galinha (INTERNATIONAL CHICKEN GENOME SEQUENCING CONSORTIUM, 2004), genoma bovino, de um animal da raça Hereford (*Bos taurus*) (ELSIK et al. 2009) e ovino (ARCHIBALD et al., 2010b).

Além dos *drafts* publicados dos genomas acima mencionados (animais e vegetais), uma série de pragas agrícolas, assim como endo e ectoparasitas que causam doenças e perdas econômicas em lavouras e rebanhos, também tem tido seu genoma sequenciado. Alguns exemplos incluem o nematoide do nó da raiz (*Meloidogyne incognita*), que ataca um amplo espectro de plantas e tem alto poder destrutivo (ABAD et al., 2008), o ácaro *Tetranychus urticae*, considerado um dos ácaros fitófagos mais importantes do mundo (GOTOH et al., 1993) e um dos principais ácaros praga do Brasil, em função do número de espécies vegetais atacadas e dos danos a elas causadas (MORAES; FLECHTMANN, 2008) e o besouro *Tribolium castaneum*, que causa danos a grãos secos e às farinhas (TRIBOLIUM GENOME SEQUENCING CONSORTIUM, 2008). Genomas bacterianos de importantes patógenos de plantas como o da *Xylella fastidiosa* (THE XYLELLA CONSORTIUM OF THE ORGANIZATION FOR NUCLEOTIDE SEQUENCING AND ANALYSIS, 2000), *Xhantomonas axonopodis* pv. *citri* e *X. campestris* pv. *campestris*, bactérias causadoras do cancro cítrico e podridão negra em crucíferas (SILVA et al., 2002), respectivamente, além de outras espécies de *Xhantomonas*, como *X. oryzae* pv. *oryzae* (LEE et al., 2005; SALZBERG et al., 2008), que ataca culturas de arroz e *X. albilineans* (PIERETTI et al., 2009), patógeno da cana-de-açúcar, também foram decodificados. Outros exemplos ainda incluem algumas *Pseudomonas* (BUELL et al., 2003; JOARDAR et al., 2005) e a *Ralstonia solanacearum*, bactéria que infecta as raízes das plantas (SALANOUBAT et al., 2002). Além de genomas bacterianos, o de fungos patogênicos, como alguns do gênero *Phytophthora* (HAAS et al., 2009; TILLER et al., 2006), *Neurospora crassa* (GALAGAN et al., 2003) e o *Fusarium* (JEONG et al., 2013; MA et al., 2014), também foram publicados.

Importantes patógenos que atacam animais de produção também tiveram o seu genoma sequenciado, como a *Salmonella enterica* (CHIU et al., 2005), *Brucella abortus* (HALLING et al., 2005), *Mycobacterium avium* (LI et al., 2005), *Dichelobacter nodosus* (MYERS et al., 2007), *Clostridium perfringens* (MYERS et al., 2006), *Brucella suis* (PAULSEN et al., 2002) e a *Corynebacterium pseudotuberculosis* (SOARES et al., 2013) e uma série de outros mais.

O sequenciamento do genoma de patógenos animais e vegetais permite-nos desvendar os mecanismos responsáveis pela patogenicidade das espécies e também desenvolver testes moleculares para o seu diagnóstico. Uma vez conhecidos os mecanismos patogênicos, ações de controle podem ser implementadas. Nesse sentido, a Embrapa, por meio da RGAI (CAETANO et al., 2012), pretende sequenciar o genoma de 25 cepas de lentivírus de pequenos ruminantes (LVPRs), que causam consideráveis perdas econômicas em rebanhos caprinos e ovinos no Brasil. No caso da decodificação dos genomas das espécies animais e vegetais, além de sua importância no avanço das pesquisas biológicas, a existência de um genoma referência facilita a identificação dos transcritos sintetizados por esse organismo e a análise dos transcriptomas, como veremos a seguir.

## 2.2 Genômica comparativa

O crescente acúmulo de informações sobre espécies de interesse econômico depositadas nos bancos de dados, juntamente com o desenvolvimento de ferramentas de análise adequadas, tem permitido a realização de análises comparativas com dados de organismos modelo, facilitando a descoberta de genes envolvidos em características fenotípicas economicamente importantes.

De acordo com Sharma et al. (2014), o sequenciamento do genoma da *Arabidopsis*, planta modelo da família Brassicaceae, revolucionou nosso conhecimento no campo da biologia de plantas e

tornou-se um marco nos estudos envolvendo a genômica comparativa. Hoje, a existência de bancos de dados integrados acessados via web, contendo uma série de informações genéticas obtidas em estudos com *Arabidopsis*, além de ferramentas de análise, tem nos permitido obter avanços importantes em estudos envolvendo genômica comparativa. Dentre esses bancos, podemos citar o TAIR (SWARBRECK et al., 2008), considerado o repositório de dados e de ferramentas de análise mais importante relacionado à pesquisa com *Arabidopsis* (MOCHIDA; SHINOZAKI, 2010), o SIGnal, mantido pelo Salk Institute Genomic Analysis Laboratory<sup>1</sup> e o RARGE – RIKEN *Arabidopsis* Genome Encyclopedia (SAKURAI et al., 2005). Outros bancos de dados que incluem dados de *Arabidopsis* são o Brassica Genome Gateway<sup>2</sup>, BRAD<sup>3</sup>, Phytozome<sup>4</sup>, PlantGDB<sup>5</sup>, EnsemblPlants<sup>6</sup> e ChloroplastDB<sup>7</sup>. Além dos bancos de dados de *Arabidopsis*, informações genéticas e genômicas de plantas de interesse econômico também podem ser acessadas e utilizadas em estudos de genômica comparativa: SOL Genomics Network<sup>8</sup>, Gramene<sup>9</sup>, PLAZA<sup>10</sup>, GreenPhylDB<sup>11</sup>, BarleyBase<sup>12</sup> e PlantTribes<sup>13</sup>. Os bancos acima citados compreendem dados de plantas, no entanto, dados referentes a animais e micro-organismos também existem e estão disponíveis publicamente, apesar de não citados aqui.

### 2.3 Análise de transcriptomas

O RNA-Seq, metodologia que utiliza o sequenciamento de nova geração na análise de transcriptomas, tem sido utilizado em estudos da expressão gênica de várias plantas de importância econômica, como o arroz (LU et al., 2010), milho (HANSEY et al., 2012), cevada (MAYER et al., 2012), laranja doce (XU et al., 2013) e cana-de-açúcar (CARDOSO-SILVA et al., 2014; FERREIRA et al., 2014), e animais de produção, como bovinos de corte (BALDWIN et al., 2012; LI et al., 2011; PATEL et al., 2013) e leite (MCCABE et al., 2012), suínos (ESTEVE-CODINA et al., 2011; RAMAYO-CALDAS et al., 2012), aves (PERUMBAKKAN et al., 2013), caprinos (GENG et al., 2013, LING et al., 2014) e ovinos (ZHANG et al., 2013).

A análise de transcriptomas é essencial para se conhecer os elementos funcionais de um genoma e os constituintes moleculares de células e tecidos específicos. Sua aplicação tem como objetivos específicos catalogar todos os transcritos identificados (RNAm, RNAs não codificadores e pequenos RNAs), determinar a estrutura transcricional dos genes (3' e 5' UTR, sítio de iniciação da transcrição, padrão de processamento dos éxons e modificações pós-transcripcionais) e quan-

tificar as alterações nos níveis de expressão de cada transcrito ao longo do desenvolvimento e sob diferentes condições (WANG et al., 2009). Na agricultura, essa última aplicação permite, por meio do monitoramento das mudanças no nível de expressão gênica, identificar vias biológicas alteradas em situações onde uma perturbação externa é inserida no sistema, como por exemplo, a infecção por um patógeno, infestação por um parasita, alterações na dieta, restrição hídrica e outros tipos de estresse.

A análise de transcriptomas por meio da técnica de RNA-Seq revolucionou os estudos na área, por apresentar várias vantagens quando comparada a outras utilizadas até então para a análise de transcriptomas, como os microarranjos (MARTIN; WANG, 2011; WANG et al., 2009). O RNA-Seq fornece uma medida mais precisa do nível de transcritos, assim como suas isoformas, e promove uma acurada quantificação da expressão diferencial, capaz de gerar novos conhecimentos sobre mecanismos moleculares de uma dada característica de interesse, em adição à identificação de novos transcritos. O RNA-Seq permite também a identificação sistemática de SNPs em regiões transcritas, os quais podem ser utilizados na detecção de expressão alelo-específica e também como marcadores moleculares. Como exemplo dessa aplicação, citamos o trabalho de Cardoso-Silva et al. (2014) que aplicaram a tecnologia de RNA-Seq a 6 genótipos de cana-de-açúcar, contrastantes para o teor de sacarose. Os autores identificaram SNPs exclusivos de cada genótipo, os quais possuem uma alta probabilidade de associação com as características de interesse econômico particulares a cada um deles. Estratégias como essa, que têm por objetivo a busca de SNPs em regiões codificadoras, por meio do sequenciamento em larga escala do transcriptoma de tecidos alvo de cultivares ou variedades contrastantes para características de interesse, têm sido largamente utilizadas em projetos em andamento na Embrapa, envolvendo culturas como o trigo, soja, milho, café e cana-de-açúcar, e as características avaliadas incluem aumento de produtividade, resistência a doenças e pragas, tolerância à seca e melhoria da qualidade do produto final, entre outros.

Uma outra abordagem que utiliza a tecnologia de RNA-Seq, certamente a mais conhecida e empregada, foi recentemente utilizada por Nishiyama Junior, et al. (2014), que sequenciaram e compararam o transcriptoma de 3 cultivares comerciais de cana-de-açúcar e dos 2 genótipos parentais (*S. officinarum* e *S. spontaneum*), identificando diferenças na expressão gênica que possibilitarão, segundo os autores, associar transcritos espécie-específicos à produção de biomassa e outras características importantes. Os transcritos específicos identificados nesse tipo de estudo podem ser utilizados como marcadores moleculares, e serem empregados como ferramentas em programas de seleção assistida por marcadores, prestando um grande auxílio a avanços no melhoramento genético da cultura avaliada.

Importantes descobertas na área animal têm sido feitas utilizando-se a tecnologia de RNA-Seq. Genes relacionados a características de qualidade da carne em suínos (JUNG et al., 2012; RAMAYO-CALDAS et al., 2012) e gado de corte (LEE et al., 2013; SHENG et al., 2014) têm sido identificados. Essa característica tem sido alvo de estudos da Embrapa, que lidera projetos utilizando RNA-Seq com o objetivo de identificar os mecanismos que determinam a maciez da carne bovina, um atributo bastante apreciado pelos consumidores e que não é típico de animais *Bos indicus*, espécie a qual pertence o Nelore, principal raça de gado de corte criada do Brasil, como já foi citado. Esses animais apresentam tipicamente carne menos macia do que os animais da espécie *Bos taurus* e trabalhos têm sido conduzidos pela Embrapa Pecuária Sul, pela Embrapa Pecuária Sudeste e pela Embrapa Gado de Corte no sentido de comparar as duas subespécies

<sup>1</sup> Disponível em: <<http://signal.salk.edu/>>.

<sup>2</sup> Disponível em: <<http://brassica.bbsrc.ac.uk/>>.

<sup>3</sup> Disponível em: <<http://brassicadb.org/brad/>>.

<sup>4</sup> Disponível em: <<http://www.phytozome.net/>>.

<sup>5</sup> Disponível em: <<http://www.plantgdb.org/>>.

<sup>6</sup> Disponível em: <<http://plants.ensembl.org/index.html>>.

<sup>7</sup> Disponível em: <<http://chloroplast.cbio.psu.edu/>>.

<sup>8</sup> Disponível em: <<http://solgenomics.net/>>.

<sup>9</sup> Disponível em: <<http://www.gramene.org/>>.

<sup>10</sup> Disponível em: <<http://bioinformatics.psb.ugent.be/plaza/>>.

<sup>11</sup> Disponível em: <<http://www.greenphyl.org/>>.

<sup>12</sup> Disponível em: <<http://www.plexdb.org/plex.php?database=Barley>>.

<sup>13</sup> Disponível em: <<http://fgp.huck.psu.edu/tribedb/>>.

ou mesmo animais extremos para essa característica dentro de cada raça e identificar os genes responsáveis pela maciez.

A busca por genes e/ou mecanismos envolvidos na resistência de animais a doenças e parasitas responsáveis por perdas econômicas no setor também tem sido objeto de estudos na Embrapa. A comparação do transcriptoma de caprinos e ovinos resistentes e susceptíveis ao parasita *Haemonchus contortus*, principal helminto que ataca os rebanhos no Brasil (COSTA et al., 2000), assim como o estudo do transcriptoma do próprio parasita, estão em andamento em projetos conduzidos na Embrapa Caprinos e Ovinos. Outro parasita que causa sérios prejuízos econômicos ao Brasil e ao mundo, o carrapato bovino *Rhipicephalus microplus*, tem sido tema de muitas pesquisas na Embrapa, que buscam formas de controle por meio da compreensão dos mecanismos ativados pelo hospedeiro e/ou pelo ácaro no momento da interação e também na identificação de genes diferencialmente expressos em animais resistentes e susceptíveis. A Embrapa Pecuária Sul tem desenvolvido importantes projetos na área, juntamente com a Embrapa Pecuária Sudeste e a Embrapa Gado de Corte. Estudos acerca da identificação de genes relacionados a problemas ósseos em frangos de corte, que causam dificuldades de locomoção nas aves e consequente redução do crescimento, também têm sido investigados por meio de RNA-Seq pela Embrapa Suínos e Aves.

#### 2.4 O estudo de comunidades microbianas

As pesquisas envolvendo o estudo de microbiomas também foram alavancadas com o surgimento das novas tecnologias de sequenciamento, as quais trouxeram inúmeras vantagens em relação aos métodos até então existentes, baseados em cultura de micro-organismos *in vitro*. A utilização de NGS tem possibilitado o conhecimento da composição da comunidade microbiana, por meio do sequenciamento de regiões hipervariáveis do gene 16S RNAr de bactérias e região Internal Transcribed Spacer (ITS) de fungos, e também de uma quantidade bastante grande de genes presentes nas comunidades microbianas, por meio do sequenciamento do DNA total delas extraído (metagenomas). A metagenômica tem se mostrado mais eficiente do que os métodos culturais principalmente pelo maior poder de identificação das cepas e por tornar possível a caracterização e o monitoramento de alterações na dinâmica de uma comunidade microbiana como um todo, e não apenas de indivíduos em particular. O conhecimento dos membros e dos genes presentes em um microbioma permite-nos inferir o papel de cada um na comunidade. No entanto, além de identificar quem está presente no microbioma, saber o que fazem é igualmente importante e nos possibilita entender a dinâmica das relações entre os micro-organismos e também a relação entre microbiota e hospedeiro. Essa informação pode ser obtida por meio do sequenciamento dos transcritos expressos pelos membros da comunidade microbiana e obtenção do metatranscriptoma. A técnica de RNA-Seq, já discutida neste capítulo, tem sido utilizada com esse objetivo.

No caso dos estudos genômicos, duas abordagens de estudo de comunidades microbianas podem ser adotadas, utilizando-se as novas tecnologias de sequenciamento: a) análise de amplicons do gene 16S RNAr, onde uma ou mais regiões hipervariáveis do gene 16S RNAr de membros individuais da comunidade é amplificada, por meio de reações em cadeia de polimerase, do inglês Polymerase Chain Reaction (PCR), utilizando iniciadores complementares a regiões conservadas do gene (que flanqueiam as regiões hipervariáveis) e sequenciada, com posterior comparação das sequências com aquelas depositadas em bancos de dados existentes, classificação e cálculo da abundância relativa e b) análise do metagenoma, por meio do sequenciamento em larga escala

do DNA extraído de toda a comunidade microbiana, o que permite a identificação dos genes que compõem o microbioma e também da diversidade da população, embora de maneira menos precisa do que a obtida com a classe anterior (OVIEDO-RONDÓN, 2009). O sequenciamento de amplicons do gene 16S RNAr é a abordagem mais acessível e utilizada em estudos que visam à identificação de membros de uma comunidade microbiana, sendo considerado um bom marcador para se caracterizar a composição filogenética de uma amostra, identificar novas espécies ou mesmo grupos filogenéticos desconhecidos (GABOR et al., 2007). O fato de vários organismos possuírem sequências do 16S RNAr idênticas, mas funções diferentes, não nos permite utilizar essa abordagem para inferir adequadamente as funções dos membros de uma comunidade microbiana. Essa limitação da técnica pode ser superada com a análise de metagenomas e/ou metatranscriptomas.

Dada a importância dos micro-organismos na agricultura, são vários os exemplos de aplicação da análise de sequências de 16S RNAr e de metagenomas nos estudos de comunidades microbianas. Em recente revisão publicada por Kao-Kniffin et al. (2013), os autores citaram estudos onde a abordagem metagenômica foi utilizada na identificação de genes e compostos com ação inseticida, e genes de resistência a herbicidas, para utilização no controle de pragas em sistemas agrícolas. De acordo com os autores, o isolamento do DNA diretamente das amostras ambientais, sem passar pelo processo de cultivo em meio de cultura, aumenta muito a diversidade de micro-organismos recuperada das amostras e, conseqüentemente, a possibilidade de identificação de compostos de interesse.

Por se tratar do ambiente que abriga a maior biodiversidade do planeta (MOCALI; BENEDETTI, 2010), o estudo de comunidades microbianas do solo apresenta um potencial imenso para a agricultura: os micro-organismos do solo desempenham um papel crítico na regulação da fertilidade do solo, saúde das plantas e na ciclagem do carbono, nitrogênio e outros nutrientes (FIERER et al., 2012a). Nesse sentido, vários estudos acerca da caracterização da microbiota de diferentes tipos de solos têm sido realizados (FIERER et al., 2007; FIERER et al., 2012b; GROS et al., 2006; SCHLOSS; HANDELSMAN, 2006). O efeito das práticas agrícolas sobre as comunidades microbianas do solo também tem sido alvo de estudos, que já identificaram o pH e a disponibilidade de água como os principais fatores abióticos que afetam as populações (KÖBERL et al., 2011), além da quantidade de N depositada no solo (FIERER et al., 2012b).

Acredita-se que as plantas recrutam micro-organismos benéficos do solo, a partir de suas rizosferas, para neutralizar o ataque de patógenos (COOK et al., 1995). Ainda, existem certos tipos de solos, conhecidos como supressivos, que abrigam comunidades microbianas que impedem que os agentes patógenos causem danos às plantas. Assim, a caracterização de microbiomas da rizosfera de plantas sob condições diversas, e aquela presente em solos supressivos, é uma outra área que tem se beneficiado dos estudos metagenômicos e com imenso potencial para gerar resultados importantes para a agricultura, como a identificação de novas bactérias fixadoras de N<sub>2</sub>, supressoras de patógenos e promotoras do crescimento.

No mesmo contexto, na área animal estudos têm sido realizados com o objetivo de se identificar micro-organismos benéficos, que aumentem a produtividade e o bem estar animal. Nesse sentido, a caracterização de microbiomas do trato gastrointestinal tem sido foco de bastante atenção, uma vez que, além do seu papel na digestão, a microbiota do trato gastrointestinal exerce influência também na proteção contra patógenos, detoxificação e modulação do sistema imune do hospedeiro.

deiro (AMIT-ROMACH et al., 2004; OVIEDO-RONDÓN, 2009). A identificação de micro-organismos com baixa produção de metano em ruminantes também tem sido objeto de estudos, uma vez que a emissão entérica de metano tem sido apontada como responsável por cerca de 15% da emissão global de metano, de acordo com o Painel Intergovernamental sobre Mudanças Climáticas (IPCC) (ROSS et al., 2013; SHI et al., 2014).

A Embrapa tem liderado e atuado em parceria em vários projetos que utilizam a metagenômica na caracterização de microbiomas e para identificar micro-organismos com potencial biotecnológico. Dentre eles podemos citar o estudo do microbioma da rizosfera e folhas de plantas resistentes à seca e a patógenos, com participação da Embrapa Meio Ambiente, além do estudo de microbiomas do trato intestinal e sistema digestório de frangos de corte, da glândula mamária e rúmen de ovinos e sistema respiratório de suínos, como parte dos projetos em desenvolvimento na RGAII.

### 3 Bioinformática e a utilização de marcadores moleculares no melhoramento genético animal e vegetal

#### 3.1 Uso de marcadores moleculares do tipo SNP em genética animal aplicada

Polimorfismos de base única (SNPs) são variações no genoma em que um único nucleotídeo - A, T, C, G, difere entre membros de uma população (FOULKES, 2009). Em geral, assume-se que SNPs são bialélicos, indicando a existência de duas possíveis bases no correspondente *locus*, e que cada base deve ter uma frequência mínima de 1% na população. Embora sua distribuição não seja homogênea ao longo do genoma (ex: eles ocorrem mais frequentemente em regiões não codificantes que em regiões codificantes), existem milhões de polimorfismos deste tipo distribuídos ao longo do genoma de espécies de interesse zootécnico, o que faz dele um marcador genômico muito atraente.

A partir da montagem dos genomas de espécies de interesse zootécnico e o subsequente mapeamento de haplótipos (ex: genoma bovino (BOVINE HAPMAP CONSORTIUM, 2009; ELSIK, 2009), foram desenvolvidas tecnologias para descoberta e genotipagem em massa de centenas de milhares de marcadores SNP a um custo que favorece sua utilização em estudos de associação e mapeamento genético, seleção genômica, detecção de doenças genéticas e/ou polimorfismos associados a características de produção, ensaios diagnósticos para confirmação de paternidade, identificação individual (rastreadabilidade) (CAETANO, 2009).

O desenvolvimento de todas essas aplicações envolve uma série de etapas experimentais e de análise de dados, exigindo infraestrutura e especialistas em diferentes áreas como biologia molecular, zootecnia, veterinária, estatística e ciência da computação. Na sequência, é apresentada uma visão geral das aplicações em seleção genômica e estudos de associação genômica ampla e as iniciativas em andamento na Embrapa.

#### 3.2 Seleção Genômica (GS)

O objetivo do melhoramento genético é alcançar melhores níveis de produção, produtividade e/ou qualidade do produto final, em sintonia com o sistema de produção e as exigências do mercado (ROSA et al., 2013). Para isso, diversas características expressas nos animais da população em

seleção e relacionadas com o objetivo do programa de melhoramento genético, são monitoradas, como por exemplo: eficiência reprodutiva, eficiência alimentar, qualidade da carcaça e da carne para gado de corte, produção de leite, percentagem de gordura e proteínas no leite para gado de leite, tamanho da ninhada, peso ao desmame para suínos, entre outras (BOURDON, 2000).

Os valores mensurados dessas características, denominadas fenótipos (P), para um animal dependem basicamente de sua genética (G: genótipo), do ambiente no qual ele está exposto (E: ambiente) e da expressão dos genótipos quando expostos a diferentes condições ambientais. De forma sucinta, essa relação pode ser expressa como:  $P = G + E + G \cdot E$  (ROSA et al., 2013).

Já as avaliações genéticas têm por objetivo identificar os animais em uma população sob seleção que sejam geneticamente superiores para que, se usados na reprodução, transmitam aos seus descendentes sua superioridade, dessa forma, alterando o desempenho médio da população (MARTINS, 2013). Tradicionalmente, o processo para determinar o valor genético dos animais (*Breeding Value* - BV) depende de três diferentes fontes de informação: seu desempenho, o desempenho de seus ancestrais (pedigree) e de seus descendentes (progênie).

Meuwissen et al. (2001) propuseram a utilização de marcadores SNP, que cobrem densamente todo o genoma, para avaliação genética de animais, um processo conhecido como seleção genômica. Neste caso, o valor genético do animal é denominado valor genético genômico, do inglês *Genomic Breeding Value* (GBV) e compreende o somatório dos efeitos de todos os marcadores SNP utilizados. A utilização dessa tecnologia em países desenvolvidos tem proporcionado ganhos genéticos aos programas de melhoramento genético, devido ao aumento das acurácias dos valores genéticos para características de baixa herdabilidade e redução do intervalo entre gerações para características de difícil mensuração ou aquelas obtidas tardiamente, como eficiência alimentar e dados de carcaça.

A Embrapa, atualmente, participa de diferentes esforços multi-institucionais para inclusão da tecnologia de seleção genômica em diferentes programas de melhoramento genético animal em que atua. Um desses esforços, desenvolvido pela Embrapa Pecuária Sul em associação com a Conexão Delta G e o Gensys Consultores Associados, foca na seleção de animais com características de resistência a carrapatos para as raças Hereford e Braford. O carrapato (*Rhipicephalus microplus*) é um dos principais problemas de saúde na pecuária de corte, implicando em custos de produção adicionais com acaricidas e no tratamento de doenças infecciosas transmitidas pelo parasita, além da queda de performance dos animais. Estima-se que a perda econômica com carrapatos no Brasil seja de dois bilhões de dólares anuais (GRISI et al., 2002). Desde 2012, a Embrapa Pecuária Sul e seus parceiros publicam o sumário de touros para resistência a carrapatos para as raças Hereford e Braford (CARDOSO et al., 2013).

Dentre os outros esforços da Embrapa, visando à inclusão da tecnologia de seleção genômica em programas de melhoramento genético animal, destacam-se: a) Genhol (NAPOLIS, 2012), que tem entre seus objetivos específicos desenvolver estudos, processos e validação de modelos de predição de valores genômicos para a raça holandesa no Brasil; b) Genomilk (SILVA, 2010), que tem entre seus objetivos específicos adequar, avaliar e/ou comparar diferentes métodos estatísticos para implementar a seleção genômica e integração de dados genômicos nos sistemas de avaliação genética e testes de progênie conduzidos pela Embrapa Gado de Leite; c) arranjo MaxiBife (SILVA, 2013), que prevê a inclusão das tecnologias para avaliação genômica no programa de melhoramento genético de gado de corte Geneplus-Embrapa; d) rede MP1 RGA II (CAETANO,

2012), que possui um projeto componente (Projeto Componente 3) focado na inclusão de seleção genômica em programas de melhoramento genético de bovinos, ovinos e caprinos.

### 3.3 Estudos de associação genômica amplo (GWAS)

O termo Genome-Wide Association Studies (GWAS) deve-se à ampla cobertura dos marcadores SNP ao longo do genoma, tornando-os atrativos para análises de associação genótipo:fenótipo envolvendo efeito poligênico. Estes estudos têm como objetivo identificar padrões de polimorfismos que variam sistematicamente entre indivíduos com diferentes valores de expressão de um determinado fenótipo (BALDING, 2006). Para realização de GWAS são necessários três elementos:

- 1) Uma grande quantidade de amostras;
- 2) marcadores genéticos que cubram grande parte do genoma;
- 3) métodos analíticos poderosos o suficiente para identificar sem viés a associação entre os marcadores e os fenótipos analisados (CANTOR et al., 2010).

Os procedimentos computacionais para análise de dados de experimentos de GWAS envolvem uma série de passos (ZIEGLER et al., 2008), existindo diversos métodos para identificação dos SNP associados com o fenótipo estudado. Dentre estes, os mais comumente utilizados constituem-se em testes univariados por SNP, com correção para múltiplos testes (ZIEGLER et al., 2008). Abordagens envolvendo a análise simultânea de múltiplos marcadores baseiam-se em técnicas de regressão, paramétricas ou não paramétricas, com seleção ou encolhimento (*shrinkage*) de variáveis (MOORE et al., 2010; ZIEGLER et al., 2008).

A Embrapa tem desenvolvido diversos projetos envolvendo análise de GWAS, sendo que alguns já possuem resultados publicados. Mokry et al. (2013) analisaram um conjunto de dados 400 animais da raça Canchim genotipados com o BovineHD BeadChip (ILLUMINA INC., SAN DIEGO, CA), visando identificar SNP associados com espessura de gordura. Foi encontrado um conjunto de SNPs capaz de explicar aproximadamente 50% da variância do valor genético derredado (dEBV) para espessura de gordura e um pequeno conjunto de 5 SNPs capaz de explicar 34% da dEBV. Foram encontrados diversos Quantitative Trait Loci (QTL) relacionados com gordura na vizinhança desses SNPs, bem como genes envolvidos no metabolismo de lipídeos. Tizioto et al. (2013) conduziram um estudo de GWAS com 800 animais da raça Nelore genotipados com o BovineHD BeadChip (Illumina Inc., San Diego, CA), visando identificar SNPs associados com diferentes fenótipos relacionados com qualidade de carne: força de cisalhamento Warner-Bratzler, medida em diferentes tempos de maturação, espessura de gordura, área de olho do lombo, parâmetros de cor da carne e gordura, capacidade de retenção de água, perdas no cozimento e pH do músculo. As regiões genômicas na vizinhança dos QTL encontrados e as vias das quais esses genes fazem parte diferem daqueles identificados em raças taurinas. Espera-se que esses resultados subsidiem futuros estudos de mapeamento de QTL e o desenvolvimento de modelos para predição de mérito genético para qualidade de carne para a raça Nelore.

Diversos outros projetos encontram-se em andamento na Embrapa, fazendo uso de GWAS para identificar genes e regiões gênicas associadas com características fenotípicas de interesse zootécnico. Na RGAI (CAETANO, 2012), o projeto componente 4 está conduzindo três experimentos de GWAS: (i) estudos de associação para características de presença/ausência de chifres e criptorquidia em 250 ovinos da raça Morava Nova, genotipados com o Ovine SNP50Beadchip

(Illumina Inc., San Diego, CA); (ii) estudos de associação para 85 características avaliadas para 1.000 aves da população de Referência TT, representativa da melhor linha paterna de corte do programa de melhoramento genético da Embrapa Suínos e Aves, genotipados com o Chicken SNP60BeadChip (Illumina Inc., San Diego, CA); estudos de associação para a característica de presença de chifres e batoques em rebanhos Nelore Mocho BRGN, pertencentes à Embrapa Cerrados e de rebanhos de importantes criadores de Nelore mocho, como a Guaporé S/A, utilizando 350 animais genotipados com o Bovine SNP50Beadchip (Illumina Inc., San Diego, CA).

## 4 Laboratório multiusuário de bioinformática (LMB)

Em função da crescente demanda por poder computacional e competência multidisciplinar para lidar com os grandes volumes de dados, algoritmos e ferramentas de análise diversos, a Embrapa, por meio de uma decisão estratégica, inaugurou em outubro de 2011, o LMB, com sede na Embrapa Informática Agropecuária.

O LMB, que tem como missão viabilizar soluções de bioinformática para projetos de pesquisa, desenvolvimento e inovação da Embrapa, em um ambiente colaborativo<sup>14</sup>, busca incorporar e tornar disponíveis a comunidade científica novas tecnologias para armazenamento, processamento e análise de grandes volumes de dados. O LMB atua ainda na elaboração e execução de projetos e planos de ação que necessitam de ferramentas especializadas e computação de alto desempenho, na disponibilização de procedimentos computacionais para a montagem de genomas, análise de metagenomas e de transcriptomas, e na análise de dados de marcadores moleculares e de expressão gênica, desenvolvimento e implantação de recursos computacionais, criação e administração de bancos e bases de dados. A capacitação técnica, por meio da realização de cursos e treinamentos em ferramentas usadas para análise de dados, como os softwares Galaxy e Generic Genome Browse (Gbrowse), também faz parte das atuações do LMB.

## 5 Considerações finais

Dada a importância do melhoramento genético na agricultura sustentável e na segurança alimentar no nosso país e do mundo, o uso de dados genômicos para o melhoramento de espécies animais e vegetais de interesse econômico é de fundamental importância, e faz parte das ações da Embrapa que visam garantir a demanda da produção de alimentos frente ao crescimento da população e também aos cenários futuros de mudanças climáticas. O desenvolvimento e otimização de ferramentas de bioinformática que permitam a análise dos dados gerados pelas tecnologias genômicas em crescente avanço, assim como sua disponibilização à comunidade científica, é preocupação constante da empresa, que trabalha sempre no sentido de prover as melhores soluções para o aproveitamento das informações geradas, no sentido de contribuir para o avanço no setor e manutenção do Brasil como um país de excelência em pesquisa agrícola.

<sup>14</sup> Disponível em: <<http://www.lmb.cnptia.embrapa.br>>.

## 6 Referências

- ABAD, P. et al. Genome sequence of the metazoan plant-parasitic nematode *Meloidogyne incognita*. **Nature Biotechnology**, New York, v. 26, n. 8, p. 909-915, Aug. 2008.
- AMIT-ROMACH, E.; SKLAN, D.; UNI, Z. Microflora ecology of the chicken intestine using 16S ribosomal DNA primers. **Poultry Science**, Champaign, v. 83, n. 7, p.1093-1098, July 2004.
- ARCHIBALD, A. L.; BOLUND, L.; CHURCHER, C.; FREDHOLM, M.; GROENEN, M. A.; HARLIZIUS, B.; LEE, K. T.; MILAN, D.; ROGERS, J.; ROTHSCCHILD, M. F.; UENISHI, H.; WANG, J.; SCHOOK, L. B. Pig genome sequence - analysis and publication strategy. **BMC Genomics**, London, v. 11, p. 438, July 2010. DOI: 10.1186/1471-2164-11-438.
- ARCHIBALD, A. L.; COCKETT, N. E.; DALRYMPLE, B. P.; FARAUT, T.; KIJAS, J. W.; MADDOX, J. F.; MCEWAN, J. C.; HUTTON ODDY, V.; RAADSMA, H. W.; WADE, C.; WANG, J.; WANG, W.; XUN, X. The sheep genome reference sequence: a work in progress. **Animal Genetics**, Oxford, v. 41, n. 5, p. 449-453, Oct. 2010b. DOI: 10.1111/j.13652052.2010.02100.x.
- BALDING, D. J. A tutorial on statistical methods for population association studies. **Nature Review Genetics**, London, v. 7, n. 10, p. 781-791. Oct. 2006. DOI:10.1038/nrg1916.
- BALDWIN, R. L.; LI, R.W.; LI, C.; THOMSON, J. M.; BEQUETTE, B. J. Characterization of the longissimus lumborum transcriptome response to adding propionate to the diet of growing Angus beef steers. **Physiological Genomics**, Bethesda, v. 44, n. 10, p. 543-550, May 2012.
- BOURDON, R. M. **Understanding animal breeding**. 2nd ed. Upper Saddle River: Prentice Hall, 2000. 538 p. il.
- BOVINE HAPMAP CONSORTIUM. Genome-wide survey of snp variation uncovers the genetic structure of cattle breeds. **Science**, New York, v. 342, n. 5926, p. 528-532, Apr. 2009. DOI: 10.1126/science.1167936.
- BUELL, C. R. et al. The complete genome sequence of the Arabidopsis and tomato pathogen *Pseudomonas syringae* pv. Tomato DC3000. **Proceedings of the Academy of Sciences of the United States of America**, Washington, US, v. 100, n.18, p.10181-10186, Sept. 2003.
- CAETANO, A. R. Marcadores SNP: conceitos básicos, aplicações no manejo e no melhoramento animal e perspectivas para o futuro. **Revista Brasileira de Zootecnia**, Viçosa, MG, v. 38, p. 64-71, 2009.
- CAETANO, A. R. **RGAI - Rede nacional para o desenvolvimento e adaptação de estratégias genômicas inovadoras aplicadas ao melhoramento, conservação e produção animal**. Projeto SEG: 01.11.07.002.00.00. Embrapa, 2012.
- CANTOR, R. M.; LNAE, K.; SINSHEIMER, J. S. Prioritizing GWAS results: a review of statistical methods and recommendations for their application. **American Journal of Human Genetics**, Chicago, v. 86, n. 1, p. 6-22. Jan. 2010. DOI: 10.1016/j.ajhg.2009.11.017.
- CARDOSO, F. F.; YOKOO, M. J. I.; GULIAS-GOMES, C. C.; SOLLERO, B. P.; COSTA, R. F. da; ROSO, V. M.; BRITO, F. V.; CAETANO, A. R.; AGUILAR, I. **Avaliação genômica para resistência ao carrapato de touros Hereford e Braford**. Bagé: Embrapa Pecuária Sul, 2013. 40 p. (Embrapa Pecuária Sul. Documentos, 133). Disponível em: <http://ainfo.cnptia.embrapa.br/digital/bitstream/item/100508/1/DT1332013.pdf>. Acesso em: 22 set. 2014.
- CARDOSO-SILVA, C. B. et al. De novo assembly and transcriptome analysis of contrasting sugarcane varieties. **PLoS One**, San Francisco, v. 9, n. 2, e88462, Feb. 2014. DOI:10.1371/journal.pone.0088462.
- CHIU, C. H. et al. The genome sequence of *Salmonella enterica* serovar choleraesuis, a highly invasive and resistant zoonotic pathogen. **Nucleic Acids Research**, v. 33, n. 5, p. 1690-1698, Mar. 21, 2005. DOI:10.1093/nar/gki297.
- COOK, R. J.; COOK, R. J.; THOMASHOW, L. S.; WELLER, D. M.; FUJIMOTO, D.; MAZZOLA, M.; BANGERA, G.; KIM, D. S. Molecular mechanisms of defense root disease by rhizobacteria against root disease. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, US, v. 92, p. 4197-4201, 1995.
- DONG, Y.; XIE, M.; JIANG, Y.; XIAO, N.; DU, X.; ZHANG, W.; TOSSER-KLOPP, G.; WANG, J.; YANG, S.; LIANG, J.; CHEN, W.; CHEN, J.; ZENG, P.; HOU, Y.; BIAN, C.; PAN, S.; LI, Y.; LIU, X.; WANG, W.; SERVIN, B.; SAYRE, B.; ZHU, B.; SWEENEY, D. Sequencing and automated whole-genome optical mapping of the genome of a domestic goat (*Capra hircus*). **Nature Biotechnology**, New York, v. 31, p. 135-41, Dec. 2013. DOI:10.1038/nbt.2478.
- ELSIK, C. G.; TELLAM, R. S.; WORLEY, K. C. The genome sequence of taurine cattle: a window to ruminant biology and evolution. **Science**, New York, v. 342, n. 5926, p. 522-528, Apr. 2009. DOI: 10.1126/science.1169588.
- ESTEVE-CODINA, A.; KOFLER, R.; PALMIERI, N.; BUSSOTTI, G.; NOTREDAME, C.; PÉREZ-ENCISO, M. Exploring the gonad transcriptome of two extreme male pigs with RNA-seq. **BMC Genomics**, London, v. 12, n. 552, p. 1-14, 2011. DOI:10.1186/1471-2164-12-552.
- FIERER, N.; BRADFORD, M. A.; JACKSON, R. B. Toward an ecological classification of soil bacteria. **Ecology**, Tempe, v. 88, n. 6, p. 1354-1364, June 2007.
- FIERER, N.; LAUBER, C. L.; RAMIREZ, K. S.; ZANEVELD, J.; BRADFORD, M. A.; KNIGHT, R. Comparative metagenomic, phylogenetic and physiological analyses of soil microbial communities across nitrogen gradients. **The ISME Journal**, New York, v. 6, n. 5, p. 1007-1017, May 2012a. DOI: 10.1038/ismej.2011.159.
- FIERER, N.; LEF, J. W.; ADAMS, B. J.; NIELSEN, U. N.; BATES, S. T.; LAUBER, C. L.; OWENS, S.; GILBERT, J. A.; WALL, D. H.; CAPORASO, G. Cross-biome metagenomic analyses of soil microbial communities and their functional attributes. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, US, v. 109, n. 52, p. 21390-21395, Nov. 2012b. DOI: 10.1073/pnas.1215210110.
- FOULKES, A. S. **Applied statistical genetics with R: for population based association studies**. New York: Springer Verlag, 2009. 252 p. DOI: 10.1007/978-0-387-89554-3. Ebooks.
- GABOR, E. et al. Updating the metagenomics toolbox. **Biotechnology Journal**, Weinheim, v. 2, n. 2, p. 201-206, Feb. 2007. DOI:10.1002/biot.200600250.
- GALAGAN, J. E. et al. The genome sequence of the filamentous fungus *Neurospora crassa*. **Nature**, London, v. 422, n. 6934, p. 859-868, Apr. 2003.
- GENG, R.; YUAN, C.; CHEN, Y. Exploring differentially expressed genes by RNA-Seq in cashmere goat (*Capra hircus*) skin during hair follicle development and cycling. **PLoS One**, San Francisco, v. 8, n. 4, e62704, 2013. DOI:10.1371/journal.pone.0062704.
- GOTOH, T.; BRUIN, J.; SABELIS, M.W.; MENKEN, S.B.J. Host race formation in *Tetranychus urticae*: genetic differentiation, host plant preference, and mate choice in a tomato and a cucumber strain. **Entomologia Experimentalis et Applicata**, Dordrecht, v. 68, 171-178, 1993.
- GRISI, L.; MASSARD, L. C.; BORJA, G. E. M.; PEREIRA, J. B. Impacto econômico das principais ectoparasitoses em bovinos no Brasil. **A Hora Veterinária**, Porto Alegre, v. 125, p. 8-10, Nov. 2002.
- GROS, R.; MONROZIER, L. J.; FAIVRE, P. Does disturbance and restoration of alpine grassland soils affect the genetic structure and diversity of bacterial and N<sub>2</sub>-fixing populations? **Environmental Microbiology**, Oxford, v. 8, n. 11, p. 1889-1901, Nov. 2006.
- HAAS, B. J. et al. Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans*. **Nature**, London, v. 461, p. 393-398, Sept. 2009. DOI:10.1038/nature08358.
- HALLING, S. M. et al. Completion of the genome sequence of *Brucella abortus* and comparison to the highly similar genomes of *Brucella melitensis* and *Brucella suis*. **Journal of Bacteriology**, v. 187, n. 8, p. 2715-2726, Apr. 2005. DOI:10.1128/JB.187.8.2715-2726.2005.
- INTERNATIONAL CHICKEN GENOME SEQUENCING CONSORTIUM. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. **Nature**, London, v. 432, n. 7018, p. 695-716, Dec. 2004. DOI: 10.1038/nature03154.

JEONG, H.; LEE, S.; CHOI, G. J.; LEE, T.; YUN, S.-H. Draft genome sequence of *Fusarium fujikuroi* B14, the causal agent of the bakanae disease of rice. **Genome Announcements**, Washington, D.C. v. 1, n. 1, p.1-2, Jan./Feb. 2013. DOI: 10.1128/genomeA.00035-13.

JOARDAR, V. et al. Whole-genome sequence analysis of *Pseudomonas syringae* pv. *Phaseolicola* 1448A reveals divergence among pathovars in genes involved in virulence and transposition. **Journal of Bacteriology**, Washington, US, v. 187, n. 18, p. 6488-6498, Sept. 2005.

JUNG, W. Y. et al. RNA-Seq approach for genetic improvement of meat quality in pig and evolutionary insight into the substrate specificity of animal carbonyl reductases. **PLoS One**, San Francisco, v. 7, n. 9, e42198, Sept. 2012. DOI:10.1371/journal.pone.0042198.

KAO-KNIFFIN, A. J.; CARVER, S. M.; DITOMMASO, A.; KAO-KNIFFIN, J. Advancing weed management strategies using metagenomic techniques. **Weed Science**, Champaign, v. 61, n. 2, p. 171-184, 2013. DOI: http://dx.doi.org/10.1614/WS-D-12-00114.1.

KÖBERL, M.; MÜLLER, H.; RAMADAN, E. M.; BERG, G. Desert farming benefits from microbial potential in arid soils and promotes diversity and plant health. **PLoS One**, San Francisco, v. 6, n. 9, p. e24452, Sept. 2011. DOI: 10.1371/journal.pone.0024452.

LEE, B.-M. et al. The genome sequence of *Xanthomonas oryzae* pathovar *oryzae* KACC10331, the bacterial blight pathogen of rice. **Nucleic Acids Research**, v. 33, n. 2, p. 577-586, 2005. DOI:10.1093/nar/gki206.

LEE, H. J. et al. Comparative transcriptome analysis of adipose tissues reveals that ECM receptor interaction is involved in the depot-specific adipogenesis in cattle. **PLoS One**, San Francisco, v. 8, n. 6, p. e66267, June, 2013. DOI: 10.1371/journal.pone.0066267.

LI, R. W.; RINALDI, M.; CAPUCO, A. V. Characterization of the abomasal transcriptome for mechanisms of resistance to gastrointestinal nematodes in cattle. **Veterinary Research**, Les Ulis, v. 30, n. 42, p. 114, Nov. 2011. DOI: 10.1186/1297-9716-42-114.

LING, Y. H.; XIANG, H.; LI, Y. S.; LIU, Y.; ZHANG, Y. H.; ZHANG, Z. J.; DING, J. P.; ZHANG, X. R. Exploring differentially expressed genes in the ovaries of uniparous and multiparous goats using the RNA-Seq (Quantification) method. **Gene**, Amsterdam, v. 550, n. 1, p. 148-153, Oct. 2014. DOI: 10.1016/j.gene.2014.08.008.

MA, L.-J.; SHEA, T.; YOUNG, S.; ZENG, Q.; CORBY, H. C. Genome sequence of *Fusarium oxysporum* f. sp. *Melonis* strain NRRL 26406, a fungus causing wilt disease on melon. **Genome Announcements**, Washington, D.C., v. 2, n. 4, e00730-14, July 2014. DOI:10.1128/genomeA.00730-14.

MARTIN, J. A.; WANG, Z. Next-generation transcriptome assembly. **Nature Review Genetics**, London, v. 12, p. 671-82, Oct. 2011. DOI:10.1038/nrg3068.

MARTINS, E. N. Avaliação Genética: dos dados às DEP's. In: ROSA, A. N.; MARTINS, E. N.; MENEZES, G. R. O.; SILVA, L. O. C. (Ed.). **Melhoramento Genético Aplicado em Gado de Corte**: Programa Geneplus-Embrapa. Brasília, DF: Embrapa; Campo Grande, MS: Embrapa Gado de Corte, 2013. 241 p. il.

MCCABE, M.; WATERS, S.; MORRIS, D.; KENNY, D.; LYNN, CREEVEY, C. RNA-seq analysis of differential gene expression in liver from lactating dairy cows divergent in negative energy balance. **BMC Genomics**, London, v. 13, n. 193, May 2012. DOI:10.1186/1471-2164-13-193.

METZKER, M. L. Sequencing technologies - the next generation. **Nature Review Genetics**, London, v. 11, n. 1, p. 31-46, Jan. 2010. DOI: 10.1038/nrg2626.

MEUWISSEN, T. H. E.; HAYES, B. J.; GOODDARD, M. E. Prediction of total genetic value using genome-wide dense marker maps. **Genetics**, London, v. 157, n. 4, p. 1819-1829, Apr. 2001.

MOKRY, F. B.; HIGA, R. H.; MUDADU, M. A.; LIMA, A. O.; MEIRELLES, S. L. C.; SILVA, M. V. G. B.; CARDOSO, F. F.; OLIVEIRA, M. M.; URBINATI, I.; NICIURA, S. C. M.; TULLIO, R. R.; ALENCAR, M. M.; REGITANO, L. C. A. Genome-wide association study for backfat thickness in Canchim beef cattle using Random Forest approach. **BMC Genetics**, London, v. 14, n. 47, 2013. DOI:10.1186/1471-2156-14-47.

MOORE, J. H.; ASSELBERGS, F. W.; WILLIAMS, S. M. Bioinformatics challenges for genome-wide association studies. **Bioinformatics**, Oxford, v. 26, n. 4, p. 445-455, Jan. 2010. DOI: 10.1093/bioinformatics/btp713.

MORAES, G. J.; FLECHTMANN, C. H. W. **Manual de acarologia**: acarologia básica e ácaros de plantas cultivadas no Brasil. Ribeirão Preto: Holos, 2008. 308 p.

MYERS, G. S. et al. Genome sequence and identification of candidate vaccine antigens from the animal pathogen *Dichelobacter nodosus*. **Nature Biotechnology**, v. 25, n. 5, p. 569-575, Apr. 29, 2007. DOI:10.1038/nbt1302.

MYERS, G. S. et al. Skewed genomic variability in strains of the toxigenic bacterial pathogen, *Clostridium perfringens*. **Genome Research**, v. 16, n. 8, p. 1031-1040, Aug. 2006. DOI: 10.1101/gr.5238106.

NAPOLIS, C. **Genhol**: desenvolvimento e integração de novos processos metodológicos no sistema de avaliação genética de bovinos leiteiros e sua filiação ao Interbull. Brasília, DF, Embrapa, 2012. Projeto SEG: 02.12.02.001.00.00.

NISHIYAMA JUNIOR, M. Y.; FERREIRA, S. S.; TANG, P.-Z.; BECKER, S.; SOUZA, G. M. PÖRTNER-TALIANA, A. Full-length enriched cDNA libraries and ORFeome analysis of sugarcane hybrid and ancestor genotypes. **PLoS One**, San Francisco, v. 9, n. 9, e107351, Sept. 2014. DOI: 10.1371/journal.pone.0107351.

OVIDO-RONDÓN, E. O. Molecular methods to evaluate effects of feed additives and nutrients in poultry gut microflora. **Revista Brasileira de Zootecnia**, Viçosa, MG, v. 38, p. 209-225, Oct. 2009. Número especial. DOI: http://dx.doi.org/10.1590/S1516-5982009001300022.

PATEL, A. K.; BHATT, V. D.; TRIPATHI, A. K.; SAJNANI, M. R.; JAKHESARA, S. J.; KORINGA, P. G.; JOSHI, C. G. Identification of novel splice variants in horn cancer by RNA-Seq analysis in Zebu cattle. **Genomics**, London, v. 101, n. 1, p. 57-63, Jan. 2013. DOI: 10.1016/j.ygeno.2012.10.001.

PAULSEN, I. T. et al. The *Brucella suis* genome reveals fundamental similarities between animal and plant pathogens and symbionts. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, US, v. 99, n. 20, Oct. 2002.

PERUMBAKKAM, S.; MUIR, W. M.; BLACK-PYRKOSZ, A.; OKIMOTO, R.; CHENG, H. H. Comparison and contrast of genes and biological pathways responding to Marek's disease virus infection using allele-specific expression and differential expression in broiler and layer chickens. **BMC Genomics**, London, v. 14, n. 64, Jan. 2013. DOI:10.1186/1471-2164-14-64.

PIERETTI, I. et al. The complete genome sequence of *Xanthomonas albilineans* provides new insights into the reductive genome evolution of the xylem-limited Xanthomonadaceae. **BMC Genomics**, London, v. 10, n. 616, 2009. DOI: 10.1186/1471-2164-10-616.

POP, M.; SALZBERG, S. L. Bioinformatics challenges of new sequencing technology. **Trends in Genetics**, Cambridge, v. 24, n. 3, p. 142-149, Mar. 2008. DOI: 10.1016/j.tig.2007.12.006.

RAMAYO-CALDAS, Y.; MACH, N.; ESTEVE-CONDINA, A.; COROMINAS, J.; CASTELLÓ, A.; BALLESTER, M.; ESTELLÉ, J.; IBÁÑEZ-ESCRICHE, N.; FERNÁNDEZ, A. I.; PÉREZ-INCISO, M.; FOLCH, J. M. Liver transcriptome profile in pigs with extreme phenotypes of intramuscular fatty acid composition. **BMC Genomics**, London, v. 13, n. 547, Oct. 2012. DOI:10.1186/1471-2164-13-547.

ROSA, A. N.; MENEZES, G. R. O.; EGITO, A. A. Recursos Genéticos e Estratégias de Melhoramento. In: ROSA, A. N.; MARTINS, E. N.; MENEZES, G. R. O.; SILVA, L. O. C. (Ed.). **Melhoramento Genético Aplicado em Gado de Corte**: Programa Geneplus-Embrapa. Brasília, DF: Embrapa; Campo Grande, MS: Embrapa Gado de Corte, 2013. 241 p. il.

ROSS, E. M.; MOATE, P. J.; MARETT, L.; COCKS, B. G.; HAYES, B. J. Investigating the effect of two methane-mitigating diets on the rumen microbiome using massively parallel sequencing. **Journal of Dairy Science**, Champaign, v. 96, n. 9, p. 6030-6046, Sept. 2013.

SALANOUBAT, M. et al. Genome sequence of the plant pathogen *Ralstonia solanacearum*. **Nature**, London, v. 415, n. 6871, p. 497-502, January, 2002.

SALZBERG, S. L. et al. Genome sequence and rapid evolution of the rice pathogen *Xanthomonas oryzae* pv. *Oryzae* PXO99A. **BMC Genomics**, London, v. 9, p. 204, May 2008. DOI: 10.1186/1471-2164-9-204.

SCHLOSS, P. D.; HANDELSMAN, J. Toward a census of bacteria in soil. **PLoS Computational Biology**, San Francisco, v. 2, n. 7, e92, July 2006. DOI: 10.1126/science.1178534.

SCHNABLE, P. S. et al. The B73 maize genome: complexity, diversity, and dynamics. **Science**, New York, v. 326, n. 5956, p. 1112-1115, Nov. 2009. DOI:10.1126/science.1178534.

SHARMA, A.; LI, X.; LIM, Y. P. Comparative genomics of Brassicaceae crops. **Breeding Science**, New York, v. 64, n. 13, p. 3-13, May 2014. DOI:10.1270/jsbbs.64.3.

SHENG, X.; NI, H.; LIU, Y.; LI, J.; ZHANG, L.; GUO, Y. RNA-seq analysis of bovine intramuscular, subcutaneous and perirenal adipose tissues. **Molecular Biology Reports**, Dordrecht, v. 41, n. 3, Mar. 2014. DOI: 10.1007/s11033-013-3010-8.

SHI, W. et al. Methane yield phenotypes linked to differential gene expression in the sheep rumen microbiome. **Genome Research**, Woodbury, v. 24, p. 1517-1525, June 2014. DOI: 0.1101/gr.168245.113.

SILVA, A. C. R. et al. Comparison of the genomes of two *Xanthomonas pathogens* with differing host specificities. **Nature**, London, v. 417, n. 6887, p. 459-463, May 2002.

SILVA, L. O. C. **MaxiBife** - Desenvolvimento e integração de conhecimentos de genética molecular e quantitativa visando maximizar os ganhos genéticos para produção sustentável de carne bovina: arranjo SEG. Brasília, DF: Embrapa, 2013.

SILVA, M. V. B. S. **Genomil - Seleção Genômica em Raças Bovinas Leiteiras no Brasil**. Brasília, DF: Embrapa, 2010. Projeto SEG: 02.09.07.008.00.00.

SOARES, S. C. et al. The pan-genome of the animal pathogen *Corynebacterium pseudotuberculosis* reveals differences in genome plasticity between the biovar ovis and equi strains. **PLoS One**, San Francisco, v. 8, n. 1, e53818, Jan. 2013. DOI:10.1371/journal.pone.0053818.

SWARBRECK, D.; WILKS, C.; LAMESCH, P.; BERARDIN, T. Z.; GARCIA-HERNANDEZ, M.; FOERSTER, H.; LI, D.; MEYER, T.; MULLER, R.; PLOETZ, L.; RADENBAUGH, A.; SINGH, S.; SWING, V.; TISSIER, C.; ZHANG, P.; HUALA, E. The Arabidopsis Information Resource (TAIR): gene structure and function annotation. **Nucleic Acids Research**, London, v. 36, p. 1009-1014, Jan. 2008.

THE XYLELLA CONSORTIUM OF THE ORGANIZATION FOR NUCLEOTIDE SEQUENCING AND ANALYSIS. The genome sequence of the plant pathogen *Xylella fastidiosa*. The *Xylella fastidiosa* consortium of the Organization for Nucleotide Sequencing and Analysis. **Nature**, London, v. 13, n. 406, n. 6792, p. 151-159, July 2000.

TIZIOTO, P.C.; DECKER, J.E.; TAYLOR, J.F.; SHNABEL, R.D.; MUDADU, M.M.; SILVA, F.L.; MOURÃO, G.B.; COUTINHO, L.L.; THOLON, P.; SONSTEGARD, T.S.; ROSA, A.N.; ALENCAR, M.M.; TULLIO, R.R.; MEDEIROS, S.R.; NASSU, R.T.; FEIJÓ, G.L.D.; SILVA, L.O.C.; TORRES, R.A.; SIQUEIRA, F.; HIGA, R.H.; REGITANO, L.C.A. A genome scan for meat quality traits in Nelore beef cattle. **Physiological Genomics**, Bethesda, v. 45, n. 21, p. 1012-1020, 2013.

TRIBOLIUM GENOME SEQUENCING CONSORTIUM The genome of the model beetle and pest *Tribolium castaneum*. **Nature**, London, v. 452, p. 949-955, Apr. 24, 2008. DOI:10.1038/nature06784.

WANG, Z.; GERSTEIN, M.; SNYDER, M. RNA-Seq: a revolutionary tool for transcriptomics. **Nature Review Genetics**, London, v. 10, n. 1, p. 57-63, Jan. 2009. DOI: 10.1038/nrg2484.

ZIEGLER, A.; KÖNIG, I. R.; THOMPSON, J. R. Biostatistical aspects of genome-wide association studies. **Biometrical Journal**, Berlin, v. 50, n. 1, p. 8-28, Feb. 2008. DOI: 10.1002/bimj.200710398.

XU, Q. et al. The draft genome of sweet orange (*Citrus sinensis*). **Nature Genetics**, New York, v. 45, p. 59-66, Nov. 2013. DOI:10.1038/ng.2472.



## **Fenotipagem de plantas em larga escala: um novo campo de aplicação para a visão computacional na agricultura**

Thiago Teixeira Santos  
Juliana Erika de Carvalho Teixeira Yassitepe

### **1 Introdução**

Tecnologias visando aumentar a acurácia e rapidez das mensurações fenotípicas para melhor caracterização e entendimento das funções biológicas dos organismos em resposta a diferentes condições ambientais vêm sendo alvo de intensas pesquisas nos últimos anos. Este esforço recente busca diminuir o “gargalo de fenotipagem” (*phenotyping bottleneck*), como vem sendo chamada pela comunidade científica a lacuna entre a quantidade e qualidade de dados genômicos e fenotípicos disponíveis. Técnicas correntes de sequenciamento tornaram a caracterização de genomas um processo eficiente e barato, gerando uma enorme massa de dados para pesquisa. Porém, um avanço similar ainda não ocorreu na caracterização em larga escala de fenótipos, e isso retarda o avanço no entendimento das associações genótipo versus fenótipo.

Para alcançar a mesma velocidade observada atualmente na caracterização genômica, é necessário incorporar à fenotipagem avanços nas áreas de computação, robótica, visão computacional e análise de imagens (COBB et al., 2013; FIORANI; SCHURR, 2013; FURBANK; TESTER, 2011). Com o foco neste objetivo, uma nova área do conhecimento científico está emergindo e vem sendo chamada de “fenômica”, em alusão às ciências “ômicas” que estudam os genes, transcritos, proteínas e metabolitos de um organismo. Esta nova ciência, fenômica, se concentra na aquisição de dados fenótipos em larga escala e dimensão, em um curto espaço de tempo, oferecendo com isso uma oportunidade única de visualizar, em vários níveis hierárquicos, o fenótipo de um organismo.

Visão computacional e análise de imagens são hoje um dos principais componentes em qualquer plataforma moderna de fenotipagem (FIORANI; SCHURR, 2013; FURBANK; TESTER, 2011). Técnicas baseadas em imagem são capazes de capturar grandes volumes de dados quantitativos sobre o fenótipo, através de metodologias que não são invasivas ou destrutivas. A fenotipagem de plantas é uma aplicação desafiadora à visão computacional, por apresentar problemas nas áreas de modelagem tridimensional (3D), segmentação e classificação, metrologia por imagem e processamento de nuvens de pontos. Um exemplo da aproximação entre as áreas pôde ser visto em 2014, quando a tradicional *Conferência Europeia em Visão Computacional* - European Conference on Computer Vision (ECCV) incluiu um workshop dedicado a problemas em fenotipagem de plantas (Computer Vision Problems in Plant Phenotyping) (CVPPP).

O presente capítulo apresenta uma visão geral dos avanços recentes na fenotipagem em larga escala (Seção 2) e como a visão computacional surge como ferramenta para a caracterização fenotípica não-destrutiva da parte aérea de plantas (Seção 3). O capítulo se encerra (Seção 4) apresentando cenários futuros de pesquisa nessa área.

## 2 Tecnologias e plataformas de fenotipagem de plantas em larga escala

As tecnologias que estão sendo utilizadas nesta nova era de fenotipagem de plantas empregam técnicas de espectroscopia, termografia, tomografia, fluorescência, discriminação por isótopos e imagens digitais. Como muitas dessas técnicas de aquisição de dados são não-destrutivas, é possível coletar dados em diversos momentos ao longo do crescimento e desenvolvimento da planta. Como resultado, características que nunca foram mensuradas antes ou somente em situações específicas estão começando a ser medidas com maior frequência, como fluorescência da clorofila, atividade de estômatos, temperatura foliar e arquitetura de raízes e folhas. Além disso, processos dinâmicos tais como crescimento, desenvolvimento e resposta a estresse, os quais na maioria das vezes vinham sendo caracterizados em um único e determinado momento, podem agora ser descritos ao longo do tempo e espaço (FIORANI; SCHURR, 2013).

Tecnologias baseadas em espectroscopia de absorção estão entre as novas tecnologias de fenotipagem em larga escala mais utilizadas nos laboratórios. Dentre elas se destaca a espectroscopia de infravermelho próximo (NIRS na sigla em inglês), comumente empregada para identificar a composição química de uma amostra. A identificação de um composto por esta técnica é possível devido ao fato de as ligações químicas das substâncias possuírem frequências de vibração específicas, as quais correspondem a níveis de energia da molécula. Amostras de referência são usadas para construção de curvas de calibração validadas, que serão utilizadas na inferência sobre a composição química de amostras desconhecidas baseadas em seus espectros NIRS. Curvas de calibração acuradas são essenciais para o sucesso da aplicação de NIRS e a obtenção dessas curvas requer um trabalho prévio envolvendo a análise de uma ampla diversidade fenotípica para as características que se deseja medir. Exemplos nos quais NIRS vem permitindo um rápido *screening* de germoplasmas incluem a caracterização da composição química de grãos (JIANG et al., 2007; OSBORNE, 2006; TALLADA et al., 2009), detecção de infecção por fungos e toxinas derivadas de fungos em grãos (BERARDO et al., 2005; PEARSON; WICKLOW, 2006) e a determinação da composição química de raízes, caules e tubérculos (LEBOT et al., 2009; MURRAY et al., 2008a, 2008b; WALSH et al., 2000).

Outra tecnologia que está contribuindo extensivamente para fenômica em ambientes controlados é a captura e análise de imagens. Imagens podem ser adquiridas usando técnicas de refletância clássicas tais como fotografia de luz e microscopia ou por técnicas mais modernas como termografia, fluorescência, tomografia, ressonância magnética e ressonância nuclear (BERGER et al., 2010; HOULE et al., 2010; MONTES et al., 2007b). Tecnologias baseadas em imagens podem ser usadas para medir características bioquímicas, fisiológicas e morfológicas tais como atividade fotossintética, atividade estomática, composição química foliar e arquitetura de folhas e raízes (CLARK et al., 2011; GENTY; MEYER, 1995; LENK et al., 2007; MERLOT et al., 2002; ROMANO et al., 2011).

Além de serem aplicadas isoladamente, as técnicas baseadas em espectroscopia e imagens podem ser combinadas para permitir uma análise mais compreensível das características mensuradas (CHAEERLE et al., 2007; GOETZ et al., 1985). Neste contexto, uma das tecnologias mais promissoras, imagens hiperespectrais, reúne análises de imagens e espectros ao longo de uma ampla extensão de comprimentos de ondas em duas ou três dimensões do espaço (GOWEN et al., 2007). Imagens hiperespectrais vêm sendo utilizadas para detectar contaminação biológica em grãos (DEL FIORE et al., 2010; WILLIAMS et al., 2009) e na análise de crescimento e desenvolvimento de plantas (WALTER; SCHURR, 2005). Similarmente, imagens com fluorescência da clorofila em um contexto espacial também vêm sendo usadas para investigar resposta a estresses abióticos (KONISHI et al., 2009) e bióticos (ROLFE; SCHOLLES, 2010).

Apesar das interessantes possibilidades que as tecnologias de fenotipagem utilizadas em laboratório ou ambientes controlados podem oferecer, a maioria das características de importância agrônoma precisam ser avaliadas em campo. Felizmente, várias destas recentes tecnologias podem ser adaptadas para caracterização fenotípica no campo (HATFIELD et al., 2008). Vários tipos de sensores vêm sendo desenvolvidos para serem usados no campo. Ao nível do solo, sensores estão disponíveis em equipamentos manuais tais como espectro radiômetros, medidores de clorofila, e termômetros infravermelhos (MONTES et al., 2007a, 2007b; WELLE et al., 2003, 2005) e acoplados em tratores, picadores, colheitadeiras e mesmo em pivôs de irrigação, o que permite o ganho de escala e possibilita a análise de milhares de amostras em um curto espaço de tempo. Sensores manuais portáteis estão disponíveis para detectar deficiência de nutrientes e prever conteúdo total de nitrogênio, biomassa, produtividade de grãos e variações em status fisiológicos (BABAR et al., 2006a, 2006b; OSBORNE et al., 2002a, 2002b; POSS et al., 2006). A aplicação de NIRS acoplados a tratores e picadores foi demonstrada por Montes et al. (2006, 2007b) e Welle et al. (2003, 2005) durante avaliação da qualidade, teor de matéria seca e produtividade de silagem de milho. Algumas destas características puderam ser medidas com precisão superior às dos procedimentos clássicos de NIRS em laboratório. Características de crescimento também têm sido mensuradas utilizando tecnologia de sensoriamento remoto no campo (PETERS; EVETT, 2007; WINTERHALTER et al., 2011a, 2011b). Sensores e câmeras podem também ser acoplados em torres, dirigíveis ou drones, para nomear alguns, o que permite a aquisição de dados de uma área extensa e favorece a análise de características fenotípicas dinâmicas em milhares de plantas (FURBANK; TESTER, 2011; JONES et al., 2009).

Mesmo utilizando estas novas tecnologias, muitos dos equipamentos disponíveis para fenotipagem em larga escala ainda requerem o manuseio do operador. Este processo é muitas vezes exaustivo, tedioso e custoso. Para minimizar o viés individual obtido pelas mensurações manuais e garantir escala, acurácia e precisão nas medições, a mecanização e a automação de alguns processos surgem como alternativas, principalmente envolvendo o cultivo das plantas e o processamento de imagens e dados gerados (COBB et al., 2013; FURBANK; TESTER, 2011).

Empregando os últimos avanços tecnológicos em fenotipagem bem como tecnologias robóticas e analíticas, sofisticadas plataformas de fenotipagem em configurações automatizadas e semiautomáticas vem sendo desenvolvidas, tais como: Phenopsis (GRANIER et al., 2006), Growscreen (BISKUP et al., 2009; NAGEL et al., 2012; RASCHER et al., 2011) e TraitMill™ (REUZEAU et al., 2006), além de soluções comerciais de empresas como LemnaTec, GmbH<sup>1</sup>

<sup>1</sup> Disponível em: <<http://www.lemnatec.com>>.

e Qubit Systems<sup>2</sup>. Desenhadas para câmaras de ambiente controlado, essas plataformas ajudam a automatizar várias etapas do processo de cultivo e fenotipagem, como preparo do substrato, enchimento de potes, plantio, fertilização, irrigação, aquisição e análise de dados fenotípicos (HENTEN, 2006). Sistemas de códigos de barra e identificação por radiofrequência podem também ser incluídos na plataforma, permitindo com isso identificação e rastreamento automático de amostras. Computadores de alta capacidade, bancos de dados e software também estão sendo desenvolvidos e utilizados na aquisição, armazenamento e análise de quantidades massivas de dados que tais plataformas são capazes de gerar (HARTMANN et al., 2011; LI et al., 2011; VANKADAVATH et al., 2009).

Centros de fenômica utilizando essas plataformas surgiram nos últimos anos, principalmente na Austrália e Europa. Exemplos incluem a Australian Plant Phenomics Facility (Adelaide e Canberra, Austrália), o Jülich Phenomics Centre (Jülich, Alemanha), o Leibniz Institute of Plant Genetics and Crop Plant Research (Gatersleben, Alemanha), a PhenoArch (Montpellier, França) e a Crop Design (Zwijnaarde, Bélgica). No entanto, muitas destas plataformas foram desenvolvidas para aplicações com plantas modelo ou culturas de grande importância econômica. Plataformas genéricas e soluções modulares e flexíveis que permitam a avaliação simultânea do fenótipo de múltiplas espécies e capazes de acomodar necessidades experimentais diferentes não estão ainda disponíveis (COBB et al., 2013; FIORANI; SCHURR, 2013). Buscando unir esforços no sentido de padronizar as metodologias de fenotipagem empregadas, desenvolver novos sensores, instrumentos e estruturas para acessar, gerenciar e analisar a informação tecnológica gerada, muitos destes centros se organizam em redes, como a Rede Europeia para Fenotipagem de Plantas (European Plant Phenotyping Network - EPPN)<sup>3</sup>.

## 2.1 Representação de fenótipo

Enquanto o genoma de um organismo pode ser caracterizado através de suas sequências de bases, é virtualmente impossível caracterizar totalmente seu *fenoma*. Como expresso por Cobb et al. (2013): “o fenoma de um organismo é dinâmico e condicional, o que representa um conjunto complexo de respostas a um conjunto multi-dimensional de sinais endógenos e exógenos que são integradas ao longo do desenvolvimento de um indivíduo”.

Contudo, se nos restringirmos aos aspectos macroscópicos da parte aérea da planta, um modelo 3D pode atuar como uma representação temporal concisa e versátil do estado do indivíduo, permitindo que diferentes medidas quantitativas sejam computadas, como área e angulação foliar, porcentagem de área afetada (em patologias), curvatura e enrolamento foliar ou altura. Outras medições podem ser computadas *a posteriori*, destinadas a responder outros questionamentos, formulados posteriormente dentro do mesmo experimento ou de um novo. Dado o caráter não-invasivo e não-destrutivo dos procedimentos de aquisição de imagens, o comportamento do espécime ao longo do tempo pode ser avaliado comparando os vários modelos tridimensionais construídos ao longo do tempo.

O procedimento de criação desses modelos tridimensionais é frequentemente chamado na literatura em fisiologia vegetal de digitalização de plantas (*plant digitizing*) e diversas metodologias para tal fim foram propostas ao longo dos anos.

<sup>2</sup> Disponível em: <<http://qubitsystems.com>>.

<sup>3</sup> Disponível em: <<http://www.plant-phenotyping-network.eu>>.

## 3 Visão computacional e digitalização 3D de plantas

Os primeiros esforços na digitalização 3D de plantas empregavam dispositivos mecânicos (LANG, 1973), digitalizadores sônicos (SINOQUET et al., 1991) ou rastreadores magnéticos (RAKOCEVIC et al., 2000). Tais métodos de contato foram importantes na criação de modelos funcionais e estruturais do desenvolvimento de plantas (GODIN et al., 1999; RAKOCEVIC et al., 2000) mas são incapazes de atingir o desempenho necessário à fenotipagem em alta escala. As iniciativas atuais em fenotipagem com reconstrução tridimensional se baseiam em métodos sem contato como varredura laser, câmeras ToF (*time-of-flight*) e visão estéreo.

### 3.1 Digitalização de plantas por métodos sem contato

Provavelmente o primeiro trabalho a utilizar varredura por laser na caracterização da estrutura tridimensional de plantas em fenotipagem de larga escala foi apresentado por Kaminuma et al. (2004) em seus experimentos com *Arabidopsis thaliana*. Folhas e pecíolos foram representados em 3D por malhas poligonais. Essas malhas foram utilizadas na determinação quantitativa de dois atributos de interesse: a direção da lâmina e a epinastia da folha, visando a caracterização de dois ecótipos diferentes de *Arabidopsis*. A configuração da montagem foi capaz de produzir uma boa amostragem das superfícies: a distância de aquisição e a amostragem permitiram uma resolução de 0,045 milímetros por pixel, produzindo uma nuvem densa de pontos 3D. Contudo, plantas diferentes apresentando dimensões maiores, ou até mesmo espécimes de *Arabidopsis* em um estágio mais avançado de desenvolvimento, produziam conjuntos de pontos mais esparsos.

Equipamentos de varredura laser, como o LiDAR, são caros e difíceis de serem movimentados. Uma alternativa mais barata e flexível é o uso de visão estéreo com câmeras simples. Diversos métodos na literatura utilizam visão estéreo para digitalização sem contato de plantas, a partir do conceito de triangularização em geometria projetiva. Considere duas imagens tomadas por uma câmera em duas posições diferentes (ou, alternativamente, duas câmeras posicionadas em locais diferentes). Para cada imagem, um pixel define um raio no espaço 3D - todo ponto 3D nesse raio será projetado no mesmo pixel. Se as posições e os parâmetros internos da câmera forem conhecidos previamente, um par de pontos correspondentes em diferentes imagens irá definir um par de raios e a localização do ponto em 3D estará na intersecção dos raios (o leitor interessado em um tratamento aprofundado do problema da triangularização pode recorrer a Hartley e Zisserman, 2004).

O trabalho de Ivanov et al. (1995) é possivelmente o primeiro trabalho na literatura a utilizar visão estéreo na reconstrução 3D da superfície de um cultivar para medição e análise. Os autores estimaram a posição e a orientação de folhas de milho (*Zea mays* L.) e a distribuição da área foliar. Seu sistema utilizava um par de câmeras instalado a 8,5 metros do solo em um campo de milho apresentando altura de 2,5 metros. Infelizmente, as dificuldades impostas pelo equipamento disponível na época (a fotografia digital ainda não era largamente disseminada) comprometeram os experimentos. A segmentação das folhas e a determinação das correspondências entre imagens foram realizadas manualmente, utilizando-se ampliações fotográficas. Apesar das limitações, este trabalho foi o antecessor de sistemas mais recentes, capazes de empregar avanços em computação, imageamento digital e visão computacional. Biskup et al. (2007) desenvolveram um sistema de visão estéreo baseado em duas câmeras digitais para criar modelos tridimensionais para folhas de soja, com o objetivo de analisar o ângulo de inclinação e o movimento das folhas

ao longo do dia. Dada a importância do movimento para o experimento, o sistema era capaz de processar até três imagens por segundo, recuperando assim a informação 3D necessária para o cômputo da inclinação.

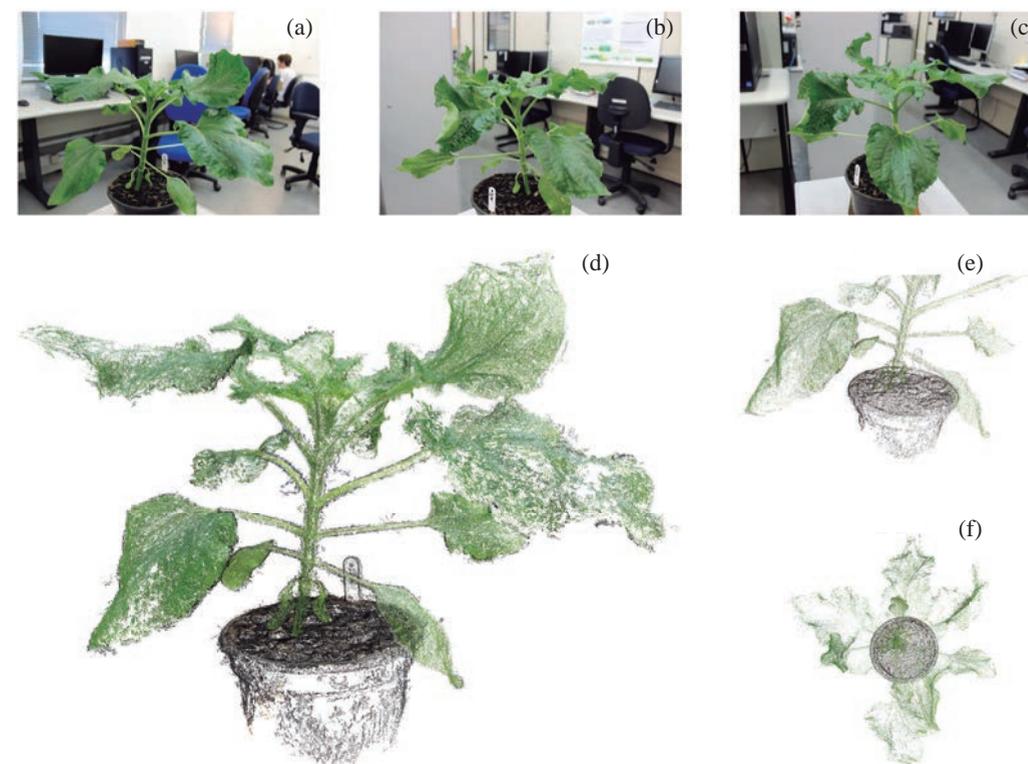
Dados de profundidade oriundos de câmeras ToF podem ser combinados a dados RGB de câmeras convencionais para produzir reconstruções 3D de folhas apesar da baixa resolução dos dispositivos ToF atuais. Song et al. (2011) combinaram dados de um par de câmeras RGB (480 x 1280 pixels) e de uma câmera ToF (64 x 48 pixels) utilizando a técnica de minimização por cortes em grafos (BOYKOV et al., 2001). O par de câmeras provê uma resolução mais alta enquanto a informação ToF auxilia na resolução de falhas no cômputo da profundidade em regiões sem textura, nas quais a determinação de correspondência entre imagens é mais propensa a falhas. Alenyà et al. (2011) empregaram uma abordagem mais simples: os pontos 3D produzidos por uma câmera ToF foram transformados para as coordenadas de referência de uma câmera RGB e, em seguida, projetados para o plano de imagem da câmera - pontos nesse plano que não apresentassem uma contraparte 3D da câmera ToF foram descartados. Apesar do método simples, os autores foram capazes de produzir bons resultados pois: a) a câmera ToF empregada apresentava uma resolução maior (200 x 200 pixels) e b) um braço robô foi utilizado na movimentação das câmeras, permitindo a aquisição de mais dados, uma abordagem de sensoriamento ativo (*sensing-for-action*) proposta pelos autores.

No arcabouço de triangularização, uma câmera do par estéreo pode ser substituída por um projetor. O mesmo princípio de intersecção de raios pode ser aplicado se a posição do projetor e seus parâmetros internos forem conhecidos. Bellasio et al. (2012) utilizaram um par câmera-projetor e um método de luz codificada (*coded-light*) para determinar a correspondência entre pixels, recuperando a superfície 3D das folhas de pimenteiros. Chéné et al. (2012) empregaram uma câmera RGB-D (um Microsoft Kinect®) para segmentar folhas e estimar sua orientação e inclinação. Tais câmeras RGB-D consistem em um par câmera-projetor e um sistema de triangularização embarcados em um único dispositivo, também utilizando luz codificada para estabelecer correspondências entre pixels (FREEDMAN et al., 2013).

Recentemente, visão estéreo múltipla *multiple-view stereo* (MVS) passou a ser empregada em digitalização de plantas como uma tentativa de solucionar problemas de oclusão encontrados na reconstrução 3D. Paproki et al. (2012) empregaram o software 3D SOM. (BAUMBERG et al., 2005) para criar modelos tridimensionais de espécimes de algodão (*Gossypium hirsutum*) e então estimar a altura de suas hastes, a largura e o comprimento das folhas. Os espécimes foram colocados em uma mesa giratória contendo um padrão de calibração utilizado pelo 3D SOM na estimação da posição relativa da câmera em cada quadro.

Santos e Oliveira (2012) empregaram o arcabouço de estruturação por movimento, *structure-from-motion* (SfM) (HARTLEY; ZISSERMAN, 2004), para recuperar o posicionamento da câmera em cada imagem. SfM estende o arcabouço de triangularização incorporando uma etapa que simultaneamente busca pelas melhores estimativas para as posições da câmera e as posições dos pontos 3D na superfície do objeto. Tal processo, *bundle-adjustment*, é um procedimento de maximização da esperança que avalia múltiplos ângulos ao mesmo tempo, minimizando o erro de reprojeção para cada imagem. Ao invés de empregar um padrão de calibração, os autores utilizaram a detecção de características visuais locais nas imagens e suas correspondências intra-imagens, como proposto por Snavely et al. (2008) no contexto de reconstrução 3D de arquitetura

de cidades, para recuperar os parâmetros da câmera (posição, orientação e parâmetros internos). Os dados da câmera e o modelo 3D inicial produzidos foram utilizados como entrada para um algoritmo de MVS, o PMVS proposto por Furukawa e Ponce (2010), o qual foi capaz de produzir nuvens de pontos 3D densas, amostrando a superfície de folhas e entrenós em experimentos com manjerição, *Ixora* (SANTOS; OLIVEIRA, 2012), hortelã (SANTOS; UEDA, 2013) e girassol (Figura 1).



**Figura 1.** Resultados para girassol produzidos a partir de 302 imagens com 3888 x 2592 pixels de resolução (a-c) pela metodologia proposta. O modelo 3D é composto por 829.722 pontos 3D (d-f).

Sirault et al. (2013) também reportam resultados em reconstrução 3D a partir do sistema PMVS de Furukawa e Ponce (2010). Os autores utilizam um conjunto de câmeras fixas e calibradas, imageando plantas em vasos posicionados em uma mesa giratória de alta precisão (2,5 milhões de passos por revolução). Sua plataforma de digitalização, PlantScan™, é uma câmara equipada com três câmeras RGB CCD, uma câmera NIR e dois varredores laser LiDAR, além de dois sensores térmicos. Seu sistema de reconstrução estéreo emprega os métodos de visão estéreo múltipla (PMVS) e colorização de voxels, *voxel colouring* (LEUNG et al., 2012), que são combinados aos dados LiDAR com o uso de algoritmos de registro RANdom SAMple Consensus (Ransac) e Iterative Closest Point (ICP). Os autores reportam que sua plataforma é capaz de digitalizar plantas de poucos centímetros até dois metros em altura e até um metro de diâmetro, permitindo assim a fenotipagem da parte aérea de diversas espécies diferentes.

Santos et al. (2014) estenderam sua metodologia inicial (SANTOS; OLIVEIRA, 2012) incorporando uma aplicação que seleciona automaticamente quadros de uma sequência de vídeo em

tempo real. O aplicativo computa características visuais locais e suas correspondências entre quadros utilizando uma unidade de processamento gráfico (GPU), selecionando um conjunto de quadros capaz de produzir boas reconstruções tridimensionais (o algoritmo garante um número mínimo de correspondências entre quadros que possibilite ao método de SfM recuperar a posição da câmera). O usuário pode mover livremente a câmera ao redor da planta, o que possibilita a resolução de oclusões.

Metodologias que empregam uma única câmera móvel apresentam duas vantagens. A flexibilidade na definição do caminho percorrido pela câmera permite um melhor tratamento de oclusão e variações no tamanho e na morfologia das plantas. O trabalho de Alenyà et al. (2011) ilustra tal característica, mostrando como o planejamento apropriado de movimentação para um braço robô é capaz de posicionar a câmera em locais onde oclusões podem ser facilmente resolvidas. A outra vantagem diz respeito ao baixo custo de tais soluções. Um experimento que requer fenotipagem de alto-desempenho pode exigir a digitalização de centenas de indivíduos por dia. Metodologias baratas de reconstrução permitem que diversos módulos de digitalização sejam montados para o experimento e empregados na paralelização do trabalho.

### 3.2 Digitalização de plantas no campo

Os métodos apresentados na seção anterior são dedicados a ambientes controlados como casas de vegetação e câmaras de crescimento. Esses ambientes permitem o controle de características ambientais e um alto nível de automação. Porém, pesquisas agrícolas geralmente requerem experimentos em campo, na tentativa de compreender como certos germoplasmas irão se comportar em condições reais de produção. Tais ambientes são mais complexos para a fenotipagem em alta escala e à reconstrução 3D de plantas devido à dificuldade de automação, à variabilidade nas condições de luz e aos problemas com movimentação devida ao vento.

Contudo, há avanços na reconstrução 3D de plantas em campo. Rovira-Más et al. (2005) acoplaram uma câmera estéreo a um helicóptero controlado por rádio. As imagens foram combinadas na produção de um mapa 3D de um campo de milho. Jay et al. (2014) acoplaram uma câmera a um eixo de translação móvel, produzindo um conjunto de imagens para linhas de plantio de beterraba em campo. Essas imagens foram utilizadas na reconstrução 3D do dossel das plantas de beterraba via SfM.

Kazmi et al. (2014) analisaram o uso de câmeras ToF na estimação de profundidade de folhas (distância entre o sensor e a superfície da folha) em ambientes de campo. Os autores mostram que tais câmeras são capazes de estimar de maneira acurada a profundidade em diversas condições de iluminação (ambiente ensolarado ou nublado). Folhas com superfícies homogêneas (sem textura), que são um problema para sistemas baseados em visão estéreo devido às dificuldades em estabelecer correspondências entre pares de imagens, não apresentam dificuldades para câmeras ToF, que usam iluminação ativa e não dependem do cômputo de correspondências intra-imagens.

### 3.3 Análise de nuvens de pontos

Os métodos apresentados anteriormente envolvendo LiDAR, câmeras ToF ou métodos de visão estéreo produzem dados 3D na forma de nuvens de pontos que amostram a superfície do objeto. Dependendo do processo de aquisição, do posicionamento dos sensores e dos algoritmos utilizados, tal amostragem pode variar em um mesmo modelo, sendo mais densa em certas partes e

mais esparsa em outras. Outros problemas podem ocorrer, como buracos causados por oclusões não resolvidas ou reflexão especular da luz. Considerando-se todos esses problemas, a nuvem de pontos precisa ser analisada conforme os seguintes passos:

- 1) A superfície da planta deve ser inferida a partir dos pontos para que medidas quantitativas como área foliar possam ser devidamente computadas (Figura 2);
- 2) O modelo deve ser segmentado em partes significativas para caracterização da estrutura da planta;
- 3) Cada segmento deve ser classificado em categorias significativas (folhas, entrenós, pecíolos, frutos, flores etc.) para que medições quantitativas sejam devidamente realizadas em cada estrutura diferente do organismo.

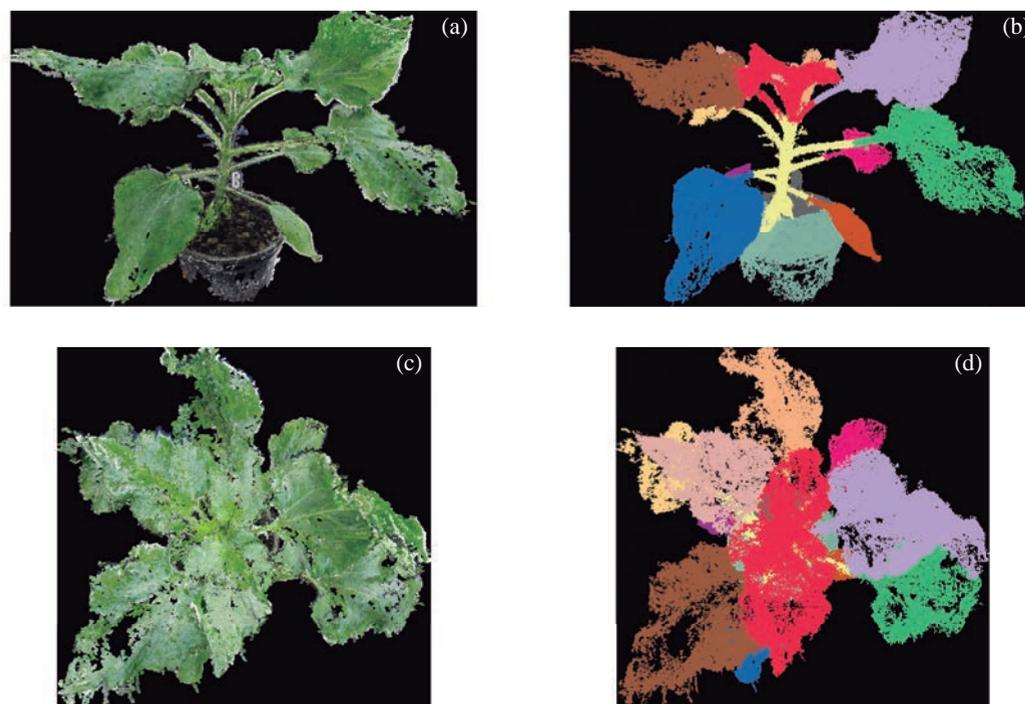


**Figura 2.** Inferindo superfícies a partir de nuvens de pontos. Nuvem de pontos produzida por SfM e MVS (a). Malha triangular produzida pelo algoritmo *ball pivoting* (b). Visualização da superfície suavizada (c).

Xu et al. (2007) produziram nuvens de pontos para árvores utilizando um dispositivo LiDAR. Pontos próximos foram conectados, estruturando a nuvem na forma de um grafo no qual cada ponto 3D corresponde a um vértice. Os grafos foram segmentados e simplificados na forma de esqueletos que representam a estrutura de troncos e galhos. Os esqueletos e as nuvens de pontos foram então empregados na estimação de parâmetros de modelos alométricos, responsáveis por definir a superfície das árvores na forma de uma malha 3D.

Santos e Ueda (2013) empregaram *clustering* de pontos para segmentar folhas em espécimes de hortelã (SANTOS; UEDA, 2013), girassol e soja (SANTOS et al. 2014), como ilustrado na Figura 3. Os segmentos obtidos são subconjuntos disjuntos da nuvem original e foram caracterizados através de medidas como comprimento e largura. Essas medidas permitiram a construção de classificadores capazes de identificar quais segmentos correspondiam às folhas da planta. Uma vez identificadas, as folhas tiveram suas superfícies caracterizadas através de *splines*, cujos parâmetros foram estimados a partir dos pontos 3D atribuídos a cada folha. Alenyà et al. (2011) procederam de forma similar: pontos foram agrupados através de algoritmos de *clustering* e um modelo quadrático foi empregado para representar superfícies curvas (folhas), sendo seus parâmetros estimados a partir dos pontos 3D através de um procedimento de otimização.

A ideia por trás dessas metodologias é que, uma vez identificada corretamente a estrutura em questão, um modelo apropriado pode ser selecionado para representá-la. Por exemplo, *splines* podem ser utilizadas para representar a superfície das folhas enquanto que cilindros generalizados são mais adequados para caracterizar estruturas como entrenós e pecíolos. Medidas de interesse, como área foliar, inclinação foliar, comprimento e largura, podem ser então estimadas a partir do modelo paramétrico (SANTOS; UEDA, 2013).



**Figura 3.** Segmentação do modelo. (a) Nuvem de pontos produzida pelo método. (b) Segmentação através de spectral clustering - cores diferentes indicam segmentos diferentes. (c) Vista superior do mesmo modelo. (d) Vista superior da segmentação.

## 4 Considerações finais

Novos equipamentos para imageamento 3D têm surgido nos últimos anos. Câmeras baseadas em métodos como ToF (KAZMI et al., 2014) ou luz estruturada (FREEDMAN et al., 2013) são hoje disponibilizadas como produtos comerciais com valores inferiores a 500 dólares. Esses equipamentos apresentam resoluções crescentes ano a ano e devem disseminar a produção de dados 3D a baixo custo. Paralelamente, avanços em visão estéreo múltipla e estruturação por movimento (NEWCOMBE et al., 2011a, 2011b; PIZZOLI et al., 2014;) irão permitir a construção de modelos 3D em tempo real, permitindo que o modelador (seja ele um operador humano ou um sistema automatizado) identifique, durante a aquisição das imagens, quais regiões da planta necessitam de mais detalhes e quais ângulos precisam ser fornecidos para completar o modelo e resolver oclusões.

Considerando que tais tecnologias irão produzir uma quantidade cada vez maior de dados tridimensionais, tornam-se cada vez mais necessárias metodologias capazes de analisar esses dados, identificando automaticamente as estruturas de interesse, realizando as medições necessárias para a caracterização do fenótipo e detalhando o comportamento dessas estruturas ao longo do tempo. A principal vantagem de métodos não-destrutivos é permitir que os indivíduos continuem se desenvolvendo ao longo do experimento e registrar esse desenvolvimento. Novos métodos em visão computacional e aprendizado de máquina são necessários para permitir o reconhecimento

automático das mesmas estruturas ao longo do tempo, por exemplo, para registrar o desenvolvimento de uma folha ao longo de um experimento.

Outro problema de interesse diz respeito à fusão de dados de múltiplos sensores. Na tentativa de obter uma caracterização mais completa do *fenômeno* de um indivíduo em um dado instante, informação oriunda de outros sensores, como termógrafos, espectrômetros e fluorímetros, poderia ser incorporada ao modelo tridimensional da planta, permitindo o registro da variação de leituras ao longo da superfície do indivíduo ou, no pior dos casos, um registro preciso da localização da leitura. A ideia é obter o *snapshot* mais completo possível do estado da planta no momento da medição.

As inovações em visão computacional, automação, robótica e imageamento multiespectral irão melhorar a qualidade, quantidade e dimensionalidade dos dados fenotípicos, permitindo uma melhor caracterização do organismo. Hoje, os custos associados com muitas destas tecnologias de fenotipagem em larga escala limitam seu uso a indústrias de biotecnologia e sementes, e a poucos projetos públicos. Entretanto, pressupondo que o fenótipo ganhará tanta notoriedade como os genótipos alcançaram na era da genômica, progressos contínuos no desenvolvimento de sensores, câmeras, metodologias de imageamento, automação, entre outros, possibilitarão a redução de custos e a expansão da sua aplicação. Se estas perspectivas se concretizarem, avanços consideráveis são esperados tanto na ciência básica envolvendo elucidação de mecanismos biológicos fundamentais como na aplicação desse conhecimento e geração de valor através de novas variedades de plantas.

## 5 Referências

- ALENYÀ, G.; DELLEN, B.; TORRAS, C. 3D modelling of leaves from color and ToF data for robotized plant measuring. In: IEEE INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION, 2011, Shanghai. **Proceedings...** Piscataway: IEEE, 2011. p. 3408-3414.
- BABAR, M. A.; GINKEL, M. van; PRASAD, B.; REYNOLDS, M. P. The potential of using spectral reflectance indices to estimate yield in wheat grown under reduced irrigation. **Euphytica**, Wageningen, v. 150, n. 1-2, p. 155-172, Aug. 2006a.
- BABAR, M. A.; REYNOLDS, M. P.; GINKEL, M. van; KLATT, A. R.; RAUN, W. R.; STONE, M. L. Spectral reflectance to estimate genetic variation for in-season biomass, leaf chlorophyll, and canopy temperature in wheat. **Crop Science**, Madison, v. 46, n. 3, p.1046-1057, 2006b. DOI: 10.2135/cropsci2005.0211.
- BAUMBERG, A.; LYONS, A.; TAYLOR, R. 3D S.O.M.—A commercial software solution to 3D scanning. **Graphical Models**, San Diego, v. 67, n. 6, p. 476-495, Nov. 2005. DOI: 10.1016/j.gmod.2004.10.002.
- BELLASIO, C.; OLEJNÍKOVÁ, J.; BELLASIO, C.; SEBELA, D.; NEDBAL, L. Computer reconstruction of plant growth and chlorophyll fluorescence emission in three spatial dimensions. **Sensors**, Basel, v. 12, n. 1, p. 1052-71, Jan. 2012. DOI:10.3390/s120101052.
- BERARDO, N.; PISACANE, V.; BATTILANI, P.; SCANDOLARA, A.; PIETRI, A.; MAROCCO, A. Rapid detection of kernel rots and mycotoxins in maize by near-infrared reflectance spectroscopy. **Journal of Agricultural and Food Chemistry**, Easton, v. 53, n. 21, p. 8128-34, Oct. 2005.
- BERGER, B.; PARENT, B.; TESTER, M. High-throughput shoot imaging to study drought responses. **Journal of Experimental Botany**, Oxford, v. 61, n. 13, p. 3519-3528, July, 2010. DOI:10.1093/jxb/erq201.

BISKUP, B.; SCHARR, H.; SCHURR, U.; RASCHER, U. A stereo imaging system for measuring structural parameters of plant canopies. **Plant, Cell & environment**, Nottingham, v. 30, n. 10, p. 1299-308, Oct. 2007.

BOYKOV, Y.; VEKSLER, O.; ZABIH, R. Fast approximate energy minimization via graph cuts. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, New York, v. 23, n. 11, p. 1222-1239, Nov. 2001. DOI: 10.1109/34.969114.

CHAERLE, L.; LEINONEN, I.; JONES, H. G.; STRAETEN, D. van der. Monitoring and screening plant populations with combined thermal and chlorophyll fluorescence imaging. **Journal of Experimental Botany**, Oxford, v. 58, n. 4, p. 773-84, 2007. DOI:10.1093/jxb/erl257.

CHÉNÉ, Y.; ROUSSEAU, D.; LUCIDARME, P.; BERTHELOOT, J.; CAFFIER, V.; MOREL, P.; BELIN, E.; CHAPEAU-BLONDEAU, F. On the use of depth camera for 3D phenotyping of entire plants. **Computers and Electronics in Agriculture**, New York, v. 82, p. 122-127, Mar. 2012. DOI: 10.1016/j.compag.2011.12.007

CLARK, R. T.; MACCURDY, R. B.; JUNG, J. K.; SHAFF, J. E.; MCCOUCH, S. R.; ANESHANSLEY, D. J.; KOCHIAN, L. V. Three-dimensional root phenotyping with a novel imaging and software platform. **Plant Physiology**, Bethesda, v. 156, n. 2, p. 455-465, Jun. 2011. DOI: http://dx.doi.org/10.1104/pp.110.169102.

COBB, J. N.; DECLERCK, G.; GREENBERG, A.; CLARK, R.; MCCOUCH, S. Next-generation phenotyping: requirements and strategies for enhancing our understanding of genotype-phenotype relationships and its relevance to crop improvement. **Theoretical and Applied Genetics**, Berlin, v. 126, n. 4, p. 867-87, May, 2013. DOI 10.1007/s00122-013-2066-0.

DEL FIORE, A.; REVERBERI, M.; RICELLI, A.; PINZARI, F.; SERRANTI, S.; FABBRI, A. A.; BONIFAZI, G.; FANELLI, C. Early detection of toxigenic fungi on maize by hyperspectral imaging analysis. **International Journal of Food Microbiology**, Amsterdam, v. 144, n.1, p. 64-71, Nov. 2010. DOI: 10.1016/j.ijfoodmicro.2010.08.001. Epub 2010 Aug 13.

FIORANI, F.; SCHURR, U. Future scenarios for plant phenotyping. **Annual Review of Plant Biology**, Palo Alto, v. 64, p. 267-91, Feb. 2013. DOI:10.1146/annurev-arplant-050312-120137.

FREEDMAN, B.; SHPUNT, A.; MACHLINE, M.; ARIELI, Y. Depth mapping using projected patterns. **Int US8493496 B2**, 2013.

FURBANK, R. T.; TESTER, M. Phenomics - technologies to relieve the phenotyping bottleneck. **Trends in Plant Science**, Oxford, v. 16, n. 12, p. 635-644, Dec. 2011.

FURUKAWA, Y.; PONCE, J. Accurate, dense, and robust multiview stereopsis. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, New York, v. 32, n. 8, p. 1362-1376, Aug. 2010. DOI: 10.1109/TPAMI.2009.161.

GENTY, B.; MEYER, S. Quantitative mapping of leaf photosynthesis using chlorophyll fluorescence imaging. **Australian Journal of Plant Physiology**, Melbourne, v. 22, n. 2, p. 277-284, 1995.

GODIN, C.; COSTES, E.; SINOQUET, H. A method for describing plant architecture which integrates topology and geometry. **Annals of Botany**, London, v. 84, n. 3, p. 343-357, May, 1999.

GOETZ, A. F. H.; VANE, G.; SOLOMON, J. E.; ROCK, B. N. Imaging spectrometry for earth remote sensing. **Science**, Washington, DC, v. 228, n. 4704, p. 1147-1153, June, 1985. DOI: 10.1126/science.228.4704.1147.

GOWEN, A. A.; O'DONNELL, C. P.; CULLEN, P. J.; DOWNEY, G.; FRIAS, J. M. Hyperspectral imaging - an emerging process analytical tool for food quality and safety control. **Trends in Food Science & Technology**, Cambridge, v. 18, n.12, p. 590-598, Dec. 2007. DOI: 10.1016/j.tifs.2007.06.001.

GRANIER, C.; AGUIRREZABAL, L.; CHENU, K.; COOKSON, S. J.; DAUZAT, M.; HAMARD, P.; TARDIE, F. Phenopsis, an automated platform for reproducible phenotyping of plant responses to soil water deficit in Arabidopsis thaliana permitted the identification of an accession with low sensitivity to soil water deficit. **New Phytologist**, Cambridge, v. 169, n. 3, p. 623-635, Jan. 2006. DOI: 10.1111/j.1469-8137.2005.01609.x.

HARTLEY, R.; ZISSERMAN, A. **Multiple view geometry in computer vision**. 2nd ed. Cambridge University Press, 2004. 670 p.

HARTMANN, A.; CZAUDERNA, T.; HOFFMANN, R.; STEIN, N.; F. SCHREIBER. HTPheno: an image analysis pipeline for high-throughput plant phenotyping. **BMC Bioinformatics**, Cambridge, v. 12, n. 1, p. 148, May, 2011. DOI:10.1186/1471-2105-12-148.

HATFIELD, J. L.; GITELSON, A. A.; SCHEPERS, J. S.; WALTHALL, C. L. Application of spectral remote sensing for agronomic decisions. **Agronomy Journal**, Madison, v.100, n.3, p. S-117-S-131, 2008. Suplemento. DOI: 10.2134/agronj2006.0370c.

HENTEN, E. J. van. Greenhouse mechanization: state of the art and future perspective. **Acta Horticulturae**, The Hague, v. 710, p. 55-70, 2006. International Symposium on Greenhouses, Environmental Controls and In-house Mechanization for Crop Production in the Tropics and Sub-Tropics, June 2006.

HOULE, D.; GOVINDARAJU, D. R.; OMHOLT, S. Phenomics: the next challenge. **Nature Reviews. Genetics**, London, v. 11, p. 855-66, Dec. 2010.

IVANOV, N.; BOISSARD, P.; CHAPRON, M.; ANDRIEU, B. Computer stereo plotting for 3-D reconstruction of a maize canopy. **Agricultural and Forest Meteorology**, Amsterdam, v. 75, n.1-3, p. 85-102, June, 1995.

JAY, S.; RABATEL, G.; GORRETTA, N. In-field crop row stereo-reconstruction for plant phenotyping. In: INTERNATIONAL CONFERENCE ON ROBOTICS, AND ASSOCIATED HIGH-TECHNOLOGIES AND EQUIPMENT FOR AGRICULTURE AND FORESTRY, 2., 2014, Madrid. **New trends in mobile robotics, perception and actuation for agriculture and forestry: proceedings...** Madrid: Polytechnic University of Madrid, 2014. p. 339-348. Edyted by Pablo Gonzalez-de-Santos and Angela Ribeiro.

JIANG, H. Y.; ZHU, Y. J.; WEI, L. M.; DAI, J. R.; SONG, T. M.; YAN, Y. L.; CHEN, S. J. Analysis of protein, starch and oil content of single intact kernels by near infrared reflectance spectroscopy (NIRS) in maize (*Zea mays* L.). **Plant Breeding**, Berlin, v. 126, n. 5, p. 492-497, Oct. 2007. DOI: 10.1111/j.1439-0523.2007.01338.x.

JONES, H. G.; SERRAJ, R.; LOVEYS, B. R.; XIONG, L.; WHEATON, A.; PRICE, A. H. 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, Nov. 2009.

KAMINUMA, E.; HEIDA N.; TSUMOTO, Y.; YAMAMOTO, N.; GOTO, N.; OKAMOTO N.; KONAGAYA, A.; MATSUI, M.; TOYODA, T. Automatic quantification of morphological traits via three-dimensional measurement of Arabidopsis. **The Plant Journal: for Cell and Molecular Biology**, Oxford, v. 38, n. 2, p. 358-365, Apr. 2004.

KAZMI, W.; FOIX, S.; ALENYÀ, G.; ANDERSEN, H. J. Indoor and outdoor depth imaging of leaves with time-of-flight and stereo vision sensors: Analysis and comparison. **ISPRS Journal of Photogrammetry and Remote Sensing**, Amsterdam, v. 88, p. 128-146, Fev. 2014. DOI: 10.1016/j.isprsjprs.2013.11.012.

KONISHI, A.; EGUCHI, A.; HOSOI, F.; OMASA, K. 3D monitoring spatio-temporal effects of herbicide on a whole plant using combined range and chlorophyll a fluorescence imaging. **Functional Plant Biology**, Victoria, v. 36, n. 11, p. 874-879, Nov. 2009. DOI: http://dx.doi.org/10.1071/FP09108.

LANG, A. R. G. Leaf orientation of a cotton plant. **Agricultural Meteorology**, Amsterdam, v. 11, p. 37-51, 1973.

LEBOT, V.; CHAMPAGNE, A.; MALAPA, R.; SHILEY, D. NIR determination of major constituents in tropical root and tuber crop flours. **Journal of Agricultural and Food Chemistry**, Easton, v. 57, n. 22, p.10539-47, Oct. 2009. DOI: 10.1021/jf902675n.

LENK, S.; CHAERLE, L.; PFÜNDEL, E. E.; LANGSDORF, G.; HAGENBEEK, D.; LICHTENTHALER, H. K.; STRAETEN, D. van der; BUSCHMANN, C. Multispectral fluorescence and reflectance imaging at the leaf level and its possible applications. **Journal of Experimental Botany**, Oxford, v. 58, n. 4, p. 807-814, Nov. 2007.

LEUNG, C.; APPLETON, B.; BUCKLEY, M.; SUN, C. Embedded voxel colouring with adaptive threshold selection using globally minimal surfaces. **International Journal of Computer Vision**, Dordrecht, v. 99, n. 2, p. 215-231, Sept. 2012.

LI, Y.-F.; KENNEDY, G.; NGORAN, F.; WU, P.; HUNTER, J. An ontology-centric architecture for extensible scientific data management systems. **Future Generation Computer Systems**, Amsterdam, v. 29, n. 2, p. 641-653, Feb. 2013.

MERLOT, S.; MUSTILLI, A. C.; GENTY, B.; NORTH, H.; LEFEBVRE, V.; SOTTA, B.; VAVASSEUR, A.; GIRAUDAT, J. Use of infrared thermal imaging to isolate Arabidopsis mutants defective in stomatal regulation. **Plant Journal: for Cell and Molecular Biology**, Oxford, v. 30, n. 5, p. 601-609, Jun. 2002.

MONTES, J. M.; MELCHINGER, A. E.; REIF, J. C. Novel throughput phenotyping platforms in plant genetic studies. **Trends in Plant Science**, Oxford, v. 12, n. 10, p. 433-6. Oct. 2007a. DOI: 10.1016/j.tplants.2007.08.006.

MONTES, J. M.; PAUL, C.; MELCHINGER, A. E. Determination of chemical composition and nutritional attributes of silage corn hybrids by near-infrared spectroscopy on chopper: evaluation of traits, sample presentation systems and calibration transferability. **Plant Breeding**, Berlin, v. 126, n. 5, p. 521-526, Oct. 2007b. DOI: 10.1111/j.1439-0523.2007.01389.x.

MONTES, J. M.; UTZ, H. F.; SCHIPPRACK, W.; KUSTERER, B.; MUMINOVIC, J.; PAUL, C.; MELCHINGER, A. E. Near-infrared spectroscopy on combine harvesters to measure maize grain dry matter content and quality parameters. **Plant Breeding**, Berlin, v. 125, n. 6, p. 591-595, Dec. 2006. DOI:10.1111/j.1439-0523.2006.01298.x.

MURRAY, S. C.; ROONEY, W. L.; MITCHELL, S. E.; SHARMA, A.; KLEIN, P. E.; MULLET, J. E.; KRESOVICH, S. Genetic improvement of sorghum as a biofuel feedstock: II. QTL for stem and leaf structural carbohydrates. **Crop Science**, Madison, v. 48, n. 6, p. 2180-2193. 2008a. DOI:10.2135/cropsci2008.01.0068.

MURRAY, S. C.; SHARMA, A.; ROONEY, W. L.; KLEIN, P. E.; MULLET, J. E.; MITCHELL, S. E.; KRESOVICH, S. Genetic improvement of sorghum as a biofuel feedstock: I. QTL for stem sugar and grain nonstructural carbohydrates. **Crop Science**, Madison, v. 48, n. 6, p. 2165-2179. Nov. 2008b.

NAGEL, K. A.; PUTZ, A.; GILMER, F.; HEINZ, K.; FISCHBACH, A.; PFEIFER, J.; SCHURR, U.; FAGET, M.; BLOSSFELD, S.; ERNST, M.; DIMAKI, C.; KASTENHOLZ, B.; ANN-KLEINERT, A.-K.; GALINSKI, A.; SCHARR, H.; FIORANI, F.; SCHURR, U. GROWSCREEN-Rhizo is a novel phenotyping robot enabling simultaneous measurements of root and shoot growth for plants grown in soil-filled rhizotrons. **Functional Plant Biology**, Victoria, v. 39, n. 11, p. 891-904. June, 2012.

NEWCOMBE, R. A.; IZADI, S.; HILLIGES, O.; MOLYNEAUX, D.; KIM, D.; DAVISON, A. J.; KOHLI, P.; SHOTTON, J.; HODGES, S.; FITZGIBBON, A. KinectFusion: Real-time dense surface mapping and tracking. In: IEEE INTERNATIONAL SYMPOSIUM ON MIXED AND AUGMENTED REALITY, 10., 2011, Basel. **Proceedings...** Piscataway: IEEE, 2011a. p. 127-136, 2011.

NEWCOMBE, R. A.; LOVEGROVE, S. J.; DAVISON, A. J. DTAM: dense tracking and mapping in real-time. In: IEEE INTERNATIONAL CONFERENCE ON COMPUTER VISION, 2011, Barcelona. **Proceedings...** Piscataway: IEEE, 2011b. p. 2320-2327. DOI:10.1109/ICCV.2011.6126513.

OSBORNE, B. G. Applications of near infrared spectroscopy in quality screening of early-generation material in cereal breeding programmes. **Journal of Near Infrared Spectroscopy**, Sussex, v. 14, n. 2, p. 93-101, 2006. DOI: 10.1255/jnirs.595.

OSBORNE, S. L.; SCHEPERS, J. S.; FRANCIS, D. D.; SCHLEMMER, M. R. 2002a. Detection of phosphorus and nitrogen deficiencies in corn using spectral radiance measurements. **Agronomy Journal**, v. 94, n. 6, p. 1215-1221, 2002a. DOI:10.1186/1471-2229-12-63.

OSBORNE, S. L.; SCHEPERS, J.S.; FRANCIS, D.D.; SCHLEMMER, M.R. Use of spectral radiance to estimate in-season biomass and grain yield in nitrogen- and water-stressed corn. **Crop science**, v. 42, n. 1, p.165-171, 2002b

PAPROKI, A.; SIRAUULT, X.; BERRY, S.; FURBANK, R.; FRIPP, J. A novel mesh processing based technique for 3D plant analysis. **BMC Plant Biology**, Cambridge, v. 12, p. 63, May, 2012. DOI:10.1186/1471-2229-12-63.

PEARSON, T. C.; D. T. WICKLOW. Detection of corn kernels infected by fungi. **Transactions of the Asabe, St. Joseph**, v. 49, n. 4, p. 1235-1246, 2006.

PETERS, R. T.; EVETT, S. R. Spatial and temporal analysis of crop conditions using multiple canopy temperature maps created with center-pivot-mounted infrared thermometers. **Transactions Of The ASABE**, St. Joseph, v. 50, n. 3, p. 919-927, May/June, 2007. <Disponível em: <http://naldc.nal.usda.gov/naldc/download.xhtml?id=2345&content=PDF>>. Acesso em: 20 out. 2014.

PIZZOLI, M.; FORSTER, C.; SCARAMUZZA, D. Remode: probabilistic, monocular dense reconstruction in real time. In: IEEE INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION, 2014, Hong Kong. **Robotics and automation: technologies enabling new economic growth**: proceedings. Hong Kong: University of Hong Kong, 2014. Não paginado.

POSS, J. A.; RUSSELL, W. B.; GRIEVE, C. M.. Estimating yields of salt- and water-stressed forages with remote sensing in the visible and near infrared. **Journal of Environmental Quality**, Madison, v. 35, n. 4, p. 1060-1071, May, 2006.

RAKOCEVIC, M.; SINOQUET, A.; VARLET-GRANCHER, C. Assessing the geometric structure of a white clover (*Trifolium repens* L.) canopy using 3-D digitising. **Annals of Botany**, London, v. 86, n. 3, p. 519-526, July, 2000. DOI:10.1006/anbo.2000.120.

RASCHE, U.; BLOSSFELD, S.; FIORANI, F.; JAHNKE, S.; JANSEN, M.; KUHN, A. J.; KUHN, A. J.; MATSUBARA, S.; MÄRTIN, L. L. A.; MERCHANT, A.; METZNER, R.; MÜLLER-LINOW, M.; NAGEL, K. A.; PIERUSCHKA, R.; PINTO, F.; SCHREIBER, C. M.; TEMPERTON, V. M.; THORPE, M. R.; DUSSCHOTEN, D. van; VOLKENBURG, E. van; WINDT, C. W.; SCHURR, U. Non-invasive approaches for phenotyping of enhanced performance traits in bean. **Functional Plant Biology**, Victoria, v. 38, n. 12, p. 968-983, 2011.

REUZEAU, C.; FRANKARD, V.; HATZFELD, Y.; SANZ, A.; CAMP, W. van; LEJEUNE, P.; BROEKAERT, W.; WILDE, C.D.; LIEVENS, K.; VRANKEN, E.; PEERBOLTE, R.; BROEKAERT, W. TraitmillTM: a functional genomics platform for the phenotypic analysis of cereals. **Plant Genetic Resources**, Cambridge, v. 4, n. 1, p. 20-24, 2006.

ROLFE, S. A.; SCHOLES, J. D. Chlorophyll fluorescence imaging of plant-pathogen interactions. **Protoplasma**, New York, v. 247, n. 3-4, p. 163-75, Dec. 2010. DOI: 10.1007/s00709-010-0203-z.

ROMANO, G.; ZIA, S.; SPREER, W.; SANCHEZ, C.; CAIRNS, J.; ARAUS, J. L.; MÜLLER, J. Use of thermography for high throughput phenotyping of tropical maize adaptation in water stress. **Computers and Electronics in Agriculture**, v. 79, n. 1, p. 67-74, Oct. 2011. DOI: 10.1016/j.compag.2011.08.011.

ROVIRA-MÁS, F.; ZHANG, Q.; REID, J. Creation of Three-dimensional Crop Maps based on Aerial Stereomages. **Biosystems Engineering**, London, v. 90, n. 3, p. 251-259, Mar. 2005. DOI:10.1016/j.biosystem-seng.2004.11.013.

SANTOS, T.; UEDA, J. Automatic 3D plant reconstruction from photographs, segmentation and classification of leaves and internodes using clustering 1. In: INTERNATIONAL CONFERENCE ON FUNCTIONAL-STRUCTURAL PLANT MODELS, 7., 2013, Saariselkä. **Proceedings...** Saariselkä: [s.n.], 2013. p. 95-97.

SANTOS, T. T.; OLIVEIRA, A. A. de. Image-based 3D digitizing for plant architecture analysis and phenotyping. In: WORKSHOP ON INDUSTRY APPLICATIONS; CONFERENCE ON GRAPHICS, PATTERNS AND IMAGES, 25., 2012, Ouro Preto. **Proceedings...** Los Alamitos: IEEE Computer Society, 2012. Não paginado. SIBGRAPI 2012.

SANTOS, T. T.; KOENIGKAN, L. V.; BARBEDO, J. G. A.; RODRIGUES, G. C. 3D plant modeling: localization, mapping and segmentation for plant phenotyping using a single hand-held camera. In: EUROPEAN CONFERENCE ON COMPUTER VISION; WORKSHOP ON COMPUTER VISION PROBLEMS IN PLANT PHENOTYPING, 2014, Zurich, 2014. Zurich. **Proceedings...** Swiss Federal Institute of Technology, 2014. p. 1-18.

- SINOQUET, H.; MOULIA, B.; BONHOMME, R. Estimating the three-dimensional geometry of a maize crop as an input of radiation models: comparison between three-dimensional digitizing and plant profiles. **Agricultural and Forest Meteorology**, Amsterdam, v. 55, n. 3-4, p. 233-249, June, 1991. DOI: 10.1016/0168-1923(91)90064-W.
- SIRAULT, X.; FRIPP, J.; PAPROKI, A.; KUFFNER, P. C.; NGUYEN, C. V.; LI, R.; DAILY, H.; GUO, J.; FURBANK, R. PlantScan™: a three-dimensional phenotyping platform for capturing the structural dynamic of plant development and growth. In: INTERNATIONAL CONFERENCE ON FUNCTIONAL-STRUCTURAL PLANT MODELS, 7., 2013, Saariselkä. **Proceedings...** Saariselkä: [s.n.], 2013. p. 45-48.
- SNAVELY, N.; SEITZ, S.; SZELISKI, R. Modeling the world from internet photo collections. *International Journal of Computer Vision*, Dordrecht, v. 80, n. 2, p. 189-210, Nov. 2008. DOI 10.1007/s11263-007-0107-3.
- SONG, Y.; GLASBEY, C.; HEIJDEN, G. Van der; POLDER, G.; J.; DIELEMAN, A. Combining stereo and Time-of-Flight images with application to automatic plant phenotyping. In: SCANDINAVIAN CONFERENCE ON IMAGE ANALYSIS, 17., 2011, Ystad. **Proceedings...** Ystad Linköping University, 2011. v. 1, p. 467-478.
- TALLADA, J. G.; PALACIOS-ROJAS, G. N.; ARMSTRONG, P. R. Prediction of maize seed attributes using a rapid single kernel near infrared instrument. **Journal of Cereal Science**, London, v. 50, n. 3, p. 381-387, Nov. 2009. DOI: 10.1016/j.jcs.2009.08.003.
- VANKADAVATH, R. N.; HUSSAIN, A. J.; BODANAPU, R.; KHARSHIING, E.; BASHA, P. O.; GUPTA, S.; SREELAKSHMI, Y.; SHARMA, S. Computer aided data acquisition tool for high-throughput phenotyping of plant populations. **Plant Methods**, London, v. 5, p. 18, Dec. 2009. DOI:10.1186/1746-4811-5-18.
- WALSH, K. B.; GUTHRIE, J. A.; BURNEY, J. W. Application of commercially available, low-cost, miniaturised NIR spectrometers to the assessment of the sugar content of intact fruit. **Australian Journal of Plant Physiology**, Melbourne, v. 27, n. 12, p.1175-1186, 2000.
- WALTER, A.; SCHURR, U. Dynamics of leaf and root growth: endogenous control versus environmental impact. **Annals of Botany**, London, v. 95, n. 6, p. 891-900, 2005. DOI: 10.1093/aob/mci103.
- WELLE, R.; GRETEN, W.; MÜLLER, T.; WEBER, G.; WEHRMANN, H. Application of near infrared spectroscopy on-combine in corn grain breeding. **Journal of Near Infrared Spectroscopy**, Sussex, v. 13, n.1, p. 69-76, 2005. DOI: 10.1255/jnirs.459.
- WELLE, R.; GRETEN, W.; RIETMANN, B.; ALLEY, S.; SINNAEVE, G.; DARDENNE, P. Near-infrared spectroscopy on chopper to measure maize forage quality parameters online. **Crop Science**, Madison, v. 43, n. 4, p. 1407-1413, June, 2003. DOI: 10.2135/cropsci2003.1407.
- WILLIAMS, P.; GELADI, P.; FOX, G.; MANLEY, M. Maize kernel hardness classification by near infrared (NIR) hyperspectral imaging and multivariate data analysis. **Analytica Chimica Acta**, Amsterdam, v. 653, n. 2, p.121-30, Oct. 2009.
- WINTERHALTER, L.; MISTELE, B.; JAMPATONG, S.; SCHMIDHALTER, U. High-throughput sensing of aerial biomass and above-ground nitrogen uptake in the vegetative stage of well-watered and drought stressed tropical maize hybrids. **Crop Science**, Madison, v. v. 51, n. 2, p. 479, 2007a. DOI: 10.2135/cropsci2010.07.0397.
- WINTERHALTER, L.; MISTELE, B.; JAMPATONG, S.; SCHMIDHALTER, U. High throughput phenotyping of canopy water mass and canopy temperature in well-watered and drought stressed tropical maize hybrids in the vegetative stage. **European Journal of Agronomy**, v. 35, n. 1, p. 22-32, June, 2011b. DOI: 10.1016/j.eja.2011.03.004.
- XU, H.; GOSSETT, N.; CHEN, B. Knowledge and heuristic-based modeling of laser-scanned trees. *ACM Transactions on Graphics*, New York, v. 26, n. 4, Oct. 2007. DOI: 10.1145/1289603.1289610.

## **Biologia computacional molecular e suas aplicações na agricultura**

José Gilberto Jardine  
Izabella Agostinho Pena Neshich  
Ivan Mazoni  
Inácio Henrique Yano  
Fábio Rogério de Moraes  
Jose Augusto Salim  
Luiz Borro  
Leticia Sayuri Nishimura  
Goran Neshich

### **1 Biologia computacional, uma nova ciência aplicada**

Uma definição abrangente para Biologia Computacional envolve qualquer técnica computacional aplicada que busca solucionar problemas da Biologia. A partir dessa definição podemos dizer que a Biologia Computacional Molecular é um ramo da biotecnologia calcada nos avanços da Biologia, da Química, da Bioquímica e Biofísica que busca, a partir de ferramentas oferecidas pela Ciência da Computação, Matemática aplicada e Estatística, oferecer uma percepção transdisciplinar de aspectos relacionados a sequências de nucleotídeos e aminoácidos, a estrutura e dinâmica de proteínas e a interação proteína-proteína, proteína-DNA, proteína-ligante.

Dentre as áreas de abrangência da Biologia Computacional, tais como bioinformática que abriga bancos de dados sobre sequências de DNA, RNA e proteínas; biomodelagem computacional, um campo da biocibernética para a construção de modelos computacionais de sistemas biológicos; simulação molecular, que lida com métodos teóricos e técnicas computacionais para modelar o comportamento de biomoléculas; biologia sistêmica, que modela redes de interação biológica, destaca-se a área que trata especificamente do planejamento e desenho de novas drogas, fármacos e agroquímicos ou agro defensivos in silico.

A agricultura dos países do Cone Sul que desempenha papel de destaque na economia mundial, sendo responsável por grande parte da renda desses países, está faminta por soluções inovadoras que propiciem aumento da produção e, principalmente, da produtividade agrícola. A geração de um inventário de novos agroquímicos e fármacos pode ser um dos mais importantes desafios perante o agronegócio desses países, representando um dos fatores fundamentais de conquista de vantagens competitivas no mercado internacional. Para ingressar nesta área é necessário lembrar que fornecer anotação funcional para a vasta quantidade de dados de sequências e estruturas e as funções das proteínas, particularmente das enzimas, geradas por tecnologias de alto desempenho em larga escala, é uma das principais tarefas da era pós genômica. Caminhando na direção do planejamento de fármacos, é fundamental a identificação dos resíduos catalíticos de uma enzima, sendo esse um passo importante no entendimento do seu papel biológico e suas aplicações,

principalmente devido ao fato de que um pequeno número de resíduos situado no sítio catalítico é responsável pela função enzimática, e participa diretamente do processo de catálise. Os arranjos espaciais, bem como as propriedades físicas, químicas e físico-químicas destes resíduos determinam a reação química catalisada pela enzima.

Os fatores que determinam a funcionalidade do sítio ativo de uma proteína são muito complexos e dependem de sua estrutura tridimensional, além das propriedades bioquímicas e biofísicas. Sítios funcionais podem ser entendidos como uma redondeza física, química e físico-química que acompanha uma função, em contraposição a um grupo de resíduos fixos que não participam diretamente, porém, providenciam o arcabouço estrutural. Um melhor entendimento deste complexo processo tem um significativo impacto no desenvolvimento de novas drogas, bem como na identificação de desordens genéticas e na engenharia de proteínas com novas funções.

Diversos métodos matemáticos e estatísticos têm sido propostos para calcular informações sobre as propriedades estruturais, físicas, químicas e físico-químicas dos aminoácidos (e.g. energia de contatos, potencial eletrostático, hidrofobicidade, densidade). Estas propriedades são denominadas descritores estruturais de proteínas, e tais propriedades permitem realizar diversas análises segundo diferenças e semelhanças entre aminoácidos de várias regiões de uma proteína. Além de descritores físicos, químicos e físico-químicos, existem descritores baseados na conservação da estrutura primária da proteína, ou seja, baseiam-se em evidências evolutivas obtidas a partir da comparação entre sequências de aminoácidos das proteínas homólogas, determinando quais aminoácidos são mais frequentes em certas posições da sequência, o que sugere sua importância na manutenção da funcionalidade da proteína. Outro descritor utilizado em sistema de predição de aminoácidos posicionados em determinados distritos proteicos é a propensão (*propensity*, em inglês), definida como sendo a porcentagem entre o número de aminoácidos de interesse e aqueles que não têm preferência para determinada posição, para cada um dos vinte tipos de aminoácidos.

O planejamento e desenvolvimento de novos fármacos e agroquímicos, modificadores de função biológica das macromoléculas cruciais para a patogenicidade de microrganismos, por exemplo, que servem para proteção das plantas, animais e humanos, é um processo bastante complexo e longo e demanda um investimento contínuo para que se possa chegar até o resultado/produto com potencial para ser oferecido ao mercado (KUNTZ, 1982).

## 2 Desenho racional de drogas, fármacos e agroquímicos

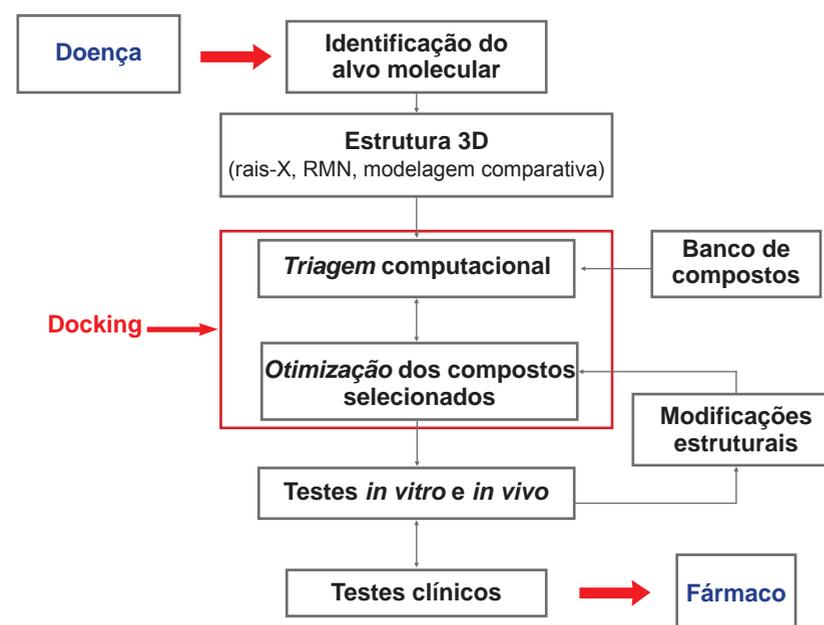
O processo para se obter novos fármacos tem mudado através dos anos. Há bem pouco tempo, novos fármacos eram resultado de testes aleatórios chamados de triagem cega (*screening*) em células, animais, plantas ou modelos destes. Esse método tradicional consiste em testar aleatoriamente várias micromoléculas sem nenhum conhecimento dos mecanismos de interação da micromolécula com a molécula ligante. Esse método é ineficiente em função de que a probabilidade é cada vez menor de se identificar um novo fármaco. A necessidade de se testar em bancada de laboratório *in vitro* e posteriormente *in vivo* milhões de compostos com, conseqüentemente, custo e tempo elevados, apesar de ter sido o método pelo qual a maioria dos medicamentos hoje disponíveis foram desenvolvidos, tornou-se num grande problema (ANDERSON, 2003).

Esses fatos implicaram a necessidade de desenvolvimento de uma metodologia sistemática objetivando o desenho racional de fármacos baseado na estrutura tridimensional do receptor das

pequenas moléculas (ligantes ou inibidoras). O desenho racional de fármacos fundamenta-se no estudo de estruturas moleculares tridimensionais da molécula receptora para o desenho de compostos protótipos, tomando como base as informações estruturais e as interações envolvidas no processo de reconhecimento molecular receptor-ligante (YANG, 2010).

Metodologias de atracamento proteína-ligante (*docking* proteína-ligante) são amplamente utilizadas tanto para a descoberta de novas substâncias bioativas, técnica conhecida como triagem virtual (*virtual screening*), quanto para o aperfeiçoamento de compostos bioativos já identificados (GUEDES et al., 2014). A metodologia de desenho racional de fármacos baseado em estrutura tridimensional consiste de seis etapas básicas, como pode ser visualizado esquematicamente na Figura 1, que são:

- 1) Escolha adequada do alvo terapêutico (proteína ou enzima) relacionada à doença cuja função deva ser bloqueada ou ativada.
- 2) Obtenção da estrutura tridimensional do biorreceptor através de técnicas experimentais, tais como difração de raios-X, ressonância magnética nuclear ou então por modelagem por homologia. Hoje existem bancos de dados de estruturas tridimensionais de moléculas de acesso livre, como o Protein Data Bank (PDB)<sup>1</sup> e bancos de dados de estruturas moleculares de moléculas ligantes, como o Maybridge e a Cambridge Structural Database (CSD).
- 3) Triagem *in silico* com *virtual screening docking* (esse procedimento não é diferente da metodologia tradicional de triagem às cegas). A diferença é que agora a busca de micromoléculas



**Figura 1.** Fluxograma com as etapas básicas do desenho racional de fármacos baseado em estrutura.

Fonte: Magalhães (2006).

<sup>1</sup> Disponível em: < <http://www.rcsb.org/pdb/home/home.do> >.

candidatas é feita com os recursos computacionais, sem gasto de materiais e a tempo reduzidíssimo.

- 4) Identificada uma molécula promissora, o próximo passo consiste na análise por um profissional especialista para identificar as alterações que serão necessárias na molécula para se obter a resposta biológica desejada. Modificações subsequentes são realizadas para aprimorar a molécula ligante e torná-la mais específica para um determinado alvo e ajustes farmacocinéticos como absorção, distribuição, metabolismo e eliminação. Nessa fase são utilizadas metodologias de *docking* mais acuradas, para a identificação da conformação de ligação das moléculas selecionadas e a otimização dos compostos modificados estruturalmente por proposição do especialista.
- 5) O penúltimo passo consiste na síntese laboratorial da molécula candidata e teste *in vitro*.
- 6) Finalmente, são requeridos testes *in vivo* para a análise do comportamento do fármaco e sua toxicidade (MAGALHÃES, 2006; MAGALHÃES et al., 2007).

### 3 Biologia computacional na Embrapa

No Laboratório do Grupo de Pesquisa em Biologia Computacional (GPBC) da Embrapa Informática Agropecuária, Unidade temática de pesquisa da Empresa Brasileira de Pesquisa Agropecuária (Embrapa), que tem como missão institucional viabilizar soluções de pesquisa, desenvolvimento e inovação em tecnologia da informação para a sustentabilidade da agricultura, foi desenvolvido o software Sting e o maior banco de dados de descritores físicos, químicos, físico-químicos, estruturais e biológicos sobre estruturas proteicas, o Sting\_DB (Neshich et al., 2003), disponível na sua versão on-line<sup>2</sup>.

Blue Star Sting é uma suíte de programas com ferramentas para a visualização e análise estrutural de proteínas. Estes programas (módulos) estão concentrados em um único pacote que visa oferecer um instrumento para estudos das macromoléculas, suas estruturas e as relações estrutura-função. Informações como posição dos aminoácidos na sequência e na estrutura, busca por padrões, identificação de vizinhança, ligações de hidrogênio, ângulos e distância entre átomos, são facilmente obtidas, além de dados sobre a natureza e volume dos contatos atômicos inter e intra-cadeias, conservação e relação entre os contatos intra-cadeia e parâmetros funcionais (Neshich et al., 2006). Principalmente pelo fato do Blue Star Sting oferecer fácil acesso a um rico repositório de características da proteína, a plataforma Sting (Neshich et al., 2004, 2005, 2006; MANCINI et al., 2004) já foi utilizada para prever classe de enzimas (BORRO et al., 2006) em análise de proteínas ligante (FERNANDEZ et al., 2003; FREITAS et al., 1997), análise de proteína mutantes (MARCELLINO et al., 1996; SIMÕES et al., 2007), análise de padrões de interação proteína-proteína, bem como em pesquisas ligadas a alguns problemas biológicos específicos (BRAGHINI et al.; DIAS-LOPES et al., 2013).

A partir de um conjunto de enzimas com seus resíduos catalíticos devidamente identificados (e.g. através de experimentos de mutagênese) como estudo de caso, um modelo pode ser construído

usando métodos de aprendizagem de máquina, para caracterizar o nano-ambiente destes resíduos segundo suas propriedades físicas, químicas e físico-químicas através de descritores de proteínas extraídos do banco de dados Sting\_DB, e assim avaliar se tais modelos possuem capacidade para predição dos resíduos catalíticos de enzimas ainda não anotadas e também de novas enzimas.

Apesar de métodos baseados em estrutura limitarem-se às proteínas com estrutura conhecida, ou a modelos estruturais de boa qualidade, tais métodos são importantes ao entendimento de como as enzimas realizam sua função, uma vez que a função das proteínas é mais conservada na estrutura terciária do que na estrutura primária.

O Grupo de Pesquisa em Biologia Computacional (GPBC) da Embrapa Informática Agropecuária executa projetos de pesquisa em bioinformática estrutural. Decorrente do trabalho realizado nos últimos cinco anos dedicado à finalização da plataforma Sting, que serve para a análise teórica das estruturas proteicas e seus complexos e para construção dos fármacos e agro defensivos/agrotóxicos embasado no conceito de *rational drug design*, obteve várias soluções inovadoras para problemas de fundamental interesse para a agricultura brasileira e demais países que se beneficiam igualmente da agricultura.

O procedimento adotado pelo GPBC consiste, inicialmente, na escolha dos alvos terapêuticos e proteínas. A escolha dos alvos proteicos e, mais especificamente, dos resíduos a serem usados como alvos em um procedimento de desenho computacional de fármacos baseado em estruturas proteicas foi o passo inicial para o desenho computacional de fármacos que resultou em quatro pedidos de patentes já registradas no Instituto Nacional da Propriedade Industrial (INPI) e que mais adiante serão utilizados para ilustrar os processos de desenho de novos fármacos. Mais importante que a definição da proteína-alvo por si só é a definição dos resíduos que serão usados como alvo para o procedimento de SBDD (do inglês Structure Based Drug Design). Usando o banco de dados Catalytic Site Atlas (CSA)<sup>3</sup> (PORTER et al., 2004) são selecionados os resíduos do sítio ativo de cada proteína alvo, caso possuam atividade enzimática e/ou, usando o software SurfV são selecionados resíduos de interface envolvidos em interação proteína-proteína. A identificação dos resíduos formadores de interface (IFR) se dá pelo cálculo da diferença entre a área acessível ao solvente (ASA) da proteína isolada e complexada usando o programa SurfV (SRIDHARAN et al., 1992).

Através do pacote de programas e bancos de dados Sting e do Java Protein Dossier é feito um refinamento de conhecimento sobre o alvo, por exemplo, buscando por cavidades próximas aos resíduos de interesse e que possam alojar um fármaco inibitório. Com o pacote de programas Sting é possível calcular o número, o tipo e, portanto, a energia total dos contatos estabelecidos entre os resíduos que se localizam proximamente aos *pockets* (bolsos) e cavidades selecionadas, bem como nas interfaces entre cadeias adjacentes, contatos esses que possivelmente deverão ser vencidos pela interação com um ligante potencial. A realização da seleção dos resíduos a serem usados como alvos terapêuticos com base em características físicas, químicas, físico-químicas, valores altos de energia de contatos, polaridade e características estruturais como área exposta ao solvente, presença em *pocket*, dentre outras de interesse como ocorrência de sítio catalítico ou próximo a ele, é efetuada usando o módulo Select do Sting Java Protein Dossier. Os programas PyMol (PyMOL MOLECULAR GRAPHICS SYSTEM, 2014) e Accelrys Discovery Studio

<sup>2</sup> Disponível em: <[www.cbi.cnptia.embrapa.br/SMS](http://www.cbi.cnptia.embrapa.br/SMS)>.

<sup>3</sup> Disponível em: <<http://www.ebi.ac.uk/thornton-srv/databases/CSA/>>.

Visualizer (ACCELRY'S SOFTWARE INC., 2014) são utilizados em conjunto ao Sting para visualização e geração de imagens moleculares.

O grupo de resíduos-alvos selecionado de acordo com as características de interesse para cada proteína-alvo deve ser analisado detalhadamente em relação à sua ocorrência e conservação entre outros organismos. São considerados alvos preferenciais aqueles que ocorrerem, na mesma posição estrutural, em organismos patogênicos, mas não em organismos não patogênicos de vida livre. Esta ausência em tais organismos deve minimizar o impacto ecológico, por exemplo, em populações bacterianas que não causam danos para o hospedeiro. Esta abordagem é efetuada através de alinhamento entre a estrutura primária de proteínas homólogas à proteína de interesse através dos programas ClustalW 2.0 (LARKIN et al., 2007), Muscle (EDGAR, 2004), T-Coffee (Tree-based Consistency Objective Function For alignment Evaluation) (NOTREDAME et al., 2000), evidenciando as similaridades e diferenças entre as proteínas e buscando as correspondências.

Selecionados os resíduos, elabora-se um pipeline de desenho de fármacos baseado na estrutura proteica (SBDD). Técnicas de SBDD podem ser aplicadas com base em possíveis alvos terapêuticos preditos.

A seguir é apresentada uma exemplificação de um dos vários caminhos (pipeline) adotados pelo GPBC para o desenho de novos fármacos.

O procedimento tem início com o mapeamento de posições favoráveis de interações para grupos funcionais (em que posições podem ser desenhadas grupos hidroxila, amina, hidrofóbicos, cíclicos, ou mesmo pequenos fragmentos de moléculas). Novos compostos podem, então, ser desenhados de modo que relevantes grupos funcionais se localizem em posições que determinem uma correta relação/posição espacial com o sítio-alvo. Após o desenho deve ocorrer a modelagem de sua estrutura tridimensional, ensaios de *docking*, escolha dos melhores ligantes, predição das bases moleculares/estruturais de sua ligação. Existe uma série de softwares úteis para o desenho de fármacos e abordagens de *screening*, como o Sprout<sup>4</sup>, muito usado para o desenho baseado em fragmentos. Este programa inclui módulos para identificar e selecionar grupos funcionais e posições nos sítios-alvo para formar fragmentos iniciais de compostos para geração de estrutura (módulo EleFANT) e, conforme estes são selecionados, são gerados esqueletos que satisfaçam as restrições estéricas do *pocket*-alvo através de crescimento de fragmentos espaçadores e conectando-os aos fragmentos iniciais (módulo SPIDeR). Por fim, é feita a substituição de átomos no esqueleto até gerar moléculas que sejam compatíveis com as propriedades eletrostáticas do sítio-alvo (módulo Marabou). As soluções podem ser agrupadas e terem escores de ligação calculados usando o módulo ALLigaTOR.

Além do desenho de compostos, pode-se também realizar a busca por compostos que se liguem nos sítios-alvo preditos por meio de *virtual screening* em larga escala (virtual High Throughput Screening- vHTS), através do uso de bancos de dados de estruturas tridimensionais de pequenas moléculas. Um exemplo de banco de dados de estruturas de pequenas moléculas é o ChEMBL<sup>5</sup>, manualmente curado. A versão ChEMBL\_19 (setembro, 2014), possui cerca de 1,64 milhão de compostos que podem ser usados em *screening* contra um sítio-alvo. A simulação da interação

proteína-ligante com o intuito de avaliar a capacidade de um ligante em formar interações fortes e de encaixar-se a um sítio-alvo é chamada *docking*. Algoritmos de *docking* tais como Dock; FlexX; AutoDock; Mol-Dock e Gold podem ser usados para realizar o *docking* na região de interesse para encontrar as moléculas que estabelecem interações mais favoráveis e que se adaptem melhor ao sítio-alvo, sendo escolhidas por meio do ranqueamento do escore de *docking*. É provável que compostos que inibam a função das proteínas envolvidas sejam encontrados dentre os melhores compostos preditos *in silico*.

Através da plataforma Discovery Studio da Accelrys, recentemente adquirida pelo GPBC, o processo de avaliação de candidatos a novos fármacos pode ser expandido para a análise da sua absorção potencial, distribuição, metabolismo e excreção pelo organismo hospedeiro e, finalmente, sua toxicidade ao organismo: *absorption, distribution, metabolism, excretion, toxicity* (ADMET). Uma análise completa, que atribui um fator favorável de ADMET ao ligante em potencial, possibilitará uma seleção mais eficaz dos compostos *leads*, ainda na fase inicial de descoberta e planejamento das novas drogas.

Em adição, pode-se valer do módulo do programa Accelrys que realiza estudos das relações quantitativas entre estrutura química e a função dos compostos escolhidos (Quantitative Structure-Activity Relationship - QSAR). Nesta abordagem estuda-se o processo pelo qual a estrutura e o arranjo de átomos-chave dos compostos *leads* se relacionam de forma quantitativa a um processo biológico como a sua ligação a um receptor ou sua reatividade química. Através da integração entre os processos de descoberta e detalhamento de alvos, estruturação de bibliotecas de compostos químicos, uso de descritores Sting de nano-ambiente que definem o *pocket*-alvo, estudos de ADMET e QSAR, acredita-se que os caminhos traçados podem convergir à descoberta e planejamento de novos fármacos.

Dentre os resultados obtidos pelo GPBC da Embrapa Informática Agropecuária, decorrentes de sua produção técnico-científica, merecem destaque quatro processos patenteados no Instituto Nacional da Propriedade Industrial (INPI):

**1) “Identificação de alvos terapêuticos para desenho computacional de drogas contra bactérias dotadas da proteína PilT”: Controle da *Xylella fastidiosa* em citros, uva e café (BR n. INPI: 020100089068).**

A *Xylella fastidiosa* é uma bactéria gram-negativa e não-flagelada que provoca várias doenças em plantas, tais como a Clorose Variegada dos Citrus (CVC), popularmente conhecida como “amarelinho” e a doença de Pierce, que afetam a citricultura e viticultura, respectivamente. A previsão de produção de laranja no Brasil, safra 2014/2015, está estimada em 289,9 milhões de caixas (de 40,9 quilos) setembro, 2014 (CITRUS BR, 2014)<sup>6</sup>. Desde a sua constatação em 1987, a CVC, ou amarelinho, passou a ser a doença mais importante da cultura dos citros, causando um prejuízo anual de 100 milhões de dólares.

O microrganismo *Xylella fastidiosa* é limitado a persistir apenas colonizando os vasos do xilema, os vasos condutores de água e sais em plantas, e no trato digestório anterior de alguns insetos, como os popularmente conhecidos como cigarrinhas, que se alimentam da seiva das plantas e servem como vetores para carreamento e inserção da bactéria nas plantas (HOPKINS; PURCELL,

<sup>4</sup> Disponível em: <<http://www.simbiosys.ca/sprout/>>.

<sup>5</sup> Disponível em: <<https://www.ebi.ac.uk/chembl/>>.

<sup>6</sup> Disponível em: <<http://www.citrusbr.com/>>.

2002). Há muitos anos tem-se testado o uso de inseticida contra os insetos-vetores desta bactéria, porém não houve suficiente eficácia.

Os mecanismos de virulência da *Xylella fastidiosa* e a forma como esta interage com plantas hospedeiras não são totalmente compreendidos. A explicação mais plausível é a formação de agregados como biofilmes que, ao colonizar vasos do xilema, causa um bloqueio ao fluxo de seiva e culmina nos sintomas da doença. Hopkins e Purcell (2002) sugeriram que a colonização e a patogenicidade da bactéria *Xylella fastidiosa* que levam à doença de Pierce estão estritamente relacionadas com a sua capacidade de se movimentar dentro dos elementos de vaso do xilema, o que permite a colonização de outras regiões da planta.

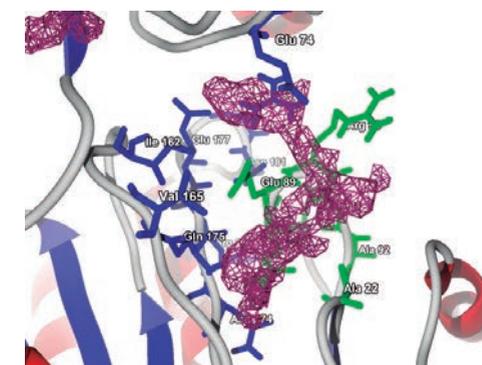
O projeto genoma da *Xylella fastidiosa* (SIMPSON et al., 2000) revelou a presença de genes que codificam proteínas envolvidas na biogênese e função do *type IV pili* (T4P). O *Twitching motility* é uma forma de movimento associado à superfície pelo qual as bactérias puxam-se rapidamente ao longo das superfícies através de ciclos de polimerização e despolimerização do Pilus (T4P). A energia necessária para o movimento é fornecida por meio de hidrólise do ATP por proteínas chamadas PilB e PilT para montagem e desmontagem, respectivamente, do Pilus. A perda de função da proteína PilT ou da PilB resulta na ausência deste tipo de motilidade, associada à privação da extensão ou retração do pilus.

Essa invenção se refere a um método para identificar regiões-alvo existentes na interface de monômeros constituintes da proteína PilT, com o objetivo de desenhar moléculas potencialmente aplicáveis no comprometimento da atividade desta proteína, controlando, assim, processos infecciosos. O método é caracterizado pela:

- 1) Seleção de, pelo menos, uma sequência de aminoácidos constituidora de monômero de PilT.
- 2) Desenvolvimento de um modelo computacional tridimensional da estrutura homo-hexamérica da PilT.
- 3) Análise computacional para determinar os resíduos de aminoácidos formadores de interface (IFR) e suas características físicas, químicas, físico-químicas e estruturais para todas as cadeias dos modelos de complexos hexaméricos gerados.
- 4) Seleção das regiões a serem usadas como alvos terapêuticos e alvos terapêuticos preferenciais na interface entre os monômeros baseada na intensidade dos seguintes parâmetros:
  - Energia de contatos interfaciais.
  - Exposição, em área em complexo.
  - Propensidade do aminoácido-alvo em estabelecer pontes de hidrogênio (como doador ou acceptor).
  - Propensidade do aminoácido-alvo em estabelecer contatos de cunho eletrostático.
  - Preferencialmente, presença em *pockets* (em isolamento ou em complexo).
  - Preferencialmente, que estejam presentes somente em organismos-alvo em relação a organismos não patogênicos.
- 5) Modelagem computacional para o desenho de moléculas potencialmente capazes de efetuar ligações e/ou interações entre regiões-alvo dos monômeros (EMBRAPA INFORMÁTICA AGROPECUÁRIA, 2010).

A pesquisa resultou na patente da metodologia para identificação de alvos terapêuticos específicos para o desenho de novos fármacos em estruturas modeladas da proteína PilT da *Xylella*

*fastidiosa* e que tem possível aplicação para organismos patogênicos que também possuem esta proteína com alto grau de similaridade de sequência e correspondência dos resíduos indicados como alvos em alinhamento de estrutura primária. Resultou também na proposição de uma lista de 54 resíduos-alvos a serem usados em processos de desenho de drogas baseado em estrutura. A Figura 2 representa a estrutural de um dos alvos terapêuticos, o resíduo Glu 89 na cadeia A do complexo XfAa1 (sobreposição estrutural do modelo da PilT de *X. fastidiosa*), mostrando os resíduos adjacentes que formam o nanoambiente no qual o fármaco a ser desenhado se ligará. Este alvo foi escolhido para representar visualmente, pois é de extrema importância por existir unicamente na XfPilT, ausente em microrganismos não patogênicos, contendo pocket adjacente, sendo essa característica muito importante ao desenho de fármaco. A estrutura coberta por uma teia nesta figura (roxo) identifica o volume e posição do ligante que interage com o alvo selecionado.



**Figura 2.** Estrutura de um alvo terapêutico, o resíduo 89 na cadeia A do complexo XfAa1 (sobreposição estrutural do modelo da PilT de *X. fastidiosa*).

Fonte: Embrapa (2010).

## 2) “Método para sugestão de mutantes que aumentam o índice de hidrofobicidade da superfície de proteínas mantendo parâmetros físico-químicos minimamente alterados no sítio catalítico”: Otimização de lipases para produção de biodiesel: (BR n. INPI: 012110000604).

Toda gordura de origem vegetal ou animal é composta, principalmente, de triglicerídeos (uma molécula de glicerol - um tri-álcool - esterificada com três moléculas de ácido graxo) e ácidos graxos livres (AGL). No processo de transesterificação para obtenção de biodiesel, os triglicerídeos presentes no óleo são transformados em moléculas menores de ésteres de ácido graxo (biodiesel) a partir de um agente transesterificante (álcool primário) e um catalisador (base ou ácido). O biodiesel pode ser obtido também usando ácidos graxos livres pelos processos de esterificação em meio preferencialmente ácido e pelo processo de craqueamento.

O álcool mais utilizado na obtenção do biodiesel é o metanol, que promove melhores rendimentos. Considerando que o Brasil é um dos maiores produtores de álcool etílico (etanol) no mundo, há um estímulo para a substituição do metanol pelo etanol, gerando um combustível agrícola totalmente independente do petróleo. A dificuldade na utilização do etanol consiste no fato de que a água é um dos agentes causadores de reações paralelas de saponificação, consumindo o catalisador e reduzindo a eficiência da reação de transesterificação. A utilização de álcool anidro é uma forma de diminuir a formação de sabões, porém, eleva em demasia os custos de produção inviabilizando esse procedimento.

O uso de catalisadores químicos (bases ou ácidos fortes) para a síntese do biodiesel possui algumas desvantagens, como a exigência de baixo teor de ácidos graxos livres e água na matéria-prima do óleo renovável. Além disso, indesejavelmente, ocorre a formação de uma emulsão devido à má solubilidade do álcool nos óleos, o que complica as etapas posteriores do tratamento. Por último, porém não menos importante, a quantidade de álcool a ser usado deve ser muito

maior que a razão molar da reação, e a evaporação/refluxo do álcool em excesso conduzem a um aumento do consumo de energia. Estas desvantagens, somadas à demanda mundial por processos limpos, renováveis e seletivos colocam em destaque a possibilidade do uso de catalisadores biológicos (enzimas) como alternativa para a síntese do biodiesel.

As lipases são enzimas que catalisam a hidrólise de triglicerídeos. Lipases de determinados microrganismos podem catalisar tanto reações de esterificações, bem como reações de transesterificação tendo, como substrato, triglicerídeos e, como reagentes, ácidos graxos de cadeia longa ou álcool primário. A utilização das lipases na produção de biodiesel é relativamente recente, porém tem se mostrado muito promissora devido a uma série de vantagens sobre os catalisadores químicos. As enzimas lipolíticas comerciais em geral foram selecionadas para fins relacionados à indústria de alimentos. Em tais processos, o meio reacional é emulsificado através do uso de detergentes, situação inviável para a produção de biodiesel, pois adiciona passos ao processo que aumentam o custo e o tempo de produção. Há que se considerar ainda que estas enzimas possuem baixa atividade catalítica em meio predominantemente apolar (hidrofóbico), contendo triglicerídeos e álcool como solventes. Como resultado, estabelece-se a necessidade do uso de grandes quantidades de catalisador e longos tempos de reação, o que torna o processo ainda mais oneroso.

Para criar enzimas mais adequadas para tal processo, é hipotetizado que uma enzima lipase com superfície mais hidrofóbica interage melhor com o substrato em um ambiente livre de solventes polares, levando a um rendimento maior na conversão de óleo em biodiesel, quando comparado com lipases naturais.

O método consistiu das seguintes etapas:

- Seleção da proteína de interesse em banco de dados públicos como o Protein Data Bank.
- Seleção de características físico-químicas e estruturais presentes no banco de dados Blue Star Sting.
- Definição de valores limites para cada uma das características selecionadas na etapa dois.
- Uso do módulo JPD do Blue Star Sting para seleção dos aminoácidos cujas características satisfazem os valores de cada um dos descritores selecionados na etapa dois.
- Modelagem por homologia com software Modeller de mutantes com mutações singulares, modificando cada um dos aminoácidos selecionados na etapa quatro, por resíduo de Valina (Alanina).
- Monitoramento da variação das propriedades físico-químicas e/ou estruturais dos aminoácidos que compõem o sítio catalítico, através de geração de arquivos em formato TGZ pelo servidor do Blue Star Sting.
- Seleção dos melhores mutantes com base no escore V4.
- Construção de modelos com mutações múltiplas baseadas nos melhores valores de V4 para as mutações singulares estudadas, utilizando o Modeller.
- Avaliação dos modelos com mutações múltiplas pelo escore V4, comparando com a estrutura nativa.
- Medida da variação da área de superfície hidrofóbica em relação à estrutura nativa da enzima selecionada visando a maximização do parâmetro SHI (EMBRAPA INFORMÁTICA AGROPECUÁRIA, 2011).

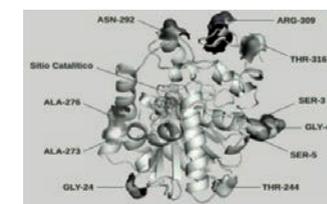
As pesquisas desenvolvidas pelo GPBC possibilitaram que a Embrapa pleiteasse junto ao INPI o registro da patente do método de engenharia de proteínas, no qual são identificadas regiões da superfície de enzimas como mostra a Figura 3, passíveis de mutações que interferem de maneira reduzida nas propriedades físicas, químicas e estruturais dos aminoácidos do sítio catalítico, com o objetivo de criar mutantes que apresentem a superfície macromolecular mais hidrofóbica e, portanto, serem empregadas para a obtenção de biodiesel por catálise biológica.

### 3) “Inibidores das enzimas poligalacturonases de fungos fitopatogênicos”: Fungicidas contra *Fusarium* e outros fungos patogênicos (BR n. INPI: 02012 000 3126).

Grande parte das doenças de plantas que causam prejuízos para a agricultura brasileira e mundial é causada por fungos fitopatogênicos, e a maioria deles de solo. Seria importante a redução do potencial patogênico destes fungos em áreas infestadas. Uma medida que tem sido utilizada é a de incorporação de matéria orgânica no solo, já que a introdução de antagonistas é uma medida de controle biológico. Entretanto, o índice de controle obtido com este método, isoladamente, está abaixo do necessário para impedir danos à cultura. O uso e desenvolvimento de cultivares resistentes seria uma melhor opção de controle destas doenças; todavia, muitos hospedeiros não apresentam resistência a esses patógenos. A funcionalidade nem sempre é possível, devido à inexistência no mercado de cultivares com todas as características desejadas. Existem métodos de controle químico contra os fungos fitopatogênicos, como, por exemplo, o realizado, até há pouco tempo, com um agrotóxico de amplo espectro, o brometo de metila utilizado nos últimos 60 anos como fumigante de solo em pré-plantio. Embora altamente eficaz, rápido, de fácil penetração no solo, amplo espectro e baixa resistência dos fungos, foi comprovado que este confere riscos para o ambiente, para o homem e para a camada de ozônio.

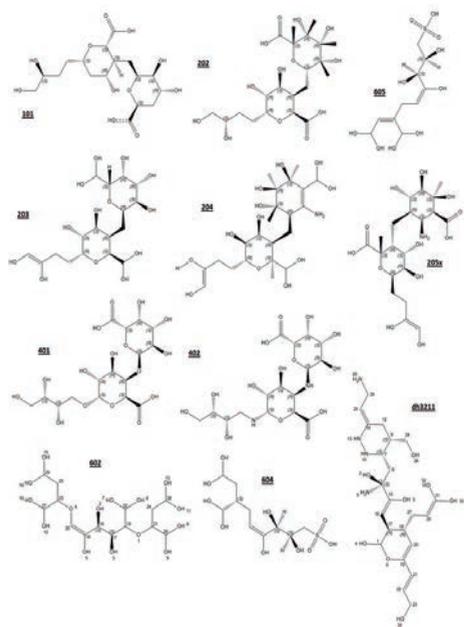
A invenção que descrevemos em seguida para ilustrar o trabalho do GPBC da Embrapa nessa área refere-se ao desenho computacional de novos compostos com potencial inibitório para enzimas endopoligalacturonases (PG) de fungos fitopatogênicos, com o intuito de evitar ou diminuir a colonização desses microrganismos nos tecidos vegetais. A PG integra um grupo de enzimas secretadas por microrganismos fitopatogênicos durante o processo de invasão dos tecidos vegetais, participando na catálise da hidrólise da pectina, culminando na desestruturação do arcabouço da parede celular, o que favorece a invasão de hifas dos fungos. Uma gama de microrganismos fitopatogênicos utiliza essas enzimas como fatores de patogenicidade que levam a doenças em uma grande variedade de plantas de interesse econômico como o trigo, a cevada, o tomate, o morango, a manga, o arroz, a cana-de-açúcar, dentre outros.

Com o objetivo de minimizar as perdas causadas por estes patógenos, foram desenhadas pequenas moléculas planejadas (Figura 4) para se ligarem com alta afinidade aos resíduos do sítio de ligação ao substrato desses patógenos. O ligante dh3211 posicionado no sítio-alvo da 1HG8 (PDB) (Figura 5), é o ligante que apresentou melhor resultado de interação entre os resíduos específicos de PG fúngicas: D194, H188 e G305, tanto na 1HG8 quanto na 2IQ7. Possui grupos químicos que conferem polaridades diferentes para cada extremidade da estrutura desta molécula o que o faz encaixar na região do sítio ligante e do catalítico através de complementaridade eletrostática.



**Figura 3.** Indicação de posição tridimensional dos aminoácidos substituídos por Valina na superfície de proteína nativa (1TCB.pdb).

Fonte: Embrapa Informática Agropecuária (2011).



**Figura 4.** Representação estrutural das moléculas desenhadas computacionalmente que obtiveram melhores resultados nos *dockings* com as PG 1HG8 e 2IQ7 (PDB) e que, provavelmente, possuirão caráter inibitório para as enzimas PG.

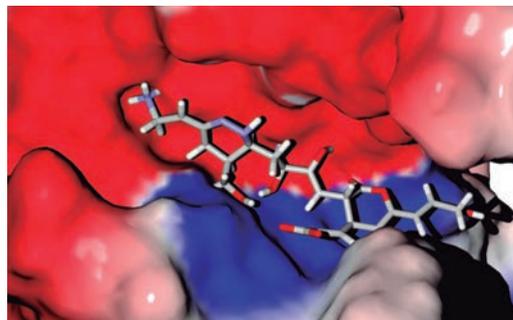
Fonte: Embrapa Informática Agropecuária (2012b).

anteras, maturação do grão de pólen, amadurecimento do fruto, formação das células do xilema e crescimento do tubo polínico (EMBRAPA, 2012).

#### 4) “Inibidores das enzimas alfa amilases de insetos”: Controle de insetos que prejudicam armazenamento de sementes (BR n. INPI: 02012 000 3145).

Os grãos estão entre os principais constituintes da base alimentar dos povos. Já sua produção é sazonal e em muitos lugares as colheitas ocorrem apenas uma vez ao ano. Para que seja possível alimentar todas as pessoas do mundo é necessário que a maior parte da produção dos principais grãos (arroz, trigo, sorgo, milho, soja e painço) seja estocada por até mais que um ano.

Para controle e minimização das perdas de grãos estocados, já existem várias estratégias, tais como desinfestação térmica, baixa umidade, baixo teor de oxigênio, resfriamento e armazenamento hermético, por exemplo, que são delineadas segundo o clima da região, tipo de grão e tipo de inseto. Ainda assim, há grandes perdas de produtos estocados causadas por pragas, que perfazem cerca de 10% da produção total, e pelo menos 50% dessa perda devem-se a insetos. Tomando o trigo como exemplo, cuja produção é estimada para a safra 2013-2014 em 711,42 milhões de toneladas, das quais 182,78 milhões de toneladas deverão ser estocadas, e considerando que o preço da tonelada seja de US\$ 496,21 em 16 de setembro de 2014 (CUSTÓDIO, 2014), com perdas de 10% desse produto em estoque alcança-se um prejuízo em torno de US\$ 9,07 bilhões (US\$ 4,13 bilhões na safra 2009-2010 por ocasião da solicitação da patente), sendo metade disso devido a insetos. Se considerarmos a crescente demanda anual de estoque de grãos



**Figura 5.** Ligante dh3211 posicionado no sítio-alvo da 1HG8 (PDB).

Fonte: Embrapa Informática Agropecuária (2012b).

A região catalítica é predominantemente negativa e na imagem aparece na cor vermelha; a região ligante, majoritariamente positiva, aparece na imagem como azul; as regiões em rosa claro são apolares.

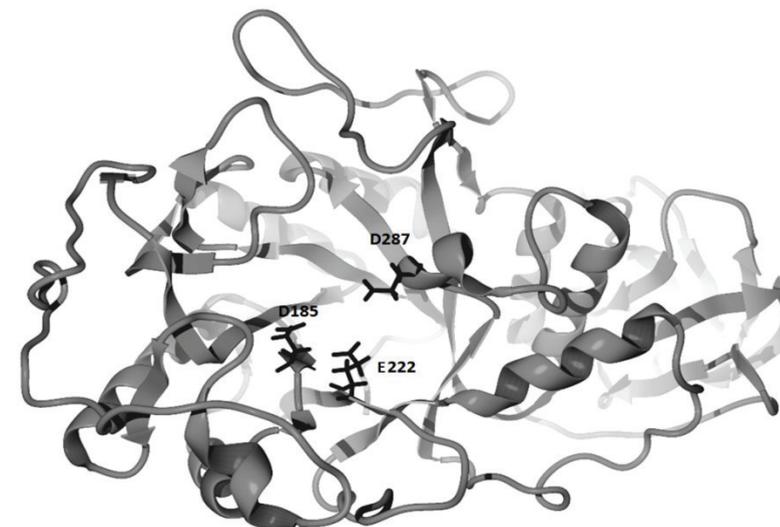
Além disso, outras funções das PG vegetais poderão ser manipuladas usando os compostos desenvolvidos neste trabalho, caso seja confirmada sua ação inibitória sobre as PG de plantas, tais como: processos de separação de células, germinação, abscisão de órgãos, deiscência das

em resposta à crescente produção agrícola e somarmos os prejuízos de todos os outros grãos, pode-se chegar a um valor exorbitante.

Logo, o interesse em reduzir perdas de grãos durante o período de estocagem é, sem dúvida nenhuma, tanto de países produtores como daqueles que importam e estocam.

A presente invenção que ilustra o trabalho do GPBC da Embrapa nesta área específica refere-se ao desenho computacional de novos compostos com potencial inibitório para enzimas alfa-amilases de insetos, com o intuito de diminuir os danos causados por eles, principalmente em produtos estocados.

Com o objetivo de diminuir as perdas agrícolas estocadas causadas pela ação de insetos, foram desenhadas estruturas que se ligam à alfa-amilase com alta afinidade teórica. Na Figura 6 podemos identificar a tríade catalítica da alfa-amilase do *Tenebrio molitor* (TMA, 1JAE.pdb) com os resíduos de Aspartato (D185 e D287) e Glutamato (E222) destacados em preto onde deve se ligar as estruturas inibidoras da enzima. Dessa forma, os insetos cuja alfa-amilase foi inibida não conseguem obter os resíduos de açúcar necessários para obtenção de energia e, em consequência, morrem de privação energética (EMBRAPA, 2012).



**Figura 6.** Visão da tríade catalítica da alfa-amilase do *Tenebrio molitor* (TMA, 1JAE.pdb) com os resíduos de Aspartato (D185 e D287) e Glutamato (E222) destacados em preto. Imagem gerada pelo Molegro Virtual Docker.

Fonte: Embrapa Informática Agropecuária (2012a).

### 3.1 Infraestrutura computacional do GPBC

O GPBC dispõe de uma infraestrutura computacional constituída de vários servidores convencionais e um servidor HPC, SGI UV200 com 128 cores, 1Tb de RAM e 72 Tb de HD que agiliza o uso das técnicas de High throughput in silico screening, drug design, ADMET, QSAR, Ludi e também possibilita buscas mais rápidas (em memória) em banco de dados Sting\_RDB expandindo a capacidade de identificação dos novos alvos para planejamento dos agroquímicos e drogas/fármacos.

Os softwares disponíveis para as pesquisas em Biologia Computacional pelo Grupo de Pesquisa são:

- Sting (Sting na versão: Blue Star Sting) - plataforma consolidada para análise de relação entre as estruturas macromoleculares e suas funcionalidades.
- Molegro - plataforma para planejamento e desenho de novos fármacos/drogas/agroquímicos que também tem forte vertente na linha de mineração de dados estruturais e é usado no laboratório de GPBC para *virtual screening* e *docking* das moléculas com alta precisão.
- Accelrys Discovery Studio - plataforma completa para análise estrutural das proteínas com módulos para varredura de alto desempenho embasada no processo de atracamento dos compostos químicos nas proteínas-alvos como também o planejamento de fármacos, análise in silico da sua potencial toxicidade, absorção, distribuição, metabolismo e excreção (ADMET)<sup>7</sup>.
- Softwares: Yassara, NAMD, VMD, Gromacs etc.

## 4 Considerações finais

O planejamento e desenvolvimento de novos fármacos e agroquímicos, que servem para proteção de plantas, animais e humanos, é um processo bastante complexo e longo, demandando investimento contínuo para que se possa chegar ao resultado/produto com potencial para ser oferecido ao mercado.

Convicta da importância de o Brasil diminuir a dependência de importação de matérias ativas para a formulação de agro defensivo e passar a produzir no País esses princípios ativos, o Grupo de Pesquisa em Biologia Computacional da Embrapa já deu os primeiros passos neste sentido e patenteou algumas tecnologias com as quais se pode desenhar os compostos químicos identificados com elevado potencial de atividade in silico contra alvos proteicos encontrados em patógenos.

Agora, é preciso avançar, na direção da síntese desses compostos desenhados, acompanhado de ensaios in vitro e in vivo, devendo, necessariamente, ocorrer uma interação permanente e retroalimentar entre o desenvolvimento in vitro e in vivo com a análise teórica (in silico) no sentido de aperfeiçoar o desenho do agro defensivo.

O produto desse esforço será um inventário de novos agroquímicos e fármacos que pode ser um dos mais importantes desafios perante o agronegócio do país, representando um dos fatores fundamentais de conquista de vantagem competitiva no mercado internacional em futuro próximo.

No final deste caminho, será necessário investir esforços no preparo do processo de produção dos novos compostos químicos em escala piloto e, posteriormente, em larga escala industrial.

<sup>7</sup> Disponível em: <<http://accelrys.com/products/discovery-studio/admet.html>>.

## 5 Referências

- ACCELRYS SOFTWARE INC. **Discovery Studio Visualizer version 4.0**. San Diego, CA, 2014.
- ANDERSON, A. C. The process of structure-based drug design. **Chemistry & Biology**, Cambridge, Mass., v. 10, n. 9, p. 787-797, Sept. 2003. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1074552103001947>>. Acesso em: 07 out. 2014.
- BORRO, L. C.; OLIVEIRA, S. R. M.; YAMAGISHI, M. E.; MANCINI, A. L.; JARDINE, J. G.; MAZONI, I.; SANTOS, E. H. dos; HIGA, R. H.; KUSER, P. R.; NESHICH, G. Predicting enzyme class from protein structure using Bayesian classification. **Genetics and Molecular Research**, Ribeirão Preto, v. 5, n. 1, p. 193-202, Mar. 2006. Disponível em: <[http://www.funpecrp.com.br/gmr/year2006/vol1-5/xm0009\\_full\\_text.htm](http://www.funpecrp.com.br/gmr/year2006/vol1-5/xm0009_full_text.htm)>. Acesso em: 07 out. 2014.
- BRAGHINI, C. A.; NESHICH, I. A. P.; NESHICH, G.; SOARDI, F. C.; MELLO, M. P. de; COSTA, V. P.; VASCONCELLOS, J. P. C. de; MELO, M. B. de. New mutation in the *myocilin* gene segregates with juvenile-onset open-angle glaucoma in a Brazilian family. **Gene**, Amsterdam, v. 523, n. 1, p. 50-57, July, 2013. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0378111913003788>>. Acesso em: 07 out. 2014.
- CITRUS BR. **Estimativa total de safra 2014/2015 e projeção de estoques - maio/2014**. Disponível em: <<http://www.citrusbr.com>>. Acesso em: 07 out. 2014.
- CUSTÓDIO, F. Trigo: USDA reporta aumento na produção mundial e preços recuam na CBOT. **Notícias Agrícolas**, Campinas, 11 jun. 2014. Disponível em: <<http://www.noticiasagricolas.com.br/noticias/usda/140724-trigo-usda-reporta-aumento-na-producao-mundial-e-precos-recuam-na-cbot.html#.VrSafldUaB>>. Acesso em: 07 out. 2014.
- DIAS-LOPES, C.; NESHICH, I. A. P.; NESHICH, G.; ORTEGA, J. M.; GRANIER, C.; CHÁVEZ-OLORTEGUI, C.; MOLINA, F.; FELICORI, L. Identification of new Sphingomyelinases D in pathogenic fungi and other pathogenic organisms. **Plos One**, San Francisco, v. 8, n. 11, Nov. 2013. Disponível em: <<http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0079240>>. Acesso em: 07 out. 2014.
- EDGAR, R. C. Muscle: multiple sequence alignment with high accuracy and high throughput. **Nucleic Acids Research**, Oxford, v. 32, n. 5, p. 1792-1797, Mar. 2004. Disponível em: <<http://nar.oxfordjournals.org/content/32/5/1792.full.pdf+html>>. Acesso em: 07 out. 2014.
- EMBRAPA INFORMÁTICA AGROPECUÁRIA. G. Neshich; I. A. Neshich; I. Mazoni; J. G. Jardine; J. A. Salim; F. R. de Moraes. **Método para previsão de mutantes que aumentem o índice de hidrofobicidade da superfície de proteínas**. BR n. PI1211604, 04 ago. 2011.
- EMBRAPA INFORMÁTICA AGROPECUÁRIA G. Neshich; I. A. P. Neshich; L. Nishimura; J. A. Salim; I. Mazoni; J. G. Jardine. **Identificação de alvos terapêuticos para desenho computacional de drogas contra bactérias dotadas da proteína PilT**. 2010, Brasil. BR n. PI 1089068, 28 out. 2010.
- EMBRAPA INFORMÁTICA AGROPECUÁRIA. G. Neshich; J. G. Jardine; I. Mazoni; I. A. P. Neshich; L. Nishimura. **Desenho computacional para novos inibidores de alfa-amilases**. BR n. PI 0003126, 09 maio 2012a.
- EMBRAPA INFORMÁTICA AGROPECUÁRIA. G. Neshich; J. G. Jardine; I. Mazoni; I. A. P. Neshich; L. Nishimura. **Inibidores das enzimas poligalacturanases de fungos fitopatogênicos**. BR n. PI 0003145, 02 fev. 2012b.
- FERNANDEZ, J. H.; HAYASHI, M. A.; CAMARGO, A. C.; NESHICH, G. Structural basis of the lisinopril-binding specificity in N- and C-domains of human somatic ACE. **Biochemical and Biophysical Research Communications**, Orlando, v. 308, n. 2, p. 219-226, Aug. 2003. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/12901857>>. Acesso em: 07 out. 2014.
- FREITAS, S. M. de; MELO, L. V. de; SILVA, M. C. M. da; VRIEND, G.; NESHICH, G.; VENTURA, M. M. Analysis of the black-eyed pea trypsin and chymotrypsin inhibitor alpha-chymotrypsin complex. **FEBS Letters**, Amsterdam, v. 409, n. 2, p. 121-127, June, 1997.

GUEDES, I. A.; MAGALHÃES, C. S. de; DARDENNE, L. E. Receptor-ligand molecular docking. **Biophysical Reviews**, Heidelberg, 2014. v. 6, n. 1, p. 75-87, Mar. 2014.

HOPKINS, D. L.; PURCELL, A. H. *Xylella fastidiosa*: cause of Pierce's disease of grapevine and other emergent diseases. **Plant Disease**, Saint Paul, v. 86, n. 10, p. 1056-1066, Oct. 2002. Disponível em: <<http://apsjournals.apsnet.org/doi/pdf/10.1094/PDIS.2002.86.10.1056>>. Acesso em: 07 out. 2014.

KUNTZ, I. D.; BLANEY, J. M.; OATLEY, S. J.; LANGRIDGE, R.; FERRIN, T. A geometric approach to macromolecular-ligand interactions. **Journal of Molecular Biology**, London, v. 161, n. 2, p. 269-288, Oct. 1982.

LARKIN, M. A.; BLACKSHIELDS, G.; BROWN, N. P.; CHENNA, R.; McGETTIGAN, P. A.; McWILLIAN, H.; VALENTIN, F.; WALLACE, I. M.; WILM, A.; LOPEZ, R.; THOMPSON, J. D.; GIBSON, T. J.; HIGGINS, D. G. Clustal W and Clustal X version 2.0. **Bioinformatics**, Oxford, v. 23, n. 21, p. 2947-2948, Nov. 2007.

MAGALHÃES, C. S. da. **Algoritmos genéticos para o problema de docking proteína-ligante**. 2006. 261 f. Tese (Doutorado em Modelagem Computacional) – Laboratório Nacional de Computação Científica - LNCC, Petrópolis. Disponível em: <[http://www.gmmsb.lncc.br/pdf/T\\_Magalhaes2006.pdf](http://www.gmmsb.lncc.br/pdf/T_Magalhaes2006.pdf)>. Acesso em: 07 out. 2014.

MAGALHÃES, C. S. de; BARBOSA, H. J. C.; DARDENNE, L. E. Métodos de docking receptor-ligante para o desenho racional de compostos bioativos. In: MORGON, N. H.; COUTINHO, K. (Org.). **Métodos de química teórica e modelagem molecular**. São Paulo: Editora Livraria da Física, 2007. p. 489-531.

MANCINI, A. L.; HIGA, R. H.; OLIVEIRA, A.; DOMINQUINI, F.; KUSER, P. R.; YAMAGISHI, M. E.; TOGAWA, R. C.; NESHICH, G. STING Contacts: a web-based application for identification and analysis of amino acid contacts within protein structure and across protein interfaces. **Bioinformatics**, Oxford, v. 20, n. 13, p. 2145-2147, Sep. 2004.

MARCELLINO, L. H.; NESHICH, G.; SÁ, M. F. G. de; KREBBERS, E.; GANDER, E. S. Modified 2S albumins with improved tryptophan content are correctly expressed in transgenic tobacco plants. **FEBS Letters**, Amsterdam, v. 385, n. 3, p. 154-158, May 1996.

NOTREDAME, C.; HIGGINS, D. G.; HERINGA, J. T-Coffee: A novel method for fast and accurate multiple sequence alignment. **Journal of Molecular Biology**, London, v. 302, n. 1, p. 205-217, Sept. 2000.

NESHICH, G.; BORRO, L. C.; HIGA, R. H.; KUSER, P. R.; YAMAGISHI, M. E.; FRANCO, E. H.; KRAUCHENCO, J. N.; FILETO, R.; RIBEIRO, A. A.; BEZERRA, G. B.; VELLUDO, T. M.; JIMENEZ, T. S.; FURUKAWA, N.; TESHIMA, H.; KITAJIMA, K.; BAVA, A.; SARAI, A.; TOGAWA, R. C.; MANCINI, A. L. The Diamond STING server. **Nucleic Acids Research**, Oxford, v. 33, Web server issue, p. W29-W35, July 2005.

NESHICH, G.; MAZONI, I.; OLIVEIRA, S. R.; YAMAGISHI, M. E.; KUSER-FALCÃO, P. R.; BORRO, L. C.; MORITA, D. U.; SOUZA, K. R.; ALMEIDA, G. V.; RODRIGUES, D. N.; JARDINE, J. G.; TOGAWA, R. C.; HIGA, R. H.; CRUZ, S. A.; VIEIRA, F. D.; SANTOS, E. H.; MELO, R. C.; SANTORO, M. M. The Star STING Server: A multiplatform environment for protein structure analysis. **Genetics and Molecular Research**, Ribeirão Preto, v. 5, n. 4, p. 717-722, 2006.

NESHICH, G.; ROCCHIA, W.; MANCINI, A. L.; YAMAGISHI, M. E. B.; KUSER, P. R.; FILETO, R.; BAUDET, C.; PINTO, I. P.; MONTAGNER, A. J.; PALADRANI, J. F.; KRAUCHENCO, J. N.; TORRES, R. C.; SOUZA, S.; TOGAWA, R. C.; HIGA, R. H. Java Protein Dossier: a novel web-based data visualization tool for comprehensive analysis of protein structure. **Nucleic Acids Research**, Oxford, v. 32, Web server issue, p. W595-W601, 2004.

NESHICH, G.; TOGAWA, R. C.; MANCINI, A. L.; KUSER, P. R.; YAMAGISHI, M. E.; PAPPAS JUNIOR, G.; TORRES, W. V.; CAMPOS, T. F. e; FERREIRA, L. L.; LUNA, F. M.; OLIVEIRA, A. G.; MIURA, R. T.; INOUE, M. K.; HORITA, L. G.; SOUZA, D. F. de; DOMINQUINI, F.; ALVARO, A.; LIMA, C. S.; OGAWA, F. O.; GOMES, G. B.; PALADRANI, J. F.; SANTOS, G. F. dos; FREITAS, E. M. de; MATTIUIZ, A. R.; COSTA, I. C.; ALMEIDA, C. L. de; SOUZA, S.; BAUDET, C.; HIGA, R. H. STING Millennium: a Web based suite of programs for comprehensive and simultaneous analysis of protein structure and sequence. **Nucleic Acids Research**, Oxford, v. 31, n. 13, p. 3386-92, July 2003.

PORTER, C. T.; BARTLETT, G. J.; THORNTON, J. M. The Catalytic Site Atlas: a resource of catalytic sites and residues identified in enzymes using structural data. **Nucleic Acids Research**, Oxford, v. 32, database issue, D129-133, Jan. 2004.

PyMOL MOLECULAR GRAPHICS SYSTEM. 2014. Disponível em: <<http://sourceforge.net/projects/pymol/>>. Acesso em: 18 set. 2014.

SIMÕES, M.; BAHIA, D.; ZERLOTINI, A.; TORRES, K.; ARTIQUENAVE, F.; NESHICH, G.; KUSER, P.; OLIVEIRA, G. Single nucleotide polymorphisms identification in expressed genes of *Schistosoma mansoni*. **Molecular and Biochemical Parasitology**, Amsterdam, v. 154, n. 2, p. 134-140, Aug. 2007.

SIMPSON, A. J. G. et al. The genome sequence of the plant pathogen *Xylella fastidiosa*. **Nature**, London, v. 406, p. 151-157, July 2000.

SRIDHARAN, S.; NICHOLLS, A.; HONIG, B. A. New vertex algorithm to calculate solvent accessible surface areas. **Biophysical Journal**, New York, v. 61, A174, 1992.

YANG, S. Y. Pharmacophore modeling and applications in drug discovery: challenges and recent advances. **Drug Discovery Today**, Kidlington, v. 15, n. 11/12, p. 444-450, June 2010. Disponível em: <[http://csmres.co.uk/cs.public.upd/article-downloads/Most%20downloaded%20article%20Q2%202010\\_a15584.pdf](http://csmres.co.uk/cs.public.upd/article-downloads/Most%20downloaded%20article%20Q2%202010_a15584.pdf)>. Acesso em: 18 set. 2014.



# Parte III

## TIC nos recursos naturais e mudanças climáticas





## TIC em agrometeorologia e mudanças climáticas

José Eduardo Boffino de Almeida Monteiro  
Aryeverton Fortes de Oliveira  
Alan Massaru Nakai

### 1 Introdução

A agricultura é a atividade econômica mais dependente das condições climáticas. Além de influenciar o crescimento, o desenvolvimento e a produtividade das culturas, o clima afeta também a relação das plantas com insetos e microrganismos, favorecendo ou não a ocorrência de pragas e doenças (SENTELHAS; MONTEIRO, 2009). Muitas práticas agrícolas de campo, como o preparo do solo, a semeadura, a adubação, a irrigação, as pulverizações, a colheita, dentre outras, dependem também de condições específicas de tempo e de umidade no solo, para que possam ser realizadas de forma adequada (PEREIRA et al., 2002).

A preocupação crescente com o aumento da população mundial, com a degradação dos recursos naturais, com as mudanças climáticas e com a sustentabilidade da agricultura tem exigido esforços no desenvolvimento de práticas agrícolas mais sustentáveis, inclusive a partir do melhor entendimento das relações entre a agricultura e o clima. Por isso, instituições governamentais, ligadas à agricultura e ao meio ambiente, têm buscado desenvolver ferramentas de informações agrometeorológicas que auxiliem no planejamento e no processo de tomada de decisão na produção agrícola, buscando maior produtividade, maior resiliência dos sistemas produtivos e menor impacto ambiental (SIVAKUMAR; MOTHAR, 2007).

Informações agrometeorológicas são aquelas que consideram os dados meteorológicos associados às necessidades dos cultivos e dos rebanhos, com o objetivo de estimar os respectivos impactos sobre estes, bem como sobre as práticas agrícolas. Por isso, os serviços de meteorologia e agrometeorologia têm muito a contribuir com a economia de seus respectivos países, por meio da divulgação e do uso eficiente de tais informações (RIJKS; BARADAS, 2000). De acordo com Sentelhas e Monteiro (2009), as informações agrometeorológicas podem ser classificadas em três graus ou níveis. As informações são de primeiro grau, quando são dados meteorológicos puros ou derivados de cálculos simples, como o balanço hídrico. De segundo grau, quando são produzidas a partir de dados meteorológicos e parâmetros específicos da cultura, indicando o estado ou a resposta da cultura à condição meteorológica observada. De terceiro grau quando indicam, além do estado ou resposta da cultura, a ação de manejo correspondente.

Nesse contexto, os recursos da área de tecnologia da informação (TI) vêm ampliando continuamente as possibilidades do monitoramento meteorológico, melhorando aplicações para transmis-

são, armazenamento e análise de dados, assim como para a produção e difusão de informações de forma didática e acessível.

## 2 Monitoramento agrometeorológico

O monitoramento agrometeorológico consiste na coleta sistemática e contínua de dados meteorológicos para a produção de informações de interesse ou uso agrícola. Sistemas que integram de forma coordenada e simultânea as funções de coleta, transmissão e processamento de dados podem fornecer informações agrometeorológicas atualizadas em tempo quase real.

Existem várias práticas agrícolas que podem se beneficiar de informações agrometeorológicas, destacando-se: o preparo do solo, a semeadura, a adubação, a irrigação, o controle fitossanitário, a colheita etc. Estimativas de produtividade, de qualidade da produção e de favorabilidade à ocorrência de doenças também necessitam de dados meteorológicos. Por isso, informações atualizadas são essenciais para o processo de tomada de decisão, ou seja, para a conversão das informações disponíveis em uma determinada ação visando maximizar a produtividade ou melhorar o aproveitamento de insumos. Decisões corretas representam maior eficiência e eficácia do sistema produtivo (SENTELHAS; MONTEIRO, 2009).

Atualmente, o Brasil conta com diversos sistemas de informações agrometeorológicas em operação, disponibilizando, basicamente, as informações de primeiro grau e algumas de segundo grau. Os mais conhecidos são o Agritempo<sup>1</sup>, da Empresa Brasileira de Pesquisa Agropecuária (Embrapa) e o Sisdagro<sup>2</sup>, do Instituto Nacional de Meteorologia (Inmet), ambos de abrangência nacional, o Ciiagro<sup>3</sup>, do Instituto Agrônomo de Campinas (IAC), no Estado de São Paulo, o Sistema de Informações Agrometeorológicas do Instituto Agrônomo do Paraná (Iapar)<sup>4</sup>, no Estado do Paraná. Existem, ainda, outros exemplos, como o Centro de Informações de Recursos Ambientais e de Hidrometeorologia de Santa Catarina (Ciram)<sup>5</sup> e a Fundação Cearense de Meteorologia e Recursos Hídricos<sup>6</sup>. Nestes sistemas, as informações contemplam diversos aspectos dos efeitos do tempo e clima na agricultura que auxiliam na previsão de safra, na definição das melhores épocas de plantio, na indicação das condições para manejo do solo, para irrigação e controle fitossanitário.

Uma das características em comum desses sistemas é que seu principal veículo de distribuição de informações são sítios na rede mundial de computadores. Desde o seu surgimento até o presente, a internet evoluiu rapidamente para se tornar um dos principais meios de comunicação do mundo contemporâneo. Várias vantagens comparativas levaram a isso como, por exemplo, os baixos custos, a interatividade, a possibilidade de integração de recursos audiovisuais, recursos gráficos, aplicativos e, em muitos países e regiões, ampla cobertura. No Brasil, no entanto, a internet ainda apresenta disponibilidade muito restrita para o grande público, tanto pela falta

de cobertura nos locais mais distantes dos grandes centros urbanos quanto pela dificuldade de acesso pelas classes mais baixas. É o que indica a pesquisa TIC Domicílios, realizada entre 2013 e 2014 pelo Centro de Estudos sobre as Tecnologias da Informação e da Comunicação (COMITÊ GESTOR DA INTERNET NO BRASIL, 2014). De acordo com a pesquisa, nas áreas urbanas, a proporção de lares com acesso à internet é de 48%, enquanto que nas áreas rurais é de 15%.

Outra característica em comum dos sistemas de informações agrometeorológicas citados é o uso da TI no gerenciamento de bancos de dados, no processamento de dados e na produção da informação, seja na forma de gráficos, mapas ou textos. Atualmente, não é mais possível dissociar os serviços de monitoramento agrometeorológico do uso intensivo de recursos de TI.

Dentre os sistemas citados, o Agritempo (EMBRAPA INFORMÁTICA AGROPECUÁRIA, 2014) não possui uma rede própria de estações meteorológicas de monitoramento. Na realidade, ele integra, em uma única base, os dados das redes de estações meteorológicas de diversas instituições diferentes, nacionais e estaduais (Inmet, CPTEC, ANA, Cemig, entre outras). Com isso, o sistema consegue compor uma base de dados com maior densidade de pontos de observação e com abrangência nacional. Operando desde 2003, o sistema foi atualizado em 2014 para uma versão mais interativa, com mais funcionalidades e de acordo com o conceito de Web 2.0 e com a opção de acesso por meio de dispositivos móveis. Algumas das tecnologias envolvidas no desenvolvimento da primeira versão do sistema Agritempo foram mantidas, como o Sistema Gerenciador de Banco de Dados PostgreSQL<sup>7</sup> e a linguagem Java. Por outro lado, outras como Ajax<sup>8</sup> e o framework DojoToolkit<sup>9</sup> foram incorporadas à nova versão. O uso da tecnologia Ajax permitiu tornar as páginas na internet mais interativas, utilizando requisições assíncronas para recuperação de informações. Já o framework DojoToolkit foi utilizado para facilitar o desenvolvimento de interfaces para páginas de internet multiplataforma.

As principais funcionalidades do sistema Agritempo estão agrupadas em “Monitoramento” e “Previsão”. Na categoria de Monitoramento, diversas variáveis medidas e calculadas podem ser visualizadas de forma espacializada, como os mapas nacionais apresentados na Figura 1. As mesmas variáveis podem também ser visualizadas em mapas estaduais, com maior detalhamento. As variáveis espacializadas medidas e calculadas são: estiagem, estiagem agrícola, disponibilidade de água no solo, necessidade de reposição por chuva, precipitação, evapotranspiração, temperatura mínima, máxima e média.

Na categoria de “Previsão”, as variáveis e indicadores derivados são obtidos a partir de modelos meteorológicos de previsão do tempo, com condições previstas com até dois dias de antecedência (48h), como ilustrado na Figura 2. Dessa forma, a partir de uma consulta ao sistema, o usuário pode decidir e planejar a execução ou não de determinadas atividades agrícolas para os próximos dois dias. As principais variáveis e informações derivadas da Previsão do Tempo disponíveis para consulta no sistema Agritempo são: condições para tratamento fitossanitário, necessidade de irrigação, condições para manejo do solo, condições para colheita, precipitação e temperatura mínima, máxima e média.

<sup>1</sup> Disponível em: <<http://www.agritempo.gov.br>>.

<sup>2</sup> Disponível em: <<http://sisdagro.inmet.gov.br>>.

<sup>3</sup> Disponível em: <<http://www.ciiagro.sp.gov.br>>.

<sup>4</sup> Disponível em: <<http://www.iapar.br>>.

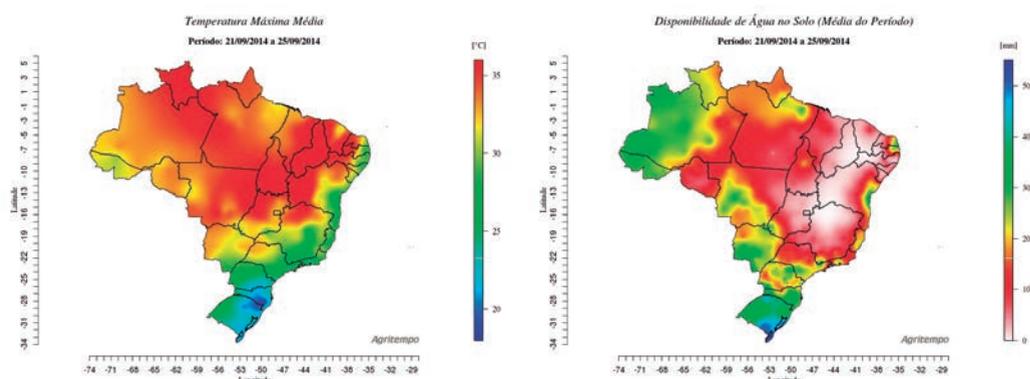
<sup>5</sup> Disponível em: <<http://ciram.epagri.sc.gov.br>>.

<sup>6</sup> Disponível em: <<http://www.funceme.br>>.

<sup>7</sup> Disponível em: <<http://www.postgresql.org/>>.

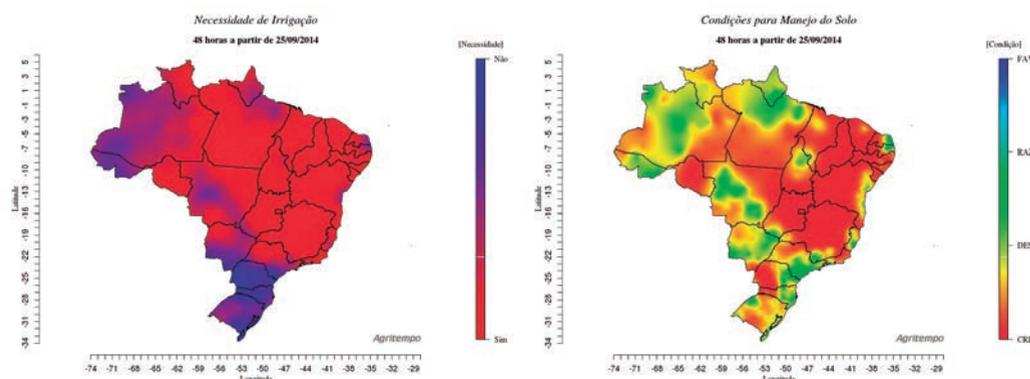
<sup>8</sup> Disponível em: <<http://www.w3schools.com/ajax/>>.

<sup>9</sup> Disponível em: <<http://dojotoolkit.org/>>.



**Figura 1.** Dados médios de temperatura máxima e disponibilidade de água no solo nos cinco dias anteriores à consulta em 26/09/2014.

Fonte: Embrapa Informática Agropecuária (2014).



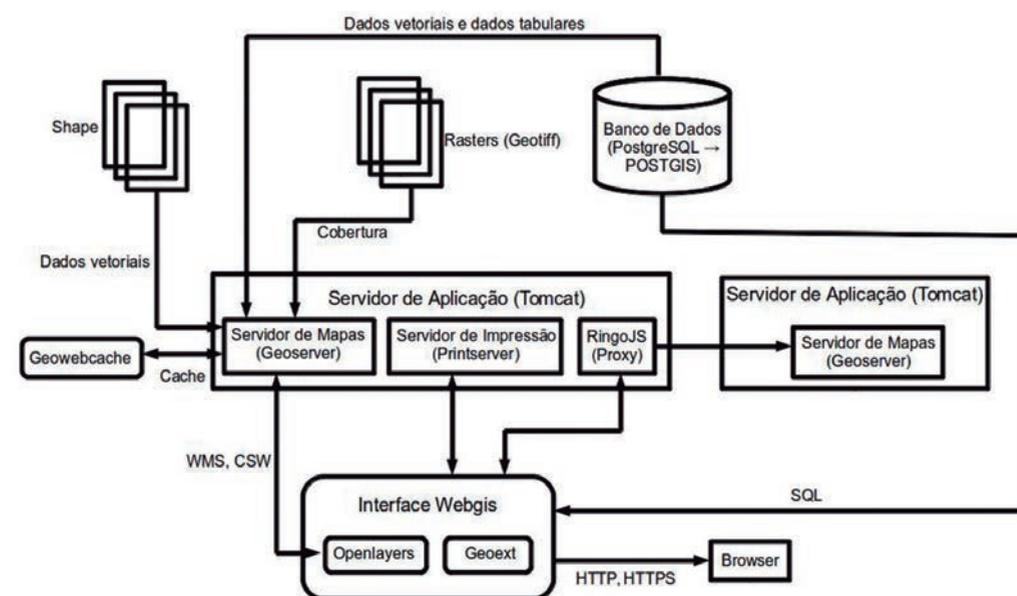
**Figura 2.** Estimativas de necessidade de irrigação e condições para manejo do solo estimados a partir de dados de previsão do tempo.

Fonte: Embrapa Informática Agropecuária (2014).

Além dessas funcionalidades, o sistema também conta com um módulo de visualização de informações geográficas (Webgis), que permite melhorar a produção de mapas temáticos, com variáveis selecionáveis. Essa interface foi desenvolvida utilizando as bibliotecas javascript OpenLayers<sup>10</sup> e Geoext<sup>11</sup>, que permitem aos usuários interagirem com os dados georreferenciados disponíveis na aplicação, possibilitando a solicitação de imagens de mapas para a aplicação servidora de Mapas (Geoserver), que implementa os serviços Web Map Service (WMS) e Catalogue Service for Web (CSW), conforme diagrama da Figura 3. Assim, cada usuário conta com a possibilidade de gerar mapas com a sua variável de interesse, dentre as seguintes opções: condições para colheita, estiagem agrícola, estiagem nos últimos cinco dias, evapotranspiração, condições para tratamento fitossanitário, previsão de geadas, necessidade de irrigação, condições para manejo do solo, disponibilidade de água no solo, precipitação acumulada semanal,

<sup>10</sup> Disponível em: <<http://openlayers.org/>>.

<sup>11</sup> Disponível em: <<http://geoext.org/>>.



**Figura 3.** Diagrama de arquitetura do módulo webgis do sistema Agritempo.

Fonte: Embrapa Informática Agropecuária (2014).

necessidade de reposição por chuvas, temperaturas máximas, mínimas e médias, localização das estações meteorológicas.

Ao clicar na variável, o servidor de mapas Geoserver é responsável por responder as requisições realizadas (WMS, CSW) e buscar nos repositórios os dados georreferenciados, sejam eles vetoriais ou matriciais. As aplicações componentes são executadas no servidor de aplicações Java, Apache Tomcat<sup>12</sup>.

Apesar da importância destes sistemas, ainda existe uma carência muito grande de ferramentas mais específicas que considerem as particularidades e necessidades de cada cultura e as condições locais, e que auxiliem os agricultores mais efetivamente em suas tomadas de decisão (SENTELHAS; MONTEIRO, 2009).

Atualmente, uma das maiores limitações à produção de estimativas e inferências mais precisas sobre as culturas é a falta de dados precisos em alta resolução espacial. Uma das carências, nesse sentido, ocorre com os solos agrícolas. O território brasileiro é muito extenso e apresenta grande variabilidade nas propriedades físicas dos perfis de solos. Essas informações são fundamentais para a determinação do potencial de armazenamento de água desses solos e, conseqüentemente, da disponibilidade de água para cultivos, pastagens e florestas. É preciso, portanto, formar uma base de dados de solos com densidade amostral suficiente que permita refinar os métodos de espacialização e elaboração de mapas situacionais em escala mais detalhada ou, ainda, que permita consultas mais específicas e pontuais com menor incerteza.

Outra limitação relevante é a carência de meios adequados que permitam modelar as condições meteorológicas na meso-escala, ou seja, nas dimensões em que a topografia condiciona o tempo

<sup>12</sup> Disponível em: <<http://tomcat.apache.org/>>.

pelas características do relevo local (exposição solar, declividade e configuração do terreno). O efeito da topografia pode ser pouco ou muito intenso dependendo das características do relevo local, condicionando a radiação solar incidente, a temperatura, a umidade, o vento e até a chuva, em algumas situações, desviando essas variáveis da média predominante do seu entorno.

Novamente, o desenvolvimento de tecnologias e sistemas para superar esses problemas passa, necessariamente, pelo uso da tecnologia da informação. No primeiro caso, aplicada ao gerenciamento e integração de bancos de dados e, no segundo caso, aplicada à implementação de ferramentas que facilitem o desenvolvimento, avaliação e uso de modelos agroambientais que, muitas vezes, requerem grande capacidade de cálculo e, também, integração a bancos de dados.

### 3 Análise de riscos climáticos na agricultura

O risco climático pode ser definido como o possível impacto negativo que um evento ou uma condição meteorológica pode causar a um bem, sociedade ou ecossistema. Uma vez que a agricultura é a atividade humana mais dependente das condições climáticas, o agronegócio é o setor mais frequentemente afetado pelos riscos climáticos.

A variabilidade natural das condições meteorológicas, caracterizadas por precipitação, temperatura, radiação solar, umidade e vento, podem causar diferentes eventos como secas, tempestades, ondas de calor, ondas de frio e subida do nível do mar. Estes eventos, por sua vez, podem gerar diversos impactos em plantações e rebanhos, como a falta ou o excesso de água, surtos de pragas e doenças, inundações de terras produtivas, incêndios de florestas naturais ou plantadas, entre outros, todos resultando em redução da produtividade agrícola (MARENGO, 2011).

O risco climático pode ser analisado através do produto da probabilidade e severidade do impacto no local em estudo. No entanto, muitas vezes, é difícil estimar o impacto preciso de uma determinada condição climática na agricultura. Por isso, o risco agroclimático tem sido tratado no Brasil mais comumente na forma de frequência de ocorrência. Nesse caso, qualquer condição particular cuja ocorrência resulte em impactos deletérios e que possa ser objetivamente caracterizada, pode ser incluída em uma estratégia de análise do risco.

O cálculo da probabilidade de ocorrência dessas condições é normalmente baseado na análise de séries temporais das variáveis envolvidas. Assim, estes estudos dependem de séries de dados meteorológicos medidos por um período longo o suficiente para que seja representativo da variabilidade natural das condições locais analisadas. As recomendações técnicas da Organização Meteorológica Mundial (OMM) definem Normais Climatológicas como valores médios calculados para um período relativamente longo e uniforme, compreendendo no mínimo três décadas consecutivas (WORLD METEOROLOGICAL ORGANIZATION, 1989). No entanto nem todas as localidades ou regiões agrícolas estudadas possuem séries longas o suficiente. Por isso, na prática, períodos mais curtos têm sido aproveitados em estudos de zoneamento. Ainda de acordo com a OMM, no caso de estações para as quais a mais recente Normal Climatológica não esteja disponível, seja porque a estação não esteve em operação durante o período de 30 anos, seja por outra razão qualquer, Normais Provisórias podem ser calculadas (WORLD METEOROLOGICAL ORGANIZATION, 1989). Normais Provisórias são médias de curto período, baseadas em observações que se estendam sobre um período mínimo de 10 anos.

As condições de risco dependem da cultura considerada e de sua fase de desenvolvimento, sendo influenciadas pelo solo e pelo manejo adotado. No Brasil, o suprimento hídrico para culturas agrícolas, inclusive pastagens, florestas e pomares, é proveniente quase que exclusivamente da chuva. Estima-se que cerca de 5% das áreas agrícolas nacionais sejam irrigadas (PAZ et al., 2000). Portanto, nas áreas não irrigadas, as culturas encontram-se sujeitas às grandes variações naturais da chuva, tanto em quantidade como em distribuição, com meses ou anos de maior ou menor oferta. Isto faz com que este seja um dos principais fatores de risco para a agricultura no Brasil. Outro fator de risco são as temperaturas extremas que, quando muito baixas ou muito altas, podem provocar estresse fisiológico ou danos diretos em plantas e animais resultando em redução de produtividade. Danos diretos provocados por vento muito intenso e granizo também são um fator de risco, mas de ocorrência muito mais localizada e esporádica.

Devido aos vários fatores envolvidos na definição de uma condição de risco agrícola e ao grande volume de dados requeridos, os procedimentos de cálculo para áreas extensas demandam considerável capacidade de processamento e armazenamento de dados, e requerem software ou sistemas computacionais apropriados.

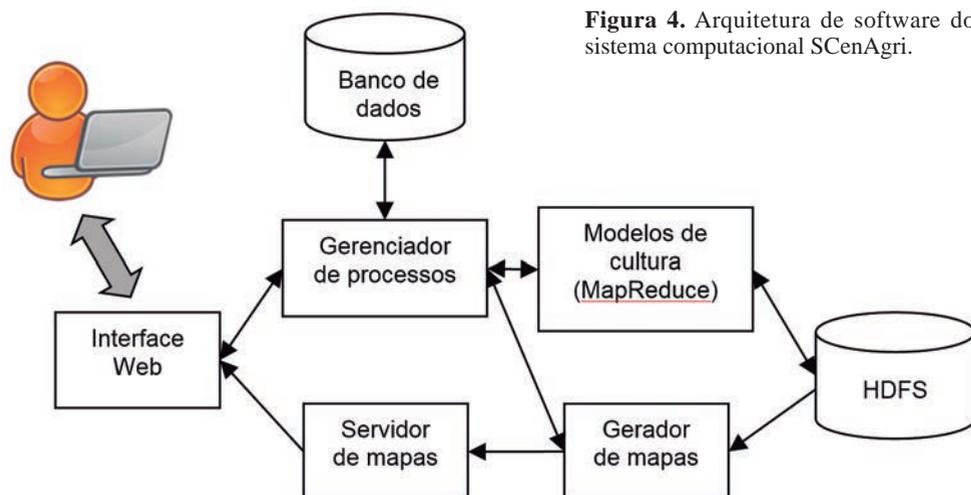
Nesse contexto, o Simulador de Cenários Agrícolas (SCenAgri) é um sistema computacional que foi desenvolvido pela Embrapa Informática Agropecuária para suprir essa necessidade. O SCenAgri provê computação de alto desempenho para simular o efeito das condições climáticas na agricultura brasileira, utilizando modelos de cultura, bancos de dados climáticos e de solos. Um dos modelos, atualmente em uso, calcula o Índice de Satisfação das Necessidades de Água (Isna) da cultura a partir das séries históricas do banco de dados meteorológicos, produzindo resultados de frequência de ocorrência de anos versus datas de plantio acima e abaixo de valores de referência relacionados ao risco de perda. O Isna pode ser diretamente relacionado a produtividade da cultura através de funções de redução de produtividade (DOORENBOS; PRUITT, 1977). Dessa forma, é possível mapear as áreas de acordo com o risco climático estimado para a cultura, em cada data de plantio. O Simulador permite que seus usuários simulem cenários agrícolas atuais, baseados nas séries de dados climáticos observados do passado até o presente, ou mesmo, cenários futuros, utilizando dados estimados de diversos modelos de projeções climáticas regionalizadas.

Este sistema opera em uma grade de computadores composta por dezenas de máquinas, e utiliza o software estatístico R<sup>13</sup> e a tecnologia Hadoop/MapReduce<sup>14</sup> para distribuir a execução das simulações entre os computadores da grade, em paralelo. Entre os vários recursos do Hadoop, dois são empregados no SCenAgri: a implementação MapReduce e o Sistema de Arquivos Distribuídos Hadoop (HDFS) (BORTHAKUR, 2007; DEAN; GHEMAYAT, 2008). MapReduce é um modelo de programação, originalmente proposto pelo Google, para processamento e geração de grandes quantidades de dados. Neste modelo, os usuários especificam o cálculo em termos de dois tipos de tarefas: mapeadores e redutores. Os mapeadores são responsáveis por executar a computação em frações de dados de entrada e geração de resultados intermediários. Os resultados intermediários são consolidados pelo redutor, que gera o resultado final. Uma camada de software subjacente que implementa o modelo de programação paraleliza

<sup>13</sup> Disponível em: <<http://www.r-project.org/>>.

<sup>14</sup> Disponível em: <[http://hadoop.apache.org/docs/r1.2.1/mapred\\_tutorial.html](http://hadoop.apache.org/docs/r1.2.1/mapred_tutorial.html)>.

automaticamente o cálculo dos mapeadores através das máquinas do cluster, conforme ilustrado na Figura 4.



**Figura 4.** Arquitetura de software do sistema computacional SCenAgri.

O HDFS é um sistema de distribuição de arquivos implementado pelo Hadoop. O HDFS replica blocos de dados e os distribui em clusters de computadores. A redundância de dados resulta em um armazenamento de dados mais confiável e de alta taxa de transferência, necessário para as operações e volume de dados processados pelo simulador. Além disso, o HDFS é projetado para ser implementado em hardware de baixo custo, ou seja, computadores comuns.

O tratamento de dados, as análises estatísticas necessárias, bem como a geração dos mapas em formato matricial (formato Geotiff) e vetorial (formato Shapefile) são executados utilizando-se o software R. O R é um software livre e muito extensível, que proporciona uma grande variedade de técnicas estatísticas. Entre as suas extensões, R fornece bibliotecas de software para manipulação e produção de dados espacializados, o que é essencial para as análises de dados e a produção dos mapas espacializados gerados pelo SCenAgri.

Por fim, uma interface web flexibiliza o acesso ao sistema por parte do usuário, que pode realizar novas simulações e recuperar resultados obtidos anteriormente. Atualmente, este sistema é restrito para atividades de pesquisa ligadas à programação da Embrapa. Fazendo uso dos dados estimados de modelos de projeções climáticas futuras, este sistema vem atendendo às necessidades de diversos estudos e avaliações de impacto sobre mudanças climáticas na agricultura.

## 4 Zoneamento agroclimático

Dentre as informações agrometeorológicas empregadas na fase de planejamento agrícola, o zoneamento agroclimático é a de uso mais difundido no Brasil.

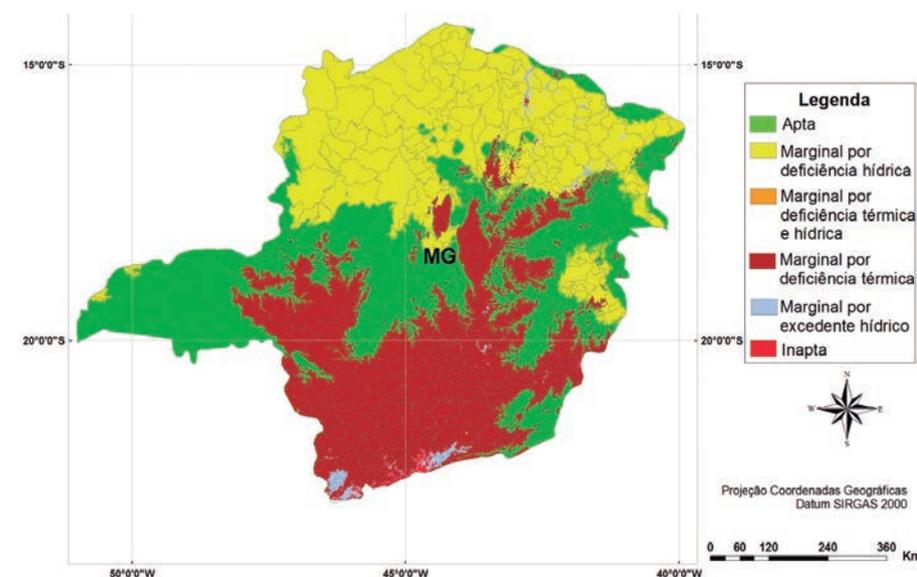
Para se alcançar uma produtividade econômica cada cultura necessita de condições favoráveis durante todo o seu ciclo vegetativo, isto é, exigem determinados limites de temperatura nas várias fases do ciclo, de uma quantidade mínima de água, e de um período seco nas fases de maturação e colheita. Um zoneamento agroclimático consiste na determinação da aptidão climática das

regiões de um país, estado ou município, considerando as exigências agroclimáticas dos cultivos e as informações climáticas do local de interesse. Como o solo é o outro componente do meio físico necessário na agricultura, pode-se considerar os aspectos edáficos de forma conjunta aos aspectos do clima, resultando em um zoneamento edafoclimático ou zoneamento ecológico das culturas. O denominado zoneamento agrícola envolve o zoneamento ecológico e o levantamento das condições socioeconômicas das regiões, para delimitar a vocação agrícola das terras. Uma vez que o clima não pode ser controlado pelo homem para se adequar às necessidades dos cultivos, essa deve ser a primeira informação a ser considerada no planejamento de um empreendimento agrícola (PEREIRA et al., 2002).

O zoneamento agroclimático pode ser empregado para a delimitação de áreas aptas, marginais ou inaptas às culturas (Figura 5), mas também para o estabelecimento das melhores épocas de semeadura com base em informações probabilísticas (Figura 6), das zonas de maturação de frutos, do risco climático associado aos impactos do déficit hídrico nas culturas, de áreas de escape de doenças, do potencial produtivo e da qualidade dos produtos (SENTELHAS; MONTEIRO, 2009).

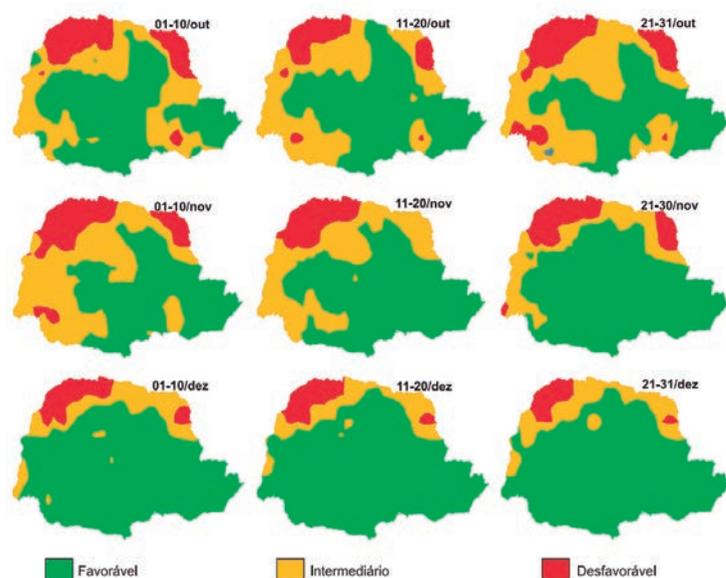
No Brasil, um dos métodos de zoneamento mais difundido é o baseado na quantificação de riscos climáticos. Esse zoneamento consiste na espacialização das informações de risco, baseado em análises de frequência que retornam a probabilidade de ocorrência de condições específicas. Essas análises de frequência consideram séries de dados meteorológicos medidos por um período longo o suficiente para que seja representativo da variabilidade natural das condições locais analisadas.

Atualmente, o Ministério da Agricultura Pecuária e Abastecimento (Mapa) do Brasil utiliza o Zoneamento Agrícola de Risco Climático (Zarc) como um instrumento de política agrícola e gestão de riscos na agricultura. O estudo é elaborado com o objetivo de minimizar os riscos



**Figura 5.** Zoneamento agroclimático para o cultivo de pinhão manso (*Jatropha curcas* L.) no Estado de Minas Gerais, considerando necessidades hídricas e térmicas da cultura.

Fonte: Yamada (2011).



**Figura 6.** Classificação de épocas de semeadura em relação ao risco hídrico à cultura da soja no estado do Paraná, em nove épocas de semeadura, para cultivar precoce (120 dias) e solo de média retenção de água (CAD= 50 mm).

Fonte: Farias et al. (2001).

relacionados a perdas agrícolas decorrentes de eventos climáticos e permite a cada município identificar a melhor época de plantio das culturas, nos diferentes tipos de solo e ciclos de cultivares. Para fazer jus a programas de seguro agrícola e à subvenção federal do seguro rural, o produtor deve observar as recomendações desse pacote tecnológico. Além disso, alguns agentes financeiros condicionam a concessão do crédito rural ao uso do zoneamento.

No estudo do Zarc são analisados os parâmetros de clima, solo e de ciclos de cultivares para, no final, ser determinada a relação de municípios indicados ao plantio de determinadas culturas, com seus respectivos calendários ou épocas de plantio. O Zoneamento Agrícola de Risco Climático foi usado pela primeira vez na safra de 1996 e é publicado na forma de portarias, no Diário Oficial da União e no site do ministério. Atualmente, os estudos de zoneamentos agrícolas de risco climático contemplam mais de 40 culturas, sendo 15 de ciclo anual e 24 de ciclo permanente, além do zoneamento para o consórcio de milho com braquiária, alcançando 24 estados brasileiros (BRASIL, 2014).

Os passos para a elaboração do zoneamento agroclimático de uma cultura envolvem a definição dos objetivos, a caracterização das exigências climáticas das culturas, a abrangência do estudo, o levantamento dos dados climáticos da região estudada e, finalmente, o processamento dos dados e produção dos resultados. A partir dos dois primeiros passos, determina-se os procedimentos de cálculo e os dados que serão necessários e, considerando a abrangência do estudo, já é possível estimar o volume aproximado de dados a serem processados.

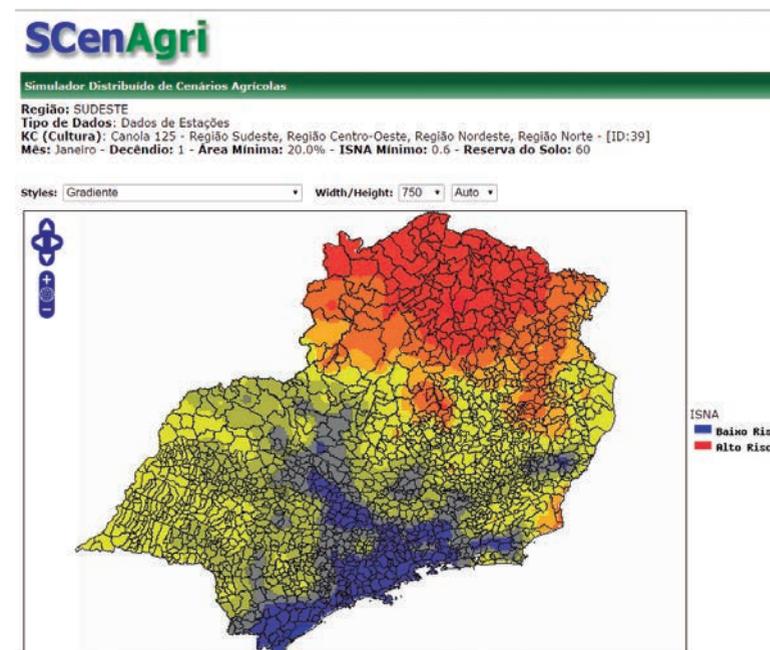
A evolução da qualidade dos estudos de zoneamento nas últimas duas décadas está relacionada não apenas aos aprimoramentos metodológicos em si mas, principalmente, com o aumento das capacidades computacionais e de processamento de dados - maior número de estações, regiões maiores, maior resolução espacial. Dessa forma, a evolução do zoneamento agrícola no Brasil

sempre esteve relacionada e se beneficiou diretamente da evolução dos recursos das tecnologias de informação e comunicação disponíveis.

Uma ferramenta de TI que merece destaque e se tornou comum nos últimos anos são os sistemas de informações geográficas (SIG), cada vez mais usados para a distribuição, processamento, análise, modelagem de dados espaciais, sendo aplicados em diversas áreas. Na elaboração de zoneamentos agroclimáticos, em particular, têm sido útil pois muitos softwares dessa classe de sistemas podem executar procedimentos diversos a partir de dados básicos e gerar informações georreferenciadas na definição de áreas propícias ao cultivo de determinadas culturas (YAMADA, 2011).

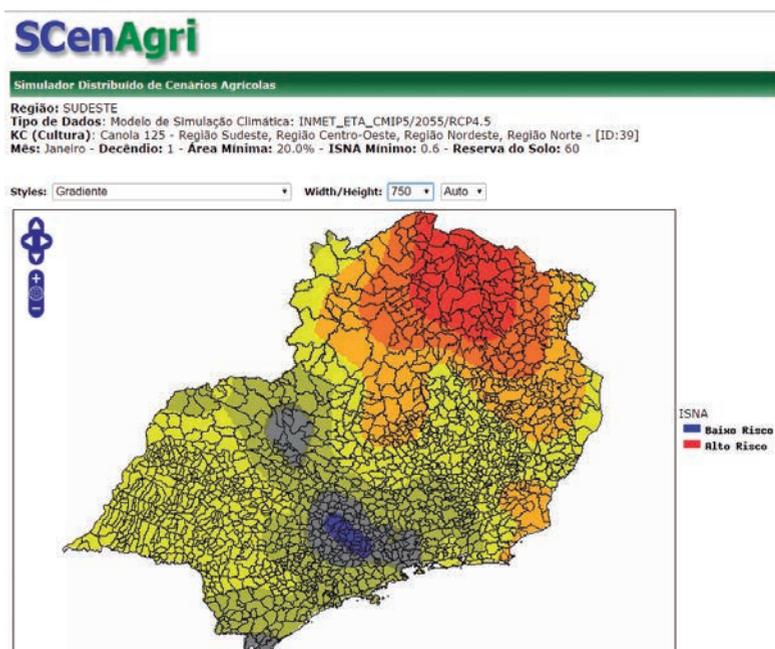
Entre as análises espaciais do SIG é possível calcular a regressão por meio da interpolação dos dados de probabilidades de ocorrência da variável climática em questão para todos os outros pontos onde não há estações meteorológicas usando-se as relações existentes entre a variável em questão, a latitude, a longitude, e a altitude (por ex. imagens SRTM como modelos digitais de elevação do terreno) para cada um dos pixels existentes no mapa, obtendo-se assim valores estimados para todas as localidades da região representada com informações a respeito do comportamento espacial da variável (ASTOLPHO, 2003; CÂMARA; MEDEIROS, 1998).

No âmbito dos assuntos relativos a zoneamento, um dos recursos disponibilizados pelo Simulador de Cenários Agrícolas (SCenAgri), já descrito no tópico Análise de Risco, é a geração de mapas com dados especializados a partir dos resultados da análise de risco para deficiência hídrica. Dessa forma, é possível identificar os municípios de alto e baixo risco quanto ao suprimento hídrico, para datas de plantio a intervalos decendiais, seguindo os mesmos critérios atualmente adotados na metodologia do Zoneamento Agrícola de Risco Climático (Zarc) (Figura 7).



**Figura 7.** Zoneamento de risco por deficiência hídrica para a cultura da canola para o período de dados 1961-1990, gerado pelo Sistema SCenAgri/Embrapa, considerando ciclo médio de 125 dias, CAD de 60 mm, e plantio no primeiro decênio de janeiro.

O SCenAgri foi criado para permitir a vinculação de conjuntos de dados históricos ou de projeções futuras - como as geradas pelos modelos de circulação geral da atmosfera - ao próprio modelo do Zarc e de outros modelos semi-empíricos ou determinísticos para a simulação de cenários agrícolas futuros. Isto tem permitido, por meio dos resultados tabulares e espaciais do SCenAgri, avaliar os possíveis impactos que as mudanças climáticas terão sobre o Zarc e sobre as culturas agrícolas em geral (Figura 8).



**Figura 8.** Zoneamento de risco por deficiência hídrica para a cultura da canola para projeções futuras 2041-2070 (modelo ETA, inicialização HadGen2ES RCP 4,5), gerado pelo Sistema SCenaAgri/Embrapa, considerando ciclo médio de 125 dias, CAD de 60 mm, e plantio no primeiro decêndio de janeiro.

## 5 Modelagem e simulação agroambiental

A modelagem é um método eficiente para representar a realidade e apoiar a tomada de decisão racional, sendo assim fundamental para a ação humana. Os sistemas Agritempo e SCenAgri, já mencionados, são exemplos de sistemas de produção de informações agroambientais que fazem uso intensivo da modelagem para a análise de dados e produção de informações.

A agropecuária se organiza em sistemas complexos, com realidades muito particulares e heterogêneas determinando seu formato. À medida que os principais interesses da agricultura passam a incluir, além dos aumentos de produtividade e de riqueza, os aspectos ambientais e sociais, um conjunto ampliado de atores requerem resultados analíticos aplicados a problemas específicos. Nesse contexto, os institutos de pesquisas são desafiados continuamente a fazer a ponte entre o conhecimento básico das ciências e a demanda prática de conhecimentos.

Atualmente, a tecnologia da informação permite que sistemas, instrumentos e dispositivos automatizados utilizem as estruturas de conhecimento criadas pela pesquisa científica em larga

escala, viabilizando o uso de dados, informações e sistemas nas cadeias produtivas agropecuárias e nas instituições que coordenam as ações produtivas. A agricultura torna-se, desta forma, uma atividade intensiva em conhecimentos, complementando os recursos de capital, trabalho e terras tradicionalmente utilizados.

Na pesquisa científica há um uso bastante ampliado de modelos (PIDD, 2010). Os modelos mais úteis capturam a essência do conhecimento disponível sobre o fenômeno, podendo ser manipulados e ajustados a um custo substancialmente menor que a experimentação. Essas estruturas analíticas podem incorporar conhecimentos gerais e específicos, partindo das definições de componentes e de seu relacionamento no sistema representado. A modelagem torna-se, desta forma, um método para analisar um fenômeno em bases qualitativas e quantitativas, em que exercícios mentais são utilizados. O esforço de modelagem torna-se também um esforço de aprendizado, o que facilita muito a difusão de conhecimentos. A utilidade da modelagem também está na especificação da informação de interesse final. Ao se perseguir a utilidade para a tomada de decisões e desenvolver métodos computacionais eficientes, todo o conjunto de informações e conhecimentos incorporados no modelo geram um produto de maior facilidade de comunicação. Dessa forma, as tecnologias de informação e comunicação proporcionam recursos importantíssimos para a viabilização do conhecimento científico da atualidade, tanto na agricultura como em outras áreas do conhecimento.

A modelagem é um fenômeno alavancado com a produção de grandes volumes de informações e capacidade de processamento, permitindo o teste de hipóteses e verificação da validade de teorias fundamentais (PIDD, 2010). O estudo do relacionamento entre as variáveis, o entendimento dos problemas de interesse e a exploração de dados tornam os modeladores potencialmente melhores, estabelecendo mais claramente relações de causa e efeito e lógicas. Este aspecto é crucial para a agricultura. A modelagem é também uma das melhores maneiras de especificar demandas adicionais de informações, em um processo custoso, que efetivamente ocorre quando há um uso justificado para atender a demanda de um usuário ou criar uma forma inovadora de encarar um problema.

A modelagem agroambiental, por exemplo, tem sido sensível à demanda de informações para gestão de políticas públicas. Informações sobre relevo, classes de cobertura vegetal e uso das terras, além do mapeamento dos solos, são utilizadas em zoneamentos de risco climático, agroecológicos e ecológico-econômicos, estabelecendo as bases para a gestão territorial. Os modelos agrometeorológicos têm concentrado esforços de especialistas na caracterização de elementos do clima, dos solos, do manejo agropecuário e da fisiologia das culturas, tornando-se uma ferramenta essencial para o conhecimento do clima, para o zoneamento agrícola - com a identificação das melhores regiões e épocas de cultivo; para o monitoramento agrometeorológico - com diversas aplicações, desde a favorabilidade à ocorrência de doenças, ao manejo de irrigação; e para projetar os impactos das mudanças climáticas na agricultura.

Em um sistema de produção real, outros elementos interferem diretamente nas possibilidades de produção, como máquinas, recursos naturais e elementos de integração da agricultura com cadeias produtivas, mercados e as instituições de um país. Uma fronteira essencial neste contexto é a integração de conhecimentos e modelos biofísicos e econômicos. As pesquisas básica e aplicada buscam relações lógicas essenciais para objetivos específicos, evitando-se a introdução de elementos desnecessários para a compreensão de um sistema e tomada de decisão.

## 6 Mudanças climáticas e agricultura

Segundo Nordhaus (2010), a ciência do aquecimento global chegou a um consenso sobre a alta probabilidade de haver um aquecimento substancial do planeta neste século. As ações tomadas para conter as emissões têm sido limitadas, desde o primeiro acordo de Kyoto, em 1997, e pouco progresso ocorreu na reunião de Copenhague, em dezembro de 2009. As projeções indicam aquecimento mesmo com redução de emissões, e ações devem ser planejadas e priorizadas o quanto antes. O autor indica que o preço para a tonelada de carbono, para manter o objetivo de conter o aumento da temperatura global média em 2 °C ou menos, firmado em Copenhague, deveria ser US\$ 59 por tonelada (a preços de 2005), um preço bastante elevado em comparação com o valor eficaz de US\$ 5 por tonelada. É bastante improvável portanto, que o objetivo de conter a elevação de temperatura estabelecido em Copenhague será atingido.

As mudanças no clima são observáveis e tendem a atingir especialmente os países em desenvolvimento, segundo Rosenzweig e Parry (1994), e esses autores utilizam e recomendam o uso de modelos de culturas como estratégias importantes para avaliar a adaptação de sistemas de produção. Os estudos sobre impactos de mudanças climáticas no mundo foram realizados extensivamente com o uso de modelos agrônômicos que representavam funções de produção, inicialmente, ou funções de respostas de culturas específicas, como Adams (1989).

A elevação da temperatura no Brasil é incontestável e eventos extremos, como secas, veranicos e tempestades apresentam sinais de agravamento. A região amazônica pode ser afetada por grandes queimadas e gerar graves problemas para toda a produção agropecuária no Brasil central, afetando o regime de chuvas e a circulação de massas de ar. O Semiárido enfrenta, em condições normais, um período de seca prolongada durante o ano, que prejudica o desempenho da maioria das culturas agrícolas, e pode ter essa situação agravada pela elevação da temperatura.

Novas condições climáticas foram observadas para as culturas agrícolas no Brasil. Novas áreas foram incorporadas na produção de soja e milho, a partir da década de 1980, e atualmente a região central do Brasil ilustra a capacidade adaptativa de sistemas de produção reais, bem como a capacidade de adoção de novas técnicas e tecnologias. Nesse contexto, o conhecimento sobre as mudanças no ambiente de produção é essencial para a definição de rotas tecnológicas promissoras em termos de aumento da produtividade e criação de alternativas efetivas para mitigar riscos e tornar a agricultura mais adaptada e capaz de produzir com elevada tecnologia.

Mesmo nas condições tecnológicas mais desenvolvidas, acredita-se que a agricultura brasileira, cuja produção em grande parte vem de áreas cultivadas tendo a chuva natural como única ou principal fonte de água, encontra-se exposta a alterações nas condições climáticas. A vulnerabilidade da agricultura foi avaliada por Pinto e Assad (2008) pela primeira vez, e agora encontra-se diante do desafio de incorporar tecnologias mitigadoras, adaptadas e ter estudos aprofundados e específicos apoiando a superação dos desafios impostos pelas mudanças climáticas.

A construção de políticas, programas, projetos e ações para a agricultura passa pelo uso intensivo de conhecimentos e informações derivados de distintas áreas das ciências, mas com grande esforço técnico para torná-los específicos e aplicáveis na tomada de decisão. O clima se transforma e as pesquisas com as questões de vulnerabilidade, mitigação e adaptação às mudanças climáticas evoluem na criação de modelos, sistemas de informações e ferramentas para orientar os tomadores de decisões. Além disso, o diagnóstico dos impactos potenciais e as alternativas de adaptação

devem ter sua robustez verificada em diferentes cenários de mudanças climáticas. Isto implica a busca por soluções de armazenamento, recuperação, processamento, transmissão e visualização de grandes volumes de dados, requerendo métodos sofisticados.

No Brasil há carência de análises que considerem custos e benefícios de ações de adaptação às mudanças climáticas mais específicas. É necessário continuar a trabalhar com o tema e a oferecer alternativas para produtores rurais e formuladores de políticas públicas. Em uma perspectiva de aliar estratégias de mitigação como externalidades positivas, Seo (2013) defende que adaptações inteligentes, que reduzem os danos do aquecimento global e as emissões de carbono ao mesmo tempo, devem ser enfatizadas nas opções políticas. Adaptações inteligentes envolvem ativamente o setor público, coordenando a sociedade na elaboração de estratégias.

É desejável, diante das demandas acima, que estudos como os realizados por Pinto e Assad (2008) sejam ampliados e utilizados na discussão sobre os fatores que determinam a vulnerabilidade dos sistemas produtivos. A construção da capacidade de resposta dos atores depende de informações consistentes, que convençam os envolvidos das soluções mais robustas de produção diante das incertezas do clima futuro. A gestão do risco, quantificado a partir de bons conteúdos e sistemas de informações, depende de um esforço crítico de modelagem econômica, caracterizando o manejo e alternativas de produção de alimentos. Linhas distintas de modelagem climática, biofísica, agroambiental e econômica estão desenvolvendo ferramentas para tratar dos impactos de mudanças climáticas, e esforços de produção técnica e intercomparação de modelos devem ser o caminho para a pesquisa na área. A melhoria dos modelos é promissora, enfrentando desafios como incorporar os efeitos de eventos climáticos extremos e integrar elementos econômicos e ambientais aos aspectos técnicos da produção.

Segundo Antle e Capalbo (2001), modelos de avaliação integrada, que utilizam modelos disciplinares interligados para avaliar sistemas complexos, naturais e humanos, estão se tornando uma metodologia padrão de análise em questões ambientais. A avaliação integrada utiliza simulações do comportamento de um sistema biofísico e as introduz em modelos econômicos. Segundo Kauffmann e Snell (1997), por exemplo, existe uma alternativa de modelagem que combina resultados de modelos estatísticos para a produtividade observada com informações de modelos de culturas. Esta opção de modelagem está sendo desenvolvida em um projeto de intercomparação de modelos na Embrapa, o AgMIP-BR. Também estão em desenvolvimento modelos derivados dos trabalhos de Mendelsohn e Dinar (2009), compreendendo as escolhas de insumos, produtos e os impactos derivados na renda e nos preços das terras.

Cada abordagem tem suas vantagens no entendimento do efeito de mudanças climáticas. Contudo, a maior incerteza nos estudos é o escopo da adaptação. A adaptação não é observada, mas existe efetivamente na realidade, e é contingenciada por uma série de escolhas. A adaptação toma um tempo longo para realização e, portanto, a evidência direta só surge após longos períodos de tempo em condições estáveis. No caso do Brasil, no entanto, as transformações recentes dos sistemas agroindustriais criam uma estrutura de produção sem comparação com o passado distante, que muda as possibilidades para o futuro.

O Simulador de Cenários Agrícolas (SCenAgri), já descrito no tópico 3. Análise de Riscos Climáticos na Agricultura e 4. Zoneamento Agroclimático, é um sistema de TIC que foi criado para permitir a vinculação de conjuntos de dados históricos ou de projeções futuras - como as geradas pelos modelos de circulação geral da atmosfera - ao próprio modelo do Zarc e de outros

modelos semi-empíricos ou determinísticos para a simulação de cenários agrícolas futuros. Isto tem permitido, por meio dos resultados tabulares e espaciais do SCenAgri, avaliar os possíveis impactos que as mudanças climáticas terão sobre o Zarc e sobre as culturas agrícolas em geral.

## 7 Considerações finais

A Tecnologia da Informação e Comunicação (TIC) é um dos componentes mais importantes do ambiente da Pesquisa, Desenvolvimento e Inovação (PDI) atual, e as organizações brasileiras têm utilizado ampla e intensamente essa tecnologia, tanto em nível estratégico como operacional.

No âmbito da PDI em agrometeorologia e mudanças climáticas, assim como em outros temas de relevância nacional, a Embrapa Informática Agropecuária ocupa uma posição muito favorável à produção de conhecimento e de inovação por estar capacitada a atuar no desenvolvimento de sistemas e aplicações específicas de TI que, apesar de necessárias, seriam inviáveis para outras organizações.

A Embrapa Informática Agropecuária compreende que um esforço significativo de análise dos sistemas produtivos, obtenção, armazenamento, recuperação e processamento dos dados obtidos deve ser a base para muitas atividades técnico científicas. Por isso, além dos esforços de construir uma equipe de desenvolvimento de software, alia competências multidisciplinares para executar as análises da realidade e escolher os melhores caminhos para a construção dos sistemas que a agricultura demanda.

## 8 Referências

ADAMS, R. M. Global climate change and agriculture: an economic perspective. **American Journal of Agricultural Economics**, Saint Paul, v. 71, n. 5, p. 1272-1279, 1989.

ANTLE, J. M.; CAPALBO, S. M. Econometric-process models for integrated assessment of agricultural production systems. **American Journal of Agricultural Economics**, Saint Paul, v. 83, n. 2, p. 389-401, 2001.

ASTOLPHO, F. **Estimativa e mapeamento de probabilidades de ocorrência de temperaturas mínimas absolutas do ar adversas à agricultura paulista**. 2003. 99 f. Dissertação (Mestrado em Agricultura Tropical e Subtropical) - Instituto Agronômico de Campinas, Campinas.

BORTHAKUR, D. **The Hadoop Distributed File System**: architecture and design. [Los Angeles, CA: The Apache Software Foundation, 2007. 14 p. Disponível em: <[http://hadoop.apache.org/docs/r0.18.0/hdfs\\_design.pdf](http://hadoop.apache.org/docs/r0.18.0/hdfs_design.pdf)>. Acesso em: 24 set. 2014.

BRASIL. Ministério da Agricultura Pecuária e Abastecimento. **Zoneamento agrícola de risco climático**. Brasília, DF, [2014]. Disponível em: <<http://www.agricultura.gov.br/politica-agricola/zoneamento-agricola>>. Acesso em: 20 set. 2014.

CÂMARA, G.; MEDEIROS J. S. de. Princípios básicos em geoprocessamento. In: ASSAD, E. D.; SANO, E. E. (Ed.). **Sistema de informações geográficas**: aplicações na agricultura. [Planaltina, DF]: EMBRAPA-CPAC, 1998. p. 1-11.

COMITÊ GESTOR DA INTERNET NO BRASIL. **TIC domicílios e empresas 2013**: pesquisa sobre o uso das tecnologias da informação e comunicação no Brasil. São Paulo, 2014. 662 p. Disponível em: <[http://www.cetic.br/media/docs/publicacoes/2/TIC\\_DOM\\_EMP\\_2013\\_livro\\_eletronico.pdf](http://www.cetic.br/media/docs/publicacoes/2/TIC_DOM_EMP_2013_livro_eletronico.pdf)>. Acesso em: 10 out. 2014.

DEAN, J.; GHEMAWAT, S. MapReduce: simplified data processing on large clusters. **Communications of the ACM**, New York, v. 51, n. 1, p. 107-113, Jan. 2008. Disponível em: <<http://doi.acm.org/10.1145/1327452.1327492>>. Acesso em: 24 set. 2014.

DOORENBOS, J.; PRUITT, W. O. **Guidelines to predicting crop water requirements**. Rome: FAO, 1977. 155 p. (FAO Irrigation and Drainage Paper, 24).

EMBRAPA INFORMÁTICA AGROPECUÁRIA. **AgriTempo - Sistema de Monitoramento Agrometeorológico**. Campinas, 2014. Disponível em: <[www.agritempo.gov.br](http://www.agritempo.gov.br)>. Acesso em: 7 out. 2014.

FARIAS, J. R. B.; ASSAD, E. D.; ALMEIDA, I. R. de; EVANGELISTA, B. A.; LAZZAROTTO, C.; NEUMAIER, N.; NEPOMUCENO, A. L. Caracterização de risco de déficit hídrico nas regiões produtoras de soja no Brasil. **Revista Brasileira de Agrometeorologia**, Santa Maria, v. 9, n. 3, p. 415-421, 2001. Número especial.

KAUFMANN, R. K.; SNELL, S. E. A biophysical model of corn yield: integrating climatic and social determinants. **American Journal of Agricultural Economics**, Saint Paul, v. 79, n.1, p. 178-190, Feb. 1997.

MARENGO, J. A. (Coord.). **Riscos das mudanças climáticas no Brasil**: análise conjunta Brasil-Reino Unido sobre os impactos das mudanças climáticas e do desmatamento na Amazônia. São José dos Campos: INPE: MetOffice, 2011. Disponível em: <<http://www.inpe.br/noticias/arquivos/pdf/relatorioport.pdf>>. Acesso em: 20 set. 2014.

MENDELSON, R.; DINAR, A. **Climate change and agriculture**: an economic analysis of global impacts, adaptation and distributional effects. Cheltenham; Northampton: Edward Elgar, 2009. 256 p. (New horizons in environmental economics).

NORDHAUS, W. D. Economic aspects of global warming in a post-Copenhagen environment. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, D. C., v. 107, n. 26, p. 11721-11726, June, 2010.

PAZ, V. P. da S.; TEODORO, R. E. F.; MENDONÇA, F. C. Recursos hídricos, agricultura irrigada e meio ambiente. **Revista Brasileira de Engenharia Agrícola e Ambiental**, Campina Grande, v. 4, n. 3, p. 465-473, set./dec. 2000. Disponível em: <<http://www.scielo.br/pdf/rbeaa/v4n3/v4n3a25.pdf>>. Acesso em: 20 set. 2014.

PEREIRA, A. R.; ANGELOCCI, L. R.; SENTELHAS, P. C. **Agrometeorologia**: fundamentos e aplicações práticas. Guaíba: Ed. Agropecuária, 2002. 478 p.

PIDD, M. Why modelling and model use matter. **Journal of the Operational Research Society**, Oxford, v. 61, p. 14-24, Jan. 2010.

PINTO, H. S.; ASSAD, E. D. (Coord.). **Aquecimento global e a nova geografia da produção agrícola no Brasil**. [Campinas]: Embrapa Informática Agropecuária: UNICAMP, 2008. 84 p.

RIJKS, D.; BARADAS, M. W. The clients for agrometeorological information. **Agricultural and Forest Meteorology**, Amsterdam, v. 103, n. 1-2, p. 27-42, June 2000.

ROSENZWEIG, C.; PARRY, M. L. Potential impact of climate change on world food supply. **Nature**, London, v. 367, p. 133-138, Jan. 1994.

SENTELHAS, P. C.; MONTEIRO, J. E. B. de A. Agrometeorologia dos cultivos: informações para uma agricultura sustentável. In: MONTEIRO, J. E. B. de A. (Org.). **Agrometeorologia dos cultivos**: o fator meteorológico na produção agrícola. Brasília, DF: INMET, 2009. v. 1, p. 5-15.

SEO, S. N. Economics of global warming as a global public good: private incentives and smart adaptations. **Regional Science Policy & Practice**, Malden, MA, v. 5, n. 1, p. 83-95, Mar. 2013.

SIVAKUMAR, M. V. K.; MOTHA, R. P. (Ed.). **Managing weather and climate risks in agriculture**. Berlin: Springer, 2007. 288 p. il.

WORLD METEOROLOGICAL ORGANIZATION. **Calculation of monthly and annual 30-year standard normals.** [Washington, D. C.], 1989. 14 p. (WCDP, n. 10; WMO-TD, n. 341).

YAMADA, E. S. M. **Zoneamento agroclimático da *Jatropha curcas* L. como subsídio ao desenvolvimento da cultura no Brasil visando à produção de biodiesel.** 2011. 135 f. Dissertação (Mestrado em Física do Ambiente Agrícola) - Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, Piracicaba. Disponível em: <<http://www.teses.usp.br/teses/disponiveis/11/11131/tde-23052011-164645/>>. Acesso em: 24 set. 2014.

## Tecnologias da informação aplicadas aos dados geoespaciais

Júlio César Dalla Mora Esquerdo  
Sérgio Aparecido Braga Cruz  
Carla Geovana do Nascimento Macário  
João Francisco Gonçalves Antunes  
João dos Santos Vila da Silva  
Alexandre Camargo Coutinho

### 1 Introdução

A dimensão espacial permeia, direta ou indiretamente, uma série de temas relacionados ao agronegócio, sendo um fator essencial que deve sempre ser considerado nas diferentes atividades de pesquisa nos mais variados domínios. A avaliação de riscos climáticos, o levantamento, a caracterização e o monitoramento de recursos naturais, o mapeamento do uso e cobertura da terra, as análises socioeconômicas, os zoneamentos e a avaliação de cenários são alguns exemplos de atividades nas quais a pesquisa geoespacial é fundamental. A análise de dados geoespaciais, produzidos a partir das geotecnologias, pode ser vista como uma das etapas de um processo de diagnóstico científico que, a partir da formulação de um problema, define um plano de execução, orienta a coleta e a integração de dados relevantes, define estratégias de análise e dá suporte para a compreensão de fenômenos naturais ou antrópicos e a tomadas de decisão.

Nesse cenário, há uma gama de dados geoespaciais produzidos por diferentes técnicas e propósitos, e por distintos perfis de usuários. Existe um aumento sem precedentes na produção e na disponibilidade de tais dados, provenientes da maior quantidade de dispositivos sensores, pela adoção cada vez maior de padrões para representação de dados geoespaciais que viabilizam o seu compartilhamento, e pelo número crescente de dispositivos móveis que favorecem a coleta georreferenciada dos mais diferentes tipos de dados. Além do enorme volume de informações, o contexto atual na produção destes dados georreferenciados é caracterizado pela velocidade em que eles são produzidos e pela sua redundância.

Em contrapartida a esse crescimento da quantidade de dados, a sua capacidade de tratamento e análise integrada permanece associada a paradigmas caracterizados pela baixa escalabilidade, baixas capacidades de reúso e repetição, representados por soluções apoiadas em Sistemas de Informações Geográficas (SIGs) de uso local ou soluções para geoprocessamento individuais. No contexto do tratamento de dados geoespaciais, esse cenário gera impactos importantes, uma vez que esses dados são a base para o desenvolvimento e a integração de diferentes pesquisas multidisciplinares.

A necessidade de tratamento eficiente desse volume de dados geoespaciais tem promovido uma mudança no paradigma na construção de aplicações científicas, incentivando o surgimento de novas plataformas e ferramentas caracterizadas pela capacidade de processamento intensivo de dados e escalabilidade dinâmica. Novas soluções para a construção de aplicações em geopro-

cessamento, apoiadas em novas tecnologias e modelos de desenvolvimento, têm sido exploradas com o objetivo de atender a essa demanda crescente por geoprocessamento, buscando maior produtividade, associada à facilidade de uso, agilidade na obtenção de resultados e qualidade dos produtos.

Considerando essas novas tendências, este capítulo trata das Tecnologias da Informação aplicadas aos dados geoespaciais, com o objetivo de definir alguns dos principais conceitos abordados nesse tema e apresentar aspectos gerais das tecnologias envolvidas. São apresentados, também, exemplos de iniciativas da Embrapa Informática Agropecuária no desenvolvimento de soluções voltadas à organização, ao compartilhamento e ao tratamento das informações geográficas produzidas pela empresa em seus projetos de pesquisa e de serviços.

## 2 Dados geoespaciais e geotecnologias

Nos últimos anos, termos como *geoprocessamento*, *geotecnologia*, *geoinformação*, *dado geográfico*, *dado geoespacial* e *geodado* têm sido cada vez mais utilizados e referenciados, não apenas pelos estudiosos do assunto, como também por um público crescente de usuários de informações e ferramentas geográficas, nem sempre familiarizados com seus significados. Entre os fatores que têm estimulado a popularização desses termos estão a facilidade de acesso aos dados geoespaciais e a programas computacionais que permitem sua visualização e análise; a disponibilização de serviços geoespaciais em que os usuários podem realizar consultas específicas a partir de um navegador da internet; o crescente uso dos dispositivos móveis, como smartphones e tablets, munidos de equipamentos de localização por satélite e acesso à internet; entre outros. Esses termos podem ser compreendidos a partir do esquema da Figura 1.



Figura 1. Esquema conceitual do geoprocessamento.

É importante caracterizar esses termos. Dados geoespaciais, também denominados dados geográficos, pertencem a uma classe particular de dados espaciais, uma vez que eles descrevem fatos, objetos e fenômenos do globo terrestre associados à sua localização sobre a superfície terrestre, num certo instante ou período de tempo (CÂMARA et al., 1996). Os dados geoespaciais se distinguem essencialmente dos demais pela sua componente espacial, que associa a cada entidade ou fenômeno, uma localização traduzida por um sistema geodésico de referência terrestre.

O processo de aquisição, visualização, processamento, análise e/ou disponibilização de dados geoespaciais é realizado a partir de uma categoria especial de tecnologia, denominada geotecnologia. Neste contexto, tecnologias como o sensoriamento remoto, o *Global Positioning System* (GPS), a topografia, os Sistemas de Informações Geográficas, os bancos de dados geográficos,

entre outras, são classificadas como geotecnologias. Quando uma informação geoespacial é derivada de uma ou mais geotecnologias, ela é denominada geoinformação ou geodado. Por fim, o processo de aplicação de uma ou mais geotecnologias para adquirir, processar, visualizar, analisar e/ou disponibilizar dados espacialmente referenciados, com o intuito de gerar geoinformação, é denominado geoprocessamento.

### 2.1 O GPS

Entre as geotecnologias de aquisição dos dados geoespaciais está o GPS. O GPS, também conhecido por *Navigation System for Timing and Range* (Navstar), foi idealizado na década de 70 pelo Departamento de Defesa norte-americano, mas tornou-se totalmente operacional somente em 1995. Trata-se de uma constelação de satélites dispostos em órbitas específicas que fornece a um aparelho receptor sua posição geográfica em qualquer parte do planeta. Seu uso para fins civis se intensificou a partir do ano 2000, quando a “disponibilidade seletiva”, artifício que intencionalmente degradava o sinal para usos não militares, foi desativada, reduzindo os erros de posicionamento. Apesar de não ser o único sistema de posicionamento global - existem também o Galileu (europeu), Glonass (russo) e BeiDou (chinês) - o GPS é o sistema utilizado pela maior parte dos receptores. Com a redução do custo desses equipamentos, eles estão presentes em um número cada vez maior de dispositivos móveis, como telefones celulares, tablets e relógios. A presença do GPS e das redes de acesso à Internet nos dispositivos móveis tem provocado uma revolução tecnológica, trazendo uma série de benefícios aos seus usuários. Nas grandes cidades, já é comum o uso de aplicativos em dispositivos móveis desenvolvidos para as pessoas localizarem o táxi mais próximo e solicitar uma corrida, possibilitando ao passageiro a visualização de informações detalhadas sobre o veículo, bem como um perfil com foto e avaliações do taxista que pretende chamar. Existe ainda uma gama de aplicativos que, em conjunto com as informações de localização geográficas providas pelo GPS e das informações das redes móveis de Internet, auxiliam as pessoas a encontrar a melhor rota para escapar do trânsito, o posto de gasolina mais próximo para abastecer o seu veículo, a localização de uma pizzaria nas redondezas, etc.

### 2.2. O sensoriamento remoto

Outra geotecnologia de aquisição de dados geoespaciais é o sensoriamento remoto, que pode ser entendido como um processo de obtenção de imagens e dados da superfície terrestre por meio da captação e registro da energia eletromagnética refletida ou emitida pela superfície, sem que haja contato físico entre o elemento sensor e o alvo. O sensoriamento remoto pode se dar nos níveis terrestre, suborbital e orbital, sendo esse último o mais comum, onde os sensores imageadores ficam a bordo de satélites em altitudes que variam de 400 a 36.000 km. Existe um conjunto de princípios físicos que define uma base teórica necessária ao entendimento das imagens de satélite, as quais não podem ser meramente chamadas de “fotografias”, uma vez que são obtidas a partir de sensores, denominados radiômetros, capazes de captar a energia refletida e/ou emitida pela superfície terrestre em diferentes comprimentos de onda do espectro eletromagnético, não apenas no visível.

Uma imagem de satélite é oriunda de sensores que apresentam quatro tipos de resolução:

- Espacial, que define o tamanho do pixel da imagem e que representa a capacidade do sensor distinguir objetos.
- Temporal, relacionada ao tempo de revisita do sensor a um mesmo local na superfície terrestre.

- c) Espectral, que define as larguras e quantidades de faixas espectrais que o sensor é capaz de “enxergar”; e
- d) Radiométrica, relacionada à capacidade do sensor em distinguir diferentes níveis digitais, ou seja, a eficiência do sistema em detectar e registrar diferenças na energia refletida e/ou emitida pelos elementos.

Atualmente existe uma gama de produtos oriundos do sensoriamento remoto orbital, com as mais variadas resoluções espectral, temporal, radiométrica e espacial, cuja utilidade depende de sua aplicação. Por exemplo, satélites meteorológicos, como o *Geostationary Operational Environmental Satellite* (Goes) e o *National Oceanic and Atmospheric Administration* (NOAA), apresentam baixa resolução espacial (4,0 km e 1,0 km, respectivamente), mas oferecem altíssima resolução temporal (15 minutos e 12 horas, respectivamente). Tais características fazem todo o sentido quando o objetivo é monitorar o deslocamento de nuvens ou estimar a presença de frentes frias e massas de ar quente em escala continental, fenômenos que demandam pouco detalhamento espacial, mas alta periodicidade de informações. Por outro lado, se o objetivo é monitorar, por exemplo, mudanças do uso e ocupação da terra em uma cidade ou identificar focos de desflorestamento na Amazônia, essas imagens terão pouca utilidade. Para esses outros casos, devem ser utilizadas imagens provenientes de sensores com maior detalhamento espacial que, em alguns casos, chega a níveis submétricos; em contrapartida, sua periodicidade é menor, chegando a dias ou até semanas.

As primeiras pesquisas envolvendo o uso das imagens de satélites no monitoramento de recursos naturais terrestres ocorreram na década de 70, sobretudo a partir do programa *Resources Technology Satellite* (ERTS), que em 1975 passou a se chamar Landsat. Trata-se do programa mais bem sucedido de desenvolvimento de satélites de observação de recursos naturais terrestres, implementado pela *National Aeronautics and Space Administration* (Nasa), e em operação até hoje com o Landsat-8, lançado em fevereiro de 2013. Atualmente, as pesquisas envolvendo o uso de imagens de satélite encontram-se em um patamar científico elevado, não só pela diversidade de dados, mas principalmente pela política de disseminação e compartilhamento livre dessas informações pela internet, o que estimula seu uso nas mais variadas áreas do conhecimento. O Brasil foi um dos precursores dessa política de liberação de imagens, pelo Instituto Nacional de Pesquisas Espaciais (Inpe), que desenvolveu um catálogo de imagens com mecanismos de buscas simples, com o objetivo de disseminar e estimular o uso das imagens do satélite CBERS (*China-Brazil Earth Resources Satellite*), fruto da parceria técnico-institucional com a China. Outros exemplos dessa política podem ser encontrados em repositórios dos Estados Unidos, como no caso da gama de produtos pré-processados do sensor *Moderate Resolution Imaging Spectroradiometer* (Modis), e no caso das imagens do programa Landsat. Também são encontrados repositórios da Comunidade Europeia, como no caso de alguns produtos derivados do programa *Satellite Pour l'Observation de la Terre* (Spot).

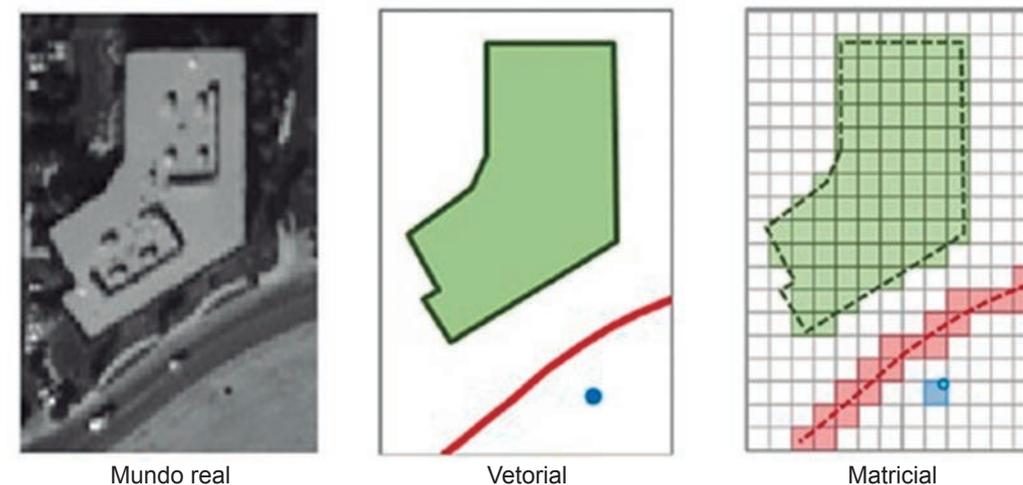
Uma das tecnologias responsáveis pela popularização do sensoriamento remoto é o Google Earth, um programa de computador desenvolvido e distribuído pela Google, cuja função é apresentar um modelo tridimensional do globo terrestre, construído a partir de um mosaico de imagens de satélite obtidas de diversas fontes, em diversas resoluções. Apesar de não ter sido o pioneiro neste tipo de serviço (o primeiro foi o Nasa World Wind), o Google Earth está em constante evolução, levando aos usuários um conjunto de inovações nunca antes vista, que inclui

atualizações regulares da base de imagens, a integração com outras bases cartográficas e outras funcionalidades.

### 2.3. O Sistema de Informações Geográficas

Considerando as geotecnologias para visualização, análise e tratamento dos dados geoespaciais, destacam-se os Sistemas de Informações Geográficas (SIG), cujo nome vem do termo em inglês *Geographic Information System* (GIS). Existem diversas definições sobre o que são SIGs, desde as mais complexas às mais simples. Pires et al. (1994) definem SIG como um sistema que realiza o tratamento computacional de dados geoespaciais, armazenando, gerenciando e recuperando informações. Estes sistemas são muito utilizados em ambientes de decisão, provendo aos usuários facilidades de combinar as informações de uma determinada região. A principal diferença entre um SIG e um sistema de informação convencional é sua capacidade de armazenar tanto os atributos descritivos dos dados, quanto as geometrias dos diferentes tipos de dados geográficos. As principais características de SIGs são: inserir e integrar, numa única base de dados, informações espaciais textuais e outras fontes de dados como imagens de satélite e dados de GPS; e oferecer mecanismos para combinar as várias informações, por meio de algoritmos de manipulação e análise, bem como para consultar, recuperar e visualizar o conteúdo da base de dados geográficos.

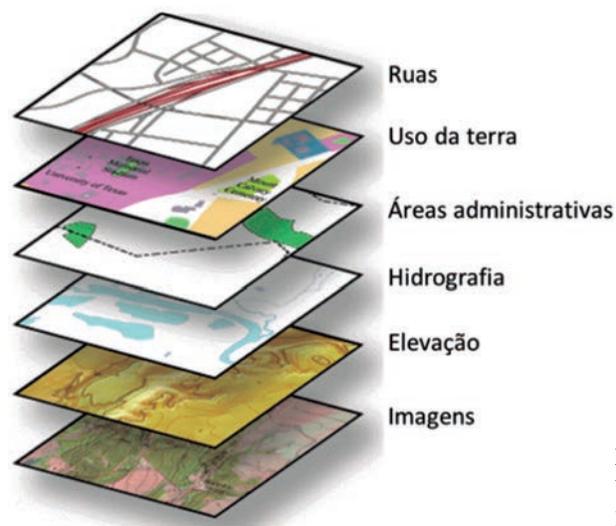
Do ponto de vista computacional, os dados geoespaciais podem ser representados dentro de um SIG a partir de dois modelos principais: o vetorial e o matricial (também chamado raster). No modelo vetorial, a apresentação de um elemento ou objeto é uma tentativa de reproduzi-lo da forma mais precisa possível do ponto de vista geométrico. Neste tipo de representação, qualquer entidade ou elemento gráfico de um mapa é reduzido a três formas básicas: pontos, linhas e polígonos. Já o modelo de representação matricial consiste no uso de uma malha quadriculada regular sobre a qual se constrói, célula a célula, o elemento que está sendo representado. A cada célula, atribui-se um código referente ao atributo estudado, de tal forma que o computador saiba a que elemento ou objeto pertence determinada célula (CÂMARA et al., 2001). A Figura 2 ilustra como uma área do mundo real pode ser representada a partir dos modelos vetorial e matricial.



**Figura 2.** O mundo real e sua representação computacional a partir dos modelos vetorial e matricial.

Fonte: Adaptado de Satar (2014).

A abordagem tradicionalmente utilizada para a organização de dados geoespaciais em um SIG é a distribuição em camadas, também denominadas *layers* ou planos de informação, em que cada nível aborda um tema distinto para uma dada região geográfica (Figura 3). Por exemplo, uma imagem de satélite de uma região é uma camada, assim como os municípios desta região, a sua geomorfologia e sua hidrologia. Cada camada é representada internamente usando estruturas lógicas próprias de cada SIG e é armazenada em arquivos distintos, de acordo com o formato do sistema utilizado.



**Figura 3.** Representação das camadas de temas em um SIG.

Fonte: Adaptado de ArcGIS Resources (2014).

## 3 Organização e disponibilização dos dados geoespaciais

### 3.1 Características dos dados geoespaciais

Dados geoespaciais possuem três características fundamentais: espaciais, não-espaciais e temporais. As características espaciais informam a posição geográfica do fenômeno e a sua geometria, correspondendo à pergunta “onde?”. As não-espaciais descrevem o fenômeno e suas propriedades e informam sobre o “como?”. Por fim, as características temporais informam o tempo de validade do dado e suas variações ao longo do tempo e dizem respeito ao “quando”.

O dado geoespacial, por suas características, apresenta propriedades topológicas e geométricas. As topológicas são baseadas nas posições relativas dos objetos no espaço, como conectividade, orientação, adjacência e contenção. Já as geométricas são propriedades métricas que representam a geometria das entidades, obtidas a partir de suas feições geométricas primitivas, como pontos, linhas e polígonos. Considerando essas propriedades, é possível definir as relações entre dados geoespaciais, que são essenciais nas aplicações geográficas. É por meio delas que são executadas operações importantes como, por exemplo, o cálculo de distância entre pontos, a definição da área de um talhão agrícola, a localização de estradas que cruzam um determinado município, a identificação das cidades de um estado por onde passa um rio, entre outros.

As principais relações possíveis entre dados geoespaciais são: topológica, métrica e de distância, orientação, ordem e *fuzzy*. As relações topológicas são aquelas que descrevem os conceitos

de vizinhança, incidência, sobreposição, mantendo-se invariante ante a transformações, como escala e rotação. Os relacionamentos métricos e de distância consideram distâncias e direções, expressando a métrica das feições com referência a um sistema de coordenadas. Relações de orientação ou direcionais descrevem a orientação no espaço, como norte e sul, e relações de ordem são relativas à ordem total ou parcial dos objetos espaciais, como em frente a, atrás de, acima e abaixo. Por fim, tem-se as *fuzzy*, que envolve termos que não são claramente entendidos como relacionamentos padrão. Exemplos: próximo a, perto, longe, na região de. Todos os relacionamentos são importantes para o tratamento, análise e uso da informação geoespacial. No entanto, dentre todos os relacionamentos, os topológicos são os mais usados.

### 3.2 Os bancos de dados geoespaciais

Para que a informação geográfica seja tratada usando um computador, é necessária a sua representação computacional, ou seja, a adoção de estruturas e arquivos que representem essa informação no mundo computacional. Por exemplo, como representar uma determinada área florestal que está sendo alvo de desflorestamento? Diferentemente dos dados tradicionais, constituídos basicamente de atributos descritivos como nome, data e descrição, a representação de um dado geoespacial requer, também, a descrição de sua forma geométrica associada à sua posição geográfica. Assim, usando os relacionamentos espaciais, torna-se possível executar diferentes análises considerando as características espaciais dos elementos envolvidos. No exemplo citado, a área florestal poderia ser representada por um polígono e as áreas desflorestadas por outros polígonos menores. Uma relação de “está contido” indicaria quais, dentre várias áreas identificadas, fazem parte da área original. Nesse caso, atributos temporais permitiriam a manutenção de um histórico das áreas.

Para que esses dados sejam manipulados, é necessário o seu armazenamento em estruturas adequadas. Surge então o conceito de Banco de Dados Geográficos, que nada mais é do que um banco de dados que manipula dados geoespaciais. Em um banco de dados geográficos, os dados são georreferenciados, analisados por meio de operações espaciais e, geralmente, sua visualização é cartográfica. Assim como em bancos de dados tradicionais, existem, disponíveis, os Sistemas Gerenciadores de Banco de Dados (SGBD) com extensão para dados geográficos. A maior vantagem da adoção desses sistemas é que eles implementam os tipos e operadores espaciais, além das outras funcionalidades de um SGBD, como independência de dados e acesso eficiente; compartilhamento de dados e redução de redundância; integridade de dados e segurança; administração uniforme de dados; desenvolvimento de aplicações em tempo reduzido e acesso concorrente. Tudo isso facilita a tarefa do desenvolvedor de sistema de informação geográfica. As principais extensões para dados geográficos são DB2 Spatial Extender (IBM CORPORATION, 2014a), Informix Spatial (IBM CORPORATION, 2014b), OracleSpatial (ORACLE CORPORATION, 2014) e PostGIS (POSTGIS PROJECT STEERING COMMITTEE, 2014).

A maioria dos sistemas em uso trata principalmente de dados vetoriais, oferecendo meios de armazenamento e operações para sua manipulação. No entanto, muitos dados geoespaciais são disponibilizados em formato matricial, como as imagens de satélite; neste sentido, extensões para manipulação desse tipo de dado têm sido apresentadas. Um exemplo é a extensão *Raster Well-Known Text (WKT)* do PostGIS, que possibilita armazenar e analisar os dados matriciais, promovendo avanços em seu uso.

O banco de dados geoespaciais é considerado parte de um SIG, sendo responsável por prover funcionalidades de armazenamento e de manipulação dos dados. A Figura 4 ilustra a arquitetura de um SIG. Nela, os dados são armazenados num banco de dados geográficos, que é gerenciado por um SGBD-geográfico. O usuário acessa os dados via uma interface gráfica e as operações disponíveis para ele tradicionalmente são: entrada e integração de dados; consulta e análise espacial desses dados, por meio dos relacionamentos espaciais; e visualização dos dados e dos resultados das análises na forma de mapas.

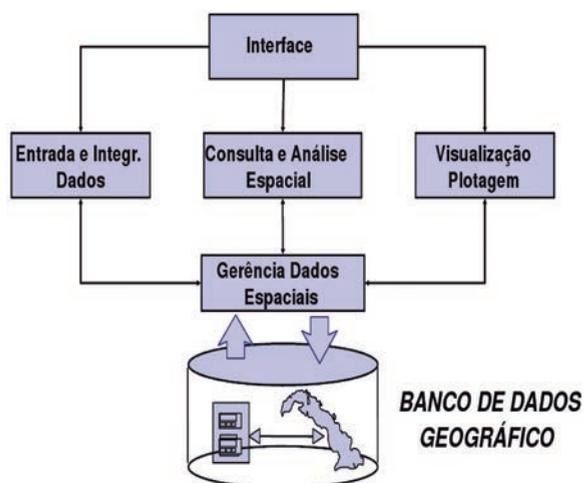


Figura 4. Arquitetura de um SIG.  
Fonte: Casanova et al. (2005).

### 3.3 Padrões e compartilhamento de dados geoespaciais

Com o aumento de sua capacidade de transmissão de dados, a *World Wide Web* (WWW) está se tornando cada vez mais o principal meio de troca de informações, incluindo os níveis pessoal, corporativo, governamental e acadêmico. Neste cenário, problemas relacionados à baixa eficiência no acesso, ao armazenamento e ao uso dos dados compartilhados se tornam cada vez mais relevantes e são resultado de vários fatores. A produção dos dados em ambientes computacionais heterogêneos, com diferentes propósitos e modelos da realidade, resultam em dados expressos em diferentes formatos, e com diferentes semânticas. As dificuldades de compartilhamento e uso efetivo dos dados nestas condições definem o problema de interoperabilidade de dados (STUCKENSCHMIDT; HARMELEN, 2004). O consórcio *Open Geospatial Consortium* (OGC) (OPEN GEOSPATIAL CONSORTIUM, 2014) vem realizando esforços para superação deste problema no contexto da produção de dados geoespaciais.

O OGC é um consórcio formado por cerca de 490 membros, envolvendo instituições acadêmicas, empresas privadas e instituições governamentais, que participam de forma consensual na definição de padrões públicos relacionados à produção de dados geoespaciais. Estes padrões permitem incorporar a resolução de problemas de interoperabilidade de dados geoespaciais em soluções de TI sob diferentes plataformas computacionais, fornecendo, ao desenvolvedor, ferramentas que possibilitam que serviços geoespaciais e informações geoespaciais complexas possam ser acessados e utilizados por diferentes tipos de aplicações.

As especificações OGC são a base para solução dos problemas de interoperabilidade por diferentes tecnologias e ferramentas utilizadas durante o processo de gestão de dados geoespaciais.

Além disso, envolvem desde padrões de formato para troca de dados geoespaciais até padrões para armazenamento, processamento e acesso aos dados.

#### 3.3.1 Padrões OGC voltados aos bancos de dados geoespaciais

A especificação OGC - *Simple Feature Access* (SFA) define um modelo padrão para representação de geometrias georreferenciadas bidimensionais. Esse modelo especifica tanto as estruturas de dados para representação das geometrias quanto as operações possíveis para cada tipo de dado geoespacial. O principal impacto dessa especificação pode ser verificado pela padronização no armazenamento e acesso a dados geoespaciais pelos diferentes SGBDs disponíveis do mercado. A extensão PostGIS do SGBD PostgreSQL (THE POSTGRESQL GLOBAL DEVELOPMENT GROUP, 2014) é o principal exemplo de aderência ao padrão SFA no âmbito dos softwares livres. Nesse SGBD, os tipos de dados geoespaciais, funções e suporte a consultas, considerando relacionamentos geoespaciais, seguem a especificação SFA. O mesmo padrão é adotado por mais 50 soluções para armazenamento de dados geoespaciais implementadas por mais de 20 empresas, dentre as quais podemos destacar as extensões geoespaciais já citadas dos SGBD Oracle, DB2 e Informix.

#### 3.3.2 Padrões OGC voltados aos serviços geoespaciais

O modelo OGC *Reference Model* (ORM) (PERCIVALL, 2003) descreve um ambiente de computação distribuída baseado na infraestrutura da WWW, o qual tem como objetivo definir uma plataforma padrão sob a qual são elaboradas as especificações OGC. Essas especificações descrevem aspectos relativos à representação e ao processamento de dados geoespaciais, tais como, formatos padrões para representação de dados obtidos de sensores, representação de objetos georreferenciados, formatos para intercâmbio de dados geoespaciais, além da arquitetura e integração de serviços geoespaciais.

Alguns exemplos de tipos de serviços especificados na infraestrutura proposta pela ORM são:

- 1) *Web Map Service* (WMS): serviço que gera mapas dinamicamente a partir de dados georreferenciados, sejam na forma matricial ou vetorial. Um WMS é implementado como uma aplicação Web que pode ser acessada em uma *Uniform Resource Locator* (URL) por meio de um navegador. O formato da URL reflete os parâmetros necessários para execução das operações pelo serviço.
- 2) *Web Feature Service* (WFS): serviço previsto no ORM provendo acesso remoto a dados geoespaciais vetoriais. Este serviço possibilita, além da consulta, operações para criação, remoção, atualização e seleção de dados. O acesso a esse serviço possibilita a troca de dados geoespaciais entre instituições diretamente por meio de ferramentas de SIG.
- 3) *Web Coverage Service* (WCS): serviço que permite a recuperação de dados geoespaciais no formato *raster*. O WCS fornece essas informações em formato adequado para uso em modelos científicos e visualizações, por exemplo.
- 4) *Web Processing Service* (WPS): este serviço define uma estrutura geral padrão para que instituições que desejem oferecer funcionalidades de geoprocessamento possam disponibilizá-las pela Internet. As funcionalidades podem ser desde simples combinações de planos de informação até a execução de modelos sofisticados sobre dados geoespaciais. Para descrever esta ampla gama de necessidades, o WPS é especificado como um mecanismo geral de interface entre os clientes e servidores do serviço de geoprocessamento. Esta descrição não impõe nenhum

formato específico e nem número de entradas e saídas que um serviço necessita e nem mesmo o modo de interação com o servidor. Ao invés disto, a especificação WPS apresenta elementos de como estas necessidades específicas podem ser descritas e invocadas pela da web.

- 5) *Catalogue Service (CS)*: este serviço de catalogação fornece uma visão integrada de todos os demais serviços OGC. Os serviços de dados, visualização e processamento são implementados e publicados de forma independente. Os serviços são descritos em um catálogo gerenciado pelo serviço CS, o qual permite a busca considerando diferentes características de um serviço desejado e, desta forma possibilita a sua localização.

O consórcio OGC é responsável também pela especificação da linguagem *Geography Markup Language (GML)* que permite a modelagem, transporte e armazenamento de informações geográficas. Um dado expresso em GML corresponde a um documento *eXtensible Markup Language (XML)*, cuja estrutura reflete o modelo de entidades genérico elaborado pela OGC com o objetivo de possibilitar o intercâmbio de informações geográficas entre diferentes aplicações em diferentes domínios. Essa linguagem é um dos principais recursos, possibilitando a troca de dados geoespaciais vetoriais entre instituições. A atuação do consórcio OGC tem sido fundamental para viabilizar o compartilhamento e o uso de dados geoespaciais em nível mundial e suas especificações devem sempre ser consideradas durante o desenvolvimento de novas soluções ou ferramentas de TI que envolvam o tratamento de dados geoespaciais.

### 3.3.3 Infraestrutura Nacional de Dados Espaciais (Inde)

No Brasil, o Governo Federal vem implementando medidas que visam integrar os dados geoespaciais produzidos por diversas instituições nacionais. A principal delas é a Infraestrutura Nacional de Dados Espaciais (Inde), instituída em 2008, com o propósito de catalogar, integrar e harmonizar dados geoespaciais existentes nas instituições do governo brasileiro, produtoras e mantenedoras desse tipo de dado, de maneira que possam ser facilmente localizados, explorados e acessados para os mais diversos usos, por qualquer cliente que tenha acesso à Internet. Os dados geoespaciais devem ser catalogados por seus respectivos metadados, publicados pelos produtores/mantenedores desses dados. O acesso aos dados deve se realizar por meio do geoportal denominado SIG Brasil (BRASIL, 2014a).

De acordo com o Decreto N°6.666 de 27/11/2008, a Inde tem como principais objetivos:

- 1) Promover o adequado ordenamento na geração, armazenamento, acesso, no compartilhamento, na disseminação e no uso dos dados geoespaciais.
- 2) Promover a utilização, na produção dos dados geoespaciais pelos órgãos públicos das esferas federal, estadual, distrital e municipal, dos padrões e normas homologados pela Comissão Nacional de Cartografia - Concar.
- 3) Evitar a duplicidade de ações e o desperdício de recursos na obtenção de dados geoespaciais, por meio da divulgação da documentação (metadados), dos dados disponíveis nas entidades e nos órgãos públicos das esferas federal, estadual, distrital e municipal.

O Perfil de Metadados Geoespaciais do Brasil (Perfil MGB) é um padrão de metadados baseado na norma ISO 19115, criado com o objetivo de obter um subconjunto mínimo de metadados para descrever dados geoespaciais gerados no país. O MGB Foi elaborado pela Comissão Nacional de Cartografia (Concar), órgão colegiado do Ministério do Planejamento, Orçamento e Gestão e é o perfil de metadados adotado pela Inde.

## 4 Iniciativas da Embrapa Informática Agropecuária em soluções geoespaciais

### 4.1 Natdata - Plataforma de Informação de Recursos Naturais dos Biomas Brasileiros

Natdata é uma plataforma para integração de dados sobre os recursos naturais dos biomas brasileiros. É resultado de um projeto conduzido pela Embrapa Informática Agropecuária, com recursos de royalties pagos pela Monsanto, com o objetivo de prover a Empresa Brasileira de Pesquisa Agropecuária (Embrapa) de um ambiente tecnológico e integrado para a gestão da informação de recursos naturais dos biomas brasileiros, visando a geração de inteligência competitiva para a agricultura nacional, fornecendo subsídios para as ações de diagnóstico e gestão de forma mais precisa, especialmente nas áreas estratégicas de zoneamento ambiental, caracterização e manejo da biodiversidade, conservação do solo e da água. O projeto surgiu da demanda da Embrapa em desenvolver uma plataforma que pudesse organizar e integrar as diferentes bases geoespaciais produzidas e mantidas pela empresa, nem sempre disponíveis aos demais usuários. Diversos fatores contribuíam para essa indisponibilidade dos dados: a) por serem produzidos por entidades diferentes, eles geralmente encontravam-se dispersos geograficamente, armazenados nas instituições que os produziram; b) alguns deles ainda não encontravam-se organizados em banco de dados, estando em muitos casos armazenados sob a forma de planilhas, textos e até documentos em papel; e c) cada instituição adotava formato específico, o que dificultava seu uso em conjunto.

A principal dificuldade em organizar essa imensa massa de dados está relacionada à sua heterogeneidade em muitos aspectos: linguagem, formato, espacialidade, semântica e uso, características que dificultam sua integração. No entanto, dados de recursos naturais têm um ponto em comum: seu atributo geoespacial, que descreve suas características espaciais e referenciam alguma localidade na superfície na Terra e pode ser armazenado numa estrutura denominada espacial. Exemplos desse tipo de dado incluem informações sobre recursos naturais, como solos, clima e vegetação, essenciais para responder a questões do tipo “Qual a forma de intensificar a produção de cana-de-açúcar na região de Piracicaba com a manutenção dos biomas envolvidos?” ou “Qual o padrão climático esperado para uma região de determinado tipo de solo no bioma Cerrado?”. Essas questões são básicas para o estabelecimento de políticas públicas de intensificação agrícola com a conservação ambiental. Respostas a essas questões envolvem o uso combinado de dados provenientes de diferentes domínios.

O acesso integrado às diferentes informações existentes garante agilidade e qualidade na resposta gerada. Esse acesso requer um entendimento único do dado sendo usado/trocado. Ou seja, é necessário que se garanta o significado semântico comum do dado. A garantia da interoperabilidade semântica foi apontada por Sciore et al. (1994) como sendo a chave para o sucesso na integração de dados, motivadora de várias pesquisas em Banco de Dados, uma das áreas da Tecnologia de Informação.

A plataforma NatData fornece todos os dados necessários por meio de um sistema único, usando um formato comum, tornando mais fácil para os usuários a sua busca. Além disso, a maioria dos sistemas semelhantes disponíveis oferece apenas instrumentos de pesquisa tradicionais, com base em palavras-chave que, no contexto dos recursos naturais, não é suficiente por conta da busca por informações considerando questões diferentes. Neste sentido, a plataforma também

oferece ferramentas de pesquisa para recuperar informações considerando recursos geoespaciais, como um bioma, uma região geográfica ou até um determinado local informado pelo usuário.

A plataforma está sendo implementada usando o SGBD PostGreSQL/PostGIS, a tecnologia Java e considerando os padrões definidos pela OGC, abrigando hoje dados de clima, solos e biodiversidade. Considerando os resultados já obtidos, percebe-se a importância de uma plataforma que integre os diferentes tipos de dados num único local. Além disso, a facilidade de exportação dos dados em diferentes formatos aumenta sua usabilidade, permitindo sua adoção em diversos ambientes e setores decisórios que buscam a sustentabilidade. A plataforma prevê, como trabalhos futuros, a melhoria no compartilhamento da informação geográfica e a disponibilização de um conjunto de serviços que permitam a execução de análises e simulações, baseados em *workflows* científicos.

#### 4.2 Sistema Interativo de Suporte ao Licenciamento Ambiental (Sisla)

Na maioria dos estados brasileiros, as análises dos processos de licenciamento ambiental são realizadas de forma analógica (como o uso de mapas impressos), demandando tempo considerável até a sua conclusão. Além da morosidade desses procedimentos manuais, um mesmo processo pode percorrer várias gerências dentro do órgão ambiental, uma vez que para a autorização de um empreendimento há necessidade de se analisar as conformidades ambientais em relação às Unidades de Conservação, Áreas de Preservação Permanente (APPs), Reservas Legais, Áreas Indígenas, entre outros. Visando contribuir com a operacionalização da política governamental de forma eficiente, rápida e transparente, o Estado de Mato Grosso do Sul, por meio do Instituto de Meio Ambiente de Mato Grosso do Sul (Imasul), firmou uma parceria com a Embrapa Informática Agropecuária, com o intuito de reunir um conjunto de informações geográficas úteis à regularização e ao licenciamento ambiental e desenvolver um sistema informatizado de consultas, denominado Sistema Interativo de Suporte ao Licenciamento Ambiental (Sisla).

O Sisla é um sistema de informação geográfica para a Web (disponível no sítio do Imasul no endereço: <http://sisla.imasul.ms.gov.br>), que permite ao usuário obter a análise espacial e a emissão de relatório do entorno do seu empreendimento, a partir do cruzamento de informações sobre a declividade do terreno, a extensão dos biomas, as áreas protegidas (Unidades de Conservação e APPs) e as áreas indígenas de forma rápida, pela internet. O sistema foi implementado com rotinas PHP (PHP, 2014), MapServer (MAPSERVER, 2014), I3GEO (BRASIL, 2014b), banco de dados PostgreSQL/PostGIS e JavaScript (MOZILLA DEVELOPER NETWORK, 2014) e é um dos mais importantes resultados do Projeto GeoMS, coordenado pela Embrapa Informática Agropecuária, financiado pelo Governo de Mato Grosso do Sul.

Um usuário que possui informações georreferenciadas (como um ponto, uma linha ou um polígono) do seu empreendimento, pode obter a análise espacial do seu entorno em menos de dois minutos. Esse fato permitiu ao Estado de Mato Grosso do Sul passar da análise espacial do licenciamento ambiental de forma analógica, manual, por vezes demorada, para uma análise na forma digital, via internet, rápida, segura e precisa. Essa modificação de procedimento promoveu um impacto positivo na qualidade das análises de regularização ambiental e autorização de licenciamento, principalmente para implantação de empreendimentos agropecuários, auxiliando na sustentabilidade da agricultura e na conservação ambiental. Outro impacto importante que pode ser considerado é o baixo custo do desenvolvimento e implantação do Sisla, proporcionado pela parceria técnica e financeira e pelo uso de softwares livres e imagens de satélite gratuitas. Além

de permitir download de mapas e imagens de satélite, essa ferramenta auxilia na implantação das ações exigidas pelo Novo Código Florestal para todo o território nacional.

Com uma média mensal de acessos em torno de 2,7 mil e mais de 3.000 usuários cadastrados, já foram registrados aproximadamente 200 mil acessos ao sistema, desde a sua implantação em outubro de 2008. Entre os usuários estão órgãos públicos, como Ministérios, Secretarias Estaduais de Meio Ambiente, universidades, Polícia Ambiental, instituições de pesquisa; e privados, como Organizações Não-Governamentais, bancos, empresas de consultoria ambiental, cooperativas, sindicatos, produtores rurais, entre outros.

#### 4.3 WebGIS Amazônia Legal

A implantação harmônica de políticas de gestão territorial e ordenamento do uso e ocupação da terra é uma necessidade premente no Brasil, em particular na Amazônia Legal, foco de amplas preocupações ambientais, geopolíticas e de desenvolvimento do país. O Zoneamento Ecológico-Econômico (ZEE), importante instrumento de planejamento territorial, tem como objetivo viabilizar o desenvolvimento sustentável a partir da compatibilização do desenvolvimento socioeconômico com a conservação ambiental. O ZEE vem sendo realizado em todos os estados da Amazônia Legal, sob a coordenação geral do Ministério do Meio Ambiente (MMA), que mobiliza para apoio do Consórcio ZEE Brasil, um conjunto de empresas públicas de notória especialização em suas respectivas áreas de atuação. No entanto, não há uma uniformidade no nível e parâmetros mínimos para todos os estados, sobretudo ao se comparar metodologias e critérios em que a execução não foi coordenada pelo Consórcio. Além disso, as diretrizes de uso e ocupação são discrepantes, principalmente nas áreas fronteiriças, requerendo discussão para se estabelecer um mínimo de uniformidade para ser colocada em prática. Neste sentido, o Projeto “Uniformização do Zoneamento Ecológico-econômico da Amazônia Legal e Integração com Zoneamentos Agroecológicos da Região”, financiado pela Finep - Inovação e Pesquisa, tem por objetivo articular os ZEEs dos Estados da Amazônia Legal, orientados para uma visão macrorregional e integrá-los com os Zoneamentos Agroecológicos. Além do estabelecimento de parâmetros mínimos para uniformizar e integrar os ZEEs desses estados e a compatibilização das legendas e diretrizes de uso e ocupação na região, o projeto prevê a disponibilização de uma infraestrutura de dados espaciais e o desenvolvimento do WebGIS Amazônia Legal, uma ferramenta Web capaz de dar suporte à manutenção e disponibilização do conteúdo gerado com a aquisição, a análise e a unificação das informações do projeto, além de oferecer subsídios para tomadas de decisão, especialmente em assuntos que dizem respeito à gestão ambiental.

Nesse sistema, desenvolvido a partir de ferramentas de código livre, os dados são armazenados em um banco de dados PostgreSQL, o qual utiliza as funcionalidades da extensão PostGIS para implementar as consultas espaciais. A utilização da plataforma i3GEO (BRASIL, 2014b) possibilita, além da interatividade de uma ferramenta Web Mapping, reunir características adicionais que permitem a análise geográfica de temas, desenho vetorial e outras ferramentas que auxiliam o usuário na exploração de detalhes do mapa, exibindo diversas informações pela mudança de parâmetros de visualização. O sistema inclui ainda uma ferramenta de geração de relatórios de análise espacial, capaz de cruzar um dado vetorial inserido pelo usuário (polígono, ponto ou linha) com os diversos mapas contidos no WebGIS, obtendo-se como resultado um relatório com as análises espaciais realizadas, descrevendo de forma detalhada, distância, interseção e classes de cada tema selecionado que contém ou estão contidos no dado fornecido pelo usuário.

#### 4.4 Banco de Produtos MODIS

As imagens de satélite têm sido uma fonte importante de informações para estudos dos ecossistemas, pois proveem a necessária visão sinótica e temporal da superfície terrestre. Atualmente, existe uma série de sensores disponíveis aos usuários, gerando dados da superfície terrestre com diferentes detalhamentos e periodicidades. As políticas de incentivo ao compartilhamento de dados, aliadas ao desenvolvimento de sistemas web de distribuição, têm facilitado o acesso do público em geral às imagens de satélite, permitindo o desenvolvimento de estudos nos mais variados temas. Um exemplo são as imagens do sensor Modis, principal instrumento a bordo das plataformas orbitais Terra e Aqua, administradas pela Nasa.

Uma das vantagens no uso dos dados Modis é a alta periodicidade de revisita, por conta de sua alta resolução temporal, representando uma fonte importante de informações para a construção de séries temporais de imagens e sua aplicação nos mais diversos estudos sobre a superfície terrestre. Apesar de existirem repositórios que oferecem produtos pré-processados derivados do sensor Modis, como o MOD13Q1 e o MYD13Q1 (NASA, 2014), que disponibilizam imagens de índices vegetativos da cobertura terrestre, esses dados são compartilhados em formatos pouco usuais, numa projeção cartográfica especial, e em recortes espaciais de 10 x 10 graus de longitude/latitude, denominados tiles.

Para facilitar o acesso a esses produtos no Brasil, a Embrapa Informática Agropecuária iniciou o desenvolvimento do Banco de Produtos MODIS na Base Estadual Brasileira, com o intuito de armazenar e disponibilizar ao usuário imagens já prontas para uso, em recortes estaduais, sem a necessidade de qualquer processamento complementar. No desenvolvimento do banco, além do SGBD PostgreSQL, foi utilizado o ambiente GeoNetwork (GEONETWORK OPENSOURCE, 2014), uma ferramenta web de código livre para a documentação, edição e disseminação de metadados geográficos. O GeoNetwork é um ambiente padronizado e descentralizado para a gestão de informação espacial, desenhado para proporcionar acesso a bancos de dados georreferenciados, produtos cartográficos e metadados relacionados, obtidos a partir de uma variedade de fontes de dados. O Banco de Produtos MODIS pode ser acessado pelo endereço [www.modis.cnptia.embrapa.br](http://www.modis.cnptia.embrapa.br) e, atualmente, disponibiliza quase 50 mil imagens da série histórica completa dos índices vegetativos NDVI e EVI (derivados dos produtos MOD13Q1 e MYD13Q1), em recortes estaduais, com atualização periódica a cada 8 dias, conforme disponibilização dos dados pelo repositório da Nasa.

#### 4.5 TerraClass - Sistema de Monitoramento do Uso e Cobertura da Terra nas Áreas Desflorestadas da Amazônica Legal

Constituindo um dos principais focos das atenções internacionais e nacionais, sobretudo pela sua importância absoluta e relativa no montante Global de ecossistemas florestais ainda preservados, a Amazônia Legal é uma região cuja política de ocupação histórica determinou um padrão de intervenção humana incompatível com os atuais princípios de sustentabilidade ambiental, social e econômica.

Desde o final da década de 1980 os desflorestamentos na Amazônia Legal vêm sendo monitorados anualmente pelo Projeto Prodes “Monitoramento da Floresta Amazônica Brasileira por Satélite” e o conhecimento acumulado possibilitou a definição e monitoramento dos impactos de políticas públicas que tinham como objetivo alterar o estado de emergência crônica associado a esse fenômeno.

Em 2008, atendendo a uma demanda do Governo Federal que buscava identificar quais eram os destinos dados às áreas desflorestadas da Amazônia Legal, como uma maneira de aprofundar o conhecimento sobre o fenômeno dos desflorestamentos e de promover, além das políticas de comando e controle para contenção dos desflorestamentos, uma agenda positiva para orientar e estimular o bom uso das terras das áreas já desflorestadas, foi articulado o projeto TerraClass “Sistema de Monitoramento do Uso e Cobertura da Terra nas Áreas Desflorestadas da Amazônica Legal” (COUTINHO et al., 2013).

Os avanços percebidos na área de Tecnologia da Informação e Comunicação das últimas décadas, tanto em termos do aumento da capacidade de armazenamento e processamento de dados e informações, quanto na sua capacidade de comunicação e compartilhamento, tornou possível o enfrentamento, o desafio de desenvolver e implementar um sistema de monitoramento da dinâmica de uso e cobertura de uma região tão vasta e de acesso extremamente complexo.

A adoção de geotecnologias relacionadas à aquisição, ao processamento e à disponibilização de dados geográficos, possibilitou a integração de diferentes metodologias de processamento de dados de sensores remotos orbitais, para a geração sistemática de mapas sobre o uso e cobertura da terra de toda a região e, mais importante do que isto, têm viabilizado a sua publicação integral por meio da internet.

Com a ampliação da série histórica de dados do Projeto TerraClass e a ampliação da sua área de abrangência, incluindo também as áreas antropizadas do bioma Cerrado, para atender nova solicitação do Governo Federal, surge uma nova demanda, relacionada à disponibilização de serviços geográficos pela internet, para potencializar ainda mais a adoção dos dados e informações gerados pelo projeto. Considerando a importância estratégica dessa demanda, a Embrapa Informática Agropecuária está desenvolvendo um novo ambiente computacional que permita, além de organizar, armazenar e disponibilizar dados e informações sobre esses biomas, oferecer serviços geográficos via web, para potencializar o atendimento das demandas dos diferentes tipos de usuários desses produtos.

## 5 Visão de futuro sobre as geotecnologias e dados geoespaciais

A visão de futuro é a base de todo planejamento estratégico e é o que orienta as atividades cotidianas de uma organização. Por meio de exercícios de imaginação do futuro, as chances de identificar oportunidades e de antecipar as barreiras e tendências são maiores.

Em outubro de 2011, a Organização das Nações Unidas (ONU) promoveu um fórum de especialistas em Gestão Global da Informação Geoespacial - *Initiative on Global Geospatial Information Management* (GGIM), que contou com a participação do Brasil, quando foram criados grupos de trabalhos para elaborar uma visão de futuro de 5 a 10 anos do setor de geotecnologias. O interesse em saber como as novidades se relacionariam com o desenvolvimento humano, levando em consideração o crescimento econômico, a sustentabilidade ambiental, a gestão de desastres e o bem-estar social, estimularam essa iniciativa (CARPENTER; SNELL, 2014).

No documento produzido pelo comitê da ONU, que reúne os especialistas no gerenciamento das informações geoespaciais globais, foram elencadas quatro principais preocupações (CARPENTER; SNELL, 2014):

- 1) No processo de evolução das informações geoespaciais, considerar a passagem da visão comum para a gestão de informação geoespacial nacional e mundial.
- 2) Investir em plataformas abrangentes para melhorar a partilha e divulgação nacional e global de dados geoespaciais.
- 3) Elaborar códigos de conduta para a informação geoespacial para melhorar a confiança do público.
- 4) Evolução de um relacionamento mutuamente benéfico entre o governo e as entidades não governamentais.

O GGIM identificou, ainda, cinco temas para desenvolvimento:

- 1) Acompanhar as tendências tecnológicas, bem como as diretrizes para a criação, manutenção e gestão de dados geoespaciais.
- 2) A necessidade de se evoluir quanto aos aspectos legais e políticos das informações geoespaciais.
- 3) Estabelecer responsabilidades ao longo da produção e disseminação, bem como treinamentos.
- 4) O papel dos setores privados e não governamentais.
- 5) O futuro papel dos governos na transmissão de dados e a gestão.

Com a evolução dos dispositivos de conexão sem fio, o grande mercado das geotecnologias não será mais o de produção de mapas, mas o advento dos serviços baseados em localizações, que engloba desde atendimentos de emergência em estradas até monitoramento de filhos adolescentes, passando por comunicação eficiente entre equipes de campo. O mercado atual de SIG contabiliza dois milhões de usuários, enquanto que o mercado para serviços remotos, baseados em posicionamento, chegará a pelo menos 10 vezes isso em poucos anos, ocupado por um novo tipo de indústria, capaz de unir conhecimento na área de telecomunicações e GPS com a inevitável demanda por fornecimento de dados inerentes a um ambiente móvel.

A ciência está respondendo a esses desafios com grande empenho e muita pesquisa. Os termos mapas e cartografia estão sendo substituídos por ontologia, incerteza e modelos espaço-temporais. A palavra ontologia diz respeito aos problemas de representar o conhecimento geográfico no computador. Não se trata mais de escolher entre vetores ou matrizes, mas buscar a equivalência semântica de conceitos. A pesquisa atual baseia-se no uso do paradigma de estatística espacial para representação e propagação de incertezas. O grande desafio dos modelos espaço-temporais é nos libertar da visão estática dos mapas dos SIG já que ações estão continuamente modificando o mundo à nossa volta.

As perspectivas científicas e tecnológicas na área de informação espacial são extremamente estimulantes e desafiadoras. Terá sucesso nesses novos paradigmas quem souber incluir, em sua visão do mundo geográfico, uma abordagem baseada na localização e na dinâmica das ações humanas (CÂMARA, 2000).

## 6 Considerações finais

A evolução das geotecnologias, a velocidade na geração dos dados geoespaciais e a massiva quantidade de geoinformação produzida impõem desafios enormes para o tratamento, organização e disponibilização desse volume de dados. As perspectivas futuras indicam que os dados geoespaciais estarão cada vez mais presentes no dia-a-dia das pessoas, aumentando a demanda por serviços geoespaciais e soluções para a melhoria de processos.

A demanda por processos eficientes para tratamento e organização de uma massa crescente de dados geoespaciais tem estimulado uma mudança de paradigma no que se refere à construção de aplicações científicas, incentivando o surgimento de novas plataformas e ferramentas de alta capacidade de processamento.

As capacidades de armazenamento e de processamento de dados geoespaciais devem ser incrementadas para que seja possível superar os desafios representados pelo chamado *Big Data*, definido como um conjunto de dados caracterizado pelo seu grande volume, sua alta velocidade de atualização e pela sua abrangência ou variedade de temas. Este tipo de dado se tornará cada vez mais presente nas atividades de pesquisa, sendo um resultado da utilização crescente de dados de sensores de maior resolução espacial e temporal, pela maior facilidade de acesso a produtos geoespaciais, e pela necessidade de uma análise integrada destes dados com dados não-espaciais.

## 7 Referências

- ARCGIS RESOURCES. **Key aspects of GIS**. 2014. Disponível em: <<http://resources.arcgis.com/en/home/>>. Acesso em: 2 out. 2014.
- BRASIL. Infraestrutura Nacional de Dados Espaciais. **SIG Brasil - O Portal Brasileiro de Dados Geoespaciais**. 2014a. Disponível em: <<http://www.inde.gov.br/a-inde/>>. Acesso em: 22 set. 2014.
- BRASIL. Portal do Software Público Brasileiro. **i3Geo**. 2014b. Disponível em: <[http://www.softwarepublico.gov.br/ver-comunidade?community\\_id=1444332](http://www.softwarepublico.gov.br/ver-comunidade?community_id=1444332)>. Acesso em: 22 set. 2014.
- CÂMARA, G. Perspectivas ao Norte do Equador. **InfoGEO**, Curitiba, v. 16, nov. 2000. Disponível em: <<http://mundogeo.com/blog/2000/12/01/perspectivas-ao-norte-do-equador/>>. Acesso em: 22 set. 2014.
- CÂMARA, G.; MONTEIRO, A. M. V. Conceitos básicos em Ciência da Geoinformação. In: CÂMARA, G.; DAVIS, C.; MONTEIRO, A. M. V. (Ed.). **Introdução à Ciência da Geoinformação**. São José dos Campos: Inpe, 2001. Disponível em: <<http://mtc-m12.sid.inpe.br/col/sid.inpe.br/sergio/2004/04.22.07.43/doc/publicacao.pdf>>. Acesso em: 22 set. 2014.
- CÂMARA, G.; CASANOVA, M.; MEDEIROS, C. B.; MAGALHÃES, G.; HEMERLY, A. **Anatomia de Sistemas de Informação Geográfica**. Campinas: Unicamp, 1996. 193 p. Trabalho apresentado na Escola de Computação, julho 1996, Campinas.
- CARPENTER, J.; SNELL, J. **Future trends in geospatial information management: the five to ten year vision**. New York: United Nations Initiative on Global Geospatial Information Management, 2013. 40 p. Disponível em: <<http://ggim.un.org/docs/Future-trends.pdf>>. Acesso em: 22 set. 2014.
- CASANOVA, M.; CÂMARA, G.; DAVIS, C.; VINHAS, L.; QUEIROZ, G. R. (Ed.). **Bancos de dados geográficos**. Curitiba: MundoGeo, 2005. Disponível em: <<http://www.dpi.inpe.br/livros/bdados/>>. Acesso em: 3 out. 2014.

- COUTINHO, A. C.; ALMEIDA, C. A.; VENTURIERI, A.; ESQUERDO, J. C. D. M.; SILVA, M. **Uso e cobertura da terra nas áreas desflorestadas da Amazônia Legal**. TerraClass 2008. São José dos Campos: Inpe; Brasília-DF Embrapa, 2013. 107 p. il.
- GEONETWORK OPENSOURCE. **GeoNetwork opensource**. 2014. Disponível em: <<http://geonetwork-opensource.org/index.html>>. Acesso em: 10 out. 2014.
- IBM CORPORATION. **DB2 spatial extender for Linux, Unix and Windows**. 2014a. Disponível em: <<http://www-03.ibm.com/software/products/en/db2spaext>>. Acesso em: 20 set. 2014.
- IBM CORPORATION. **Informix spatial software**. 2014b. Disponível em: <<http://www-01.ibm.com/software/data/informix/spatial/>>. Acesso em: 20 set. 2014.
- MAPSERVER. Open Source Web Mapping. **MapServer**. 2014. Disponível em: <<http://mapserver.org/>>. Acesso em: 25 set. 2014.
- MOZILLA DEVELOPER NETWORK. **JavaScript**. 2014. Disponível em: <<https://developer.mozilla.org/pt-BR/docs/Web/JavaScript>>. Acesso em: 20 set. 2014.
- NASA. Land processes Distributed Active Archive Center. **MODIS Data Products Table**. 2014. Disponível em: <[https://lpdaac.usgs.gov/products/modis\\_products\\_table](https://lpdaac.usgs.gov/products/modis_products_table)>. Acesso em: 20 set. 2014.
- OPEN GEOSPATIAL CONSORTIUM. Disponível em: <<http://www.opengeospatial.org/>>. Acesso em: 15 ago. 2014.
- ORACLE CORPORATION. **Oracle spatial and graph**. 2014. Disponível em: <<http://www.oracle.com/technetwork/database/options/spatialandgraph/overview/spatialandgraph-1707409.html>>. Acesso em: 20 set. 2014.
- PERCIVALL, G. (Ed.). **OGC reference model reference**. Version: 0.1.3. Reference number: OGC 03-040. 2003. Disponível em: <<http://rap.opengeospatial.org/orm.php>>. Acesso em: 3 out. 2014.
- PHP. **Hypertext preprocessor**. 2014. Disponível em: <<http://php.net/>>. Acesso em: 20 set. 2014.
- PIRES, M. F.; MEDEIROS, C. M. B.; SILVA, A. B. Modelling geographic information systems using an object-oriented framework. In: BAEZA-YATES, R. **Computer Science 2**. New York: Springer, 1994. p. 199-211. DOI: 10.1007/978-1-4757-9805-0\_18.
- POSTGIS PROJECT STEERING COMMITTEE. **PostGIS - spatial and geographic objects for PostgreSQL**. 2014. Disponível em: <<http://postgis.net/>> Acesso em: 25 set. 2014.
- SATAR, M. **Bab IV. Pengenalan ArcGIS**. 2014. Disponível em: <<https://musnanda.wordpress.com/2014/04/07/bab-iv-pengenalan-arcgis>>. Acesso em 03 out. 2014.
- SCIORE, E.; SIEGEL, M.; ROSENTHAL, A. Using semantic values to facilitate interoperability among heterogeneous information systems. **ACM Transactions Database Systems**, 19, v. 2, p. 254-290. 1994.
- STUCKENSCHMIDT, H.; HARMELEN, F. Van. **Information sharing on the semantic web**. New York: Springer, 2004. 276 p.
- THE POSTGRESQL GLOBAL DEVELOPMENT GROUP. **PostgreSQL: the world's most advanced open source database**. 2014. Disponível em: <<http://www.postgresql.org/>> Acesso em: 25 set. 2014.

# Parte IV

**TIC na automação das cadeias produtivas**





## TIC na segurança fitossanitária das cadeias produtivas

Jayme Garcia Arnal Barbedo  
Carlos Alberto Alves Meira

### 1 Introdução

Com a intensificação da indústria agropecuária, têm crescido os desafios e as preocupações relacionadas à segurança sanitária dos alimentos produzidos. A circulação de volumes cada vez maiores desse tipo de mercadoria exige que as medidas necessárias para garantir sua segurança sanitária sejam implementadas de maneira rápida, eficiente e barata. O controle manual tradicionalmente utilizado muitas vezes não é capaz de atender a esses requisitos. Como resultado, tecnologias de informação e comunicação têm sido cada vez mais utilizadas para:

- 1) Aumentar o grau de automação e, conseqüentemente, a velocidade dos processos de controle fitossanitário.
- 2) Identificar problemas sanitários tão cedo quanto possível, minimizando possíveis prejuízos econômicos, ambientais e sociais.
- 3) Identificar, a partir de variáveis ambientais e históricas, áreas potencialmente sujeitas a problemas sanitários, antes mesmo destes se manifestarem.

Este capítulo trata especificamente dos dois últimos itens. Na Seção 2, são mostradas iniciativas voltadas ao diagnóstico de doenças em plantas, explorando tecnologias como processamento digital de imagens e sistemas especialistas. A Seção 3, por sua vez, apresenta iniciativas voltadas à construção de modelos de previsão e sistemas de alerta de doenças de culturas agrícolas.

### 2 Diagnose de doenças

Desde os primórdios da agropecuária, uma das principais preocupações dos produtores tem sido evitar prejuízos ligados à manifestação e proliferação de doenças em suas lavouras, viveiros e rebanhos. A fim de que o combate a essas doenças seja eficaz, é essencial que o diagnóstico seja não apenas correto, mas tão rápido quanto possível. Na maioria das vezes, o monitoramento das condições sanitárias é feito visualmente, através da ida a campo de especialistas capazes de reconhecer possíveis problemas que estejam ocorrendo. Essa estratégia tem alguns problemas potenciais:

- Especialistas nem sempre estão disponíveis para realizar o monitoramento com a frequência necessária, especialmente em locais remotos.
- O custo associado a esse monitoramento por especialistas pode ser proibitivo para pequenos produtores.
- Ainda que haja disponibilidade de mão de obra e recursos, pode ser muito difícil fazer um monitoramento minucioso, especialmente no caso de grandes propriedades.

Como resultado, esforços vêm sendo despendidos na criação de ferramentas computacionais que auxiliem no combate aos problemas sanitários. O objetivo dessas ferramentas é reduzir o tempo necessário para que as primeiras ações sejam colocadas em andamento, especialmente na ausência de um especialista.

Esta seção apresentará alguns dos principais avanços alcançados nos últimos anos nas técnicas para detecção, quantificação e classificação de doenças na agropecuária, com ênfase naqueles baseados em imagens digitais dos sintomas das doenças, os quais são a grande maioria. Será dado particular destaque às ações que têm sido realizadas na Empresa Brasileira de Pesquisa Agropecuária (Embrapa).

## 2.1 Sistemas automáticos

Sistemas automáticos são aqueles que realizam todas as operações automaticamente com base nas imagens digitais a eles submetidos, com pouco ou nenhum envolvimento por parte do usuário. Este alto grau de autonomia possui algumas vantagens:

- O usuário não necessita de qualquer conhecimento específico a respeito do problema, já que toda a informação necessária para o diagnóstico está contida no algoritmo.
- Computadores não se cansam, portanto é possível fazer um grande número de avaliações ininterruptamente, o que possibilita monitoramento permanente.
- Computadores não estão sujeitos a ilusões de ótica que frequentemente acometem avaliadores humanos.
- Em geral possuem baixo custo de operação.

A principal desvantagem deste tipo de sistema é sua dependência em relação à qualidade da base de dados de imagens digitais utilizada no seu desenvolvimento. O resultado disso é que, embora seja possível conferir um certo grau de flexibilidade ao algoritmo, este só será capaz de lidar com situações para as quais ele foi treinado. Assim, é possível afirmar que um sistema automático para detecção, medição ou identificação de doenças só poderá ser tão bom quanto a base de dados utilizada. Na verdade, a falta de bases de dados abrangentes é o principal problema enfrentado atualmente nas pesquisas deste tipo, já que as técnicas de processamento de imagens e inteligência computacional estão suficientemente maduras para permitir o desenvolvimento de métodos realmente efetivos. Por esse motivo, há várias iniciativas em andamento para a construção de bases de dados mais completas, inclusive na Embrapa, conforme será descrito mais adiante.

A seguir é apresentado um breve estado da arte, bem como as iniciativas realizadas na Embrapa, relacionados aos temas de detecção, quantificação e classificação de doenças. É importante ressaltar que plantas têm recebido mais atenção nesses temas que animais, motivo pelo qual apenas plantas são tratadas nas seções apresentadas a seguir.

### 2.1.1 Detecção de doenças

Métodos para detecção de doenças em plantas normalmente são voltados ao monitoramento, em tempo real, da lavoura. Assim, o objetivo é detectar, o mais cedo possível, a presença de uma ou mais doenças capazes de trazer prejuízos à produção.

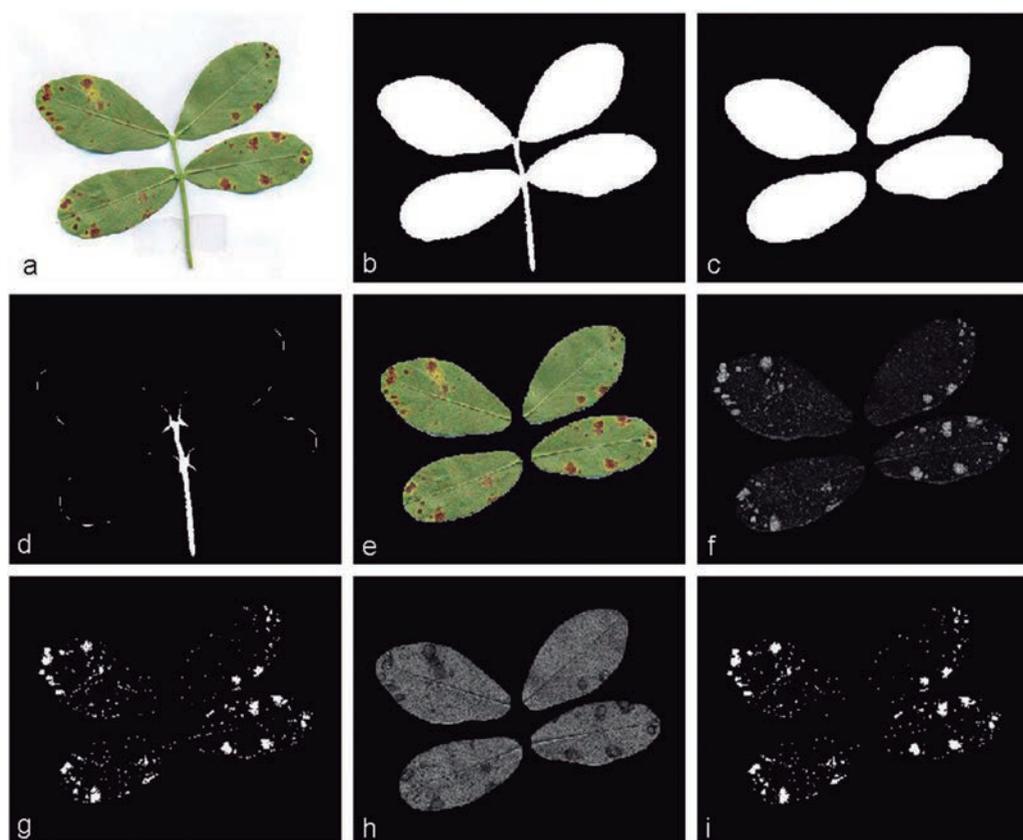
Embora o monitoramento e detecção de doenças seja de grande importância, não há na literatura muitas propostas voltadas especificamente para este fim. Isso é provavelmente decorrente do fato de que métodos para classificação de doenças, os quais normalmente empregam técnicas mais sofisticadas, se prestem também a este tipo de aplicação.

Dentre os métodos de detecção de doenças em plantas encontrados na literatura, destacam-se aqueles propostos por Sena Junior et al. (2003) e por Story et al. (2010). O primeiro é baseado na segmentação da imagem por limiarização e na contagem dos objetos detectados, enquanto o segundo realiza uma análise de regressão duplamente segmentada com base em atributos de cor e textura. Embora não haja na Embrapa nenhuma iniciativa especificamente dedicada a este fim, dentre os resultados a serem gerados dentro do projeto “Diagnóstico automático de doenças em plantas usando imagens digitais” (BARBEDO, 2013f) está um classificador de doenças que, com algumas poucas adaptações, terá a capacidade de realizar esse tipo de monitoramento. Mais informações a respeito deste projeto, o qual tem financiamento da Fundação de Amparo à Pesquisa do Estado de São Paulo (Fapesp) e está sob liderança da Embrapa Informática Agropecuária, serão fornecidas mais adiante, na parte de classificação de doenças.

### 2.1.2 Medição da severidade de doenças

A medição da severidade das doenças é muito importante no contexto do controle sanitário. Em geral, essa medição é feita de maneira subjetiva por especialistas, os quais usam escalas padronizadas para emitir um escore que indica o grau de severidade dos sintomas visíveis. Em alguns casos, faz-se também a medição manual das áreas usando um padrão quadriculado. Pesquisas demonstram que há diferenças significativas nos escores atribuídos pelos especialistas, levando a resultados inconsistentes (BOCK et al., 2010). Além disso, o processo de medição manual das áreas é cansativo, demorado e sujeito a erros. Por esses motivos, a automação da medição da severidade dos sintomas tem recebido grande atenção nos últimos anos. Um levantamento completo dos métodos propostos na literatura pode ser encontrado em Barbedo (2013a), sendo que serão citados aqui apenas aqueles com resultados mais relevantes. Esta seção focará nas três principais maneiras de se medir a severidade das doenças: estimando a área das lesões, contando o número de lesões, ou analisando as características de cor e textura das lesões.

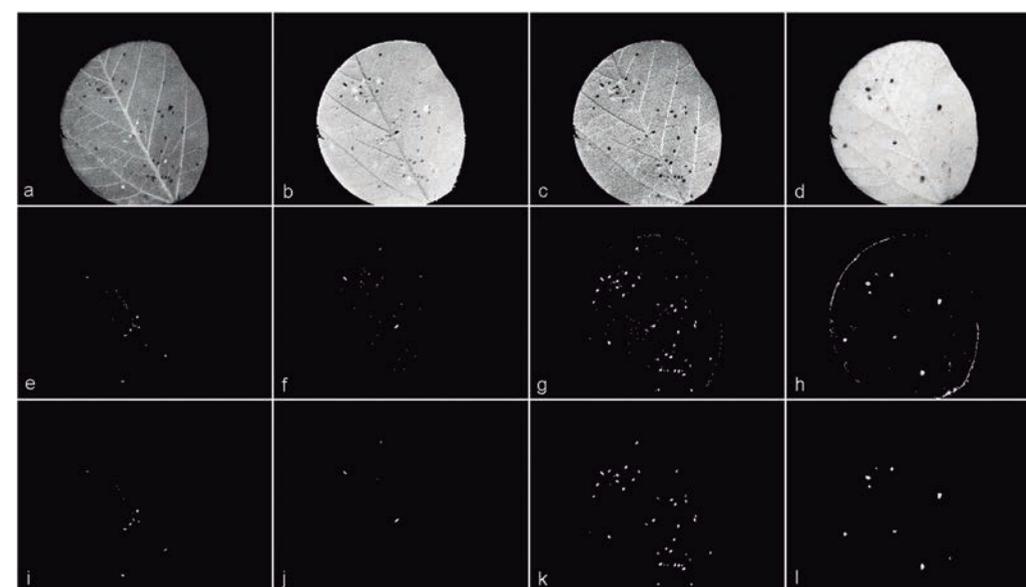
A estimativa da área das lesões é, por uma larga margem, a estratégia mais utilizada para determinar a severidade da doença. Dentre as ferramentas deste tipo recentes, pode-se citar aquelas propostas por Lloret et al. (2011), a qual é baseada em limiarização e morfologia matemática, e por Contreras-Medina et al. (2012), a qual é baseada em análise de cor para realizar a segmentação das lesões. Na Embrapa, a medição da área de lesões foi estudada no contexto do projeto “Contagem semi-automática de objetos genéricos em imagens digitais” (Contag), liderado pela Embrapa Informática Agropecuária. Esses estudos geraram um método para medição de lesões inteiramente baseado em conhecimento especialista e operações morfológicas, método este descrito em detalhes em Barbedo (2013a, 2014b). A Figura 1 apresenta um resumo de todos os passos que compõem o método desenvolvido. Este método será implementado como um módulo do sistema para contagem e medição de áreas de objetos, a ser disponibilizado na internet até o início de 2015.



**Figura 1.** Passos do algoritmo. a) Imagem típica presente na base de dados. b) Máscara binária antes do ajuste fino. c) Imagem aberta morfologicamente. d) Imagem contendo as localizações do caule e do pecíolo antes da eliminação de objetos espúrios. e) Imagem contendo apenas as regiões de interesse. f) Representação da imagem no canal amarelo do espaço de cor CMYK. g) Primeira estimativa para a localização das lesões. h) Representação da imagem no canal magenta do espaço de cor CMYK. i) Estimativa final para a localização das lesões.

A contagem de elementos ou estruturas que caracterizam a doença é também uma abordagem comum. Um exemplo desse tipo de estratégia pode ser encontrado em Boissard et al. (2008). Como mencionado anteriormente, o projeto Contag foi criado exatamente para lidar com o problema da contagem de objetos, embora seu escopo seja mais amplo que a simples contagem de estruturas relacionadas a doenças. Os resultados desse projeto foram aplicados, majoritariamente, a problemas de contagem de microrganismos (BARBEDO, 2012, 2013b, 2013c, 2013d). O algoritmo foi também aplicado à contagem de moscas brancas em folhas, as quais podem causar doenças e são uma importante fonte de perdas na agricultura (BARBEDO, 2014c). A Figura 2 fornece um exemplo de uso deste método, o qual é também inteiramente baseado em regras ad-hoc, geradas a partir de conhecimento de especialistas, e morfologia matemática. Todas essas ferramentas de contagem estarão também disponíveis como aplicativo web a partir do início de 2015.

O uso de características de cor e textura para medir a severidade de doenças é mais incomum que as outras duas estratégias, provavelmente devido às dificuldades em se garantir que as imagens capturem com fidelidade as características visuais dos sintomas, e também devido à falta de ho-



**Figura 2.** Ilustração dos passos do algoritmo para contagem de moscas brancas em folhas.

mogeneidade dessas características. Até onde vai o conhecimento dos autores, o único método a usar este tipo de abordagem foi proposto por Zhou et al. (2011).

### 2.1.3 Classificação de doenças

A classificação de doenças é, em geral, um problema mais difícil que aqueles apresentados até aqui, uma vez que, além de detectar a doença, o algoritmo deve tentar identificá-la. O problema se torna mais difícil à medida que um maior número de doenças é considerado. Apesar de haver um grande número de métodos para este fim, a grande maioria deles possui algum tipo de restrição em relação às espécies e doenças consideradas, bem como às condições sob as quais as imagens são capturadas (BARBEDO, 2013a). Dentre os métodos mais bem-sucedidos propostos na literatura, pode-se citar Camargo e Smith (2009), baseado em máquinas de vetores suporte; Huang (2007), baseado em redes neurais do tipo MLP; Pydipati et al. (2006), baseado em análise de discriminante e Xu et al. (2011), baseado em lógica nebulosa.

O projeto Digipathos pretende contribuir para o avanço desta área através de duas ações descritas a seguir.

a) Criação de uma base de dados contendo imagens de doenças de pelo menos 20 espécies com valor comercial no Brasil. Como comentado anteriormente, um método automático de diagnóstico somente pode ser tão abrangente quanto a base de dados usada no seu desenvolvimento, e as pesquisas na área sofrem com a falta de bases de dados estruturadas, completas e com imagens de boa qualidade. A fim de facilitar a submissão das imagens e organizar as imagens enviadas pelos pesquisadores envolvidos no projeto, foi criado um repositório web usando a plataforma DSpace (SMITH et al., 2003). A tela inicial do repositório é mostrada na Figura 3.

Além do repositório, foi criada uma planilha para Excel e OpenOffice, na qual os pesquisadores responsáveis podem, através do clique de um botão, listar todas as imagens presentes no diretório, visualizá-las, e preencher seus metadados.

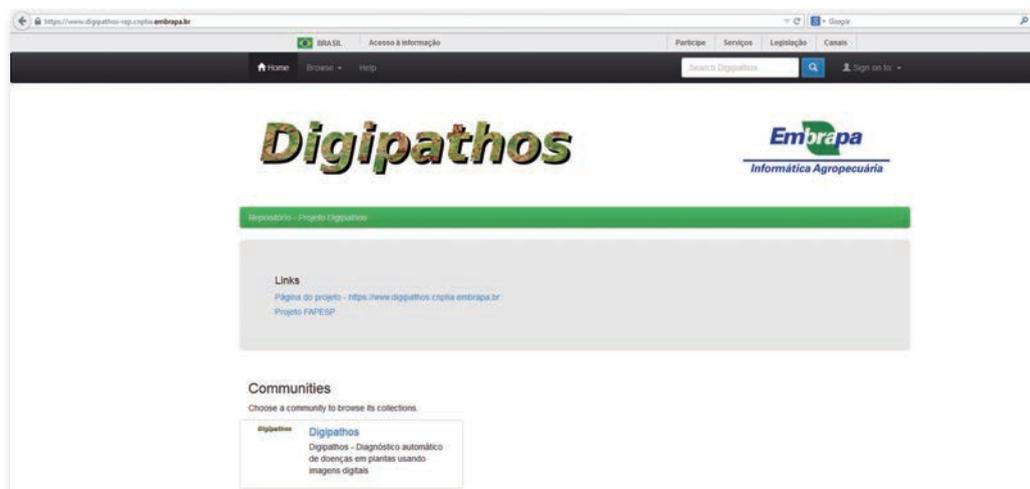


Figura 3. Tela inicial do repositório de imagens de sintomas de doenças em plantas.

Atenção especial foi dada também à captura das imagens. Embora se pretenda utilizar imagens capturadas sob condições normais ao longo do projeto, as etapas iniciais de desenvolvimento do método se beneficiariam de imagens com condições mais controladas de iluminação, e corrigidas em termos de cor e geometria. Por esse motivo, foi criado um dispositivo, consistindo de uma caixa e de um padrão (Figura 4), o qual, além de garantir condições homogêneas de captura, permite que se utilize um software para correções geométricas e de cor das imagens. Detalhes a respeito do dispositivo podem ser encontrados no website do projeto<sup>1</sup>.

b) Desenvolvimento de método para identificação de doenças em plantas. Esse método deverá se basear em técnicas de aprendizado de máquina, reconhecimento de padrões, morfologia matemática e conhecimento especialista. O aplicativo baseado nesse método terá duas ver-

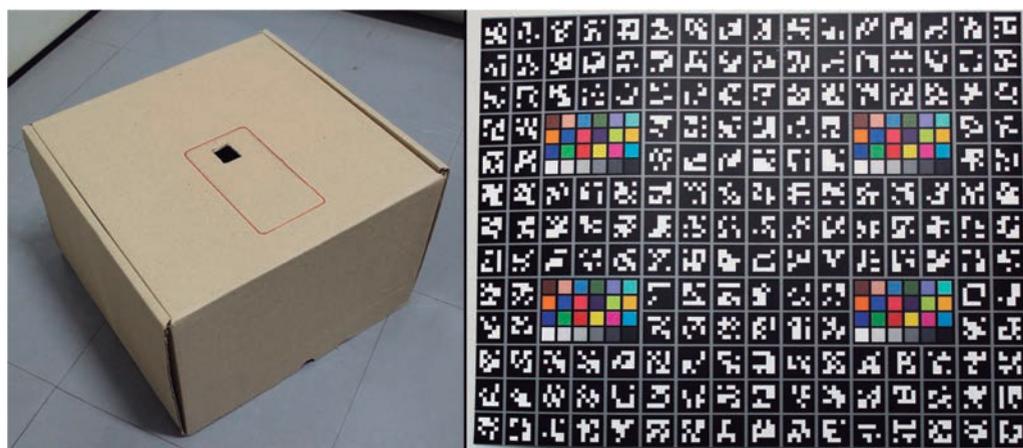


Figura 4. Caixa e padrão usados na captura das imagens de sintomas de doenças, no contexto do projeto Digipathos.

<sup>1</sup> Disponível em: <<https://www.agropediabrasilis.cnpia.embrapa.br/web/digipathos>>.

sões, uma web, a ser usada em computadores com conexão com a internet, e uma móvel, a ser embarcada em aparelhos celulares e tablets, não dependendo da existência de uma rede para funcionar.

## 2.2 Sistemas semi-automáticos

Sistemas semi-automático são aqueles que, além de terem uma parte computacional que realiza certas operações que levem ao diagnóstico, dependem também de uma participação humana para funcionarem corretamente. Pode-se dividir estes sistemas em dois tipos, os quais são apresentados nas subseções a seguir.

### 2.2.1 Correção manual dos resultados

Neste tipo de abordagem, a parte automática do sistema fornece uma resposta, a qual, se apresentar falhas ou erros visíveis, pode ser corrigida pelo usuário. Evidentemente, este tipo de abordagem não serve para a classificação de doenças, uma vez que se o usuário sabe que houve erro; então, ele já saberia a resposta de antemão, não havendo necessidade de usar o programa. Contudo, em análises quantitativas é possível, por exemplo, identificar manualmente objetos que o algoritmo não detectou, ou corrigir as bordas das áreas com sintomas detectadas. A interface do programa criado no contexto do projeto Contag permite ao usuário realizar este tipo de correção, conforme exemplificado na Figura 5.

Não é comum, na literatura, encontrar métodos ou sistemas originais que permitam a correção manual por parte do usuário. Por outro lado, ferramentas genéricas de processamento digital de imagens, tais como o ImageJ<sup>2</sup>, o Assess<sup>3</sup> e o SigmaScan Pro<sup>4</sup>, naturalmente dependem das

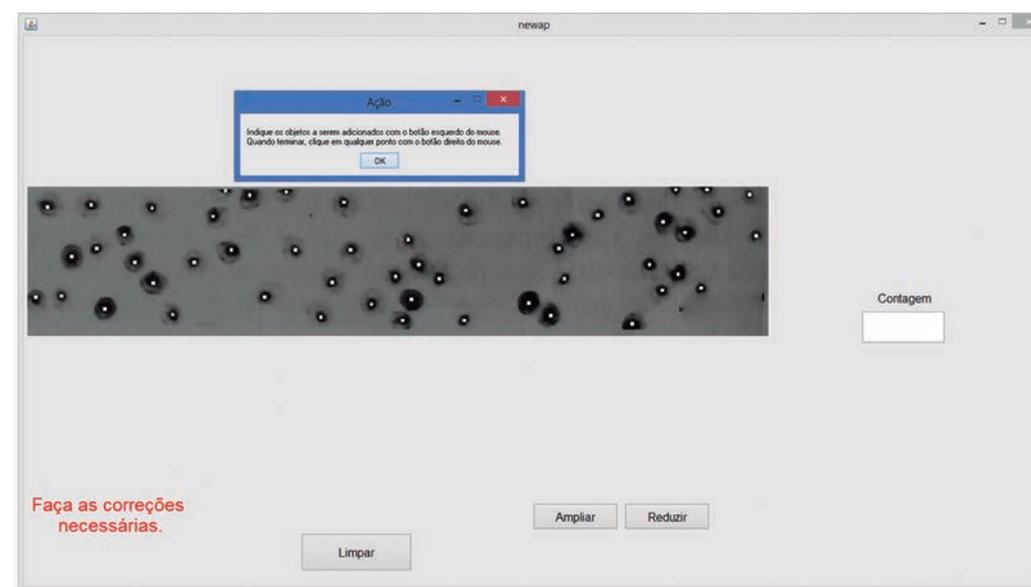


Figura 5. Exemplo de correção manual presente no sistema desenvolvido no contexto do projeto Contag.

<sup>2</sup> Disponível em: <<http://imagej.nih.gov/ij/>>.

<sup>3</sup> Disponível em: <<https://www.apsnet.org/apsstore/shopapspress/Pages/43696m5.aspx>>.

<sup>4</sup> Disponível em: <<http://www.sigmaplot.com/products/sigmaScan/sigmaScan.php>>.

entradas e intervenções feitas pelos usuários. Tais ferramentas normalmente são aplicadas em uma grande variedade de problemas, incluindo aqueles relacionados à agricultura. Um estudo comparando abordagens manuais, semi-automáticas e automáticas foi apresentado por Barbedo (2013e).

### 2.2.2 Sistemas especialistas

Sistemas especialistas são ferramentas baseadas em um conjunto de regras que tentam representar, da maneira mais acurada possível, todo o universo de possibilidades do problema que se deseja resolver. No caso específico de doenças em plantas, inicia-se com todo o universo de culturas e doenças para as quais o sistema foi treinado e, através de perguntas, as possibilidades vão sendo sucessivamente refinadas, até que se chegue à resposta desejada.

Este tipo de sistema é bastante popular em diversas áreas de aplicação. Estes começaram a ser utilizados na agricultura na década de 1980. Desde então, um grande número de sistemas foi proposto, como o Plant/ds, para diagnóstico de doenças da soja (MICHALSKI et al., 1982), o MoreCrop<sup>5</sup> para previsão e manejo da ferrugem, e o TomEx, desenvolvido no Brasil com foco em doenças do tomateiro (POZZA et al., 1997). Na verdade, pode-se encontrar na literatura centenas de sistemas especialistas diferentes, usando as mais diferentes técnicas e abordagens. Normalmente esses sistemas são desenvolvidos tendo uma única cultura particular em mente e, ainda mais importante, o conjunto de regras utilizado normalmente só é válido para as condições climáticas, ambientais e epidemiológicas típicas do país ou região no qual se pretende utilizar o sistema. Como consequência, são raras as comparações entre sistemas especialistas. Esse fato, por sua vez, faz com que não seja possível determinar o estado da arte desta área, já que cada sistema resolve um pequeno subconjunto do problema mais geral do diagnóstico de doenças em plantas.

Na Embrapa, foram desenvolvidos dois sistemas diferentes:

- **Sistema Diagnose Virtual:** desenvolvido sob liderança da Embrapa Informática Agropecuária, este sistema possui uma infraestrutura única na área de sanidade para diagnóstico de doenças de plantas via internet, a fim de subsidiar os agricultores, agrônomos e técnicos agrícolas em suas decisões sobre o manejo de doenças (MASSRUHÁ; LIMA, 2011; MASSRUHÁ et al., 2007a, 2007b, 2007c; SILVA et al., 2011). Visa possibilitar o uso racional de agrotóxicos, o que ajuda a evitar mais danos à saúde e ao meio ambiente, além de reduzir os custos da produção. O sistema possui dois módulos: a) módulo especialista, no qual fitopatologistas têm a responsabilidade de fornecer ao sistema uma base de conhecimento sobre doenças de uma determinada cultura; b) módulo produtor: os produtores rurais fornecem as informações sobre os sintomas encontrados em sua cultura, e obtêm como resultado a indicação das prováveis desordens e as medidas de controle viáveis. Atualmente, o sistema é capaz de fornecer diagnóstico para as culturas do arroz, feijão, milho, soja, tomate e trigo. A Figura 6 mostra a tela inicial do sistema, o qual pode ser acessado pelo endereço<sup>6</sup>
- **Sistema Uzum:** desenvolvido pela Embrapa Uva e Vinho, este sistema visa servir como um guia para auxiliar na identificação de possíveis agentes causadores de sintomas em videiras, fornecendo informação útil para um rápido diagnóstico inicial. Após o usuário fornecer as respostas solicitadas nas telas da interface, o sistema fornece uma lista de possíveis problemas,

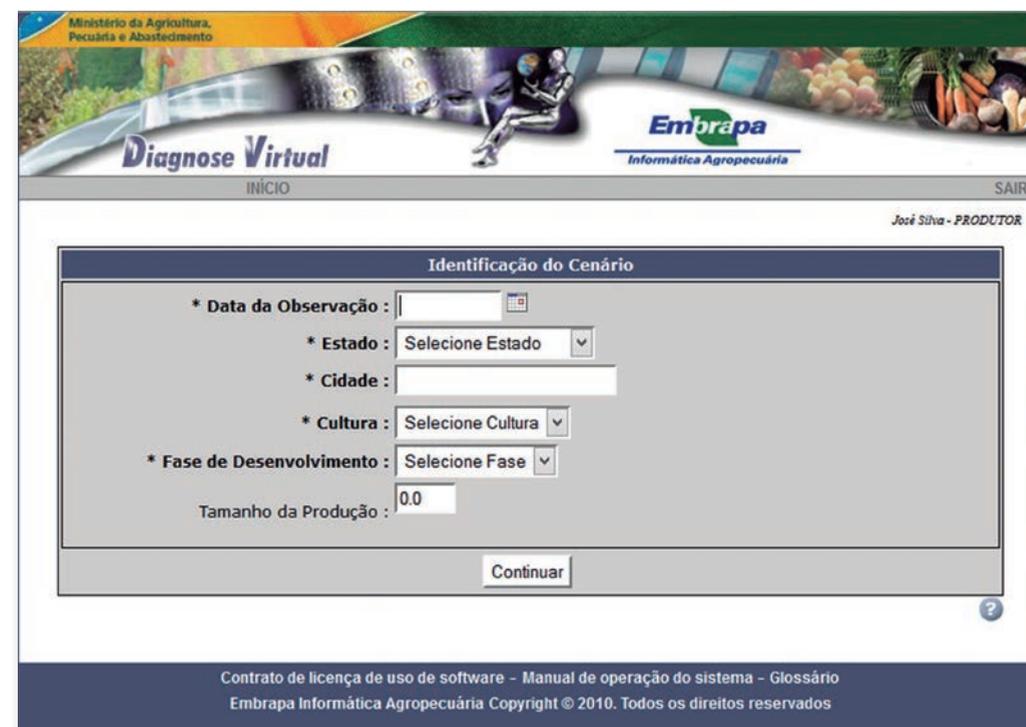


Figura 6. Tela inicial do módulo produtor do sistema Diagnose Virtual.

com as respectivas probabilidades. O sistema tem registrados 37 doenças, pragas e distúrbios fisiológicos, estando disponível na internet<sup>7</sup>. Mais detalhes podem ser encontrados em Fialho et al. (2012), e uma imagem da tela inicial é mostrada na Figura 7.



Figura 7. Tela inicial do sistema especialista para diagnóstico de doenças em videiras Uzum.

Normalmente, sistemas especialistas usados para diagnóstico de doenças em plantas precisam considerar uma ampla gama de problemas, fazendo com que seja necessária a criação de um extenso conjunto de regras, as quais devem se relacionar de maneira coerente a fim de resultar em um bom diagnóstico. Um extenso conjunto de regras significa que o processo para se chegar a um diagnóstico acurado pode demandar que um grande número de perguntas seja respondido

<sup>6</sup> Disponível em: <<http://www.diagnose.cnptia.embrapa.br/diagnose/>>.

<sup>7</sup> Disponível em: <<http://www.cnpuv.embrapa.br/tecnologias/uzum/>>.

pelo usuário. Este fato aumenta consideravelmente a probabilidade de erros, uma vez que uma única pergunta respondida de maneira errônea pode fazer com que o sistema se desvie irremediavelmente da resposta correta.

Uma maneira de se limitar o universo inicial de possibilidades é fornecer ao sistema especialista informações objetivas capazes de eliminar diagnósticos improváveis antes mesmo do usuário começar a responder as perguntas. Tendo esse objetivo em mente, está em andamento um projeto para acoplar o sistema automático baseado em imagens digitais a ser desenvolvido no projeto Digipathos, com um novo sistema especialista, o qual aproveitará muitos dos conceitos usados no sistema Diagnose Virtual para criar uma ferramenta com abrangência ainda maior. Assim, o sistema baseado em imagens será capaz de reduzir o número de respostas possíveis, e o sistema especialista será responsável por dirimir qualquer incerteza advinda da parte baseada em imagens. O novo sistema deverá ser implementado em duas versões, uma web e uma a ser embarcada em dispositivos móveis.

### 2.3 Conclusão

O uso da tecnologia para auxiliar no diagnóstico de doenças na agropecuária tem experimentado significativos avanços nas últimas décadas. Dada a importância deste tema, a Embrapa tem adotado diversas iniciativas para desenvolver tecnologias próprias e contribuir para o avanço do tema, como pode ser comprovado por ferramentas como o Diagnose Virtual e o Uzum, e projetos como o Contag e o Digipathos. Contudo, os desafios a serem vencidos ainda são consideráveis, ensejando que novas iniciativas sejam levadas adiante num futuro próximo.

## 3 Sistemas de alerta de doenças de plantas

Um sistema de previsão de doença de planta é aquele que prevê o aparecimento ou um aumento na intensidade de uma doença baseado em informação sobre o ambiente, a cultura e/ou o patógeno (CAMPBELL; MADDEN, 1990). Esse aparecimento ou aumento futuro da doença é frequentemente baseado na observação de períodos críticos ocorridos, o que acaba causando certa confusão. Com respeito aos sintomas, a previsão é anterior ao fato, mas com respeito à infecção, a previsão é posterior. Zadoks (1984) sugeriu o termo aviso de doença ou alerta de doença, para evitar problemas de terminologia e enfatizar que a mensagem aos produtores é mais importante do que a sua origem técnica.

Prever doenças de plantas é importante por duas razões principais: economia e segurança. A questão econômica é reduzir o custo de produção por meio de aplicações oportunas de medidas de controle, geralmente na forma de fungicidas. Segurança envolve não apenas a cultura, reduzindo efeitos tóxicos sobre as plantas, mas também o ambiente externo, reduzindo a exposição de agrotóxicos a outras espécies de plantas, aos trabalhadores e aos consumidores (HARDWICK, 2006).

Os alertas auxiliam os produtores a determinar a necessidade e o momento de aplicar técnicas de controle de doenças (CAMPBELL; MADDEN, 1990). Um alerta proporciona indicação de quando é provável que a doença vá se tornar crítica e, portanto, ter impacto econômico. Para algumas doenças, é importante ser capaz de prever a primeira ocorrência, enquanto para outras um certo nível de doença pode ser tolerado, particularmente em partes da planta que possuem pouca contribuição para a produção ou a qualidade (HARDWICK, 2006).

Sistemas de previsão de doenças de plantas podem ser classificados de várias maneiras, de acordo com o tipo de informação usada para fazer a previsão ou com a abordagem conceitual para a previsão. Especificamente, previsores podem ser classificados com respeito a (CAMPBELL; MADDEN, 1990): se informação da cultura, da doença, do patógeno ou do ambiente, ou uma combinação dessas, são usadas para fazer as previsões; se as previsões são pré-plantio ou pós-plantio; se informação empírica ou fundamental foi usada no desenvolvimento do sistema; e se características específicas das epidemias, tais como inóculo primário, inóculo secundário ou taxa de aumento da doença, servem de base para as previsões. Uma ampla revisão bibliográfica de sistemas de previsão de doenças de plantas, abrangendo importantes doenças de várias culturas agrícolas, encontra-se em Reis e Bresolin (2004).

Um sistema de alerta, para que tenha sucesso, precisa ser adotado e implementado pelos produtores, devendo haver a percepção de que é possível obter benefícios específicos e tangíveis com o seu uso. Atributos que asseguram o sucesso incluem (CAMPBELL; MADDEN, 1990): confiabilidade, simplicidade para implementar, importância da doença, utilidade do alerta, disponibilidade aos produtores, aplicabilidade a várias doenças/pragas e eficiência de custo.

Segundo Bourke (1970), para desenvolver um sistema de alerta é preciso que a doença satisfaça quatro requisitos: a doença causa perdas economicamente significativas na qualidade ou na quantidade da produção; a doença varia entre cada estação de cultivo; medidas de controle da doença estão disponíveis e são economicamente viáveis; e há informação suficiente a respeito da natureza da dependência da doença em relação às condições meteorológicas.

Muitos estudos a respeito do papel das condições meteorológicas nas doenças de plantas indicam que a doença é mais afetada pelas condições microclimáticas no dossel das plantas do que pelas condições macroclimáticas medidas a uma certa distância da cultura em estações meteorológicas padrão. Entretanto, condições macroclimáticas produzem o microclima - é possível usar regras para determinar relacionamentos entre o macro e o microclima - e existe um limite na extensão com que o microclima pode facilitar o desenvolvimento da doença sob condições macroclimáticas desfavoráveis (COAKLEY, 1988). Ele chegou ainda a sugerir que tentativas de relacionar dados macroclimáticos com doenças podem ter alcançado resultados limitados, em parte, por causa da dificuldade de analisar grandes quantidades de dados sem o auxílio de computador.

Anteriormente, desvantagens de usar dados microclimáticos de dentro dos campos de cultivo no desenvolvimento e/ou implementação de sistemas de alerta incluíam períodos de tempo relativamente curtos e poucos locais para os quais estavam disponíveis os dados, alto custo de coleta desses dados e frequência considerável de perda de dados devido a falhas nos equipamentos (COAKLEY, 1988). Nos últimos anos, os instrumentos manuais deram lugar às estações meteorológicas automáticas, que podem ser instaladas próximas aos campos de cultivo. A cada geração, essas estações tornam-se mais sofisticadas, confiáveis e de menor preço. Mesmo assim, exigir de produtores a instalação e a manutenção dessas estações meteorológicas e o gerenciamento dos dados de cada campo não é prático nem economicamente viável.

Como alternativa a essa questão, esforços têm sido feitos para desenvolver e validar sistemas de alerta que utilizem dados de redes regionais de estações meteorológicas (GENT; SCHWARTZ, 2003) e tecnologias para a obtenção de dados de locais específicos sem sensores in loco (MAGAREY et al., 2001). Outra tendência é utilizar dados meteorológicos estimados a partir de modelos de previsão do tempo para antecipar ainda mais os alertas, permitindo um tempo

maior para a tomada de decisão e para a aplicação de medidas de controle (SHTIENBERG ; ELAD, 1997).

O foco principal desta seção é apresentar modelos de previsão e um sistema de alerta da ferrugem do cafeeiro desenvolvidos pela Embrapa Informática Agropecuária. Outra iniciativa semelhante da Embrapa é o SisAlert, um sistema web de previsão de doenças de plantas baseado em modelos de simulação modulares e genéricos para prever o estabelecimento de doenças a partir de dados meteorológicos obtidos de estações meteorológicas automáticas e de prognósticos de tempo de curto prazo (FERNANDES et. al., 2011). Esse sistema está operacional para duas doenças do trigo e cinco doenças da cultura da maçã, além de ter modelos de outras culturas agrícolas em avaliação e em desenvolvimento.

### 3.1 Modelos de previsão de doenças de plantas

Modelos representam a percepção (ou imaginação) da realidade de forma simbólica e simplificada. Modelagem e simplificação são essenciais ao processo científico. A solução ideal é o modelo contemplar os aspectos essenciais do sistema real pertinentes ao problema em questão. A complexidade ou simplicidade do modelo deve estar em acordo com o seu propósito - enquanto a simplicidade facilita o entendimento do modelo, a complexidade pode permitir maior acurácia na descrição do sistema (CAMPBELL et al., 1988).

Modelos podem ser classificados em dois grupos, dependendo da abordagem de desenvolvimento. O primeiro tipo de modelo é chamado fundamental ou mecanístico. O desenvolvimento desses modelos parte de um conceito, hipótese ou teoria, em vez de um conjunto de dados. Um modelo consistente com o conceito é elaborado e depois são realizados experimentos para testar a sua acurácia (CAMPBELL; MADDEN, 1990).

Esses modelos são derivados de tentativas de compreensão da realidade, sendo que essa compreensão pode ser obtida a partir de experimentos prévios - em laboratório, câmara de ambiente controlado, casa de vegetação ou campo - ou de princípios biológicos (CAMPBELL et al., 1988).

Modelos de previsão fundamentais são geralmente simples e baseados em um ou poucos componentes do ciclo da doença, nos quais a infecção é o componente que prevalece. Contudo, é possível considerar o ciclo completo da doença por meio de uma abordagem de análise sistêmica e modelos de simulação (MADDEN; ELLIS, 1988).

O segundo tipo de modelo é chamado empírico ou correlativo. Os modelos empíricos são desenvolvidos a partir da coleta e análise de dados atuais e históricos sobre níveis da doença e outros fatores bióticos e abióticos (MADDEN; ELLIS, 1988). Eles descrevem um relacionamento observado entre duas ou mais variáveis do conjunto de dados, normalmente derivados a partir do ajuste dos dados a um modelo aceitável (CAMPBELL; MADDEN, 1990). Conhecimento teórico relacionado a mecanismos básicos não é exigido (CAMPBELL et al., 1988).

Modelos empíricos estão relacionados com uma previsão apenas por estação de cultivo ou podem envolver múltiplas previsões. Os primeiros são úteis quando é importante prever o inóculo inicial ou o nível inicial da doença, e o seu desenvolvimento geralmente requer observações de vários anos e/ou locais. Aqueles que envolvem múltiplas previsões são úteis quando a doença pode aumentar rapidamente durante a estação de crescimento, ou quando o valor econômico da cultura justifica várias intervenções de controle. O seu desenvolvimento requer observações feitas sobre a doença, o ambiente e/ou outros fatores por toda a estação de cultivo. Em geral, mais

de dois anos ou locais precisam ser estudados e anos adicionais são necessários para a validação (MADDEN; ELLIS, 1988).

Coakley (1988), baseado em sua experiência, sugeriu um mínimo de oito a doze anos de registro de dados, de campos com fontes naturais de inóculo, para identificar com segurança quais podem ser os fatores climáticos de influência no desenvolvimento de uma doença. Sugeriu também, quando se tem menos de oito anos de registro, que dados de diferentes localidades de uma região geográfica podem ser utilizados.

Madden e Ellis (1988) indicaram duas formas pelas quais modelos de previsão empíricos podem ser desenvolvidos. A primeira, chamada de qualitativa, envolve o desenvolvimento de critérios de previsão sem qualquer análise estatística formal. Alguns exemplos de modelos desenvolvidos dessa forma são: previsão da mancha preta do amendoim (JENSEN; BOYLE, 1966; PARVIN JUNIOR. et al., 1974; PEDRO JÚNIOR et al., 1994), previsão da requeima da batateira (COSTA et al., 2002; WALLIN, 1962), previsão da pinta-preta do tomateiro (MADDEN et al., 1978) e previsão da queima das folhas da cenoura (SOUZA et al., 2002).

A segunda forma de desenvolvimento dos modelos de previsão, chamada de quantitativa, é baseada em análise estatística e modelagem dos dados observados. Os métodos e técnicas utilizados são variados. Em uma retrospectiva dos modelos apresentados na literatura, é possível perceber que o emprego desses métodos e técnicas acompanha a evolução nas disciplinas relacionadas com a análise de dados.

O predomínio é de métodos e técnicas estatísticos, sendo a análise de regressão a mais popular (MADDEN; ELLIS, 1988). Como exemplo, cita-se o uso de regressão linear múltipla no desenvolvimento de modelos para prever a severidade de epidemias da ferrugem asiática da soja (DEL PONTE et al., 2006). Alguns métodos matemáticos e estatísticos para a modelagem de dados epidemiológicos foram revisados por Hau e Kranz (1990).

Mais recentemente, os trabalhos publicados relatam o uso de métodos e técnicas modernos e sofisticados, não necessariamente inovadores, mas que ganharam/recuperaram visibilidade e importância nos últimos anos, como redes neurais (PAUL; MUNKVOLD, 2005), regressão logística (DE WOLF et al., 2003), árvores de decisão (MEIRA et al., 2009) e análise estatística de séries temporais (XU et al., 2000).

### 3.2 Modelagem para previsão e alerta da ferrugem do cafeeiro

A ferrugem, causada pelo fungo *Hemileia vastatrix* Berk. & Br., é a principal doença do cafeeiro em todo o mundo. No Brasil, em regiões onde as condições climáticas são favoráveis à doença, os prejuízos na produção atingem cerca de 35% em média, podendo chegar a mais de 50%. Os principais danos causados pela ferrugem são a queda precoce das folhas e a seca dos ramos, que, em consequência, não produzem no ano seguinte (ZAMBOLIM et al., 2005).

Além da importância econômica, a ferrugem atende outros requisitos, como a variação na sua intensidade entre cada estação de cultivo e a disponibilidade de medidas de controle economicamente viáveis, que justificam o desenvolvimento de modelos de previsão e um sistema de alerta ou aviso.

O cafeeiro alterna anos de alta e baixa produção (carga pendente de frutos), caracterizando um ciclo bienal. A ferrugem ataca com maior intensidade em anos de alta carga. A curva de progres-

so padrão da doença inicia-se em dezembro/janeiro, aumenta em escala logarítmica de março a abril e atinge o pico por volta de junho. A partir daí, decresce devido às baixas temperaturas e à desfolha das plantas ocasionada pela colheita, pela senescência natural e também pela severidade da doença.

A elevação da temperatura e da precipitação pode atrasar o começo da epidemia e, simultaneamente, a ocorrência de chuvas esporádicas e aumento na temperatura média entre abril e julho permitem à doença manter altos índices de incidência até agosto. Nesses casos, torna-se necessário rever e readaptar as medidas de controle (CHALFOUN et al., 2001). O controle da ferrugem envolve principalmente o uso de fungicidas protetores e sistêmicos. O uso de variedades resistentes também é importante no controle da doença (ZAMBOLIM et al., 2005).

A modelagem da ferrugem do cafeeiro para estudos epidemiológicos é encontrada na literatura científica e existem exemplos de modelos de previsão da doença, a maioria deles empíricos. O ajuste dos dados observados a equações de regressão foi a técnica mais comum de modelagem, mas existem exemplos também de modelo fundamental e de modelo empírico com abordagem qualitativa. Trabalhos mais recentes utilizam técnicas de maior visibilidade no momento atual, como as consideradas da área de mineração de dados.

### 3.2.1 Previsão com modelo fundamental

Kushalappa e Eskes (1989) propuseram o desenvolvimento de um modelo compreensivo capaz de explicar o curso de ação biológica do patógeno e integrar os vários fatores que influenciam o sistema, argumentando que esses diversos fatores que afetam o progresso da ferrugem do cafeeiro no campo não poderiam ser identificados por um experimento ou por procedimentos estatísticos clássicos, como a análise de regressão.

O sistema epidêmico da ferrugem do cafeeiro é composto de processos epidemiológicos políciclicos, que consistem de uma série de processos monocíclicos. O princípio de uma epidemia começa com um inóculo inicial e cada ciclo da doença (processo monocíclico) é formado pelos macroprocessos de esporulação, disseminação e infecção. Todos constituem os componentes estruturais do sistema epidêmico (KUSHALAPPA, 1994).

Baseado nesses aspectos, foi desenvolvido um modelo de previsão da taxa de progresso da ferrugem, considerando o inóculo inicial e fatores significativos do ambiente e do hospedeiro que influenciam o processo monocíclico de *H. vastatrix*. Denominado de “razão de sobrevivência líquida para o processo monocíclico” (RSLPM), o modelo foi formado pela integração de modelos fundamentais e empíricos desenvolvidos para cada um dos componentes estruturais da doença (KUSHALAPPA et al., 1983, 1984).

Na obtenção do modelo RSLPM, vários fatores que influenciam o progresso da ferrugem do cafeeiro, relacionados com o hospedeiro, o patógeno e o ambiente – ou, especificamente, os micro e mesoprocessos componentes dos macroprocessos – foram transformados em “equivalentes de processo” para o ambiente e o hospedeiro. Depois, os produtos multiplicativos dos equivalentes de micro e mesoprocessos foram derivados, designados de “equivalentes de processo monocíclico” para o ambiente e para o hospedeiro. Estes e o nível de inóculo foram então transformados em outro parâmetro multiplicativo, a razão de sobrevivência líquida para o processo monocíclico de *H. vastatrix* (KUSHALAPPA, 1989a). A incorporação dos três componentes do triângulo de doenças de plantas no modelo foi baseada na atividade biológica do fungo, o que fez os criadores do modelo o considerarem do tipo fundamental.

A influência do inóculo, denominada de razão de sobrevivência básica (RSB), foi quantificada com base na incidência (proporção de folhas com ferrugem) ou na severidade (proporção de área foliar com ferrugem).

A influência do ambiente, ou equivalente de processo monocíclico para o ambiente (EPMA), foi calculada pela multiplicação de sua influência nos processos de disseminação e de infecção. O equivalente de disseminação foi determinado em função da velocidade diária do vento, da quantidade de chuva diária e da densidade de plantas, enquanto o equivalente de infecção foi determinado em função da duração do molhamento foliar (em horas) e da temperatura durante esse período.

O equivalente de processo monocíclico para o hospedeiro (EPMH) foi determinado pelo equivalente de processo devido à predisposição do hospedeiro ao ataque da ferrugem por causa de alta produção.

Por fim, a razão de sobrevivência líquida para processo monocíclico foi definida como o produto final da multiplicação das influências do inóculo, do ambiente e do hospedeiro, conforme a equação 1 (KUSHALAPPA, 1989a; KUSHALAPPA et al., 1983).

$$RSLPM = RSB \times EPMA \times EPMH \quad (1)$$

Valores de RSLPM a partir de dados observados no campo, para um intervalo de 28 dias antes da data de predição (DP), foram relacionados, por análise de regressão, com taxas de infecção da ferrugem observadas 28 dias após DP, corrigidas para o crescimento do hospedeiro (KUSHALAPPA et al., 1984). Esse intervalo de predição foi escolhido por causa da média observada do período latente do fungo, de outubro a março, que foi de aproximadamente 28 dias.

Várias equações foram desenvolvidas para predizer a taxa de infecção, considerando diferentes parâmetros de área de produção de inóculo no cálculo dos valores de RSLPM (KUSHALAPPA et al., 1984). As equações que obtiveram os melhores coeficientes de determinação ( $R^2$ ) foram:

$$k'' = 0,00044 + 14,766 \times RSLPM - 2511,21 \times RSLPM^2 \quad (2)$$

$$k'' = 0,023 + 14,026 \times RSLPM - 87,382 \times RSLPM^2 \quad (3)$$

onde a equação 2 é para predizer a severidade da doença, considerando a proporção de área foliar com ferrugem como o parâmetro RSB; a equação 3 é para predizer a incidência da doença, considerando a proporção de folhas com ferrugem como o parâmetro RSB;  $k''$  é a taxa de infecção para 28 dias após DP, corrigida para o crescimento do hospedeiro; RSLPM é dada pela média diária da razão de sobrevivência líquida para o processo monocíclico durante 28 dias antes de DP. As equações 2 e 3 explicaram 76% e 64% ( $R^2$  igual a 0,76 e 0,64), respectivamente, da variação em  $k''$ .

Considerando que uma incidência de ferrugem de cerca de 10% justificaria uma aplicação de fungicida, um limite do valor de RSLPM para recomendar aplicações de fungicida foi derivado

pela substituição de  $k'' = 0,1$  na equação 3 (KUSHALAPPA, 1989b; KUSHALAPPA et al., 1984). Esse limite foi  $RSLPM = 0,0057$ , considerando a proporção de folhas com ferrugem como inóculo ( $RSB$ ). Substituições semelhantes em outras equações permitiram se chegar no limite  $RSLPM = 0,00015$ , considerando a proporção de área foliar com ferrugem como inóculo.

A partir desses limites, um sistema de alerta simples e outro mais complexo foram desenvolvidos para recomendar aplicações de fungicida para o controle da ferrugem do cafeeiro. O sistema simples se resumiu em uma tabela, que foi formada pelo agrupamento dos valores de inóculo ( $RSB$ ), de produção ( $EPMH$ ) e de condições do ambiente ( $EPMA$ ), observados no Estado de Minas Gerais, em certos intervalos convenientes.

No caso do sistema simples, em intervalos quinzenais, deve-se quantificar a incidência (percentual de folhas atacadas) ou a severidade (área foliar atacada) da ferrugem e indicar a produção como alta ou baixa; depois, basta consultar a tabela sobre a recomendação ou não de aplicação de fungicida.

No caso do sistema complexo, deve-se quantificar  $RSLPM$ , em intervalos de 14 dias, e se recomenda a aplicação de fungicida quando o limite preestabelecido é igualado ou superado –  $RSLPM \geq 0,0057$  ou  $RSLPM \geq 0,00015$ , caso se tenha quantificado a incidência ou a severidade da doença, respectivamente. Ambos os sistemas foram validados em condições de campo e foram considerados eficientes na determinação das épocas oportunas de aplicação de fungicidas (KUSHALAPPA et al., 1986).

### 3.2.2 Previsão com modelo empírico pela abordagem qualitativa

Modificações no clima, nos últimos anos, têm ocasionado alterações na severidade da ferrugem, bem como no início e no pico da doença em algumas regiões do Brasil (ZAMBOLIM et al., 2002). Diante dessas alterações, surgiu a dúvida sobre a aplicação de fungicidas sistêmicos por meio de duas pulverizações foliares ou aplicações via solo, para que se obtivesse controle racional e econômico da ferrugem.

Para se identificar, então, os períodos favoráveis à ferrugem, nos quais as plantas deveriam ser atomizadas, foi desenvolvido um sistema de previsão ou de aviso (GARÇON et al., 2004). O objetivo foi criar um sistema simples - simplicidade é um atributo importante para a aceitação do sistema, pois maiores são as chances de adoção pelos agricultores - e confiável de prever o desenvolvimento da doença no campo, determinando o momento propício para iniciar o controle químico por meio de pulverizações com fungicida sistêmico, bem como o intervalo entre as aplicações.

As variáveis meteorológicas empregadas no sistema de previsão foram o molhamento foliar diário e a temperatura média durante esse período de molhamento, obtidas em estação meteorológica colocada no meio da área experimental. Com os dados diários dessas variáveis meteorológicas calculou-se o valor de severidade da ferrugem (VSF), a partir de uma matriz de valores de severidade semelhante à idealizada por Wallin (1962) para a requeima da batateira, modificada para a ferrugem do cafeeiro (Tabela 1).

O limiar de ação para indicação do momento da pulverização foi baseado no acúmulo dos valores diários de VSF. Os limites de VSF estipulados para o teste e a validação do sistema foram 29, 34, 39 e 44, para anos de alta carga pendente de frutos (alta intensidade da ferrugem), e 49, 59, 69 e 79, para anos de baixa ou média carga pendente (baixa intensidade da ferrugem).

**Tabela 1.** Matriz para cálculo dos valores de severidade da ferrugem (VSF) do cafeeiro, com base no período de molhamento foliar e na temperatura média do período.

Molhamento foliar (h/diárias)	Temperatura (°C)						
	< 16	16-18	19-20	21-24	25-26	27-29	30
0	0*	0	0	0	0	0	0
0 < h ≤ 8	0	0	1	2	1	0	0
8 < h ≤ 17	0	1	2	3	2	1	0
17 < h ≤ 24	0	2	3	4	3	2	0
h = 24**	0	0	1	2	1	0	0

\* Valor de severidade da ferrugem (VSF) diário.

\*\* Molhamento foliar diário de 24 h, porém este dentro de um período de molhamento de mais de 48 h sem interrupção.

Obs.: se dentro de 30 dias não houver acumulado mais de 5 VSF, desconsideram os VSF acumulados até o momento.

Fonte: Garçon et al. (2004).

Em uma lavoura de alta carga pendente (101,5 sacas beneficiadas/hectare), foram recomendadas duas pulverizações com fungicida sistêmico quando o valor acumulado de VSF atingiu 29-31, igualando-se as duas aplicações do tratamento com calendário fixo. Em uma lavoura de média carga pendente (22,4 sacas beneficiadas/hectare), recomendou-se uma única pulverização quando o valor acumulado de VSF atingiu 49-51, enquanto todos os outros tratamentos demandaram duas aplicações. Portanto, o sistema baseado no VSF foi tão eficiente quanto o calendário no controle da ferrugem do cafeeiro, porém com economia de uma pulverização na lavoura com carga média de frutos.

Os resultados mostraram que, normalmente, em lavouras com baixa a média carga pendente de frutos, uma única aplicação de fungicida sistêmico, no momento oportuno, indicado com base no número de horas de molhamento foliar e na temperatura média durante o período de molhamento, como sugerido pelo modelo de aviso proposto, foi suficiente para se alcançar um controle eficiente e racional da ferrugem do cafeeiro.

### 3.2.3 Previsão com modelos empíricos pela abordagem quantitativa

A modelagem da ferrugem do cafeeiro para estudos epidemiológicos encontrada na literatura científica revela diferentes exemplos de modelos de previsão da doença, a maioria deles empíricos desenvolvidos segundo a abordagem quantitativa. O ajuste dos dados observados a equações de regressão é a técnica mais comum de modelagem utilizada (KUSHALAPPA; ESKEES, 1989; MORAES et al., 1976; ZAMBOLIM et al., 2002).

Moraes et al. (1976) observaram que o período de incubação (PI) - período decorrido da inoculação até a formação de 50% de pústulas - tendeu a encurtar nos meses mais quentes (28 dias) e tornar-se mais longo nos meses mais frios (65 dias). Os autores sugeriram a utilização da seguinte equação para a estimativa do PI:

$$y = 103,01 - 0,98 \times x_1 - 2,1 \times x_2 \quad (4)$$

onde  $y$  é a estimativa do período de incubação em dias,  $x_1$  é a temperatura média máxima e  $x_2$  a temperatura média mínima durante o período.

Essa equação foi usada para dar uma ideia de quão severa a ferrugem poderia ser durante certas estações ou meses do ano. Moraes (1983) sugeriu, com base em estimativas obtidas pela equação 4, para o período de outubro a março, os seguintes níveis de severidade de ataque da ferrugem do cafeeiro: risco alto de ataque severo, quando o PI estimado for inferior a 35 dias; risco médio de ataque severo quando o PI for estimado entre 35 e 45 dias; e pequena probabilidade de risco de ataque severo, quando o PI estimado for superior a 45 dias.

Chaves et al. (1970) citado por Kushalappa (1989b), logo que a ferrugem do cafeeiro surgiu no Brasil, consideraram desnecessárias aplicações de fungicida, de maio a agosto, devido aos períodos latentes mais longos. Kushalappa (1989b) argumentou que a razão de não haver necessidade de aplicações de fungicida nos meses mais frios seria diferente: a temperatura durante o período de molhamento foliar geralmente fica abaixo de 15°C, o que é limitante para a infecção.

Alfonsi et al. (1974) estudaram a associação entre níveis de infecção (média de pústulas por folha), variáveis climáticas e área foliar das plantas. A média das temperaturas máximas, a média das temperaturas mínimas e o total de chuvas, registrados em períodos de 15, 30 e 45 dias, foram correlacionados com os níveis de infecção observados ao final dos respectivos períodos. Os coeficientes de determinação ( $R^2$ ) obtidos entre o nível de infecção e as três variáveis climáticas, independente da área foliar, mostraram que a associação com o período de 45 dias expressou melhor a proporção de acréscimos de pústulas (cerca de 95% de explicação da variação na severidade da doença). Na literatura consultada, não há registro de sistema de alerta da ferrugem do cafeeiro baseado nas equações de regressão obtidas.

Também por análise de regressão, vários fatores meteorológicos e biológicos foram considerados para explicar a taxa de progresso da ferrugem (KUSHALAPPA; ESKES, 1989). Os fatores mais significativos foram identificados pelo critério de seleção stepwise. Como variáveis dependentes, foram consideradas a severidade da doença na data de previsão (DP) e a taxa de infecção da ferrugem para os intervalos de um a dois períodos latentes (28 dias) após DP. A equação que explicou a máxima variação (94%) na taxa de infecção foi:

$$k'' = 0,031 + 4,881 \times PAFE + 0,022 \times PNF - 0,001 \times MIN - 0,001 \times MAX - 0,001 \times CHUVA \quad (5)$$

onde  $k''$  é a taxa de infecção, corrigida para o crescimento do hospedeiro, para 56 dias depois de DP; *PAFE* é a proporção de área foliar com esporos na DP; *PNF* é a proporção de novas folhas formadas durante 14 dias antes de DP; *MIN* é a média das mínimas e *MAX* é a média das máximas temperaturas (em °C) para 14 dias antes de DP; *CHUVA* é o total de chuvas (em mm) entre 14 a 28 dias antes de DP. Não foi encontrado, na revisão bibliográfica, trabalho a respeito do uso das equações desenvolvidas em sistema de alerta da ferrugem do cafeeiro.

Correlações significativas foram observadas entre variáveis independentes usadas na formulação de equações de regressão para prever a taxa de progresso da ferrugem do cafeeiro (KUSHALAPPA et al., 1983). Como resultado, alguns parâmetros, que independentemente explicaram variação significativa na doença, foram eliminados devido à multicolinearidade. O sucesso preditivo de tais modelos depende da ocorrência futura dos diferentes parâmetros, incluindo aqueles não usados no modelo, em combinações semelhantes àquelas observadas. Isso

torna esse tipo de modelo menos estável sob condições de campo, a menos que esteja baseado em vários anos de registro de dados (KUSHALAPPA; ESKES, 1989).

Mais recente, Pinto et al. (2002) avaliaram o potencial de redes neurais para descrever a epidemia da ferrugem do cafeeiro. Elas foram empregadas para estabelecer relações entre variáveis climáticas e produção e a incidência da ferrugem. As seguintes variáveis foram utilizadas para construir as redes: precipitação pluvial; número de dias com e sem precipitação; umidade relativa do ar; horas de insolação; temperaturas média, máxima e mínima, calculadas como médias ou somatórios para os 15, 30, 45 e 60 dias anteriores às avaliações da incidência da ferrugem; e a variável produção, a qual assumiu valor '0' para as plantas antes do início da produção e '1' para as plantas em fase de produção. Séries temporais da incidência da doença, isoladamente, também foram utilizadas na elaboração de redes neurais.

A camada de entrada para as redes foi formada pelas variáveis climáticas mais a variável de produção (variáveis independentes) e a variável de saída foi a incidência da ferrugem (variável dependente). Os menores valores do erro médio de previsão (EMP = 1,17%) e do quadrado médio do desvio (QMD = 3,43) foram obtidos para a rede neural elaborada com as variáveis produção, umidade relativa, horas de insolação e temperatura mínima, relativas ao período de 30 dias anteriores à avaliação da incidência da doença. A melhor rede neural (EMP = 4,72% e QMD = 3,95) elaborada a partir das séries temporais teve como variáveis de entrada as observações da incidência da doença de quatro quinzenas anteriores à data de avaliação (PINTO et al., 2002).

As redes neurais elaboradas poderiam ser utilizadas como modelos de previsão da ferrugem do cafeeiro. Apesar do melhor desempenho do modelo que incluiu as variáveis climáticas, os autores sugeriram que o uso de séries temporais, apenas, poderia facilitar a previsão de epidemias da doença. Eles consideraram que avaliar a intensidade da doença é mais fácil para o produtor ou o agente de extensão do que coletar variáveis climáticas.

A indução de árvores de decisão é uma técnica de modelagem alternativa. O seu propósito é descobrir a estrutura preditiva do problema e/ou produzir modelos de predição precisos (BREIMAN et al., 1984). A indução não é prejudicada por multicolinearidade entre as variáveis independentes como na regressão (BUTT; ROYLE, 1990). Além disso, diversas variáveis, numéricas ou categóricas, podem ser analisadas ao mesmo tempo, sendo que o próprio algoritmo de indução se encarrega de selecionar as de maior importância.

A árvore de decisão é um modelo representado graficamente por nós e ramos, parecido com uma árvore, mas no sentido invertido (WITTEN et al., 2011). O nó raiz, no topo da estrutura, e os nós internos são nós de decisão. Cada um contém um teste sobre uma variável independente e os resultados desse teste formam os ramos da árvore. Os nós folhas, nas extremidades, representam valores de predição da variável dependente ou distribuições de probabilidade desses valores.

Paul e Munkvold (2004) usaram este tipo de modelagem para prever categorias de severidade da cercosporiose do milho em estágio avançado do cultivo. Árvores de decisão também modelaram epidemias de giberela do trigo, procurando prever se a severidade da doença seria maior ou igual a 10% (MOLINEROS et al., 2005). Baker et al. (1993) desenvolveram uma árvore de decisão para prever o risco (alto ou baixo) de mortalidade de pinus em decorrência de podridão das raízes causada por *Heterobasidion annosum*.

Com relação à ferrugem do cafeeiro, Meira et al. (2008) analisaram manifestações epidêmicas da doença com o auxílio de uma árvore de decisão. As variáveis explicativas mais importantes foram a temperatura média nos períodos de molhamento foliar, a carga pendente de frutos, a média das temperaturas máximas diárias no período de incubação e a umidade relativa do ar. A árvore de decisão demonstrou seu potencial simbólico e interpretável, auxiliando na compreensão de quais variáveis e como as interações dessas variáveis conduziram as epidemias da ferrugem no campo.

Modelos em árvore de decisão foram desenvolvidos para alerta da ferrugem do cafeeiro em lavouras com alta carga pendente de frutos e em lavouras com baixa carga (MEIRA, 2008; MEIRA; RODRIGUES, 2009; MEIRA et al., 2009). Dados de incidência mensal da doença no campo coletados durante oito anos foram transformados em valores binários considerando limites de 5 e 10 pontos percentuais (p.p.) na taxa de infecção. O alerta é indicado quando a taxa de infecção, prevista para o prazo de um mês, atingir ou ultrapassar o respectivo limite.

A acurácia do modelo para lavouras com alta carga pendente e limite de 5 p.p. foi de 81%, calculada por validação cruzada, também com valores bons e equilibrados para outras medidas de avaliação; o modelo para o limite de 10 p.p. teve acurácia de 79%, sem apresentar o mesmo equilíbrio entre as demais medidas. Os modelos para lavouras com baixa carga pendente não apresentaram o mesmo bom desempenho.

A partir do mesmo conjunto de dados preparados por Meira (2008), Cintra et al. (2011) desenvolveram modelos com um método de indução de árvores de decisão fuzzy. O método, denominado Fuzzy DT, é baseado no algoritmo clássico C 4.5 (QUINLAN, 1993), acrescentado de características interessantes da lógica fuzzy relacionadas à interpretabilidade dos modelos e à manipulação de atributos contínuos.

Valores numéricos dos atributos preditivos foram transformados em três valores linguísticos fuzzy: baixo, médio e alto. Foram utilizados conjuntos fuzzy triangulares uniformemente distribuídos nas três partições. O número de conjuntos fuzzy foi definido empiricamente. Os experimentos utilizaram o método de raciocínio fuzzy tradicional, que classifica um novo exemplo usando a classe da regra com a maior compatibilidade com ele (CINTRA et al., 2011).

Comparados com modelos correspondentes obtidos pelo método clássico, os modelos fuzzy apresentaram taxas de erro competitivas (CINTRA et al., 2011). Além disso, o uso dos termos linguísticos evitaram alterações bruscas de decisão com base em limites numéricos dos atributos preditivos, uma desvantagem do método tradicional, e permitiram uma interpretação mais natural das regras dos modelos.

Foram ainda desenvolvidos, em colaboração com grupos de pesquisa internacionais, modelos para a ferrugem do cafeeiro com outras técnicas de modelagem: equações de regressão capazes de prever intervalos em vez de pontos, usando support vector machines (SVM) (LUACES et al., 2010); modelos de predição obtidos por aprendizado não determinístico (LUACES et al., 2011); e predição usando redes bayesianas (PÉREZ-ARIZA et al., 2012).

Modelos selecionados foram validados com dados diferentes dos usados no treinamento desses modelos (DI GIROLAMO NETO et al., 2012). Uma vez que os modelos não mantiveram o bom desempenho nessa validação, iniciou-se um novo ciclo do processo de mineração de dados. O objetivo foi a evolução de modelos em árvore de decisão e a geração de novos modelos com

as técnicas máquinas de vetores suporte, redes neurais e florestas aleatórias (DI GIROLAMO NETO et al., 2014).

No novo ciclo do processo, além do acréscimo das técnicas de mineração de dados, houve outras mudanças em termos da metodologia (DI GIROLAMO NETO et al., 2014): inclusão de dados de dois outros municípios da região cafeeira Sul de Minas, com características climáticas distintas; uso de método de balanceamento de classes nos dados oriundos das lavouras de baixa carga pendente; seleção de atributos por meio de métodos objetivos e automáticos; e utilização do espaço Receiver Operating Characteristic (ROC) para avaliação, comparação e seleção dos melhores modelos.

As técnicas de máquinas de vetores suporte e de florestas aleatórias geraram os modelos com os melhores desempenhos. O procedimento de balanceamento de classes melhorou a taxa de acerto dos modelos para lavouras com baixa carga pendente. Para alta e baixa cargas pendentes, as melhores taxas de acerto foram 85,3% e 88,9%, respectivamente. Outras medidas de desempenho como sensibilidade e especificidade também apresentaram valores altos e equilibrados (DI GIROLAMO NETO et al., 2014).

### 3.3 Sistema de alerta da ferrugem do cafeeiro

Sistemas de alerta que utilizam modelos de previsão ajudam a promover o uso racional de agroquímicos, ao indicar as condições que favorecem ou deixam de favorecer uma doença ou praga, permitindo agir somente quando necessário. A dificuldade de adoção desses sistemas é um problema comum, que inclui certas restrições (CAMPBELL; MADDEN, 1990): complexidade do modelo; dificuldade de obtenção dos dados necessários ao uso do modelo; e custos de implantação e manutenção do sistema para os produtores.

A abordagem normalmente utilizada no desenvolvimento de modelos empíricos consiste em realizar experimentos de curta duração (dois a quatro anos), em que equipamentos com sensores para a obtenção de dados meteorológicos são instalados no interior da cultura. Dessa forma, os modelos obtidos carecem de representatividade quanto às condições meteorológicas, pelo pouco tempo de observação dos dados, e o uso deles provavelmente vai requerer dos produtores a instalação e manutenção de equipamentos em suas lavouras.

Outro fator que contribui para a não adoção dos modelos é a falta de planejamento das atividades seguintes à modelagem. Muitas vezes, o próprio desenvolvimento do modelo é o principal objetivo, como forma de se compreender a epidemia da doença, terminando com uma avaliação estatística do seu desempenho. Não é planejada a sua validação, nem como colocá-lo em situação de uso real.

A Embrapa Informática Agropecuária, buscando contribuir para reduzir essas dificuldades, coordenou um projeto em que modelos de predição da ferrugem foram desenvolvidos com diferentes técnicas de mineração de dados (DI GIROLAMO NETO et al., 2014), a partir de uma infraestrutura estabelecida há vários anos de monitoramento de doenças e pragas do cafeeiro. Alguns dos modelos de melhor desempenho foram selecionados e incorporados em um sistema de alerta na Web (THAMADA et al., 2013). O propósito do sistema é servir como instrumento adicional de apoio aos técnicos da Fundação Procafé na elaboração dos comentários e das recomendações fitossanitárias para o público-alvo dos seus boletins de avisos mensais.

A Fundação Procafé, sediada em Varginha-MG, inaugurou em 1998 uma estação de avisos fitossanitários para a cultura do café. Com o passar dos anos, junto com o Ministério da Agricultura, Pecuária e Abastecimento (Mapa) e outras instituições parceiras, expandiram as estações de avisos para outros municípios da região Sul de Minas e para outras regiões produtoras do estado, como o Triângulo Mineiro e o Alto Paranaíba. Mais recentemente, iniciou também atuação na região Mogiana do Estado de São Paulo.

Em cada estação de avisos, são feitas avaliações mensais da incidência de doenças e pragas do cafeeiro e são registrados dados meteorológicos por meio de estação meteorológica automática. A partir desse monitoramento, são elaborados e divulgados avisos fitossanitários em boletins mensais destinados a técnicos de assistência ao produtor, imprensa, cafeicultores, estudantes e pesquisadores.

O Sistema de Alerta da Ferrugem do Cafeeiro - SafCafe (THAMADA et al., 2013), desenvolvido pela Embrapa com o apoio do Consórcio Pesquisa Café e em parceria com a Fundação Procafé, analisa e informa o risco de epidemias da doença por meio de modelos de alerta desenvolvidos com técnicas de mineração de dados, no âmbito das estações de avisos fitossanitários da Fundação Procafé na região Sul de Minas Gerais. Os modelos incorporados no sistema obtiveram desempenho satisfatório durante a modelagem, com acurácia em torno de 83% obtida por validação cruzada.

O sistema web SafCafe é carregado com arquivos de registros horários das estações meteorológicas instaladas nas cidades de Varginha, Carmo de Minas e Boa Esperança e prepara os dados de entrada para os modelos preditivos. Esses modelos avaliam as condições propícias ou desfavoráveis ao desenvolvimento da ferrugem, a partir dos dados preparados e de acordo com certas características das lavouras, como a carga pendente de frutos (alta ou baixa) e o tipo de plantio (largo ou adensado).

Como resultado, o SafCafe avisa para o mês seguinte sobre a possibilidade ou não de aumentos além de 5 ou 10 p.p. (pontos porcentuais) na taxa de progresso da incidência da doença. A saída dos modelos é binária, emitindo alerta (1) quando é provável que, no próximo mês, a taxa de progresso da ferrugem seja maior ou igual ao limiar considerado. O não alerta (0) indica probabilidade alta de que a taxa não ultrapasse esse limiar.

No fluxo normal de uso do sistema de alerta, o técnico carrega, no início de cada mês, um arquivo contendo dados meteorológicos registrados ao longo do mês anterior e os valores do nível de incidência da ferrugem nas lavouras avaliados no final daquele mês. Os dados de incidência da doença não são usados na predição, mas são importantes para o acompanhamento da evolução da ferrugem no campo e do desempenho do sistema. Automaticamente após o carregamento, o SafCafe processa os registros para que possam ser utilizados pelos modelos e o sistema fica apto a realizar as predições.

Para realizar a predição, é necessário escolher como opções: a estação (Varginha, Carmo de Minas ou Boa Esperança); a carga pendente de frutos do cafeeiro (alta, baixa ou todas), o limiar de pontos porcentuais da taxa de progresso da ferrugem a ser considerado (5 p.p., 10 p.p. ou todas); e o mês para qual é desejado realizar a predição. Escolhidas essas opções, o processo preditivo pode ser acionado.

A Figura 8 mostra uma predição realizada para Varginha, considerando todas as cargas e limiares da taxa de progresso, para agosto de 2013. As abas situadas logo abaixo da seleção das opções de



**Figura 8.** Tela de predição do SafCafe para agosto de 2013 em Varginha; destaque da aba de carga alta, lavoura adensada e limiar de 5p.p. da taxa de progresso da ferrugem do cafeeiro.

predição representam as combinações dessas opções. A aba selecionada (Alta - 5 p.p. - Adensada) mostra que o sistema indicou que não haveria aumento acima de 5 p.p. na taxa de progresso da ferrugem para aquele mês, considerando lavoura de carga alta e plantio adensado. O sistema exhibe esse resultado textualmente e destaca-o por meio de um sinal verde (ícone maior verde à esquerda na Figura 8).

O resultado da predição do SafCafe considera um sistema de votação simples entre três modelos escolhidos e incorporados ao sistema. O alerta é positivo (sinalizado em vermelho) caso dois ou mais modelos indiquem esse resultado; o mesmo é feito para a emissão do não alerta (sinalizado em verde). No caso da Figura 8, os três modelos indicaram um não alerta para agosto de 2013 (Alta - 5 p.p. - Adensada).

Abaixo do resultado da predição para o mês escolhido, o sistema apresenta o histórico das predições nos meses anteriores, desde o início do ano agrícola (quadro inferior da tela do SafCafe na Figura 8). Além das predições anteriores, o sistema apresenta os valores de incidência e da taxa de progresso da doença que foram observados no campo, na realidade, permitindo, assim, acompanhar o desempenho do sistema com relação à acurácia (taxa de acerto) de seus alertas positivos e negativos.

Com a finalidade de adquirir a confiança necessária para que o SafCafe fosse utilizado no procedimento de elaboração dos boletins, foi realizada uma etapa de validação e confirmação do bom desempenho dos modelos incorporados ao sistema. O teste do sistema e dos seus modelos foi realizado com dados não utilizados no treinamento desses modelos durante a fase de modelagem. A avaliação foi realizada no período de três anos agrícolas (2011/2012, 2012/2013 e 2013/2014) com dados obtidos na Fazenda Experimental de Varginha.

A Tabela 2 apresenta os valores de acurácia do SafCafe nos anos agrícolas de validação, para as combinações de produtividade (carga alta e carga baixa) e tipo de plantio (adensado e

**Tabela 2.** Desempenho do SafCafe - Sistema de Alerta da Ferrugem do Cafeeiro na validação.

		Acurácia do sistema SafCafe (%)		
		Lavoura carga alta		Lavoura carga baixa
		Limiar 5 p.p.*	Limiar 10 p.p.*	Limiar 5 p.p.*
Plantio adensado	2011/2012	81,8	90,9	72,7
	2012/2013	90,9	90,9	81,8
	2013/2014	72,7	63,6	63,6
Plantio largo	2011/2012	72,7	90,9	63,6
	2012/2013	100	81,8	81,8
	2013/2014	81,8	63,6	63,6
Média	2011/2012	77,3	90,9	68,2
	2012/2013	95,5	86,4	81,8
	2013/2014	77,3	63,6	63,6
	2011 a 2014	83,3	80,3	71,2

\* Alerta emitido quando a taxa de progresso da ferrugem do cafeeiro está prevista para aumentar além do limiar definido em pontos percentuais (5 p.p. ou 10 p.p.).

largo), como também as médias em cada ano e a média geral obtida no período todo de 2011 a 2014.

A acurácia do sistema no ano 2011/2012 ficou bem próxima da obtida no treinamento dos modelos, considerando carga alta e alerta para o limiar de 5 p.p. no plantio adensado (81,8%). No caso do plantio largo, houve um decréscimo da taxa de acerto (72,7%) devido a dois falsos positivos - previsão de aumento maior ou igual a 5 p.p. que não se confirmou no campo - emitidos pelo sistema. Também para carga alta, o desempenho em 2011/2012 dos alertas para o limiar de 10 p.p. foi ótimo (90,9%), nos dois tipos de plantio.

O SafCafe obteve os melhores resultados no ano agrícola 2012/2013, com desempenho excelente para carga alta e alerta para o limiar de 5 p.p. (90,9% no plantio adensado e 100% no plantio largo) e resultados muito bons do alerta para o limiar de 10 p.p. (90,9% e 81,8%, respectivamente). Em outras palavras, o sistema previu com exatidão a evolução da taxa de progresso da ferrugem, considerando o limiar de 5 p.p., durante todo o ano agrícola de carga alta na lavoura de plantio largo.

No ano agrícola 2013/2014, o desempenho do sistema não foi tão bom para as lavouras de carga alta, principalmente com relação ao alerta para o limiar de 10 p.p. (63,6%). O sistema acompanhou razoavelmente a evolução da ferrugem com relação aos alertas para o limiar de 5 p.p. (72,7% no plantio adensado e 81,8% no plantio largo), incluindo o período em que a doença esteve pressionada pela falta de chuva (janeiro a março), mas não identificou a evolução tardia da doença a partir de abril. Nos três meses em que a taxa de progresso da ferrugem no campo ultrapassou 10 p.p. (abril, junho e julho), não foram emitidos alertas considerando esse limiar.

A acurácia do sistema para lavouras de carga baixa foi boa somente no ano 2012/2013 (81,8% nos dois tipos de plantio). Nos outros anos avaliados, o desempenho foi fraco. Uma possível explicação é que o limiar do alerta em 5 p.p. pode não estar adequado, uma vez que a evolução da ferrugem é menos acelerada em anos de carga baixa. Sugere-se, então, a hipótese de que um

limiar menor do que 5 p.p. pode permitir um melhor desempenho dos modelos de previsão e, conseqüentemente, do sistema de alerta em anos de baixa carga pendente.

Um aspecto bastante importante a ressaltar é que o sistema de alerta foi exato na identificação do início da epidemia da ferrugem nos três anos agrícolas avaliados. No ano 2011/2012, o sistema acertou o início da epidemia em janeiro de 2012, indicando acertadamente que a evolução mais acentuada da doença não começaria em dezembro de 2011. Nos anos seguintes, 2012/2013 e 2013/2014, o sistema indicou corretamente o início das epidemias em dezembro. A confiabilidade desse tipo de informação é importante para ajudar o produtor a decidir pelo melhor momento de iniciar o controle da ferrugem do cafeeiro nas lavouras.

Conclui-se que o sistema de alerta da ferrugem do cafeeiro apresentou desempenho bastante satisfatório nos anos agrícolas avaliados, especialmente para lavouras com carga pendente alta e considerando os alertas para o limiar de 5 p.p. Dessa forma, acredita-se que o sistema SafCafe vai ser útil aos responsáveis pela elaboração dos comentários e das recomendações divulgados nos boletins de avisos fitossanitários da Fundação Procafé.

## 4 Considerações finais

Este capítulo mostrou como a aplicação de uma série de tecnologias pode contribuir significativamente para a redução de perdas e riscos associados a problemas sanitários que afetam virtualmente todas as áreas da agropecuária. Apesar dos avanços significativos já alcançados, é importante observar que a história da aplicação das TIC a problemas zoofitossanitários é muito recente, havendo ainda muitas oportunidades de pesquisa a serem exploradas no futuro. Com a evolução dos recursos computacionais disponíveis, tecnologias com custo computacional e financeiro proibitivos passam a ser opções viáveis, ampliando ainda mais as possibilidades de avanços.

Esforços de pesquisa futuros deverão se concentrar na criação e aplicação de novas tecnologias aos (muitos) problemas para os quais ainda não há uma solução satisfatória e estender tecnologias bem-sucedidas já desenvolvidas a problemas similares que ocorram em outras lavouras, criadouros e rebanhos. É importante enfatizar que muitas dessas tecnologias estão sendo planejadas, desenvolvidas e adaptadas de modo a explorar a crescente popularização e o aumento do poder computacional de dispositivos móveis, expandindo significativamente a aplicabilidade e o alcance das novas ferramentas desenvolvidas.

## 5 Referências

ALFONSI, R. R.; ORTOLANI, A. A.; PINTO, H. S.; PEDRO JUNIOR, M. J.; BRUNINI, O. Associação entre nível de infecção da ferrugem do cafeeiro, variáveis climáticas e área foliar, observadas em *Coffea arabica* L. In: CONGRESSO BRASILEIRO DE PESQUISAS CAFEIEIRAS, 2., 1974, Poços de Caldas. **Resumos...** Rio de Janeiro: IBC-GERCA, 1974. p. 80-83.

BAKER, F. A.; VERBYLA, D. L.; HODGES, C. S.; ROSS, E. W. Classification and regression tree analysis for assessing hazard of pine mortality caused by *Heterobasidion annosum*. **Plant Disease**, St Paul, v. 77, n. 2, p. 136-139, 1993.

- BARBEDO, J. G. A. Automatically measuring early and late leaf spot lesions in peanut plants using digital image processing. In: CONGRESSO BRASILEIRO DE AGROINFORMÁTICA, 9., 2013, Cuiabá. **Agroinformática: inovação para a sustentabilidade do agronegócio brasileiro: anais**. Cuiabá: Universidade Federal de Mato Grosso, 2013a. Não paginado. SBIAgro 2013.
- BARBEDO, J. G. A. Automatic method for counting and measuring whiteflies in soybean leaves using digital image processing. In: CONGRESSO BRASILEIRO DE AGROINFORMÁTICA, 9., 2013, Cuiabá. **Agroinformática: inovação para a sustentabilidade do agronegócio brasileiro: anais**. Cuiabá: Universidade Federal de Mato Grosso, 2013b. Não paginado. SBIAgro 2013.
- BARBEDO, J. G. A. Automatic object counting in Neubauer chambers. In: SIMPÓSIO BRASILEIRO DE TELECOMUNICAÇÕES, 31., 2013, Fortaleza. [Anais...]. Rio de Janeiro: Sociedade Brasileira de Telecomunicações, 2013c. Não paginado.
- BARBEDO, J. G. A. An algorithm for counting microorganisms in digital images. **Revista IEEE América Latina**, v. 11, p. 1354-1359, 2013d.
- BARBEDO, J. G. A. A Study on the use of semi - automatic systems for counting objects in digital images. **International Journal of Computer and Information Technology**, v. 2, p. 1022-1028, Nov. 2013e.
- BARBEDO, J. G. A. **Digipathos**. 2013f. Disponível em: <<https://www.agropediabrasilis.cnptia.embrapa.br/web/digipathos>>. Acesso em: 2 set. 2014.
- BARBEDO, J. G. A. Computer-aided disease diagnosis in aquaculture: current state and perspectives for the future. **Revista Innovar**, v. 1, n. 1, p. 19-32, 2014a.
- BARBEDO, J. G. A. An automatic method to detect and measure leaf disease symptoms using digital image processing. **Plant Disease**, St Paul, 2014b. Em fase de elaboração.
- BARBEDO, J. G. A. Using digital image processing for counting whiteflies on soybean leaves. **Journal of Asia-Pacific Entomology**, Taiwan, v. 17, p. 685-694, 2014c.
- BARBEDO, J. G. A. Method for counting microorganisms and colonies in microscopic images. In: INTERNATIONAL CONFERENCE ON COMPUTATIONAL SCIENCE AND ITS APPLICATIONS, 12., 2012, Salvador. **Proceedings...** Piscataway, NJ : IEEE, 2012. p. 83-87. ICCSA 2012.
- BOCK, C. H.; POOLE, G. H.; PARKER, P. E.; GOTTSWALD, T. R. Plant disease severity estimated visually, by digital photography and image analysis, and by hyperspectral imaging. **Critical Reviews in Plant Sciences**, Boca Raton, v. 29, n. 1-3, p. 59-107, Jan./Jun. 2010. DOI: 10.1080/07352681003617285.
- BOISSARD, P.; MARTIN, V.; MOISAN, S. A cognitive vision approach to early pest detection in greenhouse crops. **Computers and Electronics in Agriculture**, Amsterdam, v. 62, n. 2, p. 81-93, July. 2008. DOI: 10.1016/j.compag.2007.11.009.
- BOURKE, P. M. A. Use of weather information in the prediction of plant disease epiphytotics. **Annual Review of Phytopathology**, Palo Alto, v. 8, n. 1, p. 345-370, Sept. 1970.
- BREIMAN, L.; FRIEDMAN, J. H.; OLSHEN, R. A.; STONE, C. J. **Classification and regression trees**. Boca Raton: CRC Press, 1984. 358 p.
- BUTT, D. J.; ROYLE, D. J. Multiple regression analysis in the epidemiology of plant diseases. In: KRANZ, J. (Ed.) **Epidemics of plant diseases: mathematical analysis and modeling**. 2nd ed. Berlin: Springer-Verlag, 1990. p. 143180.
- CAMARGO, A.; SMITH, J. S. Image pattern classification for the identification of disease causing agents in plants. **Computers and Electronics in Agriculture**, Nwe York, v. 66, n. 1, p. 121-125, May, 2009. DOI: 10.1016/j.compag.2009.01.003.
- CAMPBELL, C. L.; MADDEN, L. V. **Introduction to plant disease epidemiology**. New York: J. Wiley, 1990. 532 p.
- CAMPBELL, C. L.; REYNOLDS, K. M.; MADDEN, L. V. Modeling epidemics of root diseases and development of simulators. In: KRANZ, J.; ROTEM, J. (Ed.) **Experimental techniques in plant disease epidemiology**. Berlin: Springer-Verlag, 1988. p. 253-265.
- CHALFOUN, S. M.; CARVALHO, V. L. de; PEREIRA, M. C. Efeito de alterações climáticas sobre o progresso da ferrugem (*Hemileia vastatrix* Berk. & Br.) do cafeeiro (*Coffea arabica* L.). **Ciência e Agrotecnologia**, Lavras, MG, v. 25, n. 5, p. 1248-1252, 2001.
- CINTRA, M. E.; MEIRA, C. A. A.; MONARD, M. C.; CAMARGO, H. A.; RODRIGUES, L. H. A. The use of fuzzy decision trees for coffee rust warning in Brazilian crops. In: INTERNATIONAL CONFERENCE ON INTELLIGENT SYSTEMS DESIGN AND APPLICATIONS, 11., 2011, Córdoba. **Proceedings...** Piscataway: IEEE, 2011. p. 1347-1352. ISDA 2011.
- COAKLEY, S. M. Variation in climate and prediction of disease in plants. **Annual Review of Phytopathology**, Palo Alto, v. 26, p. 163-181, Sept. 1988. DOI:10.1146/annurev.py.26.090188.001115.
- CONTRERAS-MEDINA, L. M.; OSORNIO-RIOS, R. A.; TORRES-PACHECO, I.; ROMERO-TRONCOSO, R. J.; GUEVARA-GONZÁLEZ, R. G.; MILLAN-ALMARAZ, J. R. Smart sensor for real-time quantification of common symptoms present in unhealthy plants. **Sensors**, Basel, v. 12, n. 1, p. 784-805, Jan. 2012. DOI: 10.3390/s120100784.
- COSTA, R. V.; ZAMBOLIM, L.; VALE, F. X. R.; MIZUBUTI, E. S. G. Previsão da queima da batateira. **Fitopatologia Brasileira**, Brasília, DF, v. 27, n. 4, p. 349-354, 2002.
- DEL PONTE, E. M.; GODOY, C. V.; LI, X.; YANG, X. B. Predicting severity of Asian soybean rust epidemics with empirical rainfall models. **Phytopathology**, St.Paul, v. 96, n. 7, p. 797-803, July, 2006. DOI: 10.1094/PHYTO-96-0797.
- DE WOLF, E. D.; MADDEN, L. V.; LIPPS, P. E. Risk assessment models for wheat Fusarium head blight epidemics based on within-season weather data. **Phytopathology**, St.Paul, v. 93, n. 4, p. 428-435, Apr. 2003. DOI: 10.1094/PHYTO.2003.93.4.428.
- DI GIROLAMO NETO, C. D.; MEIRA, C. A. A.; RODRIGUES, L. H. A. Avaliação de modelos de alerta da ferrugem do cafeeiro para lavouras com alta carga pendente de frutos. In: CONGRESSO BRASILEIRO DE PESQUISAS CAFEEIRAS, 38., 2012, Caxambu. **Anais...** Brasília, DF: Embrapa Café, 2012. p. 120-121.
- DI GIROLAMO NETO, C. D.; RODRIGUES, L. H. A.; MEIRA, C. A. A. Modelos de predição da ferrugem do cafeeiro (*Hemileia vastatrix* Berkeley & Broome) por técnicas de mineração de dados. **Coffee Science**, Lavras, MG, v. 9, n. 3, p. 408-418, jul./set. 2014.
- FERNANDES, J. M.; PAVAN, W.; SANHUEZA, R. M. V. Sisalert a generic web-based plant disease forecasting system. In: INTERNATIONAL CONFERENCE ON INFORMATION AND COMMUNICATION TECHNOLOGIES IN AGRICULTURE, FOOD AND ENVIRONMENT, 5., 2011, Skiathos Island. **Proceedings...** Skiathos: HAICTA, 2011. v. 1. p. 225-233.
- FIALHO, F. B.; GARRIDO, L. da R.; BOTTON, M.; MELO, G. W. B. de; FAJARDO, T. V. M.; NAVES, R. de L. **Diagnóstico de doenças e pragas na cultura da videira usando o sistema especialista Uzum**. Bento Gonçalves: Embrapa Uva e Vinho, 2012. 4 p. (Embrapa Uva e Vinho. Comunicado técnico, 128). Disponível em: <<http://ainfo.cnptia.embrapa.br/digital/bitstream/item/73877/1/cot128.pdf>>. Acesso em: 12 out. 2014.
- FOSTER, M.; ITO, R. P. M.; WARD, R. Detection and counting of uneaten food pellets in a sea cage using image-analysis. **Aquacultural Engineering**, London, v. 14, n. 3, p. 251-269, 1995. DOI: 10.1016/0144-8609(94)00006-M.
- GARÇON, C. L. P., ZAMBOLIM, L., MIZUBUTI, E. S. G., VALE, F. X. R.; COSTA, H. Controle da ferrugem do cafeeiro com base no valor de severidade. **Fitopatologia Brasileira**, Brasília, DF, v. 29, n. 5, p. 486-491, 2004.
- GENT, D. H.; SCHWARTZ, H. F. Validation of potato early blight disease forecast models for Colorado using various sources of meteorological data. **Plant Disease**, St Paul, v. 87, n. 1, p. 78-84, Jan. 2003. DOI: 10.1094/PDIS.2003.87.1.78.

HARDWICK, N. V. Disease forecasting. In: COOKE, B. M.; JONES, D. G.; KAYE B. (Ed.) **The epidemiology of plant diseases**. 2nd ed. Dordrecht: Springer, 2006. p. 239-267.

HAU, B.; KRANZ, J. Mathematics and statistics for analysis in epidemiology. In: KRANZ, J. (Ed.) **Epidemics of plant diseases: mathematical analysis and modeling**. 2nd ed. Berlin: Springer-Verlag, 1990. p. 12-52.

HUANG, K. Y. Application of artificial neural network for detecting Phalaenopsis seedling diseases using color and texture features. **Computers and Electronics in Agriculture**, New York, v. 57, n. 1, p. 3-11, May 2007. DOI: 10.1016/j.compag.2007.01.015.

JENSEN, R. E.; BOYLE, L. W. A technique for forecasting leafspot on peanuts. **Plant Disease Reporter**, Beltsville, v. 50, n. 11, p. 810-814, 1966.

KUSHALAPPA, A. C. Biology and epidemiology. In: KUSHALAPPA, A. C.; ESKES, A. B. (Ed.) **Coffee rust: epidemiology, resistance, and management**. Boca Raton, Florida: CRC Press, 1989a. p. 13-80.

KUSHALAPPA, A. C. Rust management: an epidemiological approach and chemical control. In: KUSHALAPPA, A. C.; ESKES, A. B. (Ed.) **Coffee rust: epidemiology, resistance, and management**. Boca Raton, Florida: CRC Press, 1989b. p. 81-139.

KUSHALAPPA, A. C. Epidemiologia da ferrugem do cafeeiro sob alta densidade de plantio: um enfoque de sistema. In: SIMPÓSIO INTERNACIONAL SOBRE CAFÉ ADENSADO, 1994, Londrina. **Anais...** Londrina: IAPAR, p. 131-147, 1994.

KUSHALAPPA, A. C.; ESKES, A. B. Advances in coffee rust research. **Annual Review of Phytopathology**, Palo Alto, v. 27, p. 503-531, Sept. 1989.

KUSHALAPPA, A. C.; AKUTSU, M.; LUDWIG, A. Application of survival ratio for monocyclic process of *Hemileia vastatrix* in predicting coffee rust infection rates. **Phytopathology**, St. Paul, v. 73, n. 1, p. 96-103, 1983. DOI: 10.1094/Phyto-73-96

KUSHALAPPA, A. C.; HERNANDEZ, T. A.; LEMOS, H. G. Evaluation of simple and complex coffee rust forecasts to time fungicide application. **Fitopatologia Brasileira**, Brasília, DF, v. 11, p. 515-26, out. 1986.

KUSHALAPPA, A. C.; AKUTSU, M.; OSEGUERA, S. H.; CHAVES, G. M.; MELLES, C. Equations for predicting the rate of coffee rust development based on net survival ratio for monocyclic process of *Hemileia vastatrix*. **Fitopatologia Brasileira**, Brasília, DF, v. 9, p. 255-271, jun. 1984.

LLORET, J.; BOSCH, I.; SENDRA, S.; SERRANO, A. A wireless sensor network for vineyard monitoring that uses image processing. **Sensors**, Basel, v. 11, n. 6, p. 6165-6196, 2011. DOI: 10.3390/s110606165.

LUACES, O.; RODRIGUES, L. H. A.; MEIRA, C. A. A.; BAHAMONDE, A. Using nondeterministic learners to alert on coffee rust disease. **Expert Systems with Applications**, New York, v. 38, n. 11, p. 14276-14283, Oct. 2011. DOI: 10.1016/j.eswa.2011.05.003.

LUACES, O.; RODRIGUES, L. H. A.; MEIRA, C. A. A.; QUEVEDO, J. R.; BAHAMONDE, A. Viability of an alarm predictor for coffee rust disease using interval regression. In: INTERNATIONAL CONFERENCE ON INDUSTRIAL ENGINEERING AND OTHER APPLICATIONS OF APPLIED INTELLIGENT SYSTEMS, 23., 2010, Cordoba. **Trends in applied intelligent systems: proceedings**. Berlin: Springer, 2010. p. 337-346. (Lecture notes in artificial intelligence, 6097; Lecture notes in computer science).

MADDEN, L. V.; ELLIS, M. A. How to develop plant disease forecasters. In: KRANZ, J.; ROTEM, J. (Ed.) **Experimental techniques in plant disease epidemiology**. Berlin: Springer-Verlag, 1988. p. 191-208.

MADDEN, L.; PENNYPACKER, S. P.; MAC NAB, A. A. FAST, a forecast system for *Alternaria solani* on tomato. **Phytopathology**, St. Paul, v. 68, p. 1354-1358, 1978.

MAGAREY, R. D.; SEEM, R. C.; RUSSO, J. M.; ZACK, J. W.; WAIGHT, K. T.; TRAVIS, J. W.; OUDEMANS, P. V. Site-specific weather information without on-site sensors. **Plant Disease**, St Paul, v. 85, n. 12, p. 1216-1226, Dec. 2001.

MASSRUHÁ, S. M. F. S.; LIMA, H. P. Uma nova abordagem para diagnóstico de doenças via web. In: In: CONGRESSO BRASILEIRO DE AGROINFORMÁTICA, 8., 2011, Bento Gonçalves. **Anais...** Florianópolis: UFSC; Pelotas: UFPel, 2011. Disponível em: <http://ainfo.cnptia.embrapa.br/digital/bitstream/item/46276/1/89850-1.pdf>. Acesso em 10 out. 2014.

MASSRUHÁ, S. M. F. S.; DUTRA, J. P.; CRUZ, S. A. B. D. **Uma abordagem orientada a objetos para ambiente de diagnóstico remoto**. Campinas: Embrapa Informática Agropecuária, 2007a. 33 p. (Embrapa Informática Agropecuária. Boletim de pesquisa e desenvolvimento, 15). Disponível em: <http://ainfo.cnptia.embrapa.br/digital/bitstream/CNPTIA/11553/1/bp15.pdf>. Acesso em: 10 out. 2014.

MASSRUHÁ, S. M. F. S.; DUTRA, J. P.; CRUZ, S. A. B.; SANDRI, S.; WAINER, J. Uma plataforma orientada a objetos para desenvolvimento de sistemas especialistas de diagnóstico de doenças via Web. In: CONGRESSO BRASILEIRO DE AGROINFORMÁTICA, 6., 2007, São Pedro, SP. **Anais...** Campinas: Embrapa Informática Agropecuária, 2007b. p. 176-180. SBIAgro 2007.

MASSRUHÁ, S. M. F. S.; DUTRA, J. P.; CRUZ, S. A. B.; SANDRI, S.; WAINER, J.; MORANDI, M. A. B. An object-oriented framework for virtual diagnosis. In: BIENNIAL CONFERENCE OF EUROPEAN FEDERATION OF IT IN AGRICULTURE, 6., 2007c, Glasgow. **Environmental and Rural Sustainability through ICT: proceedings**. Glasgow: Glasgow Caledonian University, 2007c. 6 p. Disponível em: <http://ainfo.cnptia.embrapa.br/digital/bitstream/item/101814/1/2007AA-092.pdf>. Acesso em: 10 out. 2014.

MEIRA, C. A. A. **Processo de descoberta de conhecimento em bases de dados para a análise e o alerta de doenças de culturas agrícolas e sua aplicação na ferrugem do cafeeiro**. 2008. 198 p. Tese (Doutorado) - Universidade Estadual de Campinas, Campinas.

MEIRA, C. A. A.; RODRIGUES, L. H. A. Modelos em árvore de decisão para alerta da ferrugem do cafeeiro. In: SIMPÓSIO DE PESQUISA DOS CAFÉS DO BRASIL, 6., 2009, Vitória. **Inovação científica, competitividade e mudanças climáticas: anais...** Vitória: Consórcio Pesquisa Café, 2009. Não paginado.

MEIRA, C. A. A.; RODRIGUES, L. H. A.; MORAES, S. A. Análise da epidemia da ferrugem do cafeeiro com árvore de decisão. **Tropical Plant Pathology**, Brasília, DF, v. 33, n. 2, p. 114-124, Mar./Apr. 2008.

MEIRA, C. A. A.; RODRIGUES, L. H. A.; MORAES, S. A. Modelos de alerta para o controle da ferrugem-do-cafeeiro em lavouras com alta carga pendente. **Pesquisa Agropecuária Brasileira**, Brasília, DF, v. 44, p.233-242, mar. 2009. Disponível em: <http://ainfo.cnptia.embrapa.br/digital/bitstream/item/93878/1/66.pdf>. Acesso em: 10 out. 2014.

MICHALSKI, R. S.; DAVIS, J. H.; BISHT, V. S.; SINCLAIR, J. B. PLANT/DS: an expert consulting system for the diagnosis of soybean diseases. In: EUROPEAN CONFERENCE ON ARTIFICIAL INTELLIGENCE, 1982, Orsay. **Proceedings...** Kaiserslautern: Universität Kaiserslautern, 1982. p. 82-84.

MOLINEROS, J. E.; DE WOLF, E. D.; FRANCL, L.; MADDEN, L.; LIPPS, P. Modeling epidemics of fusarium head blight: trials and tribulations. **Phytopathology**, St.Paul, v. 95, p. 71, 2005.

MORAES, S. A. **A ferrugem do cafeeiro: importância, condições predisponentes, evolução e situação no Brasil**. Campinas: Instituto Agrônomo, 1983. 50 p. (IAC. Circular, 119).

MORAES, S. A.; SUGIMORI, M. H.; RIBEIRO, I. J. A.; ORTOLANI, A. A.; PEDRO JR., M. J. Período de incubação de *Hemileia vastatrix* Berk. et Br. em três regiões do Estado de São Paulo. **Summa Phytopathologica**, Piracicaba, v. 2, n. 1, p. 32-38, 1976.

PARVIN JUNIOR., D. W.; SMITH, D. H.; CROSBY, F. L. Development and evaluation of a computerized forecasting method for *Cercospora* leaf spot of peanuts. **Phytopathology**, St. Paul, v. 64, p. 385-388, Mar. 1974.

PAUL, P. A.; MUNKVOLD, G. P. A model-based approach to preplanting risk assessment for gray leaf spot of maize. **Phytopathology**, St. Paul, v. 94, n. 12, p. 1350-1357, Dec. 2004.

PAUL, P. A.; MUNKVOLD, G. P. Regression and artificial neural network modeling for the prediction of gray leaf spot of maize. **Phytopathology**, St. Paul, v. 95, n. 4, p. 388-396, Apr. 2005. DOI: 10.1094/PHYTO-95-0388.

PEDRO JÚNIOR, M. J.; MORAES, S. A.; GODOY, I. J. Agrometeorological forecasting method for cercospora leaf spot in peanuts. **Fitopatologia Brasileira**, Brasília, DF, v. 19, n. 1, p. 69-73, 1994.

PÉREZ-ARIZA, C. B.; NICHOLSON, A. E.; FLORES, M. J. Prediction of coffee rust disease using bayesian networks. In: EUROPEAN WORKSHOP ON PROBABILISTIC GRAPHICAL MODELS, 6., 2012, Granada. **Proceedings...** Granada University, 2012, p. 259-266.

PINTO, A. C. S.; POZZA, E. A.; SOUZA, P. E.; POZZA, A. A. A.; TALAMINI, V.; BOLDINI, J. M.; SANTOS, F. S. Descrição da epidemia da ferrugem do cafeeiro com redes neuronais. **Fitopatologia Brasileira**, Brasília, DF, v. 27, n. 5, p. 517-524, out. 2002.

POZZA, E. A.; MAFFIA, L. A.; SILVA, C. A. B.; BRAGA, J. L.; CERQUEIRA, F. G. TOMEX-UFV: um sistema especialista para diagnose de doenças do tomateiro. In: CONGRESSO DA SBIAGRO, 1. AGROSOFT 97, 1997, Belo Horizonte. **Anais...** [S.l.:s.n.], 1997.

PYDIPATI, R.; BURKS, T. F.; LEE, W. S. Identification of citrus disease using color texture features and discriminant analysis. **Computers and Electronics in Agriculture**, New York, v. 52, n. 1-2, p. 49-59, Jun. 2006. DOI: 10.1016/j.compag.2006.01.004.

QUINLAN, J. R. **C4.5**: programs for machine learning. San Francisco: Morgan Kaufmann, 1993. 302 p. il. (The Morgan Kaufmann series in machine learning).

REIS, E. M.; BRESOLIN, A. C. R. Sistemas de previsão de doenças de plantas. In: REIS, E. M. (Org.) **Previsão de doenças de plantas**. Passo Fundo: UPF, 2004. p. 155-287.

SENA JUNIOR, D. G.; PINTO, F. A. C.; QUEIROZ, D. M.; VIANA, P. A. Fallarmywormdamagedmaizeplantidentificationusing digital images. **Biosystems Engineering**, London, v. 85, n. 4, p. 449-454, Aug. 2003.

SILVA, F. C.; MASSRUHÁ, S. M. F. S.; DEUS, R. S.; SANTOS, A. D.; BARBIERI, V.; CRUZ, S. A. B.; MALAVOLTA, E. A web-based expert system for diagnosis of nutritional deficiency in sugarcane. **Bioenergia em Revista: Diálogos**, Piracicaba, v. 1, n. 1, p. 26-48, jan./jun. 2011.

SHTIENBERG, D.; ELAD, Y. Incorporation of weather forecasting in integrated, biological-chemical management of *Botrytis cinerea*. **Phytopathology**, St. Paul, v. 87, n. 3, p. 332-340, Mar. 1997.

SMITH, M.; BARTON, M.; BASS, M.; BRANSCHOFOSKY, M.; MCCLELLAN, G.; STUVE, D.; TANSLEY, R.; WALKER, J. H. DSpace: an open source dynamic digital repository. **D-Lib Magazine**, v. 9, n. 1, 2003. Disponível em: <<http://www.dlib.org/dlib/january03/smith/01smith.html>>. Acesso em: 14 out. 2014.

SOUZA, R. T.; FORCELINI, C. A.; REIS, E. M.; CALVETE, E. O. Validação de dois sistemas de previsão para a queima das folhas da cenoura. **Fitopatologia Brasileira**, Brasília, DF, v. 27, n. 1, p. 87-90, jan./fev. 2002.

STORY, D.; KACIRA, M.; KUBOTA, C.; AKOGLU, A.; AN, L. Lettuce calcium deficiency detection with machine vision computed plant features in controlled environments. **Computers and Electronics in Agriculture**, New York, v. 74, n. 2, p. 238-243, Nov. 2010.

THAMADA, T. T.; GIROLAMO NETO, C. D.; MEIRA, C. A. A. Sistema de alerta da ferrugem do cafeeiro. In: SIMPÓSIO DE PESQUISA DOS CAFÉS DO BRASIL, 8., 2013, Salvador. **Pesquisa cafeeira: sustentabilidade e inclusão social**: anais. Brasília, DF: Embrapa Café, 2013. Disponível em: <<http://ainfo.cnptia.embrapa.br/digital/bitstream/item/93878/1/66.pdf>>. Acesso em: 10 out. 2014.

WALLIN, J. R. Summary of recent progress in predicting late blight epidemics in United States and Canada. **American Potato Journal**, Orono, v. 39, p. 306-312, 1962.

WITTEN, I. H.; FRANK, E.; HALL, M. A. **Data mining**: practical machine learning tools and techniques. 3rd ed. San Francisco: Morgan Kaufmann, 2011. 629 p. (Morgan Kaufmann series in data management systems).

XU, G.; ZHANG, F.; SHAH, S. G.; YE, Y.; MAO, H. Use of leaf color images to identify nitrogen and potassium deficient tomatoes. **Pattern Recognition Letters**, Amsterdam, v. 32, n. 11, p. 1584-1590, Aug. 2011. DOI: 10.1016/j.patrec.2011.04.020.

XU, X.; HARRIS, D. C.; BERRIE, A. M. Modeling infection of strawberry flowers by *Botrytis cinerea* using field data. **Phytopathology**, St. Paul, v. 90, n. 12, p. 1367-1374, Dec. 2000. DOI: 10.1094/PHYTO.2000.90.12.1367.

ZADOKS, J. C. A quarter century of disease warning, 1958 - 1983. **Plant Disease**, Saint Paul, v. 68, n. 4, p. 352-355, 1984. DOI: 10.1094/PD-68-352.

ZAMBOLIM, L.; VALE, F. X. R.; COSTA, H.; PEREIRA, A. A.; CHAVES, G. M. Epidemiologia e controle integrado da ferrugem-do-cafeeiro. In: ZAMBOLIM, L. (Ed.). **O estado da arte de tecnologias na produção de café**. Viçosa, MG: Suprema Gráfica, 2002. p. 369-449.

ZAMBOLIM, L.; VALE, F. X. R.; ZAMBOLIM, E. M. Doenças do cafeeiro. In: KIMATI, H.; AMORIM, L.; REZENDE, J. A. M.; BERGAMIN FILHO, A.; CAMARGO, L. E. A. (Ed.). **Manual de fitopatologia: doenças das plantas cultivadas**. 4. ed. São Paulo: Agronômica Ceres, 2005. v. 2. p. 165-180.

ZHOU, Z.; ZANG, Y.; LI, Y.; ZHANG, Y.; WANG, P.; LUO, X. Rice plant-hopper infestation detection and classification algorithms based on fractal dimension values and fuzzy C-means. **Mathematical and Computer Modelling**, Oxford, v. 58, n. 3-4, p. 701-709, Aug. 2011.



## AgroTIC em agricultura de precisão e automação agrícola

Ariovaldo Luchiari Junior  
Leonardo Ribeiro Queirós  
João Camargo Neto  
Aldemir Chaim

### 1 Introdução

A agricultura está incorporando tecnologias emergentes para enfrentar novos desafios que levem à mudança do padrão tecnológico a fim de produzir alimentos para uma população global estimada em nove bilhões em 2050. Tal fato exigirá um aumento de 70 por cento na produção de alimentos de forma sustentável e segura (FAO, 2009).

Nesse contexto as tecnologias da informação e comunicação aplicadas à agricultura (AgroTIC) já estão sendo utilizadas e têm uma participação crescente e importante na gestão da cadeia valor. As AgroTIC estão evoluindo de forma contínua e têm atraído e aumentado os investimentos públicos e privados em pesquisa agrícola, transferência de tecnológica, extensão rural e desenvolvimento agrícola.

AgroTIC é um termo abrangente, sendo definido como: a combinação de hardware, software e instrumentos de produção que permitam coleta, armazenamento, troca, processamento e manejo da informação e do conhecimento. De acordo com a Agência dos Estados Unidos para Cooperação Internacional (Usaid), as TIC incluem tecnologias e métodos para armazenar, manejar e processar informação (por exemplo: computadores, softwares, livros, PDAs, tablets, androides, livrarias digitais e não digitais) e para comunicar a informação (por exemplo: correio, correio eletrônico, rádio, televisão, telefones, celulares, pagers, internet etc). A redução de seu preço, acessibilidade e adaptabilidade, além de suas novas capacidades, tornaram seu uso indispensável no setor agrícola. Produtores podem utilizar a internet, telefone e outras ferramentas e tecnologias digitais para: previsão do tempo, manejo de frota de veículos, rastreamento de produtos agrícolas, informações sobre preço de insumos, serviços, produtos, acesso a mercados, variedades, técnicas de produção, serviços de armazenamento, processamento etc. Serviços especializados, disponibilizados pelos setores público e privado, como o uso de satélites ou sensores remotos, armazenamento e processamento de grandes quantidades de dados em aplicativos móveis têm sido utilizados para planejar a produção, evitar perdas de safra, gerir a cadeia financeira etc.

Estudos realizados pelo FutureFarm<sup>1</sup> confirmam esses fatos e adicionam que as AgroTIC contribuirão para o estabelecimento de práticas de manejo baseadas em robótica e automação de processos agrícolas.

<sup>1</sup> Disponível em: <<http://www.futurefarm.eu/>>.

Para exemplificar o uso das AgroTIC é possível citar que as perguntas feitas pelos agricultores podem ser respondidas de modo rápido e preciso quanto às formas de aumentar a produtividade e minimizar os riscos devido às condições climáticas. O uso das TIC tem permitido o encontro entre agricultores, especialistas e outros agentes para selecionar as melhores soluções tecnológicas em uma determinada situação ou local.

Esses exemplos representam apenas um subconjunto dos serviços de informação e comunicação que podem ser fornecidos para o setor agrícola por meio das TIC, cada vez mais comuns e acessíveis na agricultura. Tais fatos têm promovido mudanças nas formas de produção de alimentos. Com esse conjunto de técnicas é possível planejar a produção agrícola, florestal e animal e usar de forma mais eficiente os recursos naturais bióticos e abióticos, os insumos químicos e biológicos, o capital humano, a infraestrutura e o que há de mais relevante no conhecimento gerado pelas instituições de ciência e tecnologia. O uso das TIC também permite minimizar os riscos associados às mudanças climáticas e às doenças e pragas. As inovações associadas às TIC permitem produzir, de forma eficiente, aumentando a quantidade, a qualidade e atendendo, dessa forma, aos requisitos do mercado. Hoje a produção de alimentos, quer seja pela produção convencional, orgânica ou por outros protocolos, como a Produção Integrada do Ministério da Agricultura, Pecuária e Abastecimento (Mapa) e o GLOBAL-GAP, devem atender às normas para que sejam rastreáveis e tenham conformidade de certificação de qualidade, de bem estar animal e de pegada ecológica.

Novas formas de produção e gestão têm se beneficiado do uso das AgroTIC. Queirós et al. (2014) mostram como a agricultura de precisão (AP) tem se beneficiado da utilização das tecnologias da informação e comunicação na agricultura. Rusten e Ramirez (2003), citados por Queirós et al. (2014), apontam que o conhecimento tecnológico é um componente importante para o desenvolvimento do setor agrícola e que as AgroTIC aceleram o desenvolvimento do setor por organizar e facilitar a transferência do conhecimento entre os atores que atuam no setor. Afirmam, ainda, que as organizações de pesquisa terão um papel fundamental tanto na identificação de necessidades de métodos adequados de manejo e de tomada de decisões, como também na identificação de novas necessidades tecnológicas para que o uso das AgroTIC em Agricultura de Precisão e na Automação Agrícola seja mais eficaz, eficiente e mais fácil de ser utilizado. Estudos realizados pelo projeto FutureFarm indicam que após 2030 a agricultura será norteada pelo conhecimento, fato que exigirá a adoção integral das AgroTIC para a gestão racional, eficiente e efetiva dos processos produtivos.

É importante ressaltar que as AgroTIC não são isoladamente um fim para o desenvolvimento agrícola. Incertezas ainda permanecem em como tornar essas inovações replicáveis, escaláveis e sustentáveis para uma população mundial diversificada e crescente, que exige alimentos seguros, étnicos, rastreáveis e produzidos de forma sustentável. Nesse contexto, o propósito central desse capítulo é o de diagnosticar e de analisar a situação atual, as tendências evolutivas e as possíveis mudanças das AgroTIC na pesquisa agrícola, na transferência de conhecimento e no desenvolvimento da agricultura brasileira relacionado à Agricultura de Precisão e Automação Agrícola.

## 2 Estado da arte

### 2.1 Uso de padrões para integração e interoperabilidade de dados em agricultura de precisão

Estudos direcionados a padronizar o armazenamento de dados e a arquitetura de sistemas de informação distribuídos que permitam a integração desses dados, de forma simples e transparente, são de extrema importância para facilitar o uso das AgroTIC (QUEIRÓS et al., 2014). O projeto europeu FutureFarm produziu uma especificação para um sistema de informação de gestão agrícola com atenção para essas questões. Nessa especificação todos os dados devem ser documentados e armazenados na linguagem padronizada para troca de dados agroXML e a arquitetura distribuída deve ser a arquitetura Service-Oriented Architecture (SOA) (BLACKMORE; APOSTOLIDI, 2011). Nos Estados Unidos a AgGateway, uma organização sem fins lucrativos que tem por visão ser reconhecida internacionalmente por promover o uso das TIC na Agricultura de Precisão, lançou o projeto ‘padronizando o intercâmbio de dados da AP (SPADE)’. O projeto visa atender às demandas dos produtores no sentido de tornar mais amigável o uso de equipamentos e aplicativos em AP (AGGATEWAY, 2013).

No Brasil, a Rede de Agricultura de Precisão da Embrapa - Rede AP - atenta à necessidade de adoção de padrões para armazenamento e intercâmbio de dados e informações, e de uma arquitetura orientada a serviços que permita a interoperabilidade entre sistemas, mantendo a memória, a preservação, a recuperação e o intercâmbio com qualidade dos dados produzidos pelas unidades pilotos, desenvolveu um repositório de recursos de informação<sup>2</sup> que usa o perfil de metadados ‘Perfil de Metadados Geoespaciais do Brasil’ - versão homologada em 2009 pelo Comitê de Planejamento da Infraestrutura Nacional de Dados Espaciais (Concar) - para catalogar os dados geoespaciais e com arquitetura que permite a integração e interoperabilidade de aplicações. Na Figura 1a é mostrada a estrutura do banco de dados e sua integração com a camada de aplicação. Os recursos de informação digitais suportam os formatos shapefile, raster, txt, doc, xls, jpg e pdf, e estão associados a elementos de metadados. Já a camada de aplicação é composta pela integração de ferramentas de softwares livre, como banco de dados PostgreSQL<sup>3</sup>, WebGIS i3Geo<sup>4</sup> e o aplicativo para catalogação de dados geoespaciais GeoNetwork<sup>5</sup>, e de conversores de dados dos equipamentos de sensoriamento usados pela Rede AP, com a função de realizar a interface com os usuários. Na Figura 1b são ilustrados os elementos de metadados selecionados, customizados, criados e em uso, para a catalogação de dados geográficos e não geográficos (somente tabulares). Nesse diagrama os elementos de metadados foram agrupados numa generalização e especialização. A generalização do diagrama representa os elementos de metadados que devem ser preenchidos, independente do tipo de dado ser tabular ou geográfico. Já na especialização expressa, os elementos de metadados devem ser preenchidos somente para o tipo geográfico. Vale ressaltar que dois novos elementos foram criados dentro da Seção Identificação do ‘Perfil de Metadados Geoespaciais do Brasil’: “Observação” para contemplar qualquer observação ou necessidade de documentação que por ventura não possa ser expressa pelos demais elementos

<sup>2</sup> Disponível em: <<https://www.redeap.cnptia.embrapa.br>>.

<sup>3</sup> Disponível em: <<http://www.postgresql.org/>>.

<sup>4</sup> Disponível em: <<http://www.mma.gov.br>>.

<sup>5</sup> Disponível em: <<http://geonetwork-opensource.org/>>.

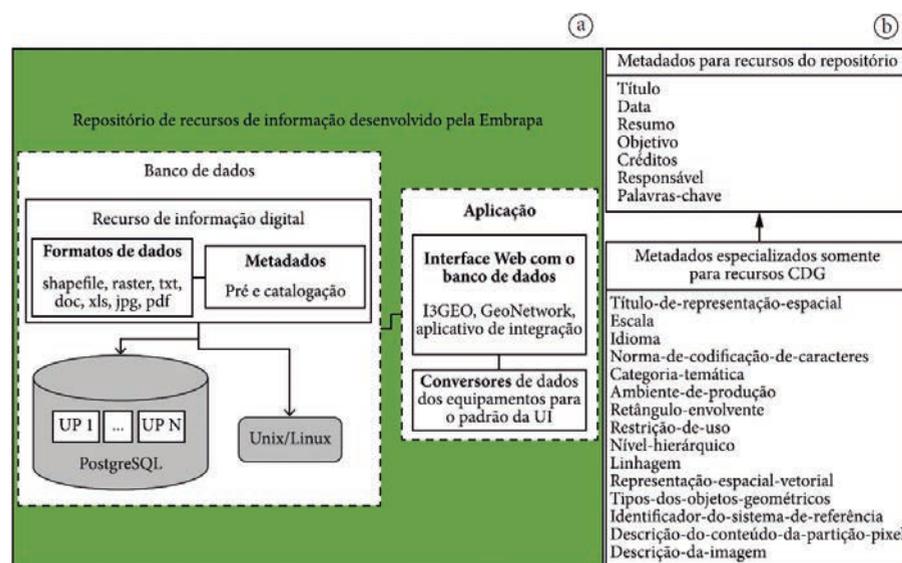


Figura 1. Repositório de recursos de informação desenvolvido pela Embrapa.

de metadados selecionados; e “Responsável pela Catalogação” para identificar o autor de documentação dos metadados. Com relação ao elemento “Observação”, a ideia é analisar a frequência de necessidade de uso desse descritor para, posteriormente, eleger ou criar novas seções ou elementos que atendam às especificidades de documentação do projeto AP (QUEIRÓS et al., 2011). O grande avanço conseguido pelo projeto é permitir a obtenção de séries históricas espaciais e temporais de dados do clima, das propriedades físicas e químicas do solo, dos parâmetros de desenvolvimento e de produtividade de plantas, sendo elemento chave não só para alimentar as novas necessidades de pesquisa, mas também para rastreabilidade e comparação entre sistemas que adotaram a AP. Considerando que o repositório da Rede AP e seus resultados permitiram o estabelecimento de padrões adequados para operacionalizar, armazenar, recuperar, intercambiar e interoperar os dados e informações obtidas nas unidades pilotos, de forma quantitativa e qualitativa, possibilitará também que essa experiência seja extrapolada para o manejo de propriedades agrícolas. Esse repositório foi concebido para atender necessidades futuras de organização e tratamento de informação.

## 2.2 Sistemas de suporte à decisão e sistemas de informação para gerenciamento integral da propriedade agrícola

A Figura 2 exemplifica o sistema de produção e os processos que nele ocorrem. Informações georreferenciadas dos atributos do solo (características físicas, químicas e biológicas) são coletadas, transmitidas e analisadas para que sejam estabelecidas as capacidades produtivas de áreas do terreno. Em função dessa análise, informações são transmitidas às máquinas e aos equipamentos para a aplicação automatizada de corretivos e fertilizantes em taxas variáveis. Em seguida, ocorre a operação de semeadura/ou plantio (mudas) automatizada com a utilização de plantas adequadas às diferentes capacidades produtivas do terreno, ou seja, para explorar a máxima capacidade produtiva do solo. Posteriormente, ocorrem as operações de manejo da cultura. Os estresses bióticos (patógenos, insetos e plantas daninhas) e abióticos (deficiências hídricas e nutricionais) podem ser



Figura 2. Ilustração das fases do sistema de produção e os processos que nele ocorrem.

determinados e georreferenciados pela utilização de sensores remotos (por satélite, avião, Vant). Todas essas informações são armazenadas e transmitidas numa linguagem padrão de intercâmbio, por exemplo AgroXML, para uma central, onde serão processadas e analisadas por um sistema específico de decisão, que encaminhará as decisões, em conformidade com o padrão ISOBUS6, para as máquinas equipadas com sistema de direção automática e equipamentos, que realizarão as operações de aplicação georreferenciadas em taxas variáveis de insumos (água, fertilizantes, defensivos, agentes de controle biológico etc). O ciclo é iniciado novamente após a colheita com a utilização de sensores de produtividade e/ou de qualidade (conteúdo de proteína, óleo ou outro parâmetro), cujos dados são enviados à central de processamento para a obtenção de mapas.

Sistemas de Suporte à Decisão (SSD) e AgroTIC são partes essenciais da Agricultura de Precisão para a coleta, manejo, análise e uso das informações espaciais e temporais.

Os SSD seguem alguns princípios essenciais:

- 1) Inteligência, que se refere à tomada de decisões baseadas em informações.
- 2) Design, refere-se ao desenvolvimento de uma solução ou ação alternativa que pode mitigar ou aliviar o problema.
- 3) Escolha, refere-se à escolha da ação ou solução mais apropriada ao problema identificado no processo de design.
- 4) Implementação da ação ou solução escolhida, no processo de design, para remediar o problema.

Uma possível maneira de se realizar essa categorização, de maneira a cobrir quase todos os aspectos da combinação TIC/Agropecuária, é baseá-la no fluxo que vai da observação do cenário vigente à tomada de decisões. Sob essa ótica, os desafios podem ser divididos em três grupos principais:

- 1) Geração de dados a partir da realidade observada.
- 2) Tratamento dos dados de modo a gerar informações relevantes.
- 3) Exploração das informações a fim de subsidiar a tomada de decisões.

O primeiro ponto é basicamente o objetivo da Agricultura de Precisão (AP), a qual vem ganhando cada vez mais espaço devido aos benefícios que esta pode trazer a toda a cadeia produtiva

(SCHEPERS et al. 2000; WOLF; WOOD, 1997). Com isso, vem sendo gerada uma grande quantidade de dados relacionados aos mais diferentes aspectos da cadeia agroindustrial.

O segundo ponto colocado acima ocupa-se exatamente do tratamento, processamento e exploração desses dados para que possam ser efetivamente transformados em informações relevantes. Dentre as tecnologias capazes de realizar essa tarefa, encontram-se métodos de modelagem baseados nos dados gerados (PAPAJORGJI; PARDALOS, 2009), técnicas de processamento digital de imagens e visão computacional capazes de explorar a grande quantidade de dados visuais que vêm sendo gerados (BARBEDO, 2013a), e técnicas de “Big Data” para os casos em que a quantidade de dados gerada é muito grande para ser processada por métodos mais convencionais (HOWE et al., 2008).

Por fim, o terceiro ponto visa explorar toda a informação gerada para ajudar nas tomadas de decisão que irão definir os rumos da lavoura. A integração de toda essa informação proveniente de diferentes fontes, de maneira que essa possa ser explorada de maneira sistêmica, é um dos maiores desafios da pesquisa agropecuária no momento. Sistemas capazes de realizar essa tarefa são comumente chamados de Sistemas de Gerenciamento de Informações da Lavoura (Farm Management Information Systems) (NIKKILA et al., 2010).

Como exemplo do uso das TIC/Agropecuária é possível citar Barbedo (2013b, 2013c) que vem desenvolvendo uma metodologia semiautomatizada, tendo como base ferramentas como aprendizado de máquina e morfologia matemática, e uma parte interativa baseada no conceito de sistemas especialistas para a detecção de doenças de plantas.

Outro caso do uso das TIC/Agropecuária são os estudos com as culturas do café (GIROLAMO NETO et al., 2014; LUACES et al., 2011; MEIRA et al., 2009) e da soja (MEGETO et al., 2014) que explora a descoberta de conhecimento em bases de dados para a análise e para formular o alerta de doenças de plantas por meio de técnicas de mineração de dados. A partir de dados de ocorrência e de incidência de doenças de culturas agrícolas e de outros dados de interesse, como, por exemplo, registros de estações meteorológicas automáticas, o objetivo é descobrir padrões e gerar modelos capazes de prever o progresso das doenças estudadas e auxiliar na tomada de decisão. Segundo os mesmos autores a tomada de decisão consiste em identificar os momentos oportunos para a aplicação de medidas de controle das doenças, visando à racionalização no uso de agroquímicos e, conseqüentemente, diminuição de custos para o produtor e menor risco de contaminação das pessoas e do ambiente. Os autores selecionam os melhores modelos que são incorporados em um sistema web de alerta para apoio a técnicos responsáveis pela emissão de boletins de avisos fitossanitários.

Para a operacionalização de Sistemas de Informação para Gerenciamento Integral da Propriedade Agrícola (FMIS), a integração, ubiquidade, alta capacidade de processamento de dados e sistemas de controle de decisões devem ser incorporados.

Queirós et al. (2014) recomendam que a integração de dados e informações obtidas por redes de sensores sem fio, ou que dados espaciais e temporais dos agroecossistemas, sejam tratados por padrões de representação e comunicação (agroXML, ISOBUS entre outros) entre sistemas numa arquitetura computacional distribuída, como o Arquitetura Orientada a Serviço (SOA). Devido à vasta quantidade de dados e informações obtidas, o processamento e análise em infraestruturas de alto desempenho computacional, como a computação em nuvens, grid, processamento paralelo, entre outros, é necessário, para o desenvolvimento de um sistema de informação de gestão agrícola automatizado, que seja robusto e confiável.

O avanço dos sistemas embarcados aliado ao custo decrescente de equipamentos digitais tem sido fecundo para realização de constantes investimentos em infraestrutura de telecomunicações em todo mundo (BALLANTYNE et al., 2010). Equipamentos como celulares, tablets, computadores pessoais - cada vez mais presentes no dia a dia das pessoas - conectados à Internet traz uma grande oportunidade de conectividade entre a ciência, produtores e demais atores relacionados ao contexto da Agricultura. Essa conectividade é facilitada quanto mais simples, autônomos e imperceptíveis forem os sistemas embarcados e equipamentos associados. A busca por não notoriedade da presença de computadores entre humanos, por meio da simplicidade de operação e maximização do funcionamento autônomo, tem sido conhecida por computação ubíqua. Torre Neto (2009) aponta como concepção da computação ubíqua a fusão dos computadores com o ambiente, a ponto de tornarem-se invisíveis para os usuários.

As tecnologias da AP tem se beneficiado dessa conectividade, em especial conectividades por meio de redes sem fio, e da computação ubíqua, nas quais sensores, redes de sensores, atuadores e sistemas de controle podem coletar dados, processá-los, realizar atuação e encaminhar informações para um computador servidor na sede da fazenda ou diretamente para algum serviço de nuvem disponível de forma autônoma e em tempo real. Como exemplo, a tecnologia de piloto automático, amplamente difundida na AP, permite que um veículo agrícola trafegue pela lavoura sem intervenção humana - sendo a coleta de informação dos sensores do motor, direção, localização espacial, entre outros, e a atuação na direção são realizadas de forma transparente e automática. Ainda nesse exemplo, o agricultor poderia monitorar em tempo real a rota realizada por meio de um aplicativo instalado em um tablet em qualquer lugar do mundo (HEST, 2013). Esses equipamentos, por atuarem de forma transparente, auxiliam o produtor a reduzir os erros e, portanto, reduz a variabilidade espacial antrópica e natural do campo.

Torre Neto (2009) tem abordado o monitoramento de controle de processos na agropecuária através do uso das inovadoras tecnologias de rede de sensores sem fio e da computação ubíqua por meio das seguintes atividades: a) a irrigação espacialmente diferenciada; b) a pulverização de precisão; c) o mapeamento da fertilidade do solo; d) a rastreabilidade animal e vegetal; e) as mudanças climáticas e os problemas fitossanitários.

### 2.3 Manejo de insumos em função da variabilidade espacial do campo

A principal hipótese para a adoção das tecnologias de Agricultura de Precisão é a existência da variabilidade espacial no campo. Luchiari Junior et al. (2001) conceituam zonas de manejo como sendo áreas do terreno de iguais produção potencial, eficiência do uso de insumos e risco de impacto ambiental. Os autores utilizaram mapas de colheita, mapas de condutividade elétrica do solo, mapas de classificação de solos, imagens do solo e de plantas para delinear zonas homogêneas de manejo e para direcionar as amostragens de solo. Shanahan et al. (2000) usaram procedimento similar para analisar o efeito de diferentes densidades de plantio em função das características do terreno e seus efeitos nos rendimentos do milho. Luchiari Junior et al. (2002) aplicaram o conceito de zonas homogêneas de manejo em solos tropicais de cerrados, cultivados com culturas anuais em plantio direto.

O uso de nutrientes dentro do conceito de zonas de manejo deve focar vários aspectos. A demanda por nutrientes pelas culturas depende de vários fatores, dentre eles a cultura e suas variedades, o potencial de rendimento e qualidade dos grãos, a distribuição de chuvas e o potencial produtivo do solo. Os principais atributos do solo que determinam seu potencial produtivo são aqueles res-

ponsáveis por manter a água no solo e fornecer nutrientes, como a textura, estrutura, agregação e sua estabilidade, relação macro e microporosidade, grau de compactação do solo e densidade. Os atributos físicos do solo muitas vezes são utilizados de forma reduzida na definição de zonas de manejo. Alguns deles, como a densidade do solo, poderiam ser incorporados no plano de manejo, pois impedimentos físicos para o crescimento radicular, mesmo em áreas onde a fertilidade química é alta, fazem com que o potencial produtivo do solo seja reduzido, bem com o potencial da cultura.

Fraisse et al. (2001) desenvolveram o primeiro aplicativo para delinear zonas homogêneas com o propósito de aplicação de nutrientes em taxa variada, o Management Zone Analyst (MZA). O software MZA requer informações quantitativas e georreferenciadas do terreno, condutividade elétrica aparente, da altitude, para que matematicamente o campo seja dividido em clusters e um número otimizado de zonas de manejo seja determinado automaticamente. O uso do software, conforme relato dos autores, tem sido usado por pesquisadores, organização de produtores, consultores e provedores de serviços agrícolas nos Estados Unidos e em 35 países ao redor do mundo. O conceito de zonas homogêneas vem evoluindo para selecionar os cultivares, densidade de plantio, práticas de manejo integradas de solo-planta-água e aplicação de agroquímicos de forma racional e criteriosa, baseada no potencial produtivo de cada zona homogênea. Bassoi et al. (2012) utilizaram essa prática em videira irrigada e mostraram que, mesmo no primeiro ciclo de cultivo, foi possível diferenciar a aplicação de água na área com o uso de sensores de umidade do solo instalados em zonas de manejo, previamente estabelecidas com base em atributos físico-hídricos do solo, e mostraram que a lâmina de água aplicada foi reduzida em algumas dessas zonas sem prejuízo à produtividade da cultura.

Speranza et al. (2011) criaram um sistema de armazenamento e recuperação de dados georreferenciados de condutividade elétrica do solo para duas profundidades de medidas - 0 a 30 cm (p1) e 0 a 90 cm (p2). Os arquivos gerados pelo equipamento durante a coleta são em formato texto e seguem um padrão específico. Cada linha do arquivo, contendo uma coleta realizada, disponibiliza a latitude, longitude e altitude do ponto coletado, além dos valores de p1 e p2, em mili Siemens por metro (mS/m). Para construção do conversor dos dados da condutividade elétrica do solo obtidas pelo equipamento VERIS foram utilizadas as linguagens PHP e JavaScript. Esse conversor foi incluído como uma ferramenta do repositório de dados da Rede Agricultura de Precisão da Embrapa (SPERANZA et al., 2011), baseado no servidor de mapas MapServer (Ambiente de software livre para construção de aplicativos espaciais na internet) e no software i3Geo (<http://www.mma.gov.br>) que é um WebGIS (aplicativo desenvolvido para o acesso e análise de dados geográficos via web) que permite a exibição dos dados coletados em forma de mapa que pode ser disponibilizado e acessado via Web. Os dados vetoriais do repositório da Rede AP são armazenados no Sistema Gerenciador de Banco de Dados (SGBD) PostgreSQL com extensão espacial PostGIS. Cada coleta de dados pode conter um ou mais arquivos e, por esse motivo, a interface da ferramenta permite o upload de vários arquivos ao mesmo tempo.

Além disso, essa interface possibilita a informação de alguns metadados pelo usuário, como Título, Data e Créditos, constituindo uma pré-catalogação da informação. Quando o upload é realizado, a ferramenta cria um novo tema que conterá a configuração básica para a exibição dos dados em forma de mapa. Cada um dos pontos de coleta é convertido no formato "POINT" disponível no PostGIS, e armazenado em uma tabela juntamente com os dados de profundidade (p1 e p2) e de altitude. O tema criado possui algumas configurações básicas geradas automaticamente,

como a classificação da p1 em cinco classes com intervalos iguais. Esse tema fica disponível na árvore de temas da unidade piloto após o upload, e os usuários do repositório poderão efetuar o download do mesmo em formato shapefile se essa permissão foi concedida pelo autor no momento do upload. A Figura 3, abaixo, mostra a sequência de passos executados, desde a coleta de dados em campo até a exibição das informações em forma de mapa. Pontos fortes dessa ferramenta são: a) a transformação de arquivos texto obtidos em campo para um formato capaz de ser visualizado e manipulado pelo usuário em forma de mapa, cuja interface de upload de arquivos pode ser visto na Figura 4; b) a relação à exibição dos dados, inicialmente os temas de condutividade elétrica do solo são criados com legenda padrão e classificação em cinco intervalos iguais referentes às medidas de p1 (0 cm a 30 cm) e permite ao usuário modificar essas configurações, variando os campos utilizados na classificação, bem como o método utilizado para a realização

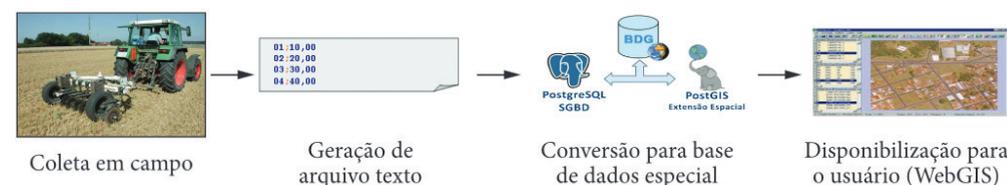


Figura 3. Coleta e disponibilização de dados de condutividade elétrica do solo.

Figura 4. Interface de upload de arquivos de condutividade elétrica do solo.

da mesma. A Figura 5 mostra um exemplo de visualização georreferenciada de dados coletados na unidade piloto de Mogi-Mirim, classificados por medidas de p2 em forma de quartis com a visualização do Google Maps disponível no repositório da Rede AP.



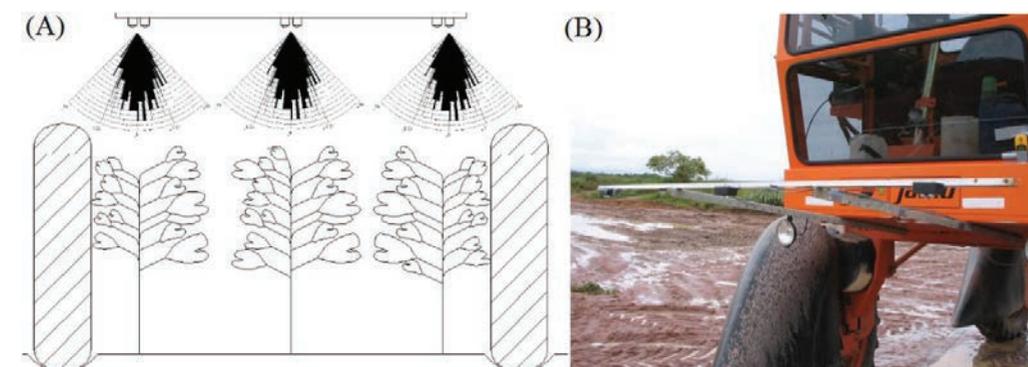
**Figura 5.** Visualização de dados de condutividade elétrica do solo no repositório da Rede AP.

Outra forma de definir zonas de manejo e o vigor de cultivos pode ser utilizando imagens multiespectrais. Luchiari Júnior et al. (2011) utilizaram esse procedimento para detectar, georreferenciar e mapear regiões de variabilidade causadas por doenças, deficiência nutricional, estresse hídrico, que refletem diretamente no vigor da planta, causando um declínio na produção de biomassa; e para mapear níveis de nitrogênio nas plantas que correlacionam resposta espectral com elevados níveis de clorofila e altas taxas de fotossínteses. Dentre esses índices podemos destacar o de vegetação de diferença normalizada na faixa do verde - Green Normalized Difference Vegetation Index (GNDVI) (SHANAHAN et al., 2001) e o de Vegetação Ajustado do Solo (SAVI) (BARET et al., 1989; HUETE, 1988; RONDEAUX et al, 1996). Estudos para estabelecimento de algoritmos para manejo de nutrientes utilizando sensores ativos de dossel, utilizando o índice de vegetação normalizada (NDVI) vêm sendo realizados no Brasil. Resende et al. (2014) destacam que um avanço importante foi a disponibilização de sensores multiespectrais. Quando acoplados a veículos agrícolas, esses sensores podem viabilizar de maneira ágil a obtenção de informações espacializadas sobre o estado nutricional das lavouras em relação ao nitrogênio (N). Se associados a equipamentos dotados de controladores de aplicação em taxa variada de fertilizantes, possibilitam o redimensionamento de dosagens de N em tempo real durante as adubações de cobertura. Resende et al. (2014) citam que no Brasil a utilização desse tipo de sensor encontra-se em fase de teste para desenvolvimento de algoritmos para aplicação de N em tempo real nas culturas da cana-de-açúcar (AMARAL; MOLIN, 2011), do milho (POVH et al., 2008; SHIRATSUCHI et al., 2011; 2014) e do trigo (GROHS et al., 2011; PIRES et al., 2014; POVH, et al., 2008).

Recentemente está crescendo a utilização de imagens multiespectrais adquiridas pelos Veículos Autônomos não Tripulados (VANT) pelas instituições de pesquisas e serviços de imageamento disponíveis no mercado por companhias privadas. As vantagens da utilização dessa tecnologia são: aquisição de imagens multiespectrais com alta resolução espacial; custo de obtenção inferior a imagens de satélites ou fotos aéreas; aquisição de imagem a qualquer instante; aquisição de imagens em tempo nublado por ser possível realizar voos abaixo da altura das nuvens; capacidade de execução de trabalhos repetitivos e perigosos em locais de difícil acesso.

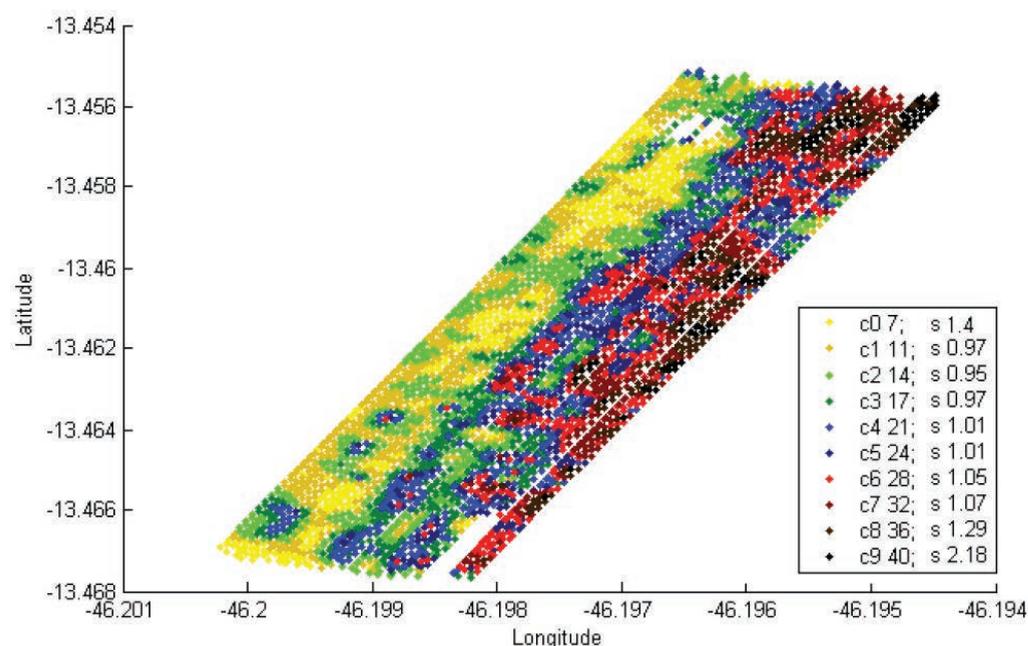
Devido aos poucos estudos sobre a aplicação de fitoreguladores (QUEIRÓS et al., 2005), propuseram e desenvolveram um sistema para automatizar o mapeamento de alturas de plantas de algodão com a intenção de auxiliar na formulação de dose variada de reguladores de crescimento. Os autores desenvolveram um protótipo baseado em conjunto de sensores ultrassônicos montados em uma barra, acoplada na parte frontal do pulverizador de maneira a não entrar em contato com as plantas sobre as fileiras de produção (Figura 6). Cada sensor estima a altura das plantas de uma fileira de produção considerando o tempo gasto entre a emissão e o retorno da onda sonora emitida na parte superior da planta. Shiratsuchi et al. (2005) justificam o desenvolvimento desse sistema com o mapeamento de uma área comercial de 50 ha de produção de algodão no município de Correntina, BA, onde constataram alta variabilidade na altura do algodão, conforme pode ser observado na Figura 7.

Segundo os autores, a justificativa de desenvolvimento do protótipo é corroborada por Thurman e Heinier (1998, 1999), onde foi conduzido experimentos em duas diferentes resoluções de amostragem (a cada 0,1 ha e 0,3 ha) de alturas de algodão em fazendas na Carolina do Norte. Os autores concluíram que a variabilidade de altura das plantas de algodão é grande o suficiente para justificar a aplicação à taxa variada de regulador de crescimento e relatam ganho de produtividade de 51-74 kg.ha<sup>-1</sup> em relação a áreas com aplicação tradicional uniforme devido à variabilidade de altura inerente do algodão. (QUEIROS et al., 2005) concluem que o sistema protótipo é viável para mapeamento automático das alturas, entretanto ressaltam a importância de mais pesquisas para o contínuo aprimoramento e adequação às realidades de produção do cotonicultor brasileiro. Uma evolução do sistema proposto por Magalhães et al. (2010) e Queiros et al. (2005, 2010), que desenvolveram um sistema de controle para aplicação à taxa variada e em tempo real de fitoreguladores na cultura do algodoeiro, foi a criação de um dispositivo de controle. Os autores consideraram a hipótese que utilizando este “dispositivo de controle com sistema hidráulico e controle eletrônico, é possível, em tempo real, variar a taxa de aplicação de fitoreguladores no algodão, baseado nas informações obtidas nos sensores de ultrassom, que detectam a altura das plantas também em tempo real; no algoritmo de crescimento da planta e na dosagem do fitoregulador estabelecido pelo usuário do sistema”. O dispositivo proposto (Figura 8) é composto de três sistemas: Hidráulico, Ultrassom e Controle. O sistema Hidráulico

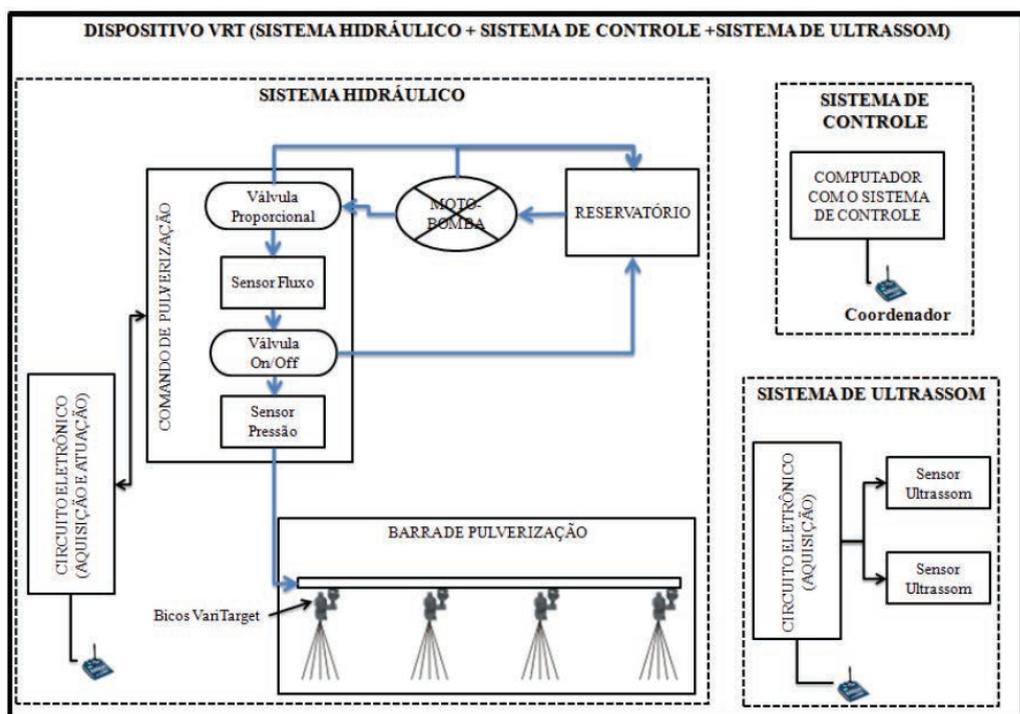


**Figura 6.** Protótipo para estimar a altura de plantas de algodão com sensores ultra-sônicos - (A) ilustração da emissão e recepção da onda sonora, (B) ilustração do arranjo de sensores montados na barra acoplada na parte frontal do pulverizador.

Fonte: Queirós et al. (2005).

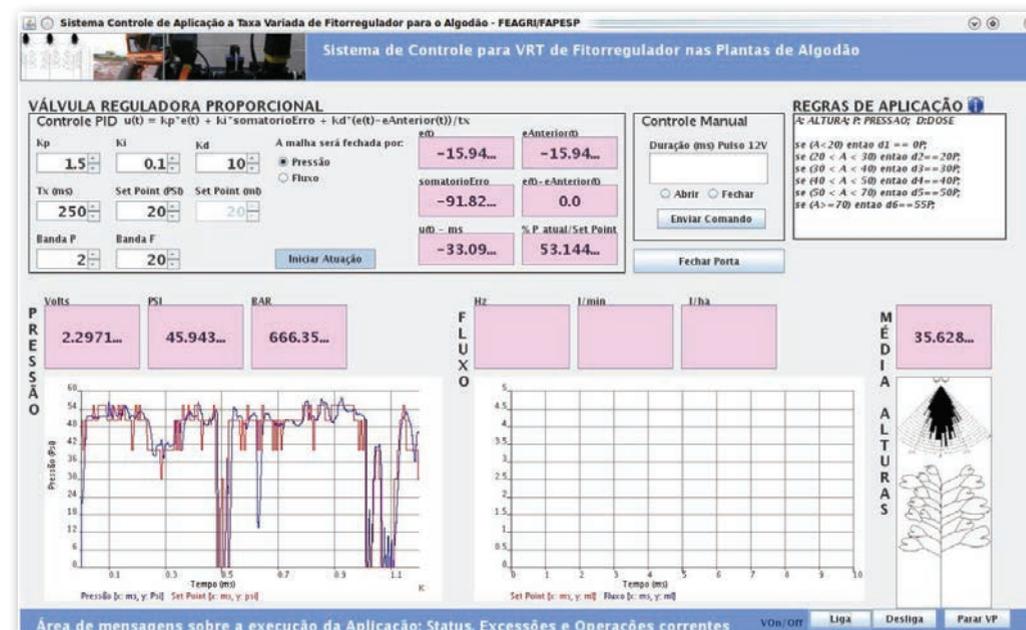


**Figura 7.** Variabilidade espacial de altura do algodão constatada numa área comercial de 50 ha no município de Correntina, BA. Cada cor corresponde a regiões de alturas diferentes sendo c0 até c9 as regiões de alturas em cm e “s” o desvio padrão de altura em cada região.



**Figura 8.** Visão geral dos componentes desses sistemas que formam o dispositivo de aplicação de fitorreguladores à taxa variada.

consiste do comando de pulverização e circuito eletrônico para aquisição de dados e atuação nos componentes desse comando. O sistema de Ultrassom é um circuito eletrônico composto por um conjunto de sensores ultrassom que realizam a medição da altura de plantas de algodão. O sistema de Controle é o núcleo do dispositivo, pois é o responsável por gerenciar as informações dos sistemas Hidráulico e de Ultrassom e, com um algoritmo de controle Proporcional-Integral-Derivativo (PID) retroalimentado pelo sensor de pressão e regras de aplicação de fitorregulador, é capaz de variar a dose em tempo real, controlando as válvulas do comando de pulverização. Os circuitos eletrônicos foram desenvolvidos com base na arquitetura microcontrolada PIC18F452 e comunicação sem fio por meio de módulos de rádio frequência que implementam o protocolo ZigBee. Para o desenvolvimento das interfaces (Figura 9) de validação dos três sistemas, adotou-se tecnologias de software livre, como a linguagem de programação Java, ambiente de desenvolvimento Netbeans<sup>6</sup> e sistema operacional Linux. Para a validação do Dispositivo proposto, foi utilizado um mapa de uma área de 7.3 ha com o registro georreferenciado de alturas de plantas de algodão e regras de aplicação de fitorregulador, considerando as velocidades de pulverização 2,5 m.s<sup>-1</sup> e 5 m.s<sup>-1</sup>. Independentemente das velocidades simuladas, houve potencial economia de aplicação de regulador de crescimento. Também foram realizadas validações para verificar se o dispositivo proposto manteria a pressão desejada, simulando o entupimento de bico de forma lenta e rápida. Os resultados mostram que em ambas as situações a pressão é mantida próxima ao valor desejado. Com os resultados obtidos, conclui-se que o dispositivo proposto atende ao objetivo de sua proposição.



**Figura 9.** Interface do Sistema de Controle do Dispositivo VRT.

<sup>6</sup> Disponível em: <<https://netbeans.org/>>.

Os autores relatam que as aplicações realizadas em laboratório pelo sistema de controle mostraram que é possível variar a dose de fitorregulador em tempo real. Entretanto, é necessário que mais pesquisas sejam realizadas com a finalidade de buscar seu aperfeiçoamento contínuo.

Para incrementar a eficiência da aplicação de agrotóxicos, Pessoa e Chaim (1999) desenvolveram um programa de computador em linguagem Qbasic que calcula o diâmetro mediano volumétrico, diâmetro mediano numérico e uniformidade de tamanho das gotas, bem como o volume de calda depositada em litros/ha. Posteriormente, outra versão do Gotas foi desenvolvida em linguagem Delphi para sistemas operacionais Windows, incorporando a análise de imagem das amostras digitalizadas por scanner. As imagens eram analisadas pelo programa, fornecendo os resultados de volume de calda depositado por hectare, tamanho de gotas (DMV e DMN), uniformidade de tamanho de gotas, densidade (número de gotas/cm<sup>2</sup>) e porcentagem de cobertura (CHAIM et al., 2002, 2006). A versão mais moderna do Gotas fornece os seguintes parâmetros:

- 1) Número de gotas encontrado na amostra. Informação para dar maior confiabilidade nos quesitos relacionados aos tamanhos das gotas, pois, quanto maior o número de gotas da amostra, mais confiável é o resultado da análise.
- 2) Número de diâmetros de gotas. Refere-se à quantidade de classes de tamanhos de gotas encontrados na amostra e está relacionada ao “fator de dispersão” e “volume de calda depositado na amostra”.
- 3) Fator de dispersão de tamanho de gotas. Esse parâmetro oferece informações das dispersões dos tamanhos das gotas dentro de uma amostra.
- 4) Volume de calda na amostra (L/ha). Esse parâmetro indica quantos litros de calda atingiram a amostra analisada.
- 5) Densidade de gotas (nº/cm<sup>2</sup>). Esse indicador fornece informações de quantas gotas atingiram uma área equivalente a um centímetro quadrado que é um parâmetro internacionalmente indicado para a calibração da deposição de gotas.
- 6) Diâmetro volumétrico D10 (mm). Representa a distribuição dos diâmetros das gotas de maneira tal que os diâmetros menores que D10 compõem 10% do volume total de líquido da amostra.
- 7) Diâmetro volumétrico D50 (mm). Representa a distribuição dos diâmetros das gotas de maneira tal que os diâmetros menores que D50 compõem 50% do volume total de líquido da amostra. Esse parâmetro também é conhecido como “diâmetro mediano volumétrico”, internacionalmente conhecido como VMD, e é muito importante para a calibração da deposição, pois expressa o padrão de diâmetro de gotas que atingiu a amostra.
- 8) Diâmetro volumétrico D90 (mm). Representa a distribuição dos diâmetros das gotas de maneira tal que os diâmetros menores que D90 compõem 90% do volume total de líquido da amostra.
- 9) Porcentagem de cobertura. Representa a porcentagem de área coberta pela mancha das gotas em relação à área total da amostra.

O programa Gotas é uma ferramenta importante para a calibração da deposição de agrotóxicos em agricultura de precisão. O acesso ao mesmo é gratuito e está disponível para download em

versão para computadores na rede de software livre para a agropecuária - Agrolivre<sup>7</sup>. As Figuras 10 e 11 ilustram a interface do sistema Gotas, disponível para execução em computador, e a interface do sistema Gotas para execução em plataforma Android, respectivamente. A versão para a plataforma Android, disponível para tablets e smartphones, estão disponíveis na loja virtual da Google - Play Store.

As condições de clima e tempo afetam todo o ciclo de produção dos cultivos. Damalgo et al. (2014) abordam a importância do uso da agrometeorologia e monitoramento climático para melhor compreensão dos processos da produção agrícola quando se utiliza tecnologias da agricultura de precisão. Ortiz (2013) descreve o “AgroClimate”<sup>8</sup>, um sistema de suporte à decisão para ser usado em práticas de manejo de sítios específicos. O Agroclimate informa sobre os riscos climáticos e ajuda na identificação das melhores práticas de manejo a serem usadas na produção agrícola para mitigar ou reduzir riscos específicos.

Por experiência, recomenda-se o sistema Agritempo<sup>9</sup> para uso em agricultura de precisão. O sistema é útil no monitoramento de variáveis meteorológicas, disponibilidade de água no solo e condições do tempo que possam prejudicar a aplicação de insumos químicos e biológicos. O sistema desenvolvido pela Embrapa e Universidade Estadual de Campinas (Unicamp) permite aos usuários o acesso, via Internet, às informações meteorológicas e agrometeorológicas de municípios e estados brasileiros. Além de informar a situação climática atual, o sistema alimenta a Rede Nacional de Agrometeorologia (RNA) do Mapa com informações básicas que orientam o zonea-

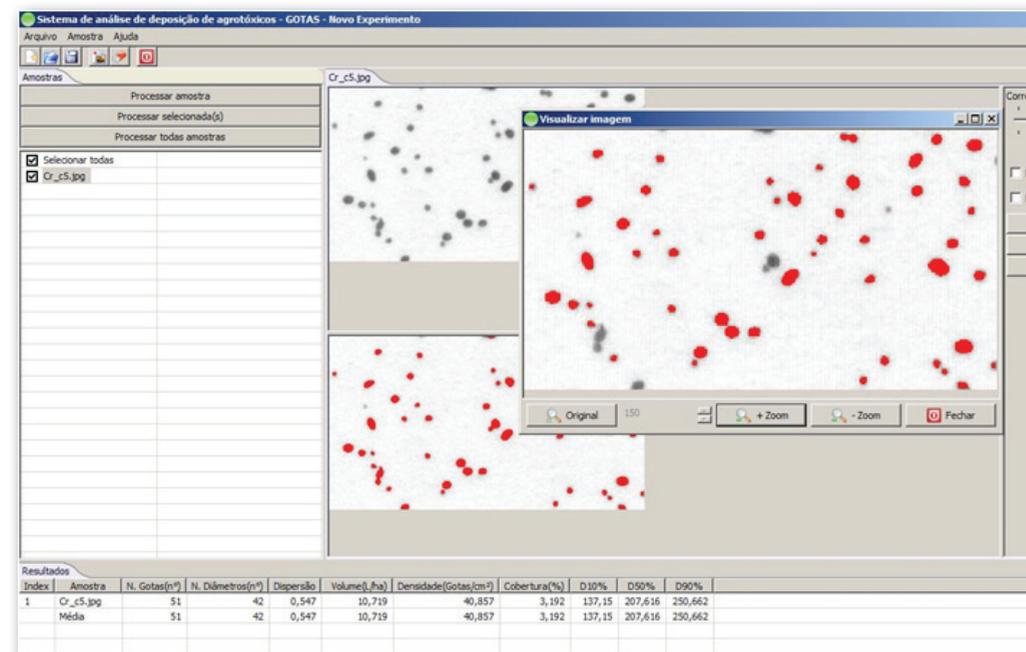


Figura 10. Interface do sistema Gotas disponível para execução em computador.

<sup>7</sup> Disponível em: <<https://repositorio.agrolivre.gov.br/>>.

<sup>8</sup> Disponível em: <[www.agroclimate.org](http://www.agroclimate.org)>.

<sup>9</sup> Disponível em: <[www.agritempo.gov.br](http://www.agritempo.gov.br)>.

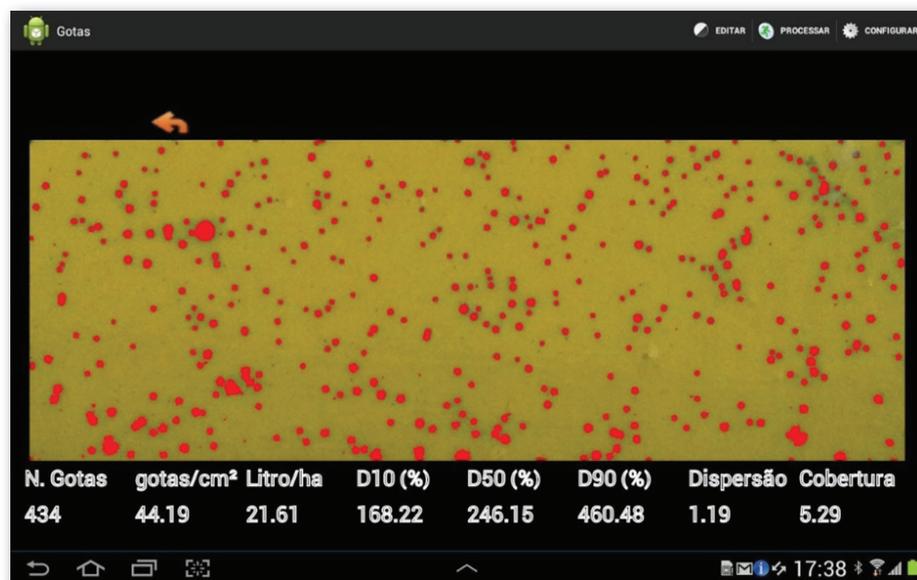


Figura 11. Interface do sistema Gotas para execução em plataforma Android.

mento agrícola brasileiro. Tal fato é importante na seleção dos melhores cultivares adaptados às zonas homogêneas de manejo definidas na propriedade rural. O sistema oferece boletins e mapas com informações sobre estiagem agrícola, distribuição temporal da precipitação pluvial, evapotranspiração, necessidades de irrigação, condições para tratamento fitossanitário, condições para manejo do solo e outras que suportam a decisão do produtor na aplicação de insumos químicos ou biológicos e no emprego de práticas agrícolas relacionadas ao manejo de água e solo.

#### 2.4 Uso de AgroTIC em processos produtivos agrícolas: formas, protocolos e normas de produção

Quando as tecnologias da Agricultura de Precisão são combinadas com as AgroTIC é possível obter, armazenar e processar informações que permitam ações de comando e controle da forma de produção. Isto permite atender, analisar, monitorar e rastrear a conformidade da produção com os requisitos de vários protocolos e normas, tais como: da Produção Integrada e da Produção Orgânica, do Mapa<sup>10</sup>, do Globalgap<sup>11</sup>, das produções agrícolas baseadas em princípios étnicos, e por contrato de produção de alimentos funcionais, entre outros.

Entretanto, o uso dessa forma de produção no Brasil não tem sido tão intenso. Furlaneto e Manzano (2010) citam o sucesso do uso de técnicas da agricultura na produção integrada e no processo de rastreabilidade do pêssego.

A coleta, registro e processamento das informações e documentação das produções certificadas estão caminhando para um sistema de manejo integrado com todos seus processos automatizados em função da tipologia e categorização dos produtores. Fundamentalmente todos os padrões e regras de produção possuem estruturas similares, que contém:

- 1) Metadados das áreas de cultivo, como, por exemplo: georreferenciamento, descrição da fisiografia, histórico de uso, cultivos a serem produzidos, métodos e formas de produção, procedimentos para análise de conformidade e de certificação, procedimentos a serem adotados em caso de não conformidade com as práticas de manejo, da qualidade do produto e com leis ambientais e trabalhistas.
- 2) Conjunto de regras: a) mandatórias em todos os casos; b) mandatórias parcialmente; c) que combinem as duas formas anteriores. Cada regra é uma descrição lógica de um fato que pode ser comprovado como “Falso” ou “Verdadeiro”, isto é, se está em conformidade ou não, seguida de uma ação corretiva para torná-la conforme. Outras categorias de regras referem-se ao estabelecimento de procedimentos operacionais padrão e da documentação de todas as operações e ações de gestão do processo produtivo.

Quando estas formas de produção são vistas dentro de um Sistema de Informação para Gerenciamento Integral da Propriedade Agrícola (FMIS) é preciso que a padronização da coleta, do armazenamento de dados e da arquitetura de sistemas de informação distribuídos permitam a integração e interoperabilidade desses dados de forma segura, rápida e simples.

Em relação ao futuro, o uso de tecnologias e processos da agricultura de precisão para atender a protocolos de certificação e rastreabilidade da produção é ainda uma incerteza crítica. Entretanto, o repositório concebido na Rede AP pode transformar-se numa tendência consolidada, com incremento do seu uso, por facilitar a organização e armazenamento de informações requeridas nas análises de conformidade constantes nos protocolos e normas de produção. Consequentemente permitirá que os produtores conquistem novos mercados, com garantia de melhores preços, devido à certificação da qualidade, segurança e origem dos produtos.

#### 2.5 Tecnologias disponíveis no mercado

Softwares para tratamento de dados e construção de mapas de colheita existem desde o início do lançamento dos monitores de colheita.

Como relatado por Hest (2013), algumas empresas já oferecem soluções que usam infraestrutura de computação em nuvem, na qual equipamentos agrícolas estão conectados por rede sem fio e as informações são disponibilizadas em tempo real e acessíveis por navegadores de Internet ou por aplicativos instalados em dispositivos móveis.

Dawson (2014) relata que a importância desse mercado de aplicativos com solução em nuvens pode ser vista pelos seguintes fatos:

- a) O big data Clima Corporation, que combina dados úteis para os agricultores juntamente com o seguro, foi recentemente comprada pela Monsanto.
- b) John Deere, DuPont e Dow Chemical uniram forças para usar dados retransmitidos de tratores para fazer recomendações aos agricultores quanto à densidade de sementes utilizadas no plantio, fertilizantes, herbicidas, fungicidas, inseticidas e outros insumos.

Ciampitti (2014) descreve as funções e finalidades de uso dos principais aplicativos para serem utilizados em dispositivos móveis. A Tabela 1 apresenta os principais apps disponíveis no mercado americano para serem utilizados em agricultura de precisão.

Um levantamento feito com empresas que atuam e disponibilizam máquinas, equipamentos e serviços no mercado brasileiro mostrou que existem tecnologias da agricultura de precisão e

<sup>10</sup> Disponível em: <www.agricultura.gov.br>.

<sup>11</sup> Disponível em: <www.globalgap.org/>.

**Tabela 1.** Aplicativos disponíveis no mercado americano para utilização em agricultura de precisão.

Ag-App Name*	iPhone	iPad	Android	Cost
Weedalert	Sim	Sim	Sim	Gratuito
Aphid Speed Scout	Sim	Sim	Sim	Gratuito
Pestbook	Sim	Sim	Não	Gratuito
Soybeans Diseases	Sim	Sim	Não	Gratuito
IPM Toolkit	Sim	Sim	Sim	Gratuito
Crop Nutrient Def	Sim	Sim	Sim	Gratuito
Fertilizer Removal	Sim	Sim	Sim	Gratuito
TankMix	Sim	Sim	Não	Gratuito
Corn N Rate Calculator	Sim	Sim	Sim	Gratuito
N price Calculator	Sim	Sim	Sim	Gratuito
Extreme Beans	Sim	Sim	Sim	Gratuito
Corn Yield Calculator	Sim	Sim	Sim	\$ 0.99
Planting Pop Calculator	Sim	Sim	Sim	Gratuito
Irrigation Calc App	Sim	Sim	Sim	\$ 1.99
CE Budgets	Sim	Sim	Sim	Gratuito

Fonte: Adaptado de Ciampitti (2014).

automação para culturas anuais, perenes e semiperenes que vão desde operações de preparo de solo até a colheita e controle de frota. Aqui não serão citados os nomes das empresas, mas serão apresentados, de forma geral, os equipamentos, atividades e serviços disponíveis. Serviços para estabelecimento de zonas de manejo já estão acessíveis, embora com uso restrito devido a sua natureza, isto é, somente uma passagem do equipamento é suficiente para estabelecer as características do terreno. Amostragens de solo georreferenciada utilizando código de barras têm sido utilizadas por produtores e provedores de serviços, e consideradas como efetivas para redução dos erros associados à coleta e análise das amostras de solo. Para o preparo do solo existem soluções na construção de terraços, curvas de níveis, canais de drenagem e nivelamento do terreno com economia de tempo e mão de obra. Soluções para o plantio em taxa variável, com comando e controle da semeadura, e da aplicação de nutrientes, associadas aos mapas de prescrição e realizadas por pilotos-automáticos e com controle de seções para evitar sobrepasses estão sendo incrementadas.

Já se encontram no mercado sensores ativos e passivos, baseados na assinatura espectral das plantas e no Índice de Vegetação por Diferença Normalizada (NDVI), além de outros equipamentos e métodos de sensoriamento contendo aplicativos baseados na web ou em dispositivos móveis para determinar, em tempo real, as necessidades de nitrogênio requeridas para as culturas de milho, trigo e algodão, determinando a presença ou ausência de plantas daninhas, insetos e doenças. Essas tecnologias já estão em utilização por produtores, provedores de serviço em várias regiões produtoras de grãos e fibras no Brasil. A localização georreferenciadas dos pontos monitorados e as doses aplicadas são enviados através do celular ou Wi-Fi para uma estação de armazenamento, controle e análise de informações para efetuar recomendações.

O uso das tecnologias da agricultura de precisão, principalmente a aplicação de insumos a taxas variadas, está se beneficiando muito do uso do piloto automático que aumenta a eficiência de uso

dos insumos, faz a aplicação nos locais pré-determinados, evita sobreposição, diminui falhas, diminui o tempo e aumenta o rendimento e a qualidade das operações e reduz a fadiga dos operadores.

Softwares e hardwares para o controle e manejo operacional da frota e equipamentos, em tempo real, já estão disponíveis para rastrear, monitorar as horas de funcionamento, diagnosticar remotamente as condições de manutenção, de uso e desempenho das máquinas, e equipamentos em atividades de preparo de solo, plantio, pulverizações, colheita, transporte e outras operações.

Um aspecto que deve ser considerado refere-se à política de propriedade e acesso aos dados, visto que o fornecimento e uso de AgroTIC para a Agricultura de Precisão e Automação de Processos Agrícolas tende a aumentar fortemente nos próximos anos.

### 3 Considerações finais

As tendências futuras indicam uma evolução e crescimento do uso das AgroTIC, eletrônica e mecatrônica nos temas de Computação Ubíqua, Tecnologia da Informação em Sensores e Equipamentos Agrícolas, Automação e Controle de Máquinas, Implementos Processos e Práticas Agrícolas. A evolução e desenvolvimento de novas tecnologias elevarão os índices de produtividade, da eficiência do uso de insumos, promoverão a redução de custos com mão de obra, a qualidade do trabalho, a segurança dos trabalhadores e a diminuição dos impactos ao meio ambiente. Estas tecnologias melhorarão a forma de gestão da propriedade, permitindo a obtenção de produtos com garantia de qualidade que atendam a protocolos de produção e requisitos de conformidade. A forma pela qual as informações e conhecimentos estão sendo disponibilizadas na web e em dispositivos móveis já estão permitindo que os atores envolvidos nas cadeias produtivas tenham acesso às informações e aos conhecimentos originados de fontes pluralísticas (GAKURU et al., 2009; GANDHI et al., 2009). Isso significa que, não somente o conhecimento gerado pelas instituições de pesquisas vem sendo utilizado, mas também o conhecimento tácito obtido por produtores, provedores de serviço e extensionistas está sendo utilizado nas inovações. A maioria das inovações já está sendo transferida de modo ubíquo, ou seja, o usuário está acessando um volume enorme de informações e utilizando formas de filtragem que selecionem as tecnologias e conhecimentos mais relevantes para sua situação.

Considerando as mudanças ocorridas na sociedade, devido aos impactos das novas Tecnologias de Informação e de Comunicação, são exigidos da Embrapa novos procedimentos e instrumentos para a organização e disponibilização das informações e dos conhecimentos de forma a culminar em transformações em direção a novos padrões tecnológicos de produção. Diante desse cenário, fica evidente a necessidade do desenvolvimento de sistemas de informação que ampliem a gestão além da propriedade, abrangendo toda a cadeia de valor.

### 4 Referências

AGGATEWAY. **Spade project**. 2013. Disponível em: <<http://www.aggateway.org/eConnectivity/Projects/CurrentOngoing/SPADE.aspx>>. Acesso em: 11 jul. 2013.

AMARAL, L. R.; MOLIN, J. P. Sensor óptico no auxílio à recomendação de adubação nitrogenada em cana-de-açúcar. **Pesquisa Agropecuária Brasileira**, Brasília, DF, v. 46, n. 12, p. 1633-1642, dec. 2011.

BALLANTYNE, P.; MARU, A.; PORCARI, E. M. Information and communication technologies — opportunities to mobilize agricultural science for development. **Crop Science**, Madison, v. 50, n. 1, p. S-63-S69, Mar./Apr. 2010. Suplemento. DOI: 10.2135/cropsci2009.09.0527.

BARBEDO, J. G. A. A digital image processing techniques for detecting, quantifying and classifying plant diseases. **SpringerPlus**, Heidelberg, v. 2, p. 1-12, Dec. 2013a. DOI:10.1186/2193-1801-2-660.

BARBEDO, J. G. A. Automatically measuring early and late leaf spot lesions in peanut plants using digital image processing. In: CONGRESSO BRASILEIRO DE AGROINFORMÁTICA, 9., 2013, Cuiabá. **Agroinformática: inovação para a sustentabilidade do agronegócio brasileiro**: anais. Cuiabá: Universidade Federal de Mato Grosso, 2013b. Disponível em: <<http://www.alice.cnptia.embrapa.br/bitstream/doc/971748/1/117625.pdf>>. Acesso em: 15 out. 2014.

BARBEDO, J. G. A. A Digital image processing-based automatic method for measuring rice panicle lengths. In: CONGRESSO BRASILEIRO DE AGROINFORMÁTICA, 9., 2013, Cuiabá. **Agroinformática: inovação para a sustentabilidade do agronegócio brasileiro**: anais. Cuiabá: Universidade Federal de Mato Grosso, 2013c. Disponível em: <http://www.alice.cnptia.embrapa.br/bitstream/doc/971740/1/117626.pdf>>. Acesso em: 15 out. 2014.

BARET, F.; GUYOT, G.; MAJOR, D. J. TSAVI: a vegetation index which minimizes soil brightness effects on LAI and APAR estimation. In: GEOSCIENCE AND REMOTE SENSING SYMPOSIUM, 12.; CANADIAN; CANADIAN SYMPOSIUM OF REMOTE SENSING, 10., 1989. Vancouver. **An economic tool for the nineties**: proceedings. New York: Institute of Electrical and Electronics Engineers, 1989. p. 1355-1358.

BASSOI, L. H.; NASCIMENTO, P. S.; COSTA, B. R. S.; SILVA, J. A.; ROCHA, M. G.; NASCIMENTO, E. F. Agricultura de precisão no manejo da irrigação em videira. In: INOVAGRI INTERNATIONAL MEETING, 1.; WORKSHOP INTERNACIONAL DE INOVAÇÕES TECNOLÓGICAS NA IRRIGAÇÃO, 4., 2012, Fortaleza. **Trabalhos apresentados...** Fortaleza: Instituto de Pesquisa e Inovação na Agricultura Irrigada, 2012. Disponível em: <<http://www.inovagri.org.br/meeting2012/wp-content/uploads/2012/06/Protocolo074.pdf>>. Acesso em: 15 out. 2014.

BLACKMORE, S.; APOSTOLIDI, K. **Futurefarm-Integration farm management information systems to support real-time management decisions and compliance of standards**: final report. 2011. Disponível em: <[http://www.futurefarm.eu/system/files/FFD8.9\\_Final\\_Report\\_4.1\\_Final.pdf](http://www.futurefarm.eu/system/files/FFD8.9_Final_Report_4.1_Final.pdf)>. Acesso em: 22 set. 2014.

CIAMPITTI, I. A. **Right app for a righ end use (RaReu)**. 2014. Disponível em: <[http://infoag.org/abstract\\_papers/papers/paper\\_260.pdf](http://infoag.org/abstract_papers/papers/paper_260.pdf)>. Acesso em: 26 set de 2014.

CHAIM, A.; CAMARGO NETO, J.; PESSOA, M. C. P. Y. **Use do programa computacional Gotas para avaliação da deposição de pulverização aérea sob diferentes condições climáticas**. Jaguariúna: Embrapa Meio Ambiente, 2006. 18p. (Embrapa Meio Ambiente. Boletim de pesquisa e desenvolvimento, 39). Disponível em: <<http://www.infoteca.cnptia.embrapa.br/bitstream/doc/15921/1/boletim39.pdf>>. Acesso em: 22 set. 2014.

CHAIM, A.; PESSOA, M.C.P.Y.; CAMARGO NETO, J.; HERMES, L. C. Comparison of microscopic method and computational program for pesticide deposition evaluation of spraying. **Pesquisa Agropecuária Brasileira**, Brasília, DF, v. 37, n. 4, p. 493-496, abr. 2002.

DAMALGO, G. A.; CUNHA, G. R. da.; PIRES, J. L. F.; SANTI, A.; FOCHESSATTO, E. Potencial de aplicação da Agrometeorologia em Agricultura de Precisão para produção de grãos. In: BERNARDI, A. C. de C.; NAIME, J. de M.; RESENDE, A. V. de; BASSOI, L. H.; INAMASU, R. Y. (Ed.). **Agricultura de precisão**: resultados de um novo olhar. Brasília, DF: Embrapa, 2014. p. 331-337. 1 CD-ROM.

DAWSON, G. **Farm app harvests big data**. modern farm. 2014. Disponível em: <<http://modernfarmer.com/2014/03/farm-app-harvests-big-data/>>. Acesso em: 26 set. 2014.

FAO. **The state of food insecurity in the world, economics crises - impacts and lessons learned**. Rome, 2009. 61 p

FRAISSE, C. W.; SUDDUTH, K. A.; KITCHEN, N. R. Delineation of site-specific management zones by unsupervised classification of topographic attributes and soil electrical conductivity. **Transactions of the ASAE**, St. Joseph, v. 44, n. 1, p. 155-166, Jan. 2001. Disponível em: <<http://www.ars.usda.gov/sp2UserFiles/Place/36221500/cswq-0051-101971.pdf>>. Acesso em: 22 set. 2014.

FURLANETO, F. B.; MANZANO, L. M. **Agricultura de precisão e a rastreabilidade de produtos agrícolas**. 2010. Disponível em: <[infobios.com/Artigos/2010\\_2/AgriculturaPrecisao/index.htm](http://infobios.com/Artigos/2010_2/AgriculturaPrecisao/index.htm)>. Acesso em: 2 jul de 2013.

GAKURU, M.; WINTERS, K.; STEPMAN, F. Inventory of innovative farmer advisory services using ICTs. In: IST AFRICA CONFERENCE & EXHIBITION, 2009, Uganda. **Proceedings...** Uganda: Ministry of ICT, 2009. p. 1-66.

GANDHI, R.; VEERARAGHAVAN, R.; TOYAMA, K.; RAMPRASAD, V. Digital green: Participatory video and mediated instruction for agricultural extension. **Information Technologies & International Development**, Los Angeles, v. 5, n. 1. p. 1-15, 2009.

GIROLAMO NETO, C.; RODRIGUES, L. H. A.; MEIRA, C. A. A. Modelos de predição da ferrugem do cafeeiro (Hemileia vastatrix Berkeley & Broome) por técnicas de mineração de dados. **Coffee Science**, Lavras, MG, v. 9, n. 3, p. 408-418, jul./set. 2014.

GROHS, D. S.; BREDEMEIER, C.; POLETT O, N.; MUNDSTOCK, C. M. Validação de modelo para predição do potencial produtivo de trigo com sensor óptico ativo. **Pesquisa Agropecuária Brasileira**, Brasília, DF, v. 46, n. 4, p. 446-449, abr. 2011.

HEST, D. **Capitalizing on the cloud: wireless connectivity in agriculture will make big gains in 2013**. 2013. Disponível em: <<https://www.onsiteag.com/news/capitalizing-on-the-cloud-wireless-connectivity-in-agriculture-will-make-big-gains-in-2013-21.html>>. Acesso em: 1 jul. 2013.

HOWE, D.; COSTANZO, M.; FEY, P.; GOJOBORI, T.; HANNICK, L.; HIDE, W.; HILL, D. P.; KANIA, R.; SCHAEFFER, M.; PIERRE, S. S.; TWIGGER, S.; WHITE, O.; RHEE, S. Y. Big data: The future of biocuration. **Nature**, v. 455, p. 47-50, Sep. 2008. DOI:10.1038/455047a.

HUETE, A. R. A soil-adjusted vegetation index (SAVI). **Remote Sensing of Environment**, New York, v. 25, n. 3, p. 295-309, Aug. 1988. DOI: 10.1016/0034-4257(88)90106-X.

LUACES, O.; RODRIGUES, L. H. A.; MEIRA, C. A. A.; BAHAMONDE, A. Using nondeterministic learners to alert on coffee rust disease. **Expert systems with applications**, New York, v. 38, n. 11, p. 14276-14283, Oct. 2011. DOI: 10.1016/j.eswa.2011.05.003.

LUCHIARI JUNIOR, A.; BORGHI, E.; AVANZI, J. C.; FREITAS, A. A.; BORTOLON, L.; BORTOLON, E. S. O.; INAMASU, R. Y. Zonas de manejo: teoria e prática. In: INAMASU, R. Y.; NAIME, J. M.; RESENDE, A. V.; BASSOI, L. H.; BERNARDI, A. C. (Ed.). **Agricultura de precisão**: um novo olhar. São Carlos, SP: Embrapa Instrumentação, 2011. p. 60-64.

LUCHIARI JUNIOR, A.; SHANAHAN, J.; LIEBIG, M.; SCHLEMMER, M.; SCHEPERS, J. S.; FRANCIS, D.; PAYTON, S. Strategies for establishing management zones for site specific nutrient management. In: INTERNATIONAL CONFERENCE ON PRECISION AGRICULTURE, 5., 2000, Minneapolis. **Proceedings...** Minneapolis: University of Minnesota, 2001. P. 1-10.

LUCHIARI JUNIOR, A.; SHANAHAN, J.; SCHEPERS, J.; FRANCIS, D.; SCHLEMMER, M.; SCHEPERS, A.; INAMASU, R.; FRANCA, G.; MANTOVANI, E.; GOMIDE, R. Crop and soil based approaches for site specific nutrient management In: CONGRESSO NACIONAL DE MILHO E SORGO, 24., 2002, Florianópolis. **Meio ambiente e a nova agenda para o agronegócio de milho e sorgo**: [palestras]. Sete Lagoas: ABMS: Embrapa Milho e Sorgo; Florianópolis: Epagri, 2002. 1 CD-ROM. Disponível em: <<http://ainfo.cnptia.embrapa.br/digital/bitstream/item/33872/1/Palestra-Crop-soil.pdf>>. Acesso em 22 set. 2014.

MAGALHAES, P. S. G.; QUEIROS, L. R.; GADANHA, C. D. Mepiquat chloride application on cotton at variable rate. In: INTERNATIONAL CONFERENCE ON PRECISION AGRICULTURE, 10., 2010, Denver. **Proceedings...** Denver: Denver Tech Center, 2010. Não paginado.

MEGETO, G. A. S.; OLIVEIRA, S. R. de M.; DEL PONTE, E. M.; MEIRA, C. A. A. Árvore de decisão para classificação de ocorrências de ferrugem asiática em lavouras comerciais com base em variáveis meteorológicas. **Engenharia Agrícola**, Jaboticabal, v. 34, n. 3, p. 590-599, maio/jun. 2014.

MEIRA, C.A.A.; RODRIGUES, L.H.A.; MORAES, S.A. de. Modelos de alerta para o controle da ferrugem-do-cafeeiro em lavouras com alta carga pendente. **Pesquisa Agropecuária Brasileira**, Brasília, DF, v. 44, n. 3, p. 233-242, mar. 2009.

NIKKILA, R.; SEILONEN, I.; KOSKINEN, K. Software architecture for farm management information systems in precision agriculture. **Computers and Electronics in Agriculture**, v. 70, n. 2, p. 328-336, Mar. 2010. DOI: 10.1016/j.compag.2009.08.013.

ORTIZ, B. V. Climate and weather data for the farmer's site specific toolkit. In: THE INFOAG CONFERENCE, 2013, Springfield. **Program guide**. 2013. Disponível em: <[http://infoag.org/abstract\\_papers/papers/abstract\\_144.pdf](http://infoag.org/abstract_papers/papers/abstract_144.pdf)>. Acesso em: 25 set. 2014.

PAPAJORGJI, P. J.; PARDALOS, P. M. Advances in modeling agricultural systems. Springer, New York, 2009. 483 p. (Springer optimization and its applications, v. 25).

PESSOA, M. C. P. Y.; CHAIM, A. Programa computacional para estimativa de perdas de herbicidas aplicados por pulverização aérea. **Pesquisa Agropecuária Brasileira**, Brasília, DF, v. 34, n.1, p. 45-56, jan. 1999.

PIRES, J. L. F.; CORASSA, G. M.; RAMBO, A. C.; KERBER, L.; PASINATO, A.; DAMALGO, G. A.; SILVA JÚNIOR, J. P. da.; SANTI, A.; SANTI, A. L.; GUARIENT, E. M.; CUNHA, G. R. da.; STRIEDER, E. F. Aplicação de nitrogênio a taxa variável em cultura de trigo: estudo de caso na Unidade Piloto de Agricultura de Precisão de Não-Me-Toque, RS. In: BERNARDI, A. C. de C.; NAIME, J. de M.; RESENDE, A. V. de; BASSOI, L. H.; INAMASU, R. Y. (Ed.). **Agricultura de precisão: resultados de um novo olhar**. Brasília, DF: Embrapa, 2014. p. 287-294. 1 CD-ROM.

POVH, F. P.; MOLIN, J. P.; GIMENEZ, L. M.; PAULETTI, V.; MOLIN, R.; SALVI, J. V. Comportamento do NDVI obtido por sensor ótico ativo em cereais. **Pesquisa Agropecuária Brasileira**, Brasília, DF, v. 43, n. 8, p. 1075-1083, ago. 2008.

QUEIRÓS, L. R. **Desenvolvimento de um sistema de controle para aplicação a taxa variada e em tempo real de fitorreguladores na cultura do algodoeiro**. 2010, 185 p. Tese (Doutorado em Engenharia Agrícola) - Faculdade de Engenharia Agrícola, Universidade Estadual de Campinas, Campinas, 2010. Orientador: Paulo Sérgio Graziano Magalhães.

QUEIRÓS, L. R.; SHIRATSUCHI, L. S.; VINHAL, C. D. N. Desenvolvimento de um sistema protótipo para o mapeamento da altura de plantas de algodão. In: SIMPÓSIO INTERNACIONAL DE AGRICULTURA DE PRECISÃO, 3., 2005, Sete Lagoas. [Anais...] Sete Lagoas: Embrapa Milho e Sorgo, 2005. p.1- 4.

QUEIRÓS, L. R.; LUCHIARI JUNIOR, A.; CAMARGO NETO, J.; MASSRUHÁ, S. M. F. S.; INAMASU, R. Y.; SPERANZA, E. A.; EVANGELISTA, S. R. M. **Análise das possibilidades e tendências do uso das tecnologias da informação e comunicação em agricultura de precisão**. In: BERNARDI, A. C. de C.; NAIME, J. de M.; RESENDE, A. V. de; BASSOI, L. H.; INAMASU, R. Y. (Ed.). **Agricultura de precisão: resultados de um novo olhar**. Brasília, DF: Embrapa, 2014. p. 97-108. 1 CD-ROM.

QUEIRÓS, L. R.; SPERANZA, E. A.; BETTIOL, G. M.; FILIPPINI ALBA, J. M.; BERNARDI, A. C. de C.; INAMASU, R. Y.; GREGO, C. R.; RABELLO, L. M. **Gestão de recursos de informação em Agricultura de Precisão**. Campinas: Embrapa Informática Agropecuária, 2011. 41 p. il. (Embrapa Informática Agropecuária. Documentos, 112).

RESENDE, A. V.; HURTADO, S. M. C.; VILELA, M. F.; CORAZZA, E. J.; SHIRATSUCHI, L. S. Aplicações da agricultura de precisão em sistemas de produção de grãos no Brasil. In: BERNARDI, A. C. de C.; NAIME, J. de M.; RESENDE, A. V. de; BASSOI, L. H.; INAMASU, R. Y. (Ed.). **Agricultura de precisão: resultados de um novo olhar**. Brasília, DF: Embrapa, 2014. p. 194-208. 1 CD-ROM.

RONDEAUX, G.; STEVEN, M.; BARET, F. Optimization of soil-adjusted vegetation indices. **Remote Sensing of Environment**, New York, v.55, n. 2, p. 95-107, Feb. 1996. DOI: 10.1016/0034-4257(95)00186-7.

RUSTEN, E.; RAMIREZ, S. **Future direction in agriculture and information and communication technologies (ICTs) at USAID**. New York: Academy for Educational Development, 2003. Disponível em: <[http://www.winrock.org/agriculture/files/ag\\_ict.pdf](http://www.winrock.org/agriculture/files/ag_ict.pdf)>. Acesso em: 22 set. 2014.

SCHEPERS, J. S.; SHANAHAN, J. F.; LUCHIARI JR, A. Precision agriculture as a tool for sustainability. In: BALAZS, E.; GALANTE, E.; LYNCH, J. M.; SCHEPERS, J. S. (Ed.) **Biological Resource Management: connecting science and policy**. Berlin ; New York : Springer, 2000. p. 129-135, 2000.

SHANAHAN, J. F.; DOERGE, T.; SYNER, C.; LUCHIARI JUNIOR, A.; JOHNSON, J. Feasibility of variable rate management of corn hybrids and seeding rates in the great plains. In: INTERNATIONAL CONFERENCE ON PRECISION AGRICULTURE, 5., 2000, Minneapolis. **Proceedings...** Minneapolis: University of Minnesota, 2000.

SHANAHAN, J. F.; SCHEPERS, J. S.; FRANCIS, D. D.; VARVEL, G. E.; WILHELM, W. W.; TRINGE, J. M.; SCHLEMMER, M. S.; MAJOR, D. J. Use of remote sensing imagery to estimate corn grain yield. **Agronomy Journal**, Madison, v. 93, p. 583-589, May/June, 2001

SHIRATSUCHI, L. S.; QUEIROS, L. R.; VINHAL, D. N. Mapeamento da altura de plantas de algodão utilizando sensor ultrassom. In: SIMPÓSIO INTERNACIONAL DE AGRICULTURA DE PRECISÃO, 3., 2005, Sete Lagoas. [Anais...]. Sete Lagoas: Embrapa Milho e Sorgo, 2005. p. 1-4. 1 CD-ROM.

SHIRATSUCHI, L. S.; VALE, W. G. do; MALACARNE, T. J.; SCHUCK, C. M.; SILVA, R. G.; OLIVEIRA JÚNIOR, O. L. de. Algoritmos para aplicações de doses variáveis de nitrogênio em tempo real para produção de milho safra e safrinha no Cerrado. Embrapa Instrumentação. In: BERNARDI, A. C. de C.; NAIME, J. de M.; RESENDE, A. V. de; BASSOI, L. H.; INAMASU, R. Y. (Ed.). **Agricultura de precisão: resultados de um novo olhar**. São Carlos: Embrapa Instrumentação, 2014. p. 224-230. 1 CD-ROM.

SHIRATSUCHI, L. S.; VILELA, M. F.; FERGUSON, R. B.; SHANAHAN, J. F.; ADAMCHUK, V. I.; RESENDE, A. V.; HURTADO, S. C.; CORAZZA, E. J. Desenvolvimento de um algoritmo baseado em sensores ativos de dossel para recomendação da adubação nitrogenada em taxas variáveis. In: INAMASU, R. Y.; NAIME, J. M.; RESENDE, A. V.; BASSOI, L. H.; BERNARDI, A. C. C. **Agricultura de Precisão: um novo olhar**. São Carlos: Embrapa Instrumentação, 2011. p. 184-188.

SPERANZA, E. A.; QUEIRÓS, L. R.; RABELLO, L. M.; GREGO, C. R.; BRANDÃO, Z. N. Armazenamento e recuperação de dados georreferenciados de condutividade elétrica do solo na Rede de Agricultura de Precisão da Embrapa. In: INAMASU, R. Y.; NAIME, J. de M.; RESENDE, A. V. de; BASSOI, L. H.; BERNARDI, A. C. de C. (Ed.). **Agricultura de precisão: um novo olhar**. São Carlos, SP: Embrapa Instrumentação, 2011. P. 46-50.

THURMAN, M. E.; HEINIGER, R. W. Evaluation of variable rate Pix (mepiquat chloride) application by soil type. In: BELTWISE COTTON CONFERENCE, 1999, Memphis. **Proceedings...** Memphis: National Cotton Council of America, 1999. p. 524-526.

THURMAN, M. E.; HEINIGER, R. W. Using GPS to scout cotton for variable rate Pix (mepiquat chloride) application. In: BELTWISE COTTON CONFERENCE, 1998, San Diego. **Proceedings...** Memphis: National Cotton Council of America, 1998. p. 1499-1503.

TORRE NETO, A. **Rede de sensores sem fio e computação ubíqua na agropecuária**. São Carlos, SP: Embrapa Instrumentação Agropecuária, 2009. 18 p. (Embrapa Instrumentação Agropecuária. Boletim de pesquisa e desenvolvimento, 31). Disponível em: <<http://www.infoteca.cnptia.embrapa.br/handle/doc/658283>>. Acesso em: 22 set. 2014.

WOLF, S. A.; WOOD, S. D. Precision farming: environmental legitimation, commodification of information, and industrial coordination. **Rural Sociology**, Auburn, v. 62, n. 2, p. 180-206, June, 1997. DOI: 10.1111/j.1549-0831.1997.tb00650.x.



## **Automação de máquinas e implementos agrícolas: eletrônica embarcada, robótica e sistema de gestão de informação**

Rafael Vieira de Sousa  
Wellington Carlos Lopes  
Ricardo Yassushi Inamasu

### **1 Introdução**

O desenvolvimento e a aplicação de tecnologia no setor agropecuário tornou-se um dos elementos fundamentais para inserção ou manutenção da competitividade comercial no mercado globalizado de países com vocação para o agronegócio. Nesse sentido, o baixo nível de aplicação de tecnologia conduz à baixa eficiência e à baixa produtividade dos processos de produção vegetal e animal e pode colocar o país em uma posição de inferioridade nas relações comerciais internacionais, principalmente no caso do Brasil que tem ainda um grande potencial para desenvolvimento do agronegócio (FAO, 2009). Associado a isso, a escassez de mão de obra, as restrições para ampliação das áreas de cultivo e a busca pela sustentabilidade econômica, ambiental e social desse setor, também geram desafios importantes para a evolução tecnológica e têm demandado inúmeros esforços de pesquisa, desenvolvimento e inovação.

Uma das tendências atuais para superação das carências apontadas é a aplicação de sistemas de automação em áreas como, por exemplo, máquinas e implementos agrícolas, irrigação, criadouros, processamento, armazenamento e transporte de produtos agrícolas, construções rurais e ambiência. Também, junto a essa evolução tecnológica, surgem novos conceitos ou práticas que procuram maximizar os seus benefícios. Entre os conceitos atuais que buscam a automação como suporte para as práticas agropecuárias destaca-se a Agricultura de Precisão (AP).

O conceito de AP foi definido de várias formas ao longo dos anos. Seu significado pode ter distintas interpretações em função do foco de análise, tecnologia disponível e método adotado. Atualmente a Comissão Brasileira de Agricultura de Precisão (CBAP) do Ministério da Agricultura, Pecuária e Abastecimento (Mapa) define a Agricultura de Precisão como “um sistema de gerenciamento agrícola baseado na variação espacial e temporal da unidade produtiva e visa ao aumento de retorno econômico, à sustentabilidade e à minimização do efeito ao ambiente”.

A escala de amostragem das variáveis envolvidas e a precisão na execução das tarefas para as novas práticas, muitas vezes, são maiores que as exigidas pela agropecuária tradicional, demandando tecnologias de informação e comunicação que as viabilizem em custo e eficiência (STEINBERGER et al., 2009). A tendência é aumentar a sua frequência de amostragem com seu limite chegando ao tratamento por planta, o que já pode ser constatado em culturas perenes como em fruticultura. Entretanto, a escala amostral para cada fator específico deve ser definida

através de um balanço que leve em conta fatores como, por exemplo, custo de tempo e financeiro em contraposição a elevação da produtividade associada ao aumento de precisão. O controle de insetos e o monitoramento do clima para o gerenciamento das tomadas de decisões para o cultivo são exemplos de variáveis que são gerenciadas em larga escala. Outros fatores como fertilidade do solo e distribuições de doenças podem variar significativamente em nível de talhão durante a semeadura e a germinação. Por esta razão, é importante observar o gerenciamento agrícola em termos das unidades espacial e temporal para tomada de decisão. Estas escalas são vistas como um fluxo contínuo de informações das áreas de atuação, desde uma planta até a propriedade como um todo.

Entre as tecnologias que podem ser consideradas como o estado da arte para automação de máquinas e implementos agrícolas destacam-se: sensores que permitem aferir variáveis agrônomicas em campo através de sensoriamento local ou remoto; sistemas de aplicação de insumos em taxa variável e sistemas que realizam sensoriamento, processamento (tomada de decisão) e atuação durante o movimento da máquina (PEETS et al., 2012). Essas tecnologias possuem um ponto de convergência que é a eletrônica embarcada, ou seja, o conjunto de sistemas eletrônicos que possuem processadores (hardware) e programas dedicados (software ou firmware) para aquisição, processamento, armazenamento e comunicação de dados. O incremento acentuado de tais sistemas eletrônicos através do desenvolvimento de monitores e controladores para máquinas e implementos agrícolas demandou a padronização dessa eletrônica embarcada, que hoje é viabilizada através da implantação da norma ISO 11783, conhecida no mercado também como Isobus (AGRICULTURAL INDUSTRY ELECTRONICS FOUNDATION, 2014; INTERNATIONAL STANDARDIZATION ORGANIZATION, 2014). A padronização é fundamental para viabilizar a eletrônica embarcada em máquinas e implementos agrícolas na medida em que evita a duplicação de instalação, elimina obsolescência por compatibilidade, possibilita intercambiabilidade, reduz custo de manutenção, libera o agricultor de fornecedores exclusivos de sistemas comerciais e pode permitir a simplificação da integração de informações com sistemas computacionais externos às máquinas.

Um dos desafios relacionados ao uso da rede Isobus é a pesquisa e desenvolvimento de controladores automáticos compatíveis com tarefas agrícolas específicas e seus respectivos implementos tendo em vista a diversidade de tratos e implementos agrícolas existentes (BACKMAN et al., 2013). Esse fato é acentuado no Brasil, pois a agropecuária nacional possui características próprias e demandam máquinas, implementos e equipamentos adequados para a nossa realidade. Junto a isso, embora haja no Brasil o entendimento da importância desse padrão, os fabricantes de máquinas e equipamentos agrícolas, e, principalmente, os fabricantes nacionais de implementos deparam-se com barreiras técnicas por não possuírem, tradicionalmente, departamentos para desenvolvimento de eletrônica embarcada e pela carência no mercado de empresas fornecedoras de eletrônica própria para aplicações agrícolas. Essas barreiras impõem às empresas a necessidade de investimentos significativos para serem superadas, e, além disso, demandam-se investimentos para formação de mão-de-obra especializada (INAMASU et al., 2011).

Além das tecnologias descritas, frutos de pesquisas e de inovações desenvolvidas nas últimas duas décadas, atualmente outro foco de pesquisa recente e importante em tecnologias para agropecuária tem sido a busca por sistemas inteligentes que tornam viável a aplicação de Robôs Agrícolas Móveis. Nos últimos anos, trabalhos de pesquisa como, por exemplo, os apresentados por Bakker et al. (2011), Cariou et al. (2009), Cheein et al. (2011), Dong et al. (2011), Griepentrog et al.

(2013) e Slaughter et al. (2008) têm gerado soluções viáveis para o desenvolvimento de sistemas inteligentes de automação e de sistemas robóticos que possibilitam operações mais precisas para reduzir custos, aumentar a produtividade e minimizar o impacto ambiental em tarefas agrícolas.

Embora na última década muitos trabalhos de pesquisa tenham proposto arquiteturas diversas de controle para sistemas automáticos inteligentes e para robôs móveis, ainda há grandes desafios para a pesquisa de arquiteturas para sistemas que atuam em ambientes pouco estruturados. A imprevisibilidade natural, a variação temporal e espacial de certos parâmetros físicos como umidade, luminosidade e temperatura, principalmente nos ambientes abertos, elevam a imprecisão natural nas medições dos sensores conduzindo a elevados graus de incerteza, como relatado por diversos autores (GRIEPENTROG et al., 2013; PETERSON et al., 2011). Essas características adversas constituem desafios a serem superados pela pesquisa e desenvolvimento de sistemas inteligentes de automação e de arquiteturas robóticas. Além disso, a capacidade de aquisição, armazenamento e processamento de dados dos sistemas de um robô demandam tecnologias de informação e comunicação muito bem estruturadas e capazes de operar dados, em volume e diversidade, superiores aos das máquinas e implementos agrícolas convencionais.

A quantidade de projetos e publicações internacionais em automação e robótica para agropecuária ganhou volume expressivo nos últimos cinco anos e ainda é crescente. Revistas científicas internacionais como a *Computers and Electronics in Agriculture* e a *Biosystems Engineering*<sup>1</sup> tem entre os temas de seu escopo a automação e a robótica. A International Federation of Automatic Control (IFAC)<sup>2</sup>, que em seus 66 anos tem sido uma das mais importantes associações internacionais de promoção da ciência e da tecnologia em automação e controle, conta hoje com comitês técnicos específicos na temática de modelagem e controle para sistemas agropecuários e ambientais (TC8.1, TC8.2 e TC8.4), com destaque para a robótica, especialmente o comitê TC8.1, que promove a *Conference on Modelling and Control in Agriculture, Horticulture and Post Harvest Industry - Agricontrol*, e em sua quarta edição, em 2013 (AGRICONTROL 2013)<sup>3</sup>, contou com sessões em instrumentação, mecatrônica e controle para sistemas agropecuários, além de três sessões específicas em robótica nessa temática.

Identifica-se projetos internacionais que concentram esforços de diversos centros de pesquisa em robótica para a agricultura, como, por exemplo, o projeto *Robot Fleets for Highly Effective Agriculture and Forestry Management* (RHEA)<sup>4</sup> que conta com integrantes de centros de pesquisa de oito países diferentes e é apoiado por comissão europeia de pesquisa do *European Union's Seventh Framework Programme*. Outra iniciativa notável é o projeto ICT-AGRI financiado pelo programa ERA-NET da Comissão Europeia por meio do *7<sup>th</sup> Framework Programme for Research*<sup>5</sup> que tem o objetivo de promover a pesquisa em tecnologias de informação e de comunicação (TIC) e em robótica para a agricultura.

Entretanto, se o panorama internacional de pesquisa em automação e robótica para a agricultura tem se tornado um tema de destaque, a realidade brasileira da pesquisa em sistemas inteligentes

<sup>1</sup> Disponível em: <[www.journals.elsevier.com/biosystems-engineering](http://www.journals.elsevier.com/biosystems-engineering)>.

<sup>2</sup> Disponível em: <[www.ifac-control.org](http://www.ifac-control.org)>.

<sup>3</sup> Disponível em: <<http://agricontrol2013.automaatioseura.com>>.

<sup>4</sup> Disponível em: <[www.rhea-project.eu](http://www.rhea-project.eu)>.

<sup>5</sup> Disponível em: <<http://db-ictagri.eu/ict-agri/content/home.php>>.

de automação e em robótica para a agricultura é bastante limitada. Nesse contexto, estabeleceu-se um cenário negativo para a agropecuária e para o agronegócio nacional, a medida em que o Brasil como potência do agronegócio mundial carece de pesquisa, desenvolvimento e inovação em tecnologias de controle e automação para os sistemas de produção vegetal e animal.

Outro desafio importante é a pesquisa de metodologias e tecnologias de informação e comunicação para integração entre robôs móveis ou máquinas e implementos agrícolas, com eletrônica embarcada compatível com o Isobus (NIKKILÄ et al., 2010), e sistemas de interpretação e manipulação de quantidades massivas e diversas de dados coletados, que auxiliem na compreensão das causas da variabilidade espacial e temporal dos processos agropecuários. Pesquisas atuais têm buscado, através do mapeamento de processos agropecuários e dos respectivos fluxos de informações, propor arquiteturas para sistemas automáticos de coleta e comunicação de dados e sistema de controle baseados na norma ISO11783, assim como propor o respectivo modelo conceitual de integração entre tais sistemas automáticos e sistemas de gestão agropecuário (NIKKILÄ et al., 2010; SØRENSEN et al., 2010).

A indústria de manufatura possui soluções bem estabelecidas baseadas em sistemas de informações gerenciais (*Management Information System* - MIS) que são parte integrante do sistema de planejamento de recursos e gestão global de empresas (*Enterprise Resource Planning* - ERP) (KARNOUSKOS; COLOMBO, 2011). Entretanto, o desenvolvimento e a aplicação de conceitos e de metodologias similares no setor agropecuário ainda apresenta grandes desafios, principalmente devido à diversidade e à variabilidade dos processos de produção vegetal e animal (SØRENSEN et al., 2010). O projeto, o desenvolvimento e a implantação de um *Farm Management Information System* (FMIS), que pode ser traduzido por Sistema de Informação de Gestão Agropecuário, para processos de produção vegetal ou animal, devem contemplar várias entradas de dados, originadas de diferentes equipamentos em diferentes formatos, incluindo formatos que atendam à norma ISO11783. Além disso, é necessário oferecer suporte ao gerenciamento de informações geográficas e capacidade de processamento de modelos biológicos, capazes de relacionar parâmetros descritivos com informações de solo, atributos biológicos dos cultivos e parâmetros climatológicos (NIKKILÄ et al., 2010; SØRENSEN et al., 2010).

Diante do panorama apresentado, pode-se verificar que há demanda premente para a pesquisa e desenvolvimento de metodologias e tecnologias agrícolas aplicadas a aquisição automática de dados e controle inteligente de operações em processos de produção vegetal e animal, e que tenham interfaces automáticas para integração das informações desses sistemas com sistemas de gestão.

## 2 Sistemas embarcados

Na década de 1980 a eletrônica começou a ser embarcada em veículos, e nos EUA, Europa e Japão surgem os primeiros protótipos e padrões para sistemas de controle automático para máquinas e implementos agrícolas. Já na década de 1990, sistemas comerciais incorporam os receptores GNSS (Sistemas de Navegação Global por Satélites - *Global Navigation Satellite Systems*) e monitor gráfico na cabine do trator para realizar mapeamento de variáveis de desempenho da máquina e variáveis agrônomicas da lavoura. O GNSS é conhecido pelo nome do sistema norte americano *Geographic Positioning System* (GPS) por este ser o primeiro sistema GNSS

implementado. Nessa década, surgem os primeiros mapas de produtividade e as telas gráficas em máquinas agrícolas, com o diferencial que o sistema era programável por meio de barramento em rede de computador embarcado.

No final da década de 1990, no Brasil, as indústrias internacionais de tratores trouxeram a eletrônica embarcada em máquinas de grande porte como as grandes colhedoras, já com capacidade para realizar mapeamento da lavoura durante a operação, ou seja, apresentaram a eletrônica embarcada em máquina para geração de mapa georreferenciado de variável e identificar a variabilidade espacial, como, por exemplo, colhedoras com sistema de mapeamento da produção agrícola. Desde então, a pesquisa em tecnologias para veículos agrícolas e a busca por inovações para atender às necessidades das novas práticas agrícolas culminou em alguns produtos comerciais que, atualmente, constituem o estado da arte das tecnologias para automação de máquinas agrícolas. Dentre essas tecnologias destacam-se a Tecnologia de Aplicação à Taxa Variável (*Variable Rate Technology* - VRT), sistemas *On-The-Go* e Piloto Automático.

As tecnologias VRT são aplicadas a máquinas agrícolas principalmente para permitir a aplicação controlada de insumos (fertilização e pulverização) ou para permitir o controle de plantio (espaçamento e quantidade de sementes) de acordo com mapa de recomendação. Normalmente necessitam de controle automático de velocidade e receptor GNSS instalado na máquina agrícola para reconhecimento da coordenada geográfica de onde se localiza. Contam com um sistema computacional de apoio prévio para estudo e geração de mapa de recomendação.

As tecnologias *On-The-Go* realizam sensoriamento, processamento (tomada de decisão) e atuação durante o movimento da máquina. Sistemas baseados nessas tecnologias não necessitam do georreferenciamento para navegação, porém incluem uma unidade de comando dinâmico que determina a aplicação através de análises em tempo real das medidas de um sensor do solo ou cultura para cada lugar dentro do campo percorrido. Frequentemente demandam controle preciso de navegação, redução no tempo de resposta do sistema e sistema computacional para gerar a recomendação imediata ao sistema de aplicação (implemento).

Os pilotos automáticos ou sistema automáticos de guiagem são tecnologias aplicadas a máquinas agrícolas para permitir a navegação autônoma e mais precisa no campo. São aplicados normalmente para viabilização de trabalho noturno, plantio, aplicação precisa de insumos e suporte para as técnicas citadas nos destaques anteriores (VRT e *On-The-Go*).

Diversos sensores para aplicação de insumos *On-The-Go* têm sido pesquisados na última década, com destaque para os sensores óticos. Tais sensores permitem realizar leitura da cor do dossel das plantas, que é utilizada para inferir a quantidade de insumo requerido pela planta. Esses sensores normalmente possuem fonte de luz que ilumina o dossel por meio de semicondutor (LED - *light-emitting diode*) e a luz refletida é captada por sensores óticos.

A Figura 1 ilustra uma possível estrutura para um sistema de sensoriamento e aplicação de insumos *On-The-Go*. Na Figura 1 pode-se observar o fluxo de dados captados por um ou mais sensores que são posteriormente armazenados (memória ou banco de dados) e processados (controlador embarcado), gerando uma ação de controle para válvulas de aplicação de insumo (atuadores).

Atualmente estão disponíveis no mercado alguns modelos de sensores óticos com a finalidade de ler o estado da cultura em tempo real para controlar a aplicação de fertilizante nitrogenado ou controlar a aplicação localizada de pesticidas. Porém, os dados dos sensores devem ainda ser cui-



**Figura 1.** Sistema de sensoriamento e aplicação *On-The-Go*.

dadosamente testados e analisados antes de uma aplicação, pois há modelos em que o algoritmo de controle de aplicação de fertilizante nitrogenado está programado em seu sistema proprietário e os ajustes são ainda de acesso restrito. Os sensores que permitem a leitura de dados possuem interface serial com liberação de dados sequenciais em tempo regular, sendo que os protocolos de comunicação ainda não são estabelecidos em normas.

### 2.1 Eletrônica embarcada

Recentes tecnologias disponíveis no mercado foram concebidas para aquisição, armazenamento e transmissão de uma grande quantidade de dados que podem ser eletronicamente adquiridos na lavoura. Sensores agrícolas, dispositivos de processamento e armazenamento, e atuadores são implementados usando padrões de comunicação para troca de informações diferentes. A eletrônica embarcada em tratores e implementos não pode ser dissociada de sistema que inclua microprocessador e, conseqüentemente, de sistema eletrônico digital com hardware e software. Tal sistema computacional é caracterizado como computador embarcado e, normalmente, possui Unidade Central de Processamento (UCP), memórias e conjunto de elementos para entradas e saídas de dados. As entradas são associadas a elementos que fornecem os dados por meio de sensores ou de dispositivos de leitura e recebimento de dados. As saídas, por meio digital, podem comandar diversos dispositivos presentes em máquinas agrícolas, desde lâmpadas, mostradores, válvulas, e até o motor e a transmissão são considerados como atuadores.

Os computadores embarcados, no âmbito de tratores e implementos, são conhecidos como Electronic Control Unit (ECU) ou Unidade de Controle Eletrônico. Com a popularização dos processadores e da tecnologia de redes de computadores embarcadas em veículos, as ECU tem sido empregadas em quantidades diversas e distribuídas na máquina para que cada qual realize diferentes tarefas, como, por exemplo, o controle da comunicação com o operador, controle do motor e transmissão, leitura e registro de tarefas em formato de mapas, leitura do posicionamento, entre outras.

A tecnologia de rede empregada em máquinas agrícolas não é como a de escritórios. As redes utilizadas nos computadores domésticos em redes com cabo são conhecidas como Ethernet. Para veículos o padrão mais popular é o Controller Area Network (CAN), pois permite a criação de

redes robustas e mais simples. Esse padrão ou protocolo também tem sido adotado para promover a interconexão entre dispositivos de automação e controle em máquinas agrícolas e viabiliza a implantação de sistemas distribuídos. Tal protocolo foi adotado compondo o padrão internacional da International Organization for Standardization (ISO) para aplicações em máquinas e implementos agrícolas, denominado ISO11783, e também conhecido atualmente no mercado como Isobus.

Para suprir as carências tecnológicas destacadas anteriormente e apoiar as empresas nacionais, está em andamento, sob a coordenação da Embrapa Instrumentação, o projeto do Fundo para o Setor do Agronegócio, CT - Agro - MAPA 01/2008, subsidiado pela Financiadora de Estudos e Projetos (Finep), processo: 01.09.0285.00, sob o título “Sistemas para Viabilizar Monitoramentos e Intervenções Localizadas”. Esse projeto congrega, além de empresas, grupos de pesquisa da Universidade de São Paulo (USP) e do Centro de Tecnologia da Informação Renato Archer (CTI), e tem como objetivo geral fomentar, no mercado nacional, as tecnologias relacionadas com o Isobus para tratores e, principalmente, para implementos agrícolas. O projeto trata da integração das tecnologias, de certa forma estabelecidas, para monitoramento do solo e de plantas para desenvolvimento de tecnologias Isobus compatíveis para o mercado nacional. Além disso, os pesquisadores participam desde 2006, junto a membros representantes da indústria de máquinas e implementos agrícolas, da Força Tarefa Isobus Brasil (FTI Brasil) para promoção da norma ISO1783 e de tecnologias relacionadas. Participam também do Comitê Brasileiro de Tratores, Máquinas Agrícolas e Florestais (CB-203), da Comissão de Estudo de Comunicação e Eletrônica Embarcada (CE-203:019.02) e da Associação Brasileira de Normas Técnicas (ABNT), responsável pelo projeto da norma ABNT NBR ISO-11783 sob o título “Tratores e máquinas agrícolas e florestais - rede serial para controle e comunicação de dados”, versão Nacional da norma ISO 11783.

Em consonância com os objetivos de pesquisa e desenvolvimento, uma das etapas do projeto foi a concepção de Unidade Eletrônica de Controle com Isobus para viabilizar a construção de sistema de automação padronizado para máquinas e implementos agrícolas (Isobus ECU ou I-ECU), incluindo também aplicação em robôs agrícolas móveis. O trabalho de pesquisa e desenvolvimento foi dividido nas seguintes etapas principais:

- Pesquisa dos elementos mínimos para projeto de sistemas embarcados ISO11783 compatíveis.
- Estudo e definição dos requisitos mínimos de sistemas mecatrônicos de robôs agrícolas móveis.
- Levantamento dos requisitos funcionais de processos comuns das práticas de AP.
- Levantamento das mensagens definidas na norma ISO11783 que atendam aos requisitos identificados em “b” e “c”.
- Especificação e projeto de unidade eletrônica de controle universal para implemento agrícola (I-ECU).
- Definição de requisitos de interface para sistema de gestão agrícola (FMIS) aplicado às práticas de AP.
- Modelagem conceitual de integração de informações para sistemas de coleta e comunicação automática de dados de máquinas agrícolas automatizadas (trator e implemento automatizado ou robô agrícola móvel).

As etapas “a” e “d” foram realizadas a partir da pesquisa de trabalhos publicados, do estudo de sistemas de aquisição e comunicação de dados e do estudo da norma ISO11783. As etapas “b” e “c” foram realizadas da mesma forma, porém incluíram levantamentos em campo por meio de visitas técnicas a cooperativas e unidades produtoras, em especial para execução da etapa “c”.

Na etapa “b” observou-se que os principais dispositivos especiais utilizados no estágio de percepção para guiagem e navegação de robôs agrícolas móveis são: câmeras; receptor GNSS, varredor ou digitalizador Laser (*Scanner Laser*); e sistema de navegação inercial ou Inertial Navigation System (INS).

A definição do FMIS na etapa “f” foi realizada a partir da pesquisa de trabalhos publicados e das informações obtidas nas etapas de “a” a “d”. A modelagem proposta na etapa “g” foi realizada a partir da sistematização de informações e da estruturação de um modelo que atenda a todos os requisitos e características dos fluxos de informação de cada subsistema.

Para a especificação e projeto da I-ECU utilizou-se as informações obtidas nos levantamentos e especificações das etapas “a” a “d”. As especificações preliminares da I-ECU orientou a aquisição do conjunto de desenvolvimento de software e de hardware i.MX28 EVK da empresa Freescale, que foi utilizado para familiarização com o processador escolhido e para definir o conjunto apropriado firmware-hardware através de implementação de testes.

A especificação da I-ECU também foi embasada em características de software e de hardware típicos de sistemas embarcados para automação com destaque para: quantidade e tipos de entradas e saídas para sensores e atuadores; quantidade e tipos de interfaces de comunicação; capacidade de processamento; necessidades do sistema operacional; capacidade de armazenamento de dados; fonte de alimentação e requisitos de consumo de energia; e características físicas do ambiente de operação (ex. temperatura, umidade, vibração e ruídos eletromagnéticos). Considerando as características citadas, os levantamentos e especificações das etapas de “a” a “d” e os resultados dos testes com o conjunto i.MX28 EVK definiu-se um esquemático básico para o circuito eletrônico da I-ECU ilustrada na Figura 2.

Como ilustrado na Figura 2, o circuito eletrônico da I-ECU baseia-se no processador i.MX287 da empresa Freescale tecnologia Advanced RISC Machine (ARM) que está conectado a quatro tipos de memórias diferentes: memória tipo NAND (4 Gb), utilizada para armazenar o sistema operacional embarcado e o firmware das aplicações ISO11783 compiladas; memória tipo DDR2 SDRAM (64 MB x 16) para armazenamento de dados temporários e operações realizadas pelo processador; memória tipo EEPROM SPI (256 kb) com interface de comunicação padrão Serial Peripheral Interface (SPI) para armazenamento de quantidade restrita de dados; e conector (*Slot*) para cartão de memória padrão SD/MMC, que pode ser inserido e removido para armazenamento e transferência de quantidade massiva de dados, inclusive para o carregamento do sistema operacional e de aplicações ISO 11783.

O hardware da I-ECU ilustrado na Figura 2 contempla as seguintes interfaces de comunicação: porta Ethernet, porta padrão USB e porta padrão RS232 para comunicação com dispositivos eletrônicos especiais que utilizam dados e parâmetros mais elaborados (maior banda), como, por exemplo, aqueles utilizados em robôs móveis (câmera, varredor laser, sistema inercial e receptor GNSS); porta padrão RS232C, utilizada como interface em computador pessoal para gravação e depuração do firmware; e portas padrão CAN que viabiliza implementação física de rede

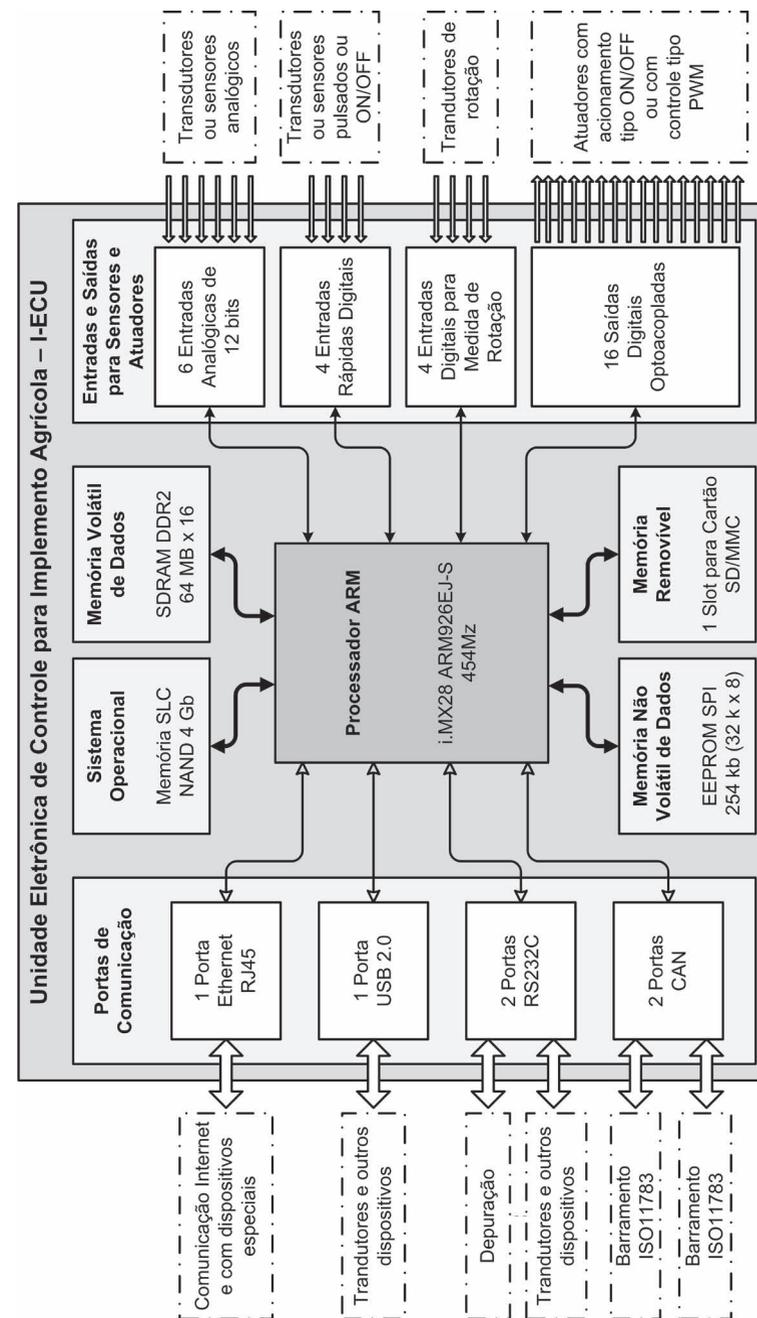


Figura 2. Diagrama funcional do circuito eletrônico da I-ECU.

ISO11783 (camada física). O hardware da I-ECU também contém entradas e saídas para sensores e atuadores com padrões comuns de mercado.

## 2.2 Norma Iso11783 - Isobus

No período em que a eletrônica embarcada começa a ser difundida em produtos comerciais nos EUA e na Europa, os agricultores passaram a demandar compatibilidade entre formato de arquivos e entre dispositivos eletrônicos. Dessa forma, surge a semente para a padronização internacional de comunicação eletrônica de dados e de sistemas de controle entre máquinas e implementos.

Atualmente, instituições de pesquisa, empresas e associações de normas concentram esforços significativos para padronização da eletrônica embarcada em máquinas e implementos agrícolas através da implantação da norma ISO11783. Observa-se também o esforço por parte de fabricantes de máquinas, implementos e de outros equipamentos para a promoção e para a implantação do Isobus. Como resultado dessas parcerias entre fabricantes, instituições de pesquisa e associações de normas, destaca-se a Agricultural Industry Electronics Foundation (AEF), uma organização internacional independente que tem o objetivo de promover não só a divulgação do padrão Isobus em diversos meios, como também criar e disponibilizar procedimentos e ferramentas para apoiar o desenvolvimento e a certificação de máquinas e equipamentos com essa tecnologia.

A norma ISO11783 tem como objetivo viabilizar a interconexão padronizada de sistemas eletrônicos embarcáveis através de um barramento de rede digital e permitir a comunicação de dados entre estes dispositivos sensores, atuadores e controladores nas máquinas e implementos agrícolas, assim como padronizar um formato de dados para interconexão com sistema externo à máquina. Como definido na norma, uma ECU pode ser responsável pela conexão de um ou mais dispositivos a um barramento de comunicação. Além disso, um dispositivo pode ser conectado a um barramento por uma ou mais ECU. O conjunto formado por uma ECU e por dispositivos conectados constitui um nó de rede padrão CAN. A ISO11783 adota o protocolo CAN como enlace de comunicação por fio e algumas especificações da norma foram baseadas na norma DIN9684 e muitas outras foram baseadas na norma SAE J1939. A documentação atual da norma ISO11783 é composta por 14 partes publicadas como padrões internacionais - International Standard (IS), elencados a seguir.

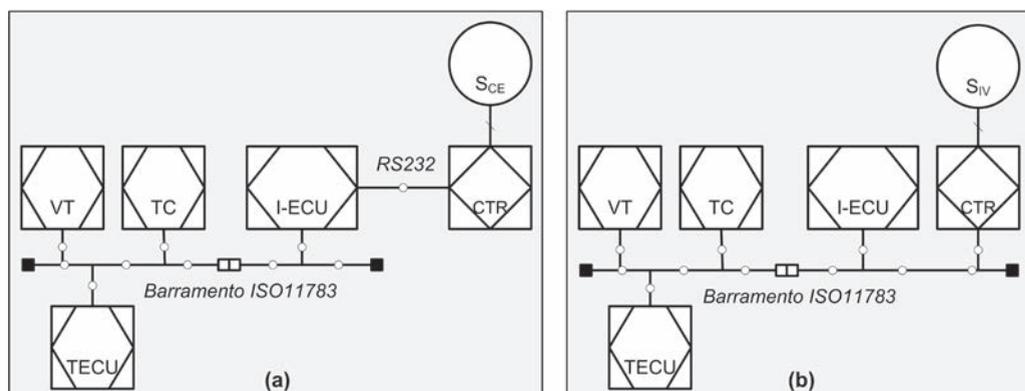
- Parte 1: Padrão Geral para Comunicação de Dados Móvel
- Parte 2: Camada Física
- Parte 3: Camada de Enlace de Dados
- Parte 4: Camada de Rede
- Parte 5: Gerenciamento de Rede
- Parte 6: Terminal Virtual
- Parte 7: Camada de Aplicação de Mensagens do Implemento
- Parte 8: Mensagens de Tomada de Potência
- Parte 9: ECU do Trator
- Parte 10: Controlador de Tarefa e Gerenciamento da Informação
- Parte 11: Dicionário de Elementos de Dados Móveis

- Parte 12: Serviços de Diagnósticos
- Parte 13: Servidor de Arquivos
- Parte 14: Controle de Sequência

Os sistemas projetados e desenvolvidos no padrão Isobus permitem uma maior conectividade entre máquinas e implementos agrícolas, independentemente dos fabricantes, além de evitar a redundância de sistemas eletrônicos (sensores, atuadores e terminais) instalados nas máquinas. Uma máquina agrícola que possua uma rede básica de comunicação no padrão ISO11783 é composta normalmente pelas seguintes unidades: Terminal Virtual (VT - *Virtual Terminal*) denominado pela AEF como Terminal Universal (UT - *Universal Terminal*); Unidade de Controle Eletrônico do Trator (TECU - *Tractor ECU*); Controlador de Tarefas (TC - *Task Controller*), e Unidade de Controle Eletrônico do Implemento (*Implement ECU*). O receptor GNSS não é obrigatório, mas sua utilização é definida na norma em compatibilidade com outra norma, a NMEA2000.

A Figura 3 apresenta um exemplo de conexão entre implementos com sensores *on-the-go* e uma máquina (trator) no padrão Isobus. Nessa Figura 3 utiliza-se a simbologia definida pela Norma ISA-S5 International Society of Automation (ISA) (em português, Sociedade Internacional de Automação) para compor o diagrama esquemático no qual a I-ECU descrita e ilustrada anteriormente na Figura 2 é utilizada como interface para um Sensor de Condutividade Elétrica do Solo ( $S_{CE}$ ) e um Sensor de Índice de Vegetação ( $S_{IV}$ ). Essas estruturas foram utilizadas no projeto mencionado anteriormente e os sensores comerciais empregados foram: sensor de condutividade elétrica do solo, desenvolvido pela Embrapa Instrumentação (RABELLO et al., 2011), e sensor ativo Crop Circle ASC 470, da empresa Holland Scientific, que emite e capta o espectro de luz refletida em comprimentos de onda específicos pela planta, calculando índices relacionados com o seu estado nutricional. Os sensores utilizados no sistema automático de aquisição possuem dispositivo controlador de registro e conversão de dados (CTR) que dispõem de porta para transferência de dados baseada no padrão RS232C. O CTR de condutividade elétrica foi integrado à I-ECU através de sua porta RS232C, porém o CTR de índice de vegetação foi integrado à I-ECU através de interface CAN, tendo em vista a aplicação de mais de um sensor na mesma operação de mapeamento em campo (rede CAN permite diversos nós conectados à I-ECU).

Na Figura 3 observa-se que a I-ECU comunica-se através do barramento ISO11783 com VT, TECU e TC. O VT especificado na Parte 6 da norma é uma ECU com tela gráfica que promove a interface entre a máquina (trator e implemento) e o operador (ser humano), e fica localizada na cabine de operação da máquina agrícola. Quando um implemento Isobus é conectado em um trator Isobus, uma ECU de implemento qualificada como Mestre do Grupo de ECU (*Working Set Master*) fica responsável por interconectar todos os dispositivos do implemento e demais ECUs relacionadas com uma aplicação agrícola ao VT através de um arquivo padrão denominado *Object Pool* (OP), transmitido do implemento (I-ECU) para o trator (VT) pela rede. O Object Pool é um arquivo binário que contém a descrição de um conjunto de objetos definidos pela norma (máscaras, botões, variáveis, indicadores gráficos, figuras etc) que são exibidos no VT para representar graficamente as funções e parâmetros da ECU do implemento responsável pelo comando da aplicação agrícola. O protocolo de tratamento da interface gráfica é organizado de maneira orientada a objeto (*object-oriented*). As informações das ECUs conectadas ao barramento são mostradas na tela física do VT e são definidas pelo elementos *data mask* (máscara de



**Figura 3.** Sistemas de coleta e comunicação automática de dados baseado na ISO11783: (a) medição de condutividade elétrica; (b) medição de índice de vegetação.

dados), *alarm mask* (máscara de alarme) e *soft-key masks* (máscaras soft-key). Os dados dessas máscaras estão contidos em objetos de definições que em conjunto compõem o *Object Pool*.

A Tecu especificada na Parte 9 é uma Unidade de Interconexão de Rede (Network Interconnection Unit - NIU). As NIUs podem ser do tipo *Repeater* (Repetidor), *Brige* (Ponte), *Router* (Roteador) ou *Gateway* (Portal), com funções semelhantes às funções destes elementos nas redes de computadores pessoais convencionais. A Tecu é, portanto, uma NIU com funções especiais e faz a interconexão do Barramento do Trator com o Barramento de Implemento. O Barramento do Trator é a sub-rede onde podem estar conectadas ECU para parâmetros veiculares associados, por exemplo, com o motor e com a transmissão. O outro tipo é a sub-rede principal de implementos, denominada Barramento de Implemento, onde podem estar conectadas ECUs para parâmetros relacionados com aplicações agrícolas, como, por exemplo, plantio ou pulverização. A Tecu deve inicializar no barramento do implemento do mesmo modo que qualquer outra ECU na rede, e o seu acesso ao VT é idêntico ao de qualquer outro implemento e dispositivo, porém existem três classe possíveis de Tecu quanto às funções: Classe 1, Classe 2 e Classe 3.

A Tecu Classe 1 possui uma interface simples com suporte para rede e fornece principalmente medições internas do trator com as mesma informações definidas na norma ISO 11786, mais antiga que a ISO11783, comum em alguns tratores. Além disso, suporta o gerenciamento de energia, armazenamento do idioma de parâmetros específicos e permite o controle da iluminação do trator. A Tecu Classe 2 possui as funções de uma Tecu Classe 1 e tem funções adicionais com dados, tais como: hora e data, distância e direção baseados no solo e roda, informações da tomada de potência mecânica (PTO), todas as mensagens da iluminação do implemento e estados de válvulas auxiliares. Isto permite um controle mais sofisticado do implemento e uma estratégia de segurança. A Tecu Classe 3 é a mais completa e possui as funções das Tecu Classe 1 e Classe 2 e, além disso, esta categoria de Tecu aceita comandos do barramento do implemento. Em particular, os comandos básicos para engate traseiro, desacoplamento da potência e o processamento do controle de válvulas auxiliares. Portanto, o implemento está hábil a controlar a potência desejada na fonte e a posição do engate.

O TC especificado na Parte 10 é uma ECU especial responsável pelo envio, recebimento e armazenamento de dados de processos agrícolas, e, além disso, contém a interface de dados para

sistemas computacionais externos à máquina agrícola. Essa ECU é responsável por gerenciar o controle de tarefas pré-determinadas, sendo que as tarefas são prescritas previamente, e são transmitidas para o trator no formato de arquivo XML (Extensible Markup Language). Para que as tarefas possam ser realizadas, um implemento Isobus compatível deve enviar as características de seus dispositivos ao TC utilizando um arquivo denominado Device Description Object Pool (DDOP), também no formato XML. Essa ECU não possui um tela gráfica e por isso pode apresentar as tarefas e as funções relacionadas no VT.

### 3 Modelo conceitual de integração de informação

A interpretação da enorme quantidade de dados coletados, a compreensão das causas da variabilidade e a proposta de estratégias sólidas para a gestão de sistemas agropecuários e agroindustriais têm demandado sistemas de informação específicos para a agropecuária, que possam melhorar as estratégias de manejo da lavoura e animal, contribuir com a gestão do agronegócio, viabilizar a regulamentação governamental e padronizar os processos que viabilizam a rastreabilidade e a certificação de produtos agropecuários. Em particular para a AP, dentre as diversas abordagens existentes relacionadas ao tema, é comum encontrar citações referentes à coleta massiva de dados, sistemas georreferenciados, geração de mapas e aplicação de insumos à taxa variável. Entretanto, essas operações quando realizadas de forma isoladas, ou sem um gerenciamento adequado, não providenciam ganhos efetivos ou, geram informações desconstruídas.

O projeto, desenvolvimento e implantação de um sistema de informação de gestão agrícola ou FMIS para processos de produção agrícola deve contemplar várias entradas de dados, originadas de diferentes equipamentos em diferentes formatos, incluindo formatos que atendam à norma ISO11783. Além disso, principalmente no contexto da AP, é necessário oferecer suporte ao gerenciamento de informações geográficas e capacidade de processamento de modelos biológicos, capazes de relacionar parâmetros descritivos com informações de solo, atributos biológicos dos cultivos e parâmetros climatológicos. Diante desse panorama, confrontando os requisitos mínimos de sistemas automação e os requisitos funcionais de processos comuns das práticas de AP com as mensagens definidas na norma ISO11783, determina-se o grupo de mensagens explicitados na Tabela 1.

**Tabela 1.** Grupo de mensagens com informações comunicadas entre um sistema automatizado embarcado e um FMIS.

Tipo de serviço	Documento da norma	Parâmetro
Operação trator-implemento	ISO11783-7	Navegação, tomada de potência, engates traseiro e dianteiro, supervisão e controle processos, comando de atuadores do implemento e leitura sensores do implemento.
Dados automotivos	ISO11783-8	Motor, transmissão, freios, caixa de marcha e sistema de combustível
Operação trator	ISO11783-9	Sistema hidráulico do implemento, tomada de potência, gerenciamento de potência e navegação (velocidade e distância baseado nas rodas e solo).
Diagnóstico de sistemas	ISO11783-12	Protocolos para diagnóstico e identificação de ECUs, versão de softwares e certificação.

Os grupos de mensagens descritos na Tabela 1 estão disponíveis pelo TC de um sistema Isobus compatível (trator, colhedora ou robô móvel) e são fundamentais para alimentar um FMIS com as informações que atendam às necessidades citadas no contexto da AP. Esses grupos contêm informações tanto para gerenciamento da operação da máquina agrícola e seu implemento, como também para gerenciamento da operação ou tarefa agrícola em cada aplicação.

Os grupos de mensagens definidos na Tabela 1 contem parâmetros a serem comunicados pelo TC entre um sistema automatizado de um conjunto trator-implemento com um FMIS. Para que as tarefas possam ser realizadas, os sistemas ISO11783 compatíveis enviam as características de seus dispositivos para o TC utilizando o arquivo DDOP. Entretanto, a norma ISO11783 não estabelece um meio específico para a transferência de dados entre a máquina agrícola e um FMIS, mas estabelece no documento ISO 11783-10 uma padronização baseada na linguagem de marcação XML que permite a criação de documentos com dados organizados hierarquicamente, tais como, textos, banco de dados ou desenhos vetoriais. São definidos no documento formatos de dados, parâmetros necessários para o controle de processos agrícolas e formatos de mensagens para comunicação entre um FMIS e o TC.

A Figura 4 mostra o modelo conceitual para a arquitetura de integração entre um sistema embarcado automatizado, baseado na norma ISO11783, e o FMIS, por meio de um terminal de acesso implantado em computador pessoal.

Na Figura 4 observa-se que a I-ECU descrita em item anterior e ilustrada na Figura 2 comunica-se através do barramento ISO11783 com o VT, com o TECU e com o TC. A I-ECU conectada à rede ISO11783 é qualificada como Mestre do Grupo, sendo responsável por interconectar todos os dispositivos do implemento e demais ECUs relacionadas à aplicação agrícola com o VT através de arquivo tipo OP trafegado na rede. De forma similar a I-ECU comunica-se com o TC através do DDOP tanto para executar alguma tarefa (comando de aplicação), assim como para comunicar variáveis agrônômicas ou dos sistemas de aplicação que forem registradas ao longo da execução da tarefa.

Os dados referentes a processos agrícolas, seguindo o fluxo ilustrado na Figura 4, podem ser transferidos para o conjunto máquina-implemento utilizando identificadores próprios para mensagens de dados de processo especificados no Dicionário de Dados (base de dados denominado ISO 11783-11)<sup>6</sup>. As mensagens descritas na Tabela 1 podem ser utilizadas para a transmissão de dados medidos ou comandos para um ou mais controladores associado à ECU, sendo que a identidade dos dados é determinada por uma entrada do Dicionário de Dados.

Como ilustrado na Figura 4, o FMIS deve possuir um elemento de conversão e comunicação de dados para acesso a serviços providos pelo FMIS. Entre as funcionalidade do FMIS estão incluídas o acesso a bancos de dados geográficos e agrícolas que armazenam as informações referentes a aquisição automática dos dados colocados na Tabela 1. O uso dessas informações dentro do processo de gestão agrícola pode trazer resultados positivos e bastante significativos, como, por exemplo: redução do consumo de combustível, redução dos custos para manutenção das máquinas, redução do tempo da máquina parada (no campo ou no pátio), melhoria na logística de distribuição das máquinas em campo e no processo de logística, monitoramento e controle mais preciso de operações agrícolas, entre outros. Além disso, é possível acompanhar a eficiência da

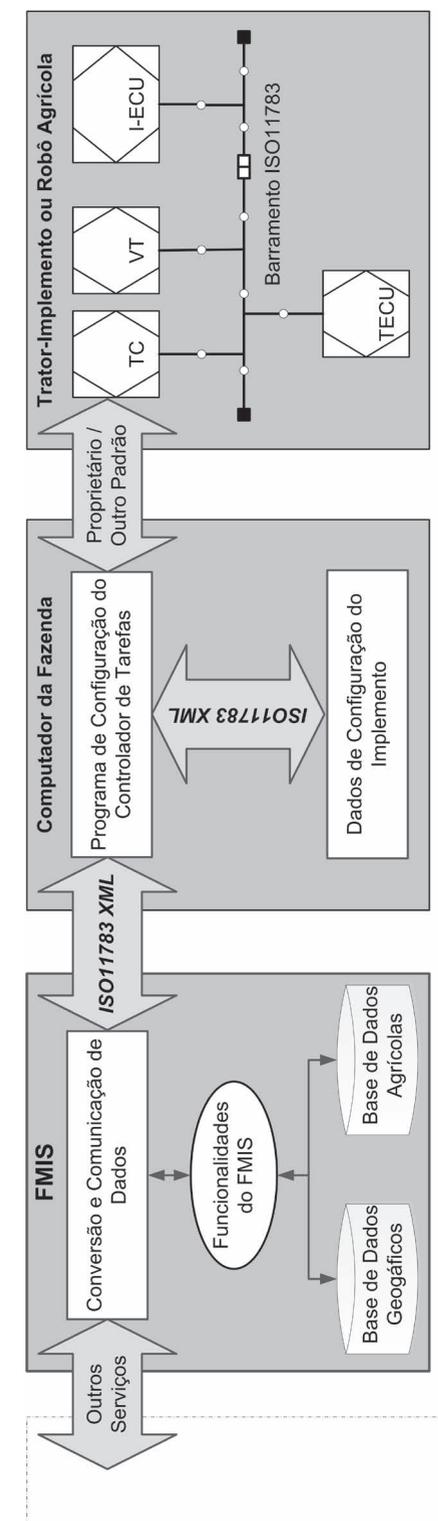


Figura 4. Arquitetura de integração entre um sistema embarcado ISO11783 e um FMIS: elementos, interfaces, interfaces e fluxo de dados.

<sup>6</sup> Disponível em: <<http://dictionary.Isobus.net/Isobus>>.

aplicação, ou seja, acompanhar parâmetros de execução de cada tarefa, como, por exemplo, valores efetivamente aplicados de insumos, valores de sementes efetivamente plantadas, registros de falhas em atuadores, entre outros.

Num sentido mais amplo, não há um limite físico do número de elementos de automação que podem ser implantados em uma máquina agrícola, seja de sensores, atuadores ou controladores, e sendo assim, desde que tais elementos sejam devidamente integrados à estrutura, como prescreve a norma ISO11783, o número de parâmetros, variáveis, comandos ou informações comunicadas com o sistema de gestão está limitado à necessidade ou à criatividade da engenharia dos sistemas agrícolas.

## 4 Considerações finais

A multidisciplinaridade envolvida no setor agrícola atual faz com que diversos métodos, baseados em distintas áreas da ciência, sejam desenvolvidos visando a uma mesma aplicação. Da mesma forma, uma determinada tecnologia não fica restrita a melhorar uma única prática, mas é empregada em múltiplas situações, auxiliando no processo como um todo. Devido ao processo contínuo de evolução das áreas que envolvem ou interagem com a instrumentação eletrônica, automação e robótica, tais como, microeletrônica, sensores, computadores, telecomunicações, tecnologia da informação e da comunicação, o potencial de inovações foi ainda muito pouco explorado.

Dentre as tecnologias que apresentam forte pressão de demanda à mecanização e automação para a agropecuária, o foco ainda é pelo aumento da produtividade através da melhora ou da viabilização de processos tais como: medidas massivas de propriedades físicas e químicas do solo; acompanhamento da produção vegetal e volume de biomassa; medição e acompanhamento de disponibilidade de água; identificação e monitoramento de pragas (doenças, plantas invasoras, insetos), e identificação e monitoramento da saúde e nutrição animal. As tecnologias que tem sido adaptadas, desenvolvidas e exploradas na agricultura para atender a essas demandas, seja através do sensoriamento local ou remoto ou para a aplicação, carecem de modelos de referência e padronização, primeiramente, para viabilizar aquisição, comunicação, integração e manipulação de dados identificados ou georreferenciados, e, posteriormente, viabilizar a transformação desses dados em informações que relacionem variáveis ou parâmetros diversos, como, por exemplo, dados de solo, atributos biológicos dos cultivos e parâmetros climatológicos, para suporte à tomada de decisão.

A pesquisa e o desenvolvimento de tais modelos de referência e padrões apresentam caráter multidisciplinar e necessitam da promoção de agrupamento de competências e habilidades nos temas e áreas com interfaces comuns para geração e integração de novos conhecimentos e informações. Além disso, se há um grande espaço para melhoria e inserção de tecnologias, existe um outro gargalo no momento que é a formação de profissionais qualificados para pesquisa, desenvolvimento e aplicação dessas tecnologias. Essa qualificação deve ser orientada tanto ao nível técnico como ao nível acadêmico, e deve ser promovida tanto por instituições públicas como privadas. Dessa forma, essas duas vertentes, criação de tecnologia e desenvolvimento, competências e habilidades, são fundamentais para o fortalecimento e a sustentabilidade da agropecuária e do agronegócio no Brasil.

## 5 Referências

- AGRICULTURAL INDUSTRY ELECTRONICS FOUNDATION. 2014. Disponível em: <<http://www.aef-online.org/>>. Acesso em: 24 set. 2014.
- BACKMAN, J.; OKSANEN, T.; VISALA, A. Applicability of the ISO 11783 network in a distributed combined guidance system for agricultural machines. **Biosystems Engineering**, London, v. 114, n. 3, p. 306-317, Mar. 2013. DOI: 10.1016/j.biosystemseng.2012.12.017.
- BAKKER, T.; ASSELT, K. VAN; BONTSEMA, J.; MULLER, J.; STRATEN, G. van. Autonomous navigation using a robot platform in a sugar beet field. **Biosystems Engineering**, London, v. 109, n. 4, p. 357-368, Aug. 2011. DOI: 10.1016/j.biosystemseng.2011.05.001.
- CARIOU, C.; LENAIN, R.; THUILOT, B.; BERDUCAT, M. Automatic guidance of a four-wheel-steering mobile robot for accurate field operations. **Journal of Field Robotics**, Hoboken, v. 26, n. 6-7, p. 504-518, July, 2009. DOI:10.1002/rob.20282.
- CHEEIN, A. F.; STEINER, G.; PAINA, G. P.; CARELLI, R. Optimized EIF-SLAM algorithm for precision agriculture mapping based on systems detection. **Computers and Electronics in Agriculture**, New York, v. 78, n. 2, p. 195-207, Sept. 2011. DOI:10.1016/j.compag.2011.07.007.
- DONG, F.; HEINEMANN, W.; KASPER, R. Development of a row guidance system for an autonomous robot for white asparagus harvesting. **Computers and Electronics in Agriculture**, New York, v. 79, n. 2, p. 216-225, Nov. 2011. DOI: 10.1016/j.compag.2011.10.002.
- FAO. **How to feed the world in 2050**. Rome, 2009. Disponível em: <<http://www.fao.org/wsfs/forum2050/wsfs-background-documents/issues-briefs/en/>>. Acesso em: 24 set. 2014.
- GRIEPENTROG, H. W.; DÜHRING, J. C. L.; PARAFOROS, D. S. Robots for field operations with comprehensive multilayer control. **KI - Künstliche Intelligenz**, Heidelberg, v. 27, n. 4, p. 325-333, June, 2013. DOI 10.1007/s13218-013-0266-z
- INAMASU, R. Y.; NAIME, J. de M.; RESENDE A. V.; BASSOI, L. H.; BERNARDI, A. C. (Ed.). **Agricultura de precisão: um novo olhar**. São Carlos, SP: Embrapa Instrumentação. 2011. 334 p. il.
- INTERNATIONAL STANDARDIZATION ORGANIZATION. **ISO11783**: tractors and machinery for agriculture and forestry - serial control and communications data network. Disponível em: <<http://www.iso.org/>>. Acesso em: 24 set. 2014.
- KARNOUSKOS, S.; COLOMBO, A. W. Architecting the next generation of service-based SCADA/DCS system of systems. In: ANNUAL CONFERENCE ON IEEE INDUSTRIAL ELECTRONICS SOCIETY, 37., 2011, Karlsruhe. p. 359-364.
- NIKKILÄ, R.; SEILONEN, I.; KOSKINEN, K. Software architecture for farm management information systems in precision agriculture. **Computers and Electronics in Agriculture**, New York, v. 70, n. 2, p. 328-336, Mar. 2010. DOI: 10.1016/j.compag.2009.08.013.
- PEETS, S.; MOUAZEN, A. M.; BLACKBURN, K.; KUANG, B.; WIEBENSOHN, J. Methods and procedures for automatic collection and management of data acquired from on-the-go sensors with application to on-the-go soil sensors. **Computers and Electronics in Agriculture**, New York, v. 81, n. 1, p. 104-112, Feb. 2012. DOI: 10.1016/j.compag.2011.11.011.
- PETERSON, G. L.; DUFFY J. P.; HOOPER, D. J. Dynamic behavior sequencing for hybrid robot architectures. **Journal of Intelligent & Robotic Systems**, Dordrecht, v. 64, n. 2, p. 179-196, Nov. 2011. DOI 10.1007/s10846-010-9535-3

RABELLO, L. M. ; INAMASU, R. Y. ; BERNARDI, A. C. de C. ; NAIME, J. M. ; MOLIN, J. P. Mapeamento da condutividade elétrica do solo - sistema protótipo. In: INAMASU, R. Y.; NAIME, J. de M.; RESENDE, A. V. de; BASSOI, L. H.; BERNARDI, A. de C. (Ed.). **Agricultura de precisão: um novo olhar**. São Carlos, SP: Embrapa Instrumentação, 2011. p. 41-45.

SLAUGHTER, D. C.; GILES, D. K.; DOWNEY, D. Autonomous robotic weed control systems: a review. **Computers and Electronics in Agriculture**, New York, v. 61, n. 1, p. 63-78, Apr. 2008. DOI: 10.1016/j.compag.2007.05.008.

SØRENSEN, C. G.; FOUNTAS, S.; NASH, E; PESONEN, L; BOCHTIS, D.; PEDERSEN, S. M; BASSO, B.; BLACKMORE, S. B. Conceptual model of a future farm management information system. **Computers and Electronics in Agriculture**, New York, v. 72, n. 1, p. 37-47, June 2010. DOI: 10.1016/j.compag.2010.02.003.

STEINBERGER, G.; ROTHMUND, M.; AUERNHAMMER, H. Mobile farm equipment as a data source in an agricultural service architecture. **Computers and Electronics in Agriculture**, New York, v. 65, n. 2, p. 238-246, Mar. 2009. DOI: 10.1016/j.compag.2008.10.005.



# Parte V

## TIC na construção das aplicações para a agricultura



## Gestão da informação e do conhecimento

Ivo Pierozzi Júnior  
Marcia Izabel Fugisawa Souza  
Tércia Zavaglia Torres  
Leandro Henrique Mendonça de Oliveira  
Leonardo Ribeiro Queiros

### 1 Introdução

Toda sociedade humana, argumenta Aranha Filho (1995), diante de seu patrimônio tecnológico e cultural, desenvolve necessariamente uma classe fundamental de tecnologias das quais não se pode prescindir - são as chamadas tecnologias de conhecimento ou tecnologias cognitivas. Tais tecnologias tanto trazem as marcas quanto também demarcam o caráter de uma cultura, e - diferentemente daquelas de cunho prático e utilitário (tecnologias, ferramentas e métodos voltados à transformação da natureza) -, compõem uma classe especial de tecnologias que visam à gestão de conhecimentos produzidos por uma sociedade, ou seja, dedicam-se a gerir a sua memória coletiva. Na sociedade contemporânea, as tecnologias que exercem a função social de tecer um novo padrão de gestão do conhecimento são as tecnologias da informação e comunicação (TIC).

As TIC se consolidam no ciberespaço que se configura como uma nova forma de organização social. O surgimento desta nova ordem social foi proporcionado pela evolução dos computadores individuais e pelo advento da rede internet que ofereceu às pessoas a possibilidade de se conectarem entre si, mesmo estando geograficamente distantes. Ciberespaço é, pois, um espaço comunicacional híbrido, composto por diversas mídias<sup>1</sup> onde todo e qualquer signo pode ser socialmente produzido, significado e sentido (SANTOS; OKADA, 2003).

Esta realidade fornece pista para a compreensão do porquê rede é a metáfora escolhida para representação do todo social (ARANHA FILHO, 1995). Na rede, aqui entendida como fluxo informacional estabelecido pelas pessoas conectadas por meio da internet, as tecnologias de conhecimento são os meios pelos quais informação e conhecimento circulam livremente. Rede é espaço topográfico, caminho, movimento, atividade que une as pessoas umas às outras, levando-as a mergulhar, navegar, pular entre as informações que lhes interessam. Pode-se afirmar que o conceito de rede é hoje uma espécie de chave-mestra que vincula três níveis de significação - a do ser, a da dinâmica da interação coletiva e a do sistema complexo. Rede é meta-ligação porque

<sup>1</sup> O conceito de mídia, neste contexto, está sendo adotado como qualquer meio que favorece e concretiza o processo de comunicação produzido pelas pessoas. Entende-se que a mídia possibilita a análise não só das questões da linguagem e das estratégias de construção de produtos culturais, mas também a compreensão da forma como as pessoas produzem os processos de comunicação e informação quando usam os meios midiáticos.

é “ao mesmo tempo o vínculo de um elemento com um todo, o vínculo entre diversos estados de um todo e o vínculo da estrutura de um todo com o funcionamento de um outro.” (MUSSO, 2004, p. 32-33).

Em rede, ligados e passando pelas informações, por entre as diversas mídias digitais, reconfigura-se a forma de ser e agir no mundo. Aranha Filho (1995) argumenta que, o que se ‘passa’ na rede não pode ser entendido unicamente como informação - pacotes de dígitos binários sequenciais; deve haver também a concepção de que, por trás da troca de informações, há um laço implícito de entendimento, uma apropriação, uma forma peculiar de interpretá-las. A rede internet nos deu mais que acesso às informações, mostrando a sua vocação relacional entre pessoas e entre pessoas e instituições. Assim, pode-se dizer que o que circula na rede é mais que informações – são atos de linguagem, são processos implícitos de aprendizagens, são formas de construção coletiva de conhecimento.

Diante dessa realidade, instituições públicas de pesquisa, desenvolvimento e inovação (PD&I), como a Embrapa, têm pela frente o desafio de encontrar novas formas de materializar a informação e o conhecimento que circulam pelas suas redes de pesquisa, pelas suas relações interinstitucionais e pelas suas interações com os diversos segmentos de públicos. Do ponto de vista organizacional, é estratégico despender esforços para buscar diversas possibilidades e soluções de identificar, gerir, disseminar e transferir as informações e conhecimentos produzidos na Embrapa. Entende-se que esta estratégia contribui para que a Empresa promova um debate em torno de novas racionalidades para pensar as demandas sociais do setor agropecuário brasileiro. Isto tanto assegura o diferencial competitivo da Embrapa neste segmento quanto contribui para evidenciar o Brasil no cenário internacional como um dos principais ‘players’ da produção de alimentos. Neste ponto reside a importância deste capítulo, cujo objetivo é registrar como a Embrapa Informática Agropecuária estuda, identifica, concebe e articula soluções de tecnologia de informação (TI) para gerir as informações e os conhecimentos que são produzidos nos diversos níveis organizacionais da Empresa.

A Embrapa Informática Agropecuária desenvolve pesquisas e aplicações nos contextos da gestão de dados (GD), gestão da informação (GI) e gestão de conhecimento (GC). Atua, ainda, na geração de metodologias e soluções web para organizar, tratar, recuperar e comunicar para a sociedade as informações e conhecimentos gerados ‘na’ e ‘pela’ Empresa. Dentre as principais áreas de estudo que contribuem para essa linha de pesquisa, encontram-se as de gestão e representação do conhecimento, ontologias, processamento de linguagem natural, terminologias, construção de espaços digitais e de conteúdo hipermediáticos.

O capítulo está dividido em seis seções, além desta introdução. Na segunda seção, é apresentado um embasamento conceitual que contextualiza a GI e a GC no âmbito das necessidades organizacionais da Embrapa, destacando como a Embrapa Informática Agropecuária contribui para atendê-las. Na seção três, discute-se como dados, informação e conhecimento são concebidos e como se alinham aos processos de GD, GI e GC da Embrapa. Nesta seção, ainda, argumenta-se sobre a necessidade de se pensar dados, informação e conhecimento a partir de uma abordagem sistêmica e holística, para além da visão dualista e reducionista, quase sempre predominante nos processos tecnológicos em curso na Empresa. Nas seções quatro e cinco, são apresentadas e discutidas as metodologias e tecnologias voltadas à interoperabilidade semântica, bem como para a criação e transformação de espaços digitais, adequando-os ao ambiente virtual de intensa

mobilidade social e tecnológica, em curso na sociedade atual. Assim, nas seções quatro e cinco aborda-se um fenômeno atual de grande impacto sociocultural e econômico, além de tecnológico, cuja centralidade reside nas pessoas. Nessas seções, ainda, as tecnologias móveis, em especial, os dispositivos como os telefones celulares, smartphones e *tablets* - enquanto instrumentos preponderantes de transformação socioeconômica e cultural - são colocados de forma crítica, para que se apresentem como solução de inclusão social, e não como simples aparatos tecnológicos com fins de adoção para alavancar e propagar o consumo acrítico de informação e conhecimento. Cabe salientar que as metodologias apresentadas nas seções quatro e cinco não são as únicas propostas de soluções de TI desenvolvidas pela Embrapa Informática Agropecuária, no escopo da sua competência. Entretanto, a escolha recaiu sobre aquelas metodologias que melhor se assentam no tripé conceitual que orienta a busca de entendimento transversal entre tecnologias, processos e pessoas. Esta tríade representa o fio condutor rumo às ‘tecnologias do conhecimento social’, como menciona Aranha Filho (1995). E, porque também é somente a partir de um entendimento indissociável sobre estes três constructos que se torna possível pavimentar as vias seguras e sustentáveis que levarão a Embrapa a compreender e gerir ativos tão estratégicos como a informação e o conhecimento. Por fim, a seção seis dedica-se à discussão dos principais resultados e conclusões em termos de ganhos com a atuação da Embrapa Informática Agropecuária na busca de soluções de TI para a GD, GI e GC para a Embrapa. A seção sete arrola as referências bibliográficas utilizadas para embasar a construção do capítulo.

## 2 GI e GC no contexto da Embrapa: concepções e necessidades

GI e GC são muitas vezes confundidas, provavelmente em função da ambiguidade, multiplicidade e transversalidade com que os termos “informação” e “conhecimento” são tomados no contexto das especificidades inerentes às diversas disciplinas, as quais respondem, substancialmente, pelos seus principais enunciados teórico-conceituais, advindos de campos científicos, como o da Informação, da Administração, da Comunicação, da Cognição, da Computação, dentre outros.

Brascher e Café (2008), analisando os processos de organização e representação da informação (OI/RI) e do conhecimento (OC/RC), aqui considerados partes dos macroprocessos de GI e GC, sugerem diferenciá-los da seguinte forma: OI/RI se aplica ao processo de organização e representação de “objetos” informacionais, enquanto OC/RC se aplica às unidades do pensamento (conceitos). Enquanto a OI/RI se ocupa em organizar conjuntos de documentos impressos ou digitais em coleções ou repositórios, a OC/RC se ocupa em construir modelos do mundo que se constituem em abstrações da realidade. Nesse aspecto, os modelos conceituais são tomados como referenciais para a concepção e execução de ações operacionais para GI.

Enquanto a GI não suscita controvérsias em relação à sua proposta conceitual e decorrente operacionalidade, o termo “gestão do conhecimento”, da forma como é amplamente empregado nos contextos das práticas administrativas e de gestão corporativa, é objeto de críticas por parte de pensadores mais puristas que preferem considerá-lo como um conjunto de atividades organizacionais que, não sendo concernentes estritamente à GI, se estendem à gestão de outras práticas de trabalho, na expectativa de que mudanças nas formas de comunicação interpessoal (incluindo aquelas mediadas pelas TIC) venham permitir melhorias no compartilhamento de informações

(WILSON, 2002). Em outras palavras, não existiria uma “gestão” do conhecimento, entendida aqui como sendo uma experiência intangível, intelectual e individualmente gerada em cada cérebro humano de forma diferente e que só é possível de ser manifestada exteriormente se seu emissor o codifica (signos, símbolos, linguagens) e o receptor, por sua vez, reconhece e processa, novamente em seu cérebro, o seu entendimento.

Desse raciocínio decorre a compreensão de como GI e GC acabam sendo indissociáveis em quaisquer contextos que possam ser empregados, sejam eles organizacionais, estritamente técnico-científicos ou acadêmicos ou amplamente sociais. No caso da Embrapa, em vários momentos, todos esses contextos se sobrepõem e à parte das discussões muito aprofundadas conceitualmente, GI e GC adquirem valor estratégico indiscutível para a Empresa, muito embora, em algumas de suas facetas, as práticas de trabalho já estabelecidas ainda sejam dispersas e desarticuladas e careçam de sistematização processual corporativa.

Enquanto o ambiente de informação global se dirige para os novos paradigmas do século XXI, alguns desafios ainda carecem de soluções como, por exemplo, as: tensões entre a adoção de padrões universais de compartilhamento de dados e informações versus variações locais, cultural e historicamente estabelecidas; tensões entre a escolha por uma infraestrutura homogeneizadora versus a posição de comunidades que insistem em suas especificidades; tensões entre os interesses do indivíduo versus os interesses de controle corporativos. Tendências mundiais no campo da gestão de ambientes informacionais em instituições de PD&I, como a Embrapa, reforçam a ênfase na construção de um arcabouço teórico-conceitual e computacional, para que se possibilite combinar dados e informações (de diferentes fontes e finalidades), moldados na harmonização e conciliação da diversidade, e estruturalmente coerentes e facilitadores para o desenvolvimento de sistemas e comunidades de informação amigáveis.

A Embrapa, ao mesmo tempo, é membro e coordenadora do Sistema Nacional de Pesquisa Agropecuária (SNPA), um conjunto de instituições públicas e privadas que, de forma cooperativa, desenvolve a pesquisa agrícola brasileira, envolvendo inúmeras áreas do conhecimento científico (PENTEADO et al., 2014). Em decorrência desse arranjo, Embrapa e SNPA interagem com inúmeros outros setores da sociedade (Estado e governos, academia, indústria, comércio, terceiro setor etc.), participando de discussões em níveis nacional e internacional, local ou global e que envolvam a agropecuária e áreas correlatas. Desse contexto de atuação, pode-se imaginar o volume incalculável e inestimável de dados, informação e de conhecimento agropecuários gerado, disseminado, re combinado e reutilizado e a conseqüente necessidade de se estabelecer processos de gestão sistêmicos para reunião, organização, tratamento, armazenamento, acesso e disseminação de tantos e diversos conteúdos.

A missão da Embrapa consiste em “viabilizar soluções de pesquisa, desenvolvimento e inovação (PD&I) para a sustentabilidade da agricultura, em benefício da sociedade brasileira” (EMBRAPA, 2008, p. 18), o que, obrigatoriamente, implica viabilizar que a informação e o conhecimento agropecuários produzidos na Empresa sejam entregues à sociedade, para que por ela possam ser apropriados. Ao desenvolver e aplicar conhecimento técnico-científico e propor soluções para a agropecuária tropical, a Embrapa envolve-se com desafios temáticos de alto grau de complexidade geoambiental e político-social, tais como: impacto das mudanças climáticas na atividade agropecuária e destas na sustentabilidade de recursos naturais; direcionamento do melhoramento genético de raças animais ou variedades de plantas frente às mudanças climáticas

ou frente a mudanças do comportamento epidemiológico de doenças ou das dinâmicas populacionais de pragas e patógenos, entre inúmeros outros. Nesse contexto, a transversalidade da informação e do conhecimento se consolida e determina que os processos de GI e GC na Embrapa adquiram importância estratégica, visando à garantia não apenas do seu compromisso social, mas também, o apoio contínuo ao desenvolvimento, à competitividade e à sustentabilidade institucionais. Informação e conhecimento estão diretamente ligados aos processos de tomada de decisão e, dessa forma, por meio dos fluxos informacionais estabelecidos nos processos corporativos, suportam todo o processo da gestão organizacional, do mesmo modo que o sangue ou a seiva sustentam os organismos vivos.

Recorrentemente, nos últimos anos, no âmbito do processo de planejamento estratégico da Embrapa Informática Agropecuária, a análise do ambiente externo tem apontado como tendências ou demandas de aplicações da computação e das tecnologias de informação (TI) para a agropecuária as questões relacionadas com melhorias e empoderamento dos processos de GI e GC. Nesse contexto ressaltam-se, por exemplo:

- Aumento na adoção de TI para obtenção, análise, disseminação e visualização de dados, informações e conhecimento pelo setor agropecuário.
- Aumento da demanda por soluções tecnológicas para resolução de problemas complexos na agropecuária que requerem equipes interdisciplinares, multi-institucionais e transnacionais.
- Aumento da oferta de dados públicos em conformidade com políticas oficiais de dados abertos.
- Incremento na utilização de inteligência computacional em aplicações do domínio agropecuário.
- Uso crescente de novos canais de comunicação e novas formas de apresentação de conteúdo, acompanhando a evolução da linguagem (Web 2.0 e Web 3.0).
- Aumento da oferta de arquiteturas para o desenvolvimento de sistemas de armazenamento de dados distribuídos, heterogêneos, em grande volume e com alta frequência de geração.

Dessa forma, a Embrapa Informática Agropecuária vem concebendo e executando estratégias computacionais que visam a aproximar e alinhar emergentes TI às demandas do setor agropecuário, cada vez mais complexas, por dados e informação de PD&I. Uma das abordagens perseguidas nesse itinerário se refere às questões de organização, armazenamento, processamento e disseminação de crescentes volumes de dados. Nesse contexto técnico-científico, emergem alternativas que somente poderão ser aplicadas como solução operacional se estiverem baseadas em conceitos, infraestrutura e suporte computacionais de alto desempenho como, por exemplo, *e-Science*, *Big Data*, *Linked Data*, *Open Linked Data*, *Web Semântica*, *Cloud Computing*, e atenderem a requisitos, tais como: interoperabilidade entre sistemas de informação, interatividade e ambientes virtuais tecnológicos para trabalho colaborativo intermediado por mídias sociais, mobilidade para acesso via dispositivos móveis, em especial, celulares e *tablets*.

GI e GC são alvos constantes de demandas, interesses e motivações, visando ao desenvolvimento, validação e implantação de TI para facilitar e empoderar processos de organização, engenharia e gestão da informação e do conhecimento gerados ou utilizados na Embrapa. Para apoiar esses processos cada vez mais exigentes de inovação e evolução, a Embrapa Informática Agropecuária tem adotado alguns referenciais teóricos que viabilizam trabalhar a integração e

as interseções entre dados, informação e conhecimento, no âmbito das atuais considerações da Ciência da Complexidade<sup>2</sup> (MORIN, 1992, 1999; SIMON, 1962; SNOWDEN, 2000).

### 3 Para além dos dados ... A informação e o conhecimento

Há uma indissociabilidade entre os conceitos de informação e conhecimento e o conceito de dados, sendo muito comum encontrar, na literatura das Ciências da Cognição, da Informação, da Administração e da Computação, tentativas de uma possível aproximação teórico-conceitual que permita, ao mesmo tempo, definir os limites entre ambos e construir um percurso contínuo que os une.

Em instituições de PD&I, como a Embrapa, a pesquisa empírica é uma prática estabelecida, sendo umas das maiores preocupações a busca de solução para problemas reais por meio da validação e apropriação do conhecimento científico pela sociedade. Aqui, a proposta sugerida pela hierarquia Dado-Informação-Conhecimento-Sabedoria (BELLINGER et al., 2004) é oportunamente absorvida, mesmo que se mantenha aquém das discussões e implicações filosóficas que dela se originam (FRICKÉ, 2009).

A representação hierárquica-piramidal é a mais comumente recuperada (BERNSTEIN, 2009; FRICKÉ, 2009; ROWLEY, 2006, 2007) para apoiar vários tipos de argumentação. Outros formatos, cognitivamente mais elaborados, representam a relação de forma linear-progressiva (BELLINGER et al., 2004; CLARK, 2004) (Figura 1). Uma curiosidade é que essas representações foram inspiradas, originalmente, nos seguintes versos de um poema americano (CHOU, 1998; HEY, 2004; ROWLEY, 2006, 2007), antecipando a discussão que, décadas depois, seria finalmente retrabalhada transferindo a reflexão filosófica para as ciências aplicadas:

*Where is the Life we have lost in living?*

*Where is the wisdom we have lost in knowledge?*

*Where is the knowledge we have lost in information?*

Parafraseando esses versos para aproximá-los da motivação real de estabelecer melhorias dos processos de GI e GC no contexto atual da Embrapa, pode-se sugerir algo como:

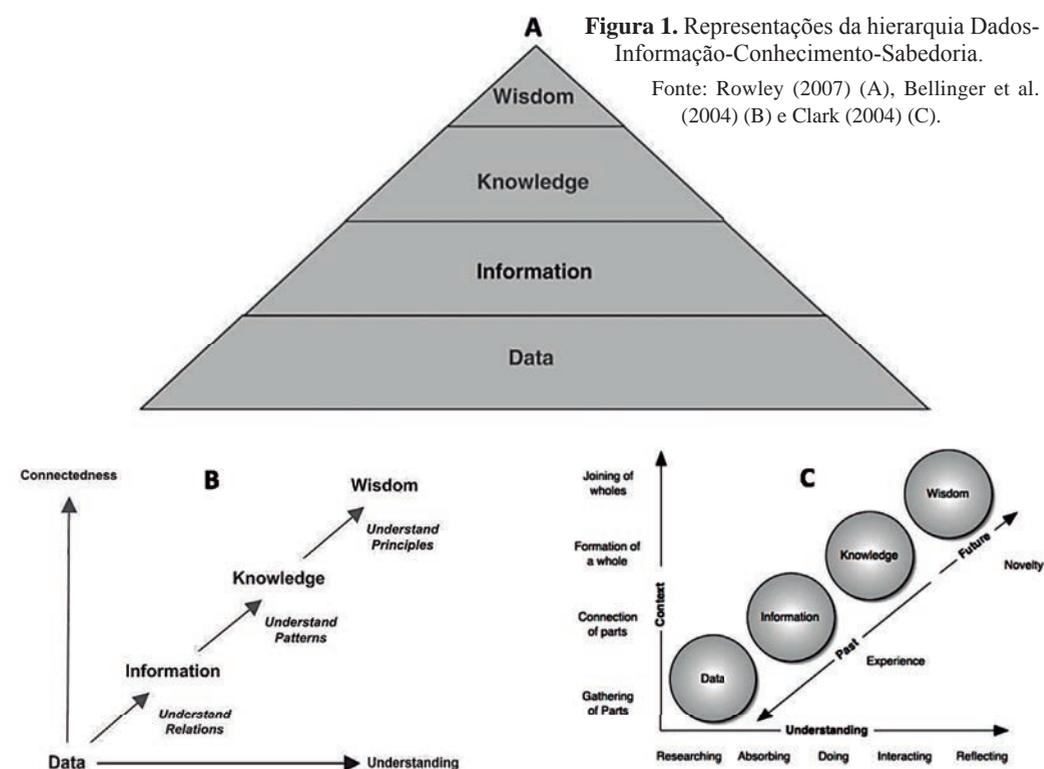
Onde estão as soluções efetivas de GI e GC que vamos perdendo ao longo do tempo em que tentamos desenvolvê-las e implantá-las?

Onde estão as boas ideias que perdemos nas pessoas, nos processos, nos departamentos, nos projetos, no país, na internet, ...?

Onde estão os resultados do nosso trabalho que perdemos nas gavetas, prateleiras, pen drives, discos rígidos, sistemas de informação...?

Embora conceitualmente instigantes, tais representações são criticadas sob o ponto de vista pragmático em relação a sua utilidade como suporte a ações operacionais para gestão de dados (GD), GI e GC (FRICKÉ, 2009). A questão fundamental referente à crítica dessas representações, como

<sup>2</sup> Ciência da Complexidade é o estudo científico dos sistemas complexos, ou seja, sistemas compostos de inúmeros elementos que interagem entre si e que produzem um comportamento global que não pode ser facilmente explicado considerando somente as interações entre os seus componentes individuais. Em outras palavras, a propriedade do sistema como um todo não é o resultado da soma das propriedades de cada um de seus componentes. Sistemas complexos incluem redes sociais, ecossistemas, cérebros, cidades, mercados, negócios etc.



base conceitual, consiste na ponderação de que se a sabedoria é tomada como uma “verdade inquestionável e irrefutável” pode ser que não seja alcançada se, eventualmente, os dados que a sustentam não forem corretos ou verdadeiros!

Apesar disso, o conceito da hierarquia Dado-Informação-Conhecimento-Sabedoria geralmente, e quase inconsistentemente, é transferido diretamente para as lógicas que suportam a concepção e construção dos processos organizacionais de GI e GC, reproduzindo-se, de forma mais ou menos automática, o pensamento linear ou reducionista de que um determinado sistema seja o resultado puro e simples da soma das propriedades dos elementos que o compõe. Nas discussões atuais, no âmbito da Ciência da Organização, essa lógica excluiria os sistemas complexos, em que a relação entre causa e efeito não é totalmente linear, ou seja, a propriedade emergente de um determinado sistema não resulta apenas da reunião das propriedades de seus componentes, mas sim da dinâmica de inter-relações entre eles, incluindo as relações de incerteza, aleatoriedade e imprevisibilidade.

Na tentativa de ajustar essa situação, propõe-se a representação apontada na Figura 2. Nota-se que, nessa nova proposição conceitual, deliberadamente, prefere-se alinhar, com base na lógica da agregação de valor cognitivo, a trajetória dado-informação-conhecimento em um circuito circular, continuamente retroalimentado em detrimento de suas convencionais representações hierárquica-piramidal ou linear-progressiva.

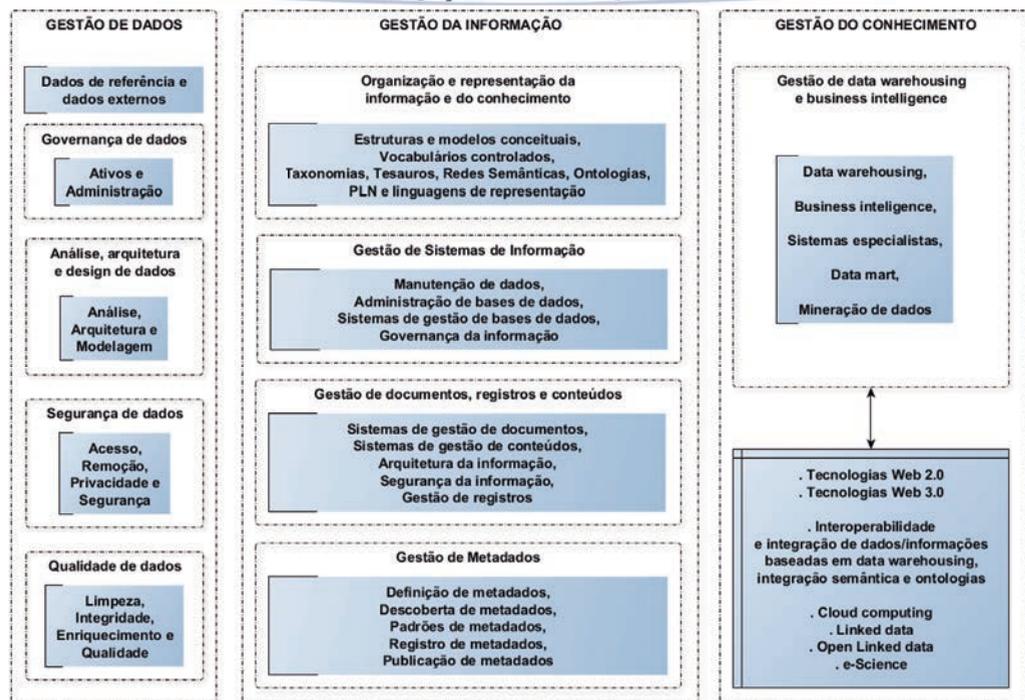
A partir dessa conceitualização, o alinhamento entre os processos de GD, GI e GC fica operacionalmente facilitado, permitindo um mapeamento das atividades, estruturas e tecnologias específicas e pertinentes a cada processo, além de mapear como cada um desses elementos se



**Figura 2.** Relação entre dado, informação e conhecimento compondo um circuito circular, contínuo e retroalimentado.

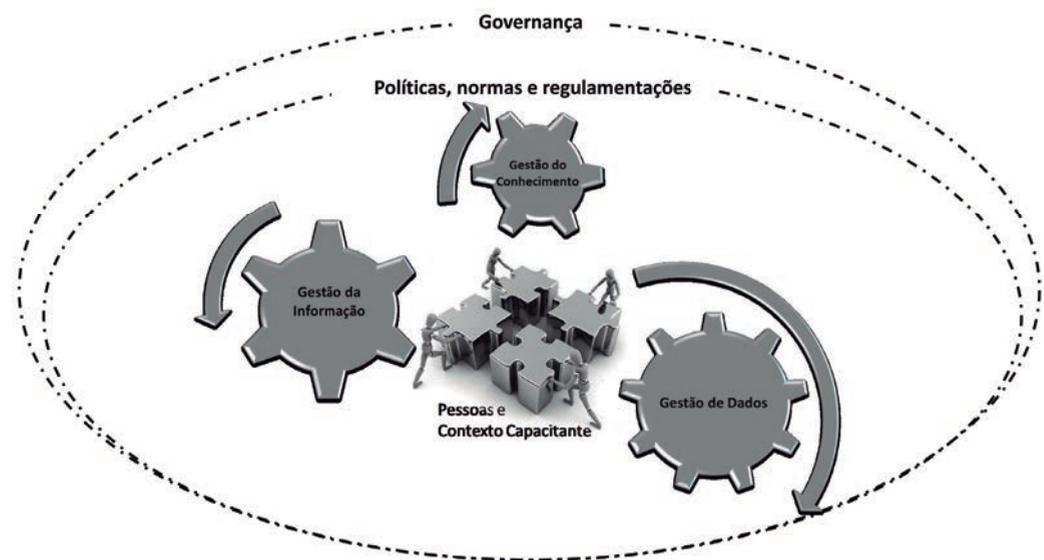
reorganiza em arranjos sucessivos, conforme proposto na relação contínua, circular e retroalimentada entre dados, informação e conhecimento (Figura 3).

No entanto, as tecnologias da informação e a computação, *per se*, não resolvem todas as questões pertinentes aos processos de gestão, havendo necessidade de associar a esse modelo os aspectos sociais e culturais que envolvem, além das pessoas, os processos, as relações de trabalho e outros fatores estruturantes que, em conjunto, viabilizam o estabelecimento do contexto capacitante ou “ba”, conforme conceitualizado na literatura (CHOO, 1998), considerado como fundamental para que de forma operacional e efetiva a conciliação entre GD, GI e GC possa ser praticada



**Figura 3.** Alinhamento dos processos de GD, GI e GC, com base na agregação de valor cognitivo e considerando as soluções tecnológicas de suporte.

(Figura 4). Na complementação desse pretendido alinhamento de processos devem ainda ser agregadas orientações de ordem corporativa (políticas, diretrizes, normas e regulamentos internos à empresa) e referenciais originários de instâncias governamentais, como a Lei de Acesso à Informação (Lei nº 12.257/2011) (BRASIL, 2011), a Política Nacional de Arquivos Públicos e Privados (Lei nº 8.159/1991) (BRASIL, 1991), a Lei da Propriedade Industrial (Lei nº 9.279/96) (BRASIL, 1996), a Lei nº 10.196/01, que altera a Lei da Propriedade Industrial (BRASIL, 2001), assim como com a Lei de Direitos Autorais (Lei nº 9.610) (BRASIL, 1998), entre outros marcos legais. Como corolário dessa sistematização, uma proposta de governança deve determinar e institucionalizar os princípios, estruturas, processos, culturas, papéis e responsabilidades relativos a dados, informação e conhecimento no contexto organizacional.



**Figura 4.** Reposicionamento da GD, GI e GC em relação a mecanismos corporativos estruturantes: governança, políticas, normas e regulamentações.

Além dos aspectos expostos anteriormente ainda há que se admitir que as instituições, dentre as quais a Embrapa, podem ser consideradas como sistemas complexos que, ao mesmo tempo, são componentes de outros sistemas complexos (sociais, econômicos e científicos) e contêm subsistemas também complexos (epistemologia do conhecimento, estruturas e processos organizacionais, relações sociais etc.). Isto impacta diretamente os processos de modelagem computacional de sistemas, porque ainda é frequente a tendência em focar apenas o pensamento reducionista. Esquemas de categorização ou classificação em uso, que embasam as arquiteturas computacionais de banco de dados e outros sistemas de informação, valorizam mais as relações hierárquicas entre conceitos, em detrimento de relações multidimensionais (de equivalência e/ou de associação), que podem agregar valor semântico e, portanto, mais significado aos dados, além de serem potencialmente ricas para a modelagem de sistemas complexos, como o agropecuário, caracterizado pelo elevado grau de multi, inter e transdisciplinaridade, englobando informações de múltiplas naturezas.

A GI na Embrapa, por quatro décadas, tem sido predominantemente focada na informação bibliográfica e documentária. Entretanto, mais recentemente, iniciativas para integração e significação

de dados e inter-relacionamento de informações têm sido realizadas, com vistas a viabilizar a construção de uma infraestrutura computacional voltada à possibilidade de gerar, gerir e extrair conhecimento.

As estruturas conceituais reducionistas e as TI (bancos e bases de dados, sistemas de informação) delas originadas têm sido desenvolvidas e aplicadas com base na premissa de que um determinado domínio de conhecimento (sistemas de produção, recursos naturais, biodiversidade, melhoramento genético etc.) deva ser decomposto em elementos constituintes e que estes, uma vez identificados, individualizados e descritos (metadados), sejam compartimentalizados nas tabelas dos modelos de entidade-relacionamento dos bancos de dados ou nos menus navegacionais dos sistemas de informação web, por exemplo. Esse tipo de abordagem condiciona as etapas subsequentes de GI (tratamento, acesso, recuperação e disseminação de informações) a uma ênfase na faceta sintática e detrimento do entendimento semântico e, assim, limitando o potencial de uma informação se transformar em conhecimento.

Identifica-se, então, a constante tensão entre o pensamento linear, analítico e reducionista, que analisa e recorta, partindo do todo para o específico e o pensamento sistêmico e holístico que, inversamente, analisa e sintetiza a partir do específico em direção à recomposição do todo. Diante desse conflito, a solução parece ser a busca do equilíbrio entre essas duas forças antagônicas, que articuladas entre si contribuem para uma melhor representação do mundo real, ao mesmo tempo em que nenhuma delas isoladamente constitui uma abordagem suficiente. No espaço entre essas duas possibilidades, interpõe-se o pensamento complexo. Nele, mais que a identificação de cada elemento componente dos sistemas e muito além do resultado de uma simples soma, valoriza-se a dinamicidade das relações entre eles.

Operacionalmente, do ponto de vista da computação, a solução vem da adoção de modelos conceituais de organização do conhecimento que permitam à cognição o livre trânsito desde as partes até o todo e vice-versa. Modelos que permitam ao máximo a construção de inter-relações e as valorizem tanto quanto cada um dos componentes do sistema, pois é exatamente no desenho das relações entre conceitos ou entidades do mundo real que se consegue incorporar as qualidades inerentes aos sistemas complexos: aleatoriedade, incerteza e imprevisibilidade.

Diante desse cenário e da necessidade constante de fazer evoluir o trabalho de organização e engenharia de dados, informação e conhecimento no contexto geral da Embrapa, a Embrapa Informática Agropecuária optou pelo desenvolvimento de abordagens contemporâneas e globalmente alinhadas, visando à interoperabilidade entre sistemas de informação e a interatividade (comunicação interpessoal mediada pelas TI), incorporando nos processos de GD, GI e GC inovações para garantir continuamente o acesso e uso dos seus conteúdos informacionais. Algumas das recentes propostas de desenvolvimento e aplicação de TI para empoderamento dos processos de GD, GI e GC agropecuários são relatadas a seguir.

## 4 Metodologias e tecnologias para a interoperabilidade semântica

Uma das soluções de TI propostas pela Embrapa Informática Agropecuária para a GI e a GC na Embrapa assenta-se na linha de pesquisa voltada para a construção e aplicação de sistemas de

organização do conhecimento (SOC)<sup>3</sup>, orientados linguisticamente e desenvolvidos na interface do referencial teórico e prático do Processamento de Linguagem Natural (PLN) e da Inteligência Artificial (IA).

Sob a denominação de SOC (SOUZA et al., 2010; ZENG, 2008) são reunidos modelos conceituais que conseguem explicitar o conhecimento por meio de representações e codificações textuais e diagramáticas. Ao proporem formalização e conceituação coletiva dos objetos do mundo real, facilitam o mapeamento e o compartilhamento de concepções e informações e, portanto, podem ser empregados nos processos de GI e GC e como ferramentas de suporte ao processo decisório. Os referenciais teóricos e metodológicos dos SOC podem ser recuperados e aplicados sob duas perspectivas: reuso de modelos já existentes e construção de novos modelos conceituais.

Aplicações de IA podem se beneficiar do PLN para agregar valor na organização e representação de domínios de conhecimento, uma vez que, na base das soluções que incluem a língua natural, está o emprego de heurísticas que considerem os níveis morfológico, sintático e/ou semântico. Do ponto de vista cognitivo, a linguagem natural, falada e escrita, é a maneira preponderante que os seres humanos usam para desenvolver e transmitir seus conhecimentos a outrem. Nesse processo, utilizam-se termos, ou seja, rótulos lexicais para designar os conceitos que se quer representar. Genericamente, a Terminologia é a parte da Linguística que reúne o conjunto de princípios e métodos adotados no processo de gestão e criação de produtos terminológicos, tais como glossários, vocabulários especializados, tesouros, para citar alguns entre os mais comuns. A Terminologia Computacional representa a simbiose da Terminologia com os processos do PLN e envolve a sistematização de métodos a partir da aplicação de ferramentas computacionais específicas e compatíveis com as tarefas terminológicas, contribuindo para o desenvolvimento e aplicação de produtos para a GC. Os SOC abrangem todos os tipos de esquemas que organizam e representam o conhecimento, incluindo desde esquemas conceituais unidimensionais, como as listas de termos e glossários, até os mais evoluídos em estrutura e função, como tesouros e ontologias. Esses últimos são passíveis de transcrição para linguagens artificiais, lidas e compreendidas pelos computadores.

Na Embrapa Informática Agropecuária, os SOC têm sido construídos por meio do alinhamento do PLN com a cognição humana, a partir de três itinerários semiautomatizados percorridos isoladamente ou em qualquer combinação entre eles: a) corpora textuais; b) estruturas conceituais e c) vocabulários, conforme apresentados na Figura 5 e detalhados como segue.

No processo de organização do conhecimento são tomadas as decisões que delimitam o domínio que será organizado e representado, além das decisões relacionadas aos itinerários metodológicos a serem seguidos, lembrando que eles não são excludentes entre si. No processo de engenharia do conhecimento são executadas as atividades operacionais pertinentes a cada um dos três itinerários, incluindo os programas utilizados para automatizar as atividades. Na ausência de um único programa que integre, em uma única interface, todas as funcionalidades que já estão informatizadas, tem sido necessário o alinhamento de aplicações disponíveis em quatro programas diferentes, para a execução de um conjunto de atividades ou de apenas algumas delas isoladamente: e-Termos; *Cmap*; *yWorks*; *NodeXL*, *TheXML*, *Protégé* e *Etecam*<sup>4</sup>. O e-Termos é um ambiente

<sup>3</sup> Tradução para o português de Knowledge Organization System (KOS).

<sup>4</sup> Extrator Automático de Termos e Estruturas Conceituais Agrícolas Multilíngues.

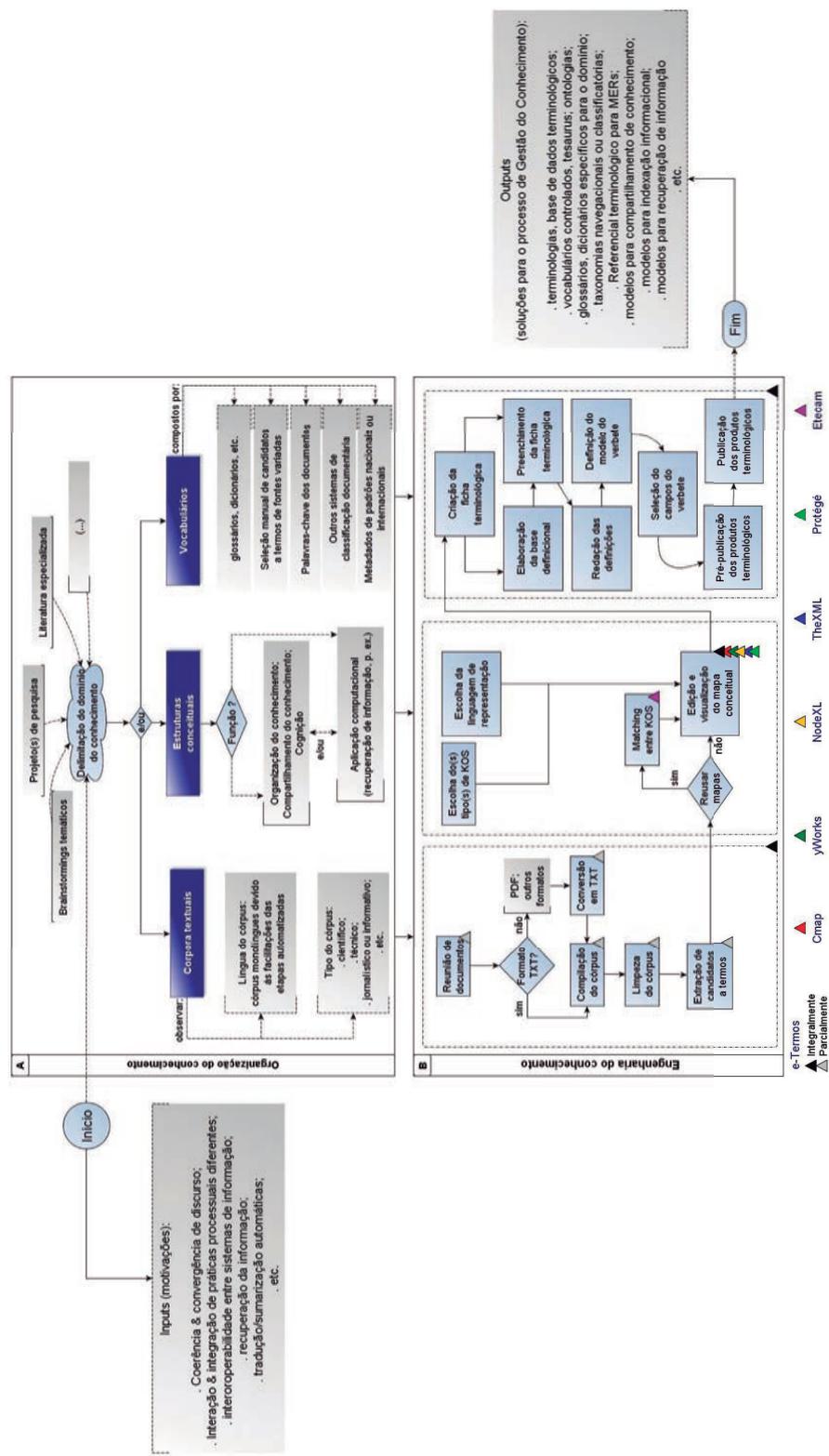


Figura 5. Etapas conceituais (A) e metodológicas (B) para concepção e construção de SOC agropecuário.

computacional colaborativo web de acesso livre e gratuito dedicado à gestão terminológica. O *Cmap*, *yWorks* e o *NodeXL* são programas gratuitos utilizados na concepção, edição, visualização e gestão de mapas conceituais. O *TheXML* é um software proprietário adquirido pela Embrapa para a construção e gestão de tesaurus. O *Protégé* também é um programa gratuito dedicado à construção de ontologias formais. O *Etecam* é uma solução mais específica, desenvolvida para comparar e reusar mapas conceituais de SOC já disponíveis, como os tesaurus. A interoperabilidade entre esses recursos tem sido feita por meio de arquivos \*.CSV, \*.TXT e arquivos gráficos ou da linguagem RDF/XML. A automatização das etapas de concepção e construção de SOC tem evoluído progressivamente. O trabalho integrado envolvendo especialistas do domínio agropecuário e de organização do conhecimento, linguística e PLN é o diferencial que tem permitido o alinhamento, a validação e o refinamento das aplicações computacionais desenvolvidas.

Nos SOC estrutural e funcionalmente mais evoluídos (tesaurus, redes semânticas e ontologias) consegue-se valorizar tanto as relações como as entidades (objetos, conceitos) que compõem em conjunto uma certa unidade de conhecimento. Essa nova visão tem contribuído para representações mais completas da complexidade do mundo real, seja ele natural, antropizado ou mesmo o ambiente organizacional de uma empresa. Na verdade, esse tipo de modelo conceitual facilita a construção de propostas de representação do mundo real de forma ontológica, evoluindo da forma convencional que o compartimentaliza ou decompõe (reducionismo), posteriormente o recompõe (holismo) e, mais recentemente, se prepara para inter-relacionar multidimensionalmente seus elementos (complexidade). Os modelos conceituais assim concebidos e desenvolvidos alinham-se à tendência de concepção e construção de SOC que se preocupam em representar sistemas complexos (OHLY, 2012) e, portanto, muito oportunos para a realidade informacional que envolve a Embrapa. Além disso, esse tipo de abordagem visa a fornecer condições para que a Embrapa insira seus conteúdos informacionais no contexto da Web Semântica, a chamada Web 3.0, uma tendência de evolução tecnológica da internet, na qual a informação digital ganha significado, os computadores ganham inteligência e os conteúdos se tornam mais pragmáticos e eficazes em termos de fornecer as informações mais precisa e eficazmente, aproximando-as mais do processo que as converte em conhecimento.

Do ponto de vista específico da organização do conhecimento, os SOC podem continuar evoluindo para integrar ou agregar funcionalidades ainda mais operacionais como, por exemplo, glossários, sistemas de categorização ou de classificação, taxonomias (científicas, de websites), tesaurus, ontologias, compondo um conjunto de recursos para organizar e representar os avanços do conhecimento sobre a agropecuária brasileira e tropical, objeto de trabalho da Embrapa. Mais que isso, no que tange ao desenvolvimento institucional, a utilização dos SOC proporciona maior visibilidade do conhecimento produzido e disseminado pela Empresa, pois outras instituições de pesquisa também tratam seus conteúdos informacionais por meio desses recursos, os quais vêm se tornando cada vez mais formalizados e construídos sob padrões internacionais.

SOC são sistemas que, elaborados a partir da lógica das linguagens naturais humanas, conseguem explicitar o conhecimento intangível, codificando-o em terminologias e, assim, constituindo-se em ferramentas para representar o conjunto informacional de um domínio específico. Além disso, esses sistemas são passíveis de ser traduzidos em linguagens computacionais (RDF, SKOS, OWL), possibilitando que o conhecimento humano possa ser lido e entendido por máquinas. Com essa vantagem tecnológica adicional, os conteúdos informacionais podem ser tratados com o objetivo de integrá-los e relacioná-los a outros conteúdos digitais,

posicionando-os mais adequadamente no itinerário da interoperabilidade e, conseqüentemente, da Web Semântica.

Em suas propostas conceituais, SOC não representa unanimidade (e é altamente desejável que assim seja). Um SOC é um recorte de representação que deve ser e estar intimamente relacionado com os contextos espaço-temporais que pretendem representar. Devem ser abertos, flexíveis e construídos de forma a permitir alterações e refinamentos continuamente; isso implica utilizar plataformas e arquiteturas computacionais condizentes com esse caráter mutável. Somente nessa acepção, os SOC conseguem ser absorvidos como ferramentas de utilidade prática.

## 5 Metodologias e tecnológicas para organização de espaços digitais móveis

Brennand e Brennand (2007) ponderam que o grande desafio da atualidade é fazer com que dados, informação e conhecimentos circulem livremente, admitindo, assim, ser necessário pensá-los como elementos fluidos que, por entre caminhos, trilhas, galerias, veios e passagens, penetram o tecido social que liga as pessoas umas às outras, permitindo-lhes, ao mesmo tempo, alcançarem objetivos comuns e gerarem inteligência coletiva para ser socializada e compartilhada com a sociedade. Nesse sentido, o próprio conceito de redes inerentes ao ciberespaço, na definição dos autores, é esclarecedor, visto que

[...] traz implícita a busca pela inteligência conectada ou inteligência coletiva, entendida como uma hipótese relativa à capacidade de um grupo de agentes cognitivos (de natureza humana, a animal ou artificial) para atingir um nível de ação otimizada. Essa capacidade compreende o desencadeamento de um processo cognitivo de aprendizagem, representação e de decisão, bem como processos sociais como: a partilha, a troca, a negociação e a auto-organização. Incorporam, ainda, processos relacionais (ou de socialização) como o reconhecimento e a competição. (BRENNAND; BRENNAND, 2007, p. 57).

Diante da imprescritível vocação da Embrapa de produzir dados, informações, tecnologias, conhecimentos e saberes, para alavancar a agricultura brasileira, é inevitável, sobretudo, o seu compromisso institucional na busca de bases metodológicas que apoiem a gestão, o processamento, a divulgação e o acesso desses ativos por ela gerados. Para tanto, uma das linhas de pesquisa da Embrapa Informática Agropecuária é dedicada ao desenvolvimento de metodologias e tecnologias para organizar espaços digitais móveis. Nela, as tecnologias e as metodologias são pensadas, estudadas, analisadas e propostas tendo como centralidade as pessoas - na capacidade que elas têm de, a partir do uso das tecnologias digitais como meios, ampliarem a percepção que possuem sobre o mundo que as cerca.

Para a Embrapa, esta linha de pesquisa representa uma nova matriz conceitual, a qual se ancora em um tipo de comunicação dialógica, sistêmica e relacional, que abriga a possibilidade de criação e de desenvolvimento de novos espaços de interlocução e conversação entre a Empresa e seus diferentes públicos. Ainda, esta linha de pesquisa é aderente e capilar aos interesses estratégicos da Empresa, na medida em que contribui, sobremaneira, para que dados, informações e conhecimentos ultrapassem o estigma de serem tratados como elementos estanques, fechados e herméticos dentro de sistemas de informações e sejam compreendidos como elementos fluidos que circulam por entre redes de pesquisas para gerar inteligência coletiva. Além disso, esta linha

de pesquisa também contribui para permitir que as redes de pesquisa e interações estabelecidas pela Embrapa para interagir com seus diversos públicos sejam concebidas como canais de promoção e de divulgação de informações e conhecimentos para a sociedade.

Nesse sentido, investimentos estão sendo despendidos pela Embrapa na criação de soluções de TI que possibilitem às pessoas - a partir da interação entre elas e das interações entre elas e as mídias digitais -, criarem novas interpretações sobre os dados, as informações e os conhecimentos gerados pela Empresa, para serem dispostos em ambientes web. Admite-se que esta é uma das alternativas viáveis para criar na Embrapa um espaço de aprendizagem organizacional que privilegie o coletivo, a cooperação e a perspectiva processual e contínua da colaboração que dará suporte à criação de uma cultura organizacional voltada para a coaprendizagem.

O conceito de coaprendizagem enfatiza a existência de uma relação mútua decorrente da interação entre as pessoas e quando esta interação é capaz de permitir que elas mudem o comportamento que possuem diante de uma determinada situação. É um conceito de grande pertinência para a Embrapa que tem nas informações e conhecimentos os insumos básicos para a produção de novos conhecimentos, sendo estes também, os ‘produtos finais’ por ela gerados. Desenvolver soluções de TI para ampliar a cognição de seus pesquisadores, para criar uma inteligência coletiva interna para considerar o saber favorece a criação de um diferencial competitivo na Embrapa.

Na Embrapa Informática Agropecuária, as metodologias e tecnologias são estudadas não apenas como um instrumento ou ferramenta que possuem um fim único, mas também como meios potenciais por onde perpassam significações, sentidos, policompreensões e aprendizagens. Por isso, cabe dizer que as metodologias e tecnologias para organizar espaços digitais web móveis são também metodologias da inteligência coletiva e das coaprendizagens. Este entendimento amplia o grau de importância que elas oferecem para o segmento agropecuário brasileiro, já que potencializa a construção de conhecimentos coletivos, o compartilhamento de informações e a interoperabilidade, especialmente, como é o caso das soluções propostas no tocante ao design de espaços e conteúdos digitais centrados ou não na aprendizagem colaborativa e social em rede (OKADA; OKADA, 2007, 2011; SOUZA, 2013).

Na Embrapa Informática Agropecuária, portanto, são construídos modelos conceituais de organização de espaços digitais web no contexto da mobilidade, voltados ao processo de transferência de tecnologia na Empresa, com foco na comunicação e disseminação de informações e tecnologias para o segmento agropecuário brasileiro. A constituição desses espaços digitais é essencial à materialização de ações de comunicação para apoiar o processo de transferência de tecnologia, em quaisquer ambientes, canais/meios. Aqui, cabe definir meios de comunicação como aqueles que servem “[...] a uma relação de comunicação na qual serão investidas operações de produção e recepção”, e acrescentar que “[...] a ideia de meio de comunicação deve satisfazer um critério que se refere ao acesso plural das mensagens disponíveis no meio.” (BARICHELLO; SCHEID, 2007, p. 4).

Ambientes comunicacionais mediados por tecnologias digitais, em especial as da Web 2.0, tornam-se imprescindíveis para a fluidez das relações de natureza dialógica e colaborativa. Nesses ambientes os indivíduos são considerados atores, já que podem exercer de forma ativa e simultânea diferentes papéis, como: emissor, receptor, produtor e coprodutor de conteúdos. Em tais espaços de comunicação midiática digital, os conteúdos de informação devem ser articula-

dos e integrados, para que facilitem a apropriação e a aprendizagem por parte dos interlocutores interessados.

Esta lógica comunicacional insere-se na perspectiva de pensar as tecnologias digitais não como meros canais de informação, mas como vias facilitadoras da compreensão da realidade, porque são capazes de permitir às pessoas a possibilidade de construir novas formas de entender o real. Sodré (2002) fala da tendência da sociedade contemporânea de virtualizar ou de telerrealizar as relações humanas a partir do entendimento de comunicação enquanto processo informacional, que se faz e se refaz por meio da ‘tecnointeração’, ou seja, de uma interação que se desenvolve no âmbito das mídias digitais virtuais.

Na Embrapa, atualmente, estão sendo reformuladas as práticas de produção e organização de conteúdos em espaços digitais, em razão:

- 1) Do uso intensivo de tecnologias, em especial, dos dispositivos móveis, que permitem às pessoas deslocarem-se através de territórios informacionais<sup>5</sup>.
- 2) Do domínio e uso simultâneo de várias mídias pelas pessoas.
- 3) Da necessidade de ajuste das linguagens híbridas a cada tipo de mídia digital, especialmente quando se quer promover a disseminação de informações e a transferência de tecnologias de forma que os indivíduos possam apreendê-los, mudando seu comportamento com vistas à adoção de novas práticas rurais (SOUZA, 2013).

Torres e Souza (2011) assinalam que conteúdos comunicacionais voltados às necessidades de disseminar e transferir tecnologia devem conter elementos atrativos e aglutinadores dos interesses dos públicos-alvo, de modo a favorecer a interlocução entre as pessoas e a Embrapa. Esta forma de organizar os conteúdos é que os tornam objetos sociais (WELLER, 2008) porque, além de servirem de ponto de partida para a promoção das interações que ocorrem entre as pessoas em espaços midiáticos, também comunicam, informam e mobilizam nas pessoas vários processos cognitivos, instigando-as a se interessarem por outros conteúdos e a compreendê-los (TORRES et al., 2013).

Nesse cenário inseriu-se a determinação da Embrapa Informática Agropecuária de propor um modelo de organização de espaços digitais para abarcar as lógicas dos processos de criação, circulação e apropriação de conteúdos web no contexto da mobilidade, de modo a assegurar uma identidade espontânea destes por parte dos indivíduos (participes), que tanto poderão produzi-los, quanto acessá-los via desktop ou dispositivos móveis. O modelo ora apresentado (Figura 6) tem a finalidade de orientar a produção de conteúdos web para incrementar as ações de disseminação de informações e de transferência de tecnologia na Embrapa. Está fundamentado em aspectos da representação do fenômeno da midiatização elaborada por Verón (1997).

O termo midiatização apareceu no início da década de 1990, em função da acelerada evolução tecnológica e das demandas sociais, as quais tem levado as pessoas a aspirarem novas formas de comunicação (STASIAK, 2013). Com base neste fenômeno, Verón (1997) propõe um esquema no qual se identificaram três fluxos que compõem a midiatização. Estes fluxos se distinguem por propiciarem a complementaridade entre emissor e receptor, a horizontalidade na troca de mensagens e a presença do feedback.



**Figura 6.** Modelo de organização de espaços digitais para a Embrapa.

Fonte: adaptado de Verón (1997).

Para Fausto Neto (2007), as tecnologias digitais agem nas relações institucionais como ‘meios-pulsão’, como consolidadoras de um novo tipo de real que se liga à produção de sentido através de meios ‘sociotécnicos’. Por isto, a midiatização é vista como estratégia lógica de aproximar as empresas de seus públicos, podendo ser considerada não apenas um meio/canal de comunicação, mas, sobretudo uma força organizativa capaz de atravessar todas as práticas sociais. Nesse sentido, no esquema traçado por Verón (1997) os processos comunicacionais não são entendidos como sendo lineares de causa e efeito e sim como sendo circuitos de feedback.

De maneira análoga, o modelo de organização de espaços digitais desenvolvido pela Embrapa Informática Agropecuária enfatiza três dimensões importantes que se complementam e dialogam horizontalmente; são elas: dimensão institucional, dimensão ambiência midiática e dimensão social. A dimensão institucional, que representa o conjunto de ações comunicacionais que precisam ser estruturadas para disponibilizar os conteúdos digitais organizados dentro de uma metodologia específica (SOUZA, 2013; TORRES; SOUZA, 2011). Basicamente, a organização dos conteúdos deve contemplar três eixos: técnico, discursivo e tecnológico. No eixo técnico, trata-se de desenvolver uma arquitetura pedagógica que favoreça, de um lado, a compreensão do conteúdo a ser disseminado e, de outro, a apropriação deles de forma que os indivíduos que os acessem possam apreendê-los (SOUZA, 2013). Para que isto ocorra deverá ser composta uma equipe técnica interdisciplinar formada por profissionais de várias áreas do conhecimento como pedagogia, comunicação, engenharia da computação, ciência da informação, pesquisadores especialistas do domínio do conhecimento etc. Esta equipe se responsabilizará por identificar, reunir, analisar, selecionar e organizar, de forma complementar e interdisciplinar, as informações técnico-científicas<sup>6</sup> produzidas pela Embrapa que servirão para construir a arquitetura pedagógica

<sup>5</sup> Áreas de controle de fluxo de informações que ocorre entre o ciberespaço e o espaço urbano e/ou rural.

<sup>6</sup> Entendidas como o conjunto de informações e conteúdos de caráter formal e explícito em algum documento interno e/ou externo produzido pelos diversos pesquisadores da Embrapa no âmbito dos vários domínios de conhecimentos, áreas temáticas e de atuação profissional que representam os resultados de pesquisas e/ou de desenvolvimento de produtos, processos, metodologias e/ou serviços.

ca dos conteúdos a serem disseminados e/ou transferidos para a sociedade. O eixo discursivo é aquele que trata de desenvolver o tratamento linguístico do conteúdo que foi organizado na arquitetura pedagógica realizada pelo eixo técnico (SOUZA, 2013). Neste eixo, ainda, estruturam-se as linguagens que deverão ser adotadas em cada tipo de mídia na qual o conteúdo será aportado. Sabe-se que cada mídia suporta um tipo de linguagem e cada indivíduo tem um estilo de aprendizagem próprio. Assim, este eixo explora o potencial comunicacional de cada uma das mídias a serem adotadas, verificando o que oferecem e como os conteúdos podem ser ajustados às possibilidades tecnológicas que cada uma delas insere. Efetuar o tratamento da linguagem discursiva de forma a conciliar o que cada mídia possibilita em termos de recursos permite que se ampliem as chances de os conteúdos serem assimilados pelos indivíduos. Por fim, no eixo tecnológico é criada a infraestrutura tecnológica que apoiará a organização dos conteúdos nas diversas mídias e a estruturação destas mídias na dimensão da ambiência midiática. É um eixo no qual se desenvolvem os aportes tecnológicos de hardware e software ou aplicativos que permitirão que as interações digitais sejam realizadas de forma ágil, segura e interoperacional.

A segunda dimensão do modelo organizacional refere-se à criação de um espaço para a promoção da ambiência midiática. Este espaço refere-se a um ambiente virtual, composto por várias mídias digitais, que suportará os conteúdos organizados na dimensão institucional. Este ambiente será integrado por um conjunto de ferramentas tecnológicas da Web 2.0 tendo a função de alavancar o desenvolvimento das interações entre os indivíduos e de permitir que criem novas formas de relação com os conteúdos. Esta dimensão deverá impulsionar a formação de redes sociais interativas, horizontais e colaborativas, servindo de instrumento para a disseminação de informações e a transferência de tecnologias. A bidirecionalidade comunicativa, tipo de comunicação ‘de muitos para muitos’, é uma das características da plataforma web. Esta plataforma permite a incorporação de ferramentas de colaboração e relacionamento interpessoal (*FaceBook, Bebo, YouTube, MySpace*, etc.), bem como a criação de blogs, wikis, compartilhamento de fotos, vídeos e lista de favoritos. Além disto, possibilita a justaposição de aplicativos para formar a base para um ambiente virtual dinâmico e interativo no qual os indivíduos aprendem coletiva e colaborativamente (WHEELER; BOULOS, 2007). Paralelamente, esses são espaços novos de visibilidade e de interação, nos quais os indivíduos têm a chance de colocarem “em debate questões de seu interesse, as quais, com frequência, dizem respeito à sua relação com as instituições”. (BARICHELLO; SHEID, 2007, p. 7).

O desenvolvimento de um ambiente virtual fundamentado neste tipo de arquitetura multifuncional e participativa contribui também para que os indivíduos modifiquem suas pautas de comportamento. Neste sentido, pode-se dizer que esta dimensão oferece oportunidades para que as pessoas (re)signifiquem as formas de interação e convivência e os conteúdos organizados, dando-lhes chances de aprenderem (TORRES et al., 2010).

A última dimensão do modelo de organização de espaços digitais é a social. Nela, estão representados todos os atores sociais que têm interesse nas informações e tecnologias desenvolvidas pela Embrapa. Ainda que estes atores tenham natureza jurídica e representem uma comunidade específica (por exemplo, uma instituição), no âmbito do modelo apresentado são intitulados de indivíduos para que seja evidenciado o caráter de ‘sujeitos ativos’ – alguém que tem domínio sobre sua ação e que se insere em complexas relações sociais propiciadoras de novas racionalidades (VERÓN, 1997). Na dimensão social, ainda, desenvolve-se o processo de interação entre os indivíduos e a Embrapa, representado na Figura 6 pelas setas de duplo sentido. Neste processo am-

bos constroem novas racionalidades sobre as informações/conteúdos que são disponibilizados, gerando um saber não institucionalizado (aqueles desenvolvidos por intermédio das dúvidas, das reflexões, das idiossincrasias, das interações sociais e das experiências pessoais – as quais se denominam conhecimentos tácitos, trazidos por todos os indivíduos que interagiram por meio do processo de comunicação digital proposto). É nesta dimensão que serão observadas o que Verón (1997) denominou de zonas de produção de coletivos. Estas zonas se constituem no palco onde são produzidas as diversas dinâmicas interacionais promovidas pelos atores que permeiam as relações ocorridas entre as três dimensões. Elas operam como centros geradores das sinapses, das ligações que ocorrerão quando os atores sociais, a partir dos conteúdos produzidos pela Empresa e dos seus próprios interesses, alterarem sua pauta de comportamento em relação à realidade. São nas zonas de produção de coletivos que os atores sociais e a Embrapa se interligarão mutuamente para ampliarem as chances de construir uma perspectiva transversal de atuação.

Este é o panorama no qual se insere o modelo de organização de espaços digitais, que aponta para a constituição de novos modos de organização social e de relacionamento entre as pessoas, e delas com as instituições sociais, no caso a Embrapa. Trata-se, pois, de um novo espaço para encontros que não mais dependem da presença dos indivíduos, mas sim de conexões e interações, as quais refletem o modo como as transformações sociotécnicas vêm ocorrendo na sociedade.

## 6 Resultados e conclusões

Neste capítulo são discutidas as soluções de TI desenvolvidas pela Embrapa Informática Agropecuária para a gestão de recursos e sistemas de dados, informação e conhecimento produzidos ‘na’ e ‘pela’ Empresa. Nesta reflexão resgata-se a matriz conceitual norteadora dos estudos e pesquisas, que se ancora nas inter-relações pessoais mediadas pelas tecnologias digitais, aqui entendidas como instrumentos que lhes dão nova dimensão à capacidade de produzir, desenvolver, ampliar, combinar e recombinar informações e conhecimentos, com vistas a criar novas interpretações sobre a realidade da agropecuária brasileira.

Duas soluções de natureza metodológica e tecnológica foram apontadas como representativas. A primeira delas é a denominada “Metodologias e tecnologias para interoperabilidade semântica”, que direciona para ações voltadas aos aspectos da interoperabilidade semântica, reunindo modelos conceituais que explicam o conhecimento por meio de representações e codificações textuais e diagramáticas. Parte-se de uma conceitualização de objetos extraída do mundo real, o que facilita o mapeamento e o compartilhamento de novas significações e informações, podendo, por esta razão, ser empregados nos processos de GI e GC, em suporte ao processo decisório. Tal solução denota relevante contribuição para a Embrapa que, doravante, pode-se dizer detentora de um aparato metodológico e tecnológico para analisar, tratar, combinar e recombinar volumes de informações sob o prisma do tratamento semântico e da cognição, o que certamente contribuirá para promover o avanço científico na área de TI e de GC. Por fim, as metodologias e tecnologias para interoperabilidade semântica também ganham importância para a Embrapa, em razão dos benefícios que podem gerar em decorrência da sua aplicação prática, considerando-se que são:

- Soluções adequadas à sua realidade porque promovem o direcionamento de conteúdos de informação ao acesso aberto, alinhando-se a uma tendência global em C&T.

- Soluções que aumentam a aplicabilidade e a efetividade no uso de dados e informações nos seus processos de PD&I.
- Soluções tecnológicas para a representação de conteúdos técnico-científicos do domínio da agropecuária brasileira, facilitando a reunião, organização, compartilhamento, disseminação de dados e informações.
- Soluções que fortalecem o posicionamento da Embrapa na internet, dado o crescente alinhamento de seus produtos digitais de informação e conhecimento às inovações propiciadas pela Web Semântica.

A segunda solução de TI construída pela Embrapa Informática Agropecuária refere-se às “Metodologias e tecnologias para organização de espaços digitais móveis”, oriundas da necessidade de se organizar e dar acesso web a grandes volumes de dados, informações e conhecimento gerados ‘na’ e ‘pela’ Empresa. Tais metodologias e tecnologias surgem, pois, no entorno das questões de mobilidade tecnológica e nelas se apoiam para desenhar novos *modus operandi* de produzir, criar, organizar, estruturar e dar acesso a conteúdos e espaços digitais móveis, que sejam compatíveis com as exigências da sociedade atual. Esta segunda solução, de viés teórico-metodológico, permite à Embrapa construir um canal de relacionamento com os diferentes públicos com os quais interage, a partir de uma premissa comunicacional dialógica, horizontal, sistêmica e participativa. Nesta lógica, ainda, esta solução metodológica também favorece a organização de dados, informações e conhecimentos, principalmente, considerando-se a dinâmica interacional que ocorre entre as pessoas quando se deparam com informações e conhecimentos organizados ou quando trocam experiências, vivências e saberes. Em espaços digitais desenhados e organizados sob este prisma são mais amplas as chances de as pessoas construírem novas interpretações e representações sobre temas que lhes interessam. De maneira geral, para a Embrapa é muito importante que tais canais e ambientes virtuais sejam assim organizados porque se inserem na perspectiva da GD, GI e GC, permitindo à Empresa:

- Aproveitar o caráter potencial das tecnologias digitais web, em especial dos dispositivos móveis, de possibilitarem às pessoas o acesso a dados, informações e conhecimentos produzidos pela Embrapa a qualquer hora e lugar.
- Desenvolver um canal comunicacional de relacionamento participativo e colaborativo com seus diversos públicos, o que contribui para o fortalecimento da sua imagem institucional, além de favorecer a disseminação de informações e conhecimentos para a sociedade.
- Estabelecer uma metodologia de organização de conteúdos digitais favorecedora das coaprendizagens.
- Explorar o potencial das linguagens híbridas em cada tipo de mídia digital, para que promovam ações alternativas de disseminação de informações e de transferência de tecnologias, permitindo às pessoas apreendê-las, desde que predispostas a mudanças.
- Criar uma cultura de aprendizagem organizacional que privilegie o coletivo, o público, a cooperação e a perspectiva processual e contínua da colaboração.

Por fim, é consensual da Embrapa Informática Agropecuária o entendimento de que as soluções de TI estudadas e aqui apresentadas não são definitivas e nem as únicas; antes, porém, representam um recorte da realidade atual na qual o Brasil e a Embrapa se integram. Ao mesmo tempo,

reconhece-se que esses estudos e propostas, enquanto soluções metodológicas e tecnológicas, representam ineditismo, o que de per se requer a continuidade rumo à busca constante de respostas a novos e antigos desafios que continuamente se apresentam. Ainda, cabe considerar que:

- 1) A Embrapa Informática Agropecuária possui competências individuais e organizacionais para o enfrentamento do desafio de fazer com que as TIC sejam entendidas e usadas para favorecer a gestão organizacional - como instrumentos gerenciais, mas e também como meios de ampliação da cognição humana.
- 2) As tecnologias digitais, parte integrante da atual cultura midiática, já se revelaram capazes de oferecer as bases para o estabelecimento de várias soluções de TI voltadas para o segmento agropecuário brasileiro.
- 3) Os avanços científicos advindos da área de conhecimento em TI são notórios. Este é um campo do conhecimento que tem se mostrado cada vez mais transversal, exigindo dos pesquisadores que nele atuam a capacidade de exercitarem a interdisciplinaridade para construir novas lógicas interpretativas, ampliando ainda mais a percepção destas tecnologias na vida das pessoas, inclusive ao ponto de seu uso se tornar imperceptível, como, por exemplo, o que há tempo ocorre com a energia elétrica.

## 7 Referências

- ARANHA FILHO, J. Tribos eletrônicas: usos e costumes. 1995. In: SEMINÁRIO PREPARATÓRIO SOBRE ASPECTOS SÓCIO-CULTURAIS DA INTERNET NO BRASIL, 1995, Rio de Janeiro. *Anais...* Rio de Janeiro: RNP, 1995. Disponível em: <<http://flanelografo.com.br/esocius/t-jayme.html>>. Acesso em: 19 set. 2014.
- BARICHELLO, E. M. da R.; SCHEID, D. Apontamentos sobre a construção da visibilidade das instituições na Internet a partir de um cenário de midiática da sociedade. *E-Compós*, São Bernardo do Campo, v. 10, p. 1-15, 2007. Disponível em: <<http://www.compos.org.br/seer/index.php/e-compos/article/view/207/208>>. Acesso em: 29 jun. 2014.
- BELLINGER, G.; CASTRO, D.; MILLS, A. Data, information, knowledge, and wisdom. In: MENTAL model musings: systems thinking. [S. l.]: Gene Bellinger, 2004. Disponível em: <<http://www.systems-thinking.org/dikw/dikw.htm>>. Acesso em: 16 set. 2014.
- BERNSTEIN, J. H. The data-information-knowledge-wisdom hierarchy and its antithesis. *North American Symposium on Knowledge Organization*, v. 2, n. 1, p. 68-75, 2009. Disponível em: <[http://arizona.openrepository.com/arizona/bitstream/10150/105414/1/NASKO2009\\_08\\_Bernstein.pdf](http://arizona.openrepository.com/arizona/bitstream/10150/105414/1/NASKO2009_08_Bernstein.pdf)>. Acesso em: 16 set. 2014.
- BRASCHER, M.; CAFÉ, L. Organização da informação ou do conhecimento? In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – IX ENANCIB, 9., 2008, São Paulo. *Anais...* São Paulo: ANCIB, 2008. Disponível em: <[http://skat.ihmc.us/rid=1KR7TM7S9-S3HDKP-5STP/BRASCHER%20CAF%C3%89\(2008\)-1835.pdf](http://skat.ihmc.us/rid=1KR7TM7S9-S3HDKP-5STP/BRASCHER%20CAF%C3%89(2008)-1835.pdf)>. Acesso em: 16 set. 2014.
- BRASIL. Lei nº 8.159, 08 de janeiro de 1991. Dispõe sobre a política nacional de arquivos públicos e privados e dá outras providências. *Diário Oficial [da] República Federativa do Brasil*, Poder Executivo, Brasília, DF, 9 de janeiro de 1991. Seção 1, p. 455.
- BRASIL. Lei nº 9.279, de 14 de maio de 1996. Regula direitos e obrigações relativos à propriedade industrial. *Diário Oficial [da] República Federativa do Brasil*, Poder Executivo, Brasília, DF, 15 de maio de 1996. Seção 1.
- BRASIL. Lei nº 9.610, de 19 de fevereiro de 1998. Altera, atualiza e consolida a legislação sobre direitos autorais e dá outras providências. *Diário Oficial [da] República Federativa do Brasil*, Poder Executivo, Brasília, DF, 20 fev. 1998. Seção 1, p. 3.

BRASIL. Lei nº 10.196, de 14 de fevereiro de 2001. Altera e acresce dispositivos à Lei nº 9.279, de 14 de maio de 1996, que regula direitos e obrigações relativos à propriedade industrial, e dá outras providências. **Diário Oficial [da] República Federativa do Brasil**, Poder Executivo, Brasília, DF, 16 de fevereiro de 2011. Seção 1, p. 4. Edição extra.

BRASIL. Lei nº 12.527, de 18 de novembro de 2011. Regula o acesso a informações previsto no inciso XXXIII do art. 5º, no inciso II do § 3º do art. 37 e no § 2º do art. 216 da Constituição Federal; altera a Lei nº 8.112, de 11 de dezembro de 1990; revoga a Lei nº 11.111, de 5 de maio de 2005, e dispositivos da Lei nº 8.159, de 8 de janeiro de 1991; e dá outras providências. **Diário Oficial [da] República Federativa do Brasil**, Poder Executivo, Brasília, DF, 18 de novembro de 2011. Edição extra.

BRENNAND, E. G. de; BRENNAND, E. J. de G. Cognição e redes abertas: a informação interativa como coração dos sistemas inteligentes. **Ciência & Cognição**, v. 10, p. 54-64, 2007. Disponível em: <<http://www.cienciasecognicao.org/revista/index.php/cec/article/view/624>>. Acesso em: 23 set. 2014.

CHOO, C. W. **The knowing organization**: how organizations use information for construct meaning, create knowledge and make decisions. New York: Oxford Press, 1998.

CHOU, P. Y. (Ed.). **Stories of the human spirit T. S. Eliot (1888-1965)**: the rock. Mountain View, CA, 1998. Disponível em: <<http://www.wisdomportal.com/Technology/TSEliot-TheRock.html>>. Acesso em: 10 set. 2014.

CLARK, D. **Understanding and performance**. 2004. Disponível em: <<http://www.nwlink.com/~donclark/performance/understanding.html>>. Acesso em: 16 set. 2014.

EMBRAPA. Secretaria de Gestão e Estratégia. **V Plano Diretor da Embrapa: 2008-2011-2023**. Brasília, DF: Embrapa, 2008.

FAUSTO NETO, A. A midiáticação jornalística do dinheiro apreendido: das fotos furtadas à fita leitora. **La Trama de La Comunicación**, v. 12, n. esp. p. 117-131, 2007. (Dossier de Estudios Semióticos). Disponível em: <<http://www.redalyc.org/pdf/3239/323927555007.pdf>>. Acesso em: 10 jun. 2014.

FRICKÉ, M. The knowledge pyramid: a critique of the DIKW hierarchy. **Journal of Information Science**, Cambridge, v. 35, n. 2, p. 131-142, Apr. 2009. Disponível em: <<http://jis.sagepub.com/content/35/2/131.full.pdf+html>>. Acesso em: 16 set. 2014.

HEY, J. **The data, information, knowledge, wisdom chain: the metaphorical link**. 2004. Disponível em: <<http://www.dataschemata.com/uploads/7/4/8/7/7487334/dikwchain.pdf>>. Acesso em: 14 set. 2014.

MORIN, E. From the concept of system to the paradigm of complexity. **Journal of Social and Evolutionary Systems**, London, v. 15, n. 4, p. 371-385, 1992.

MORIN, E. Organization and complexity. **Annals New York Academy of Sciences**, v. 879, p. 115-121, 1999.

MUSSO, P. A filosofia da rede. In: PARENTE, A. (Org.). **Tramas da rede**: novas dimensões filosóficas, estéticas e políticas de comunicação. Porto Alegre: Sulina, 2004. p. 17-38. Disponível em: <[http://search.4shared.com/postDownload/48v7xDdo/musso\\_pierre\\_-\\_a\\_filosofia\\_da\\_.html](http://search.4shared.com/postDownload/48v7xDdo/musso_pierre_-_a_filosofia_da_.html)>. Acesso em: 4 fev. 2012.

OHLY, H. P. Mission, programs, and challenges of knowledge organization. **Advances in Knowledge Organization**, v. 13, p. 25-33, 2012.

OKADA, A. COLEARN 2.0 - coaprendizagem via comunidades abertas de pesquisa, práticas e recursos educacionais. **eCurrículo**, São Paulo, SP, v. 7, n. 1, abr. 2011. Disponível em: <<http://revistas.pucsp.br/index.php/curriculum/article/viewFile/5813/4128>>. Acesso em: 23 set. 2014.

OKADA, A.; OKADA, S. Novos paradigmas na educação online com a aprendizagem aberta. In: CONFERENCE INTERNACIONAL DE TECNOLOGIAS DE INFORMAÇÃO E COMUNICAÇÃO NA EADUCAÇÃO - CHALLENGES 2007, 5., 2007, Braga, Portugal. [Anais]. Braga: Centro de Competência da Universidade do Minho, 2007. Disponível em: <<http://people.kmi.open.ac.uk/ale/papers/a10challenges2007.pdf>>. Acesso em: 23 set. 2014.

PENTEADO, M. I. de O.; FONTES, R. R.; CAMPOS, F. A. de A.; EUCLIDES FILHO, K. A trajetória do planejamento da pesquisa na Embrapa. **Cadernos de Ciência & Tecnologia**, Brasília, DF, v. 3, n. 1, p. 35-60, 2014. Disponível em: <<http://seer.sct.embrapa.br/index.php/cct/article/view/19442/12567>>. Acesso em: 13 set. 2014.

ROWLEY, J. Where is the wisdom that we have lost in knowledge? **Journal of Documentation**, London, v. 62, n. 2, p. 251-270, 2006. Disponível em: <<http://www.emeraldinsight.com/doi/pdfplus/10.1108/0022041061065332>>. Acesso em: 10 set. 2014.

ROWLEY, J. The wisdom hierarchy: representations of the DIKW hierarchy. **Journal of Information Science**, Cambridge, v. 33, n. 2, p. 163-180, 2007. Disponível em: <<http://jis.sagepub.com/content/33/2/163.full.pdf+html>>. Acesso em: 10 set. 2014.

SANTOS, E. O.; OKADA, A. L. P. A construção de ambientes virtuais de aprendizagem: por autorias plurais e gratuitas no ciberespaço. In: REUNIÃO ANUAL DA ASSOCIAÇÃO NACIONAL DE PÓS-GRADUAÇÃO E PESQUISA EM EDUCAÇÃO - ANPEd2003, 26., 2003, Poços de Caldas. [Anais...]. [Rio de Janeiro: ANPEd], 2003. Disponível em: <<http://people.kmi.open.ac.uk/ale/papers/a06anped2003.pdf>>. Acesso em: 19 set. 2014.

SIMON, H. A. The architecture of complexity. **Proceedings of the American Philosophical Society**, v. 106, n. 6, p. 467-482, 1962.

SNOWDEN, D. J. New wine in old wineskins: from organic to complex knowledge management through the use of story. **Emergence**, v. 2, n. 4, p. 50-64, 2000.

SODRÉ, M. **Antropológica do espelho**: uma teoria da comunicação linear e em rede. Petrópolis: Vozes, 2002.

SOUZA, M. I. F. **Modelo de produção de microconteúdo educacional para ambientes virtuais de aprendizagem com mobilidade**. 146 p. 2013. Tese (Doutorado - Ciências Sociais na Educação) - Faculdade de Educação, Universidade Estadual de Campinas, Campinas.

SOUZA, R. R.; TUDHOPE, D.; ALMEIDA, M. B. The KOS spectra: a tentative typology of knowledge organization systems. **Advances in Knowledge Organization**, v. 12, p. 122-128, 2010.

STASIAK, D. A comunicação organizacional sob a perspectiva da midiáticação social: uma proposta de reflexão. In: ENCONTRO ANUAL DA COMPÓS, 22., 2013, Salvador. **Anais**. Salvador, 2013. Disponível em: <[http://compos.org.br/data/biblioteca\\_2029.pdf](http://compos.org.br/data/biblioteca_2029.pdf)>. Acesso em: 10 maio 2014.

TORRES, T. Z.; PIEROZZI JÚNIOR, I.; BERNARDES, R. M.; VACARI, I. Construção de ambientes colaborativos para redes de pesquisa: modelo para a Embrapa. In: CONFERÊNCIA IADIS IBERO-AMERICANA, 2010, Algarve. **Actas ... Algarve: Iadis**, 2010. p. 339-343.

TORRES, T. Z.; SOUZA, M. I. F. Metodologia de organização de conteúdos para transferência de tecnologia na Web 2.0. In: INTERNATIONAL SYMPOSIUM ON INNOVATION AND TECHNOLOGY - ISIT 2011, 2., 2011, Lima, Perú. **Proceedings...** Lima: International Institute of Innovation and Technology, 2011. p. 28-33.

TORRES, T. Z.; SOUZA, M. I. F.; PEREIRA, N. R.; GÂMBARO, B.; MAGALHÃES, V. M. A. Communication and digital content in research network laboratories. In: Simposio sobre la Sociedad de la Información (SSI 2013), 11., Jornadas Argentinas de Informática (42 JAIIO), 42., 2013, Córdoba, Argentina. **Anales...** Córdoba: Sociad Argentina de Informática. p. 326-340. Disponível em: <<http://www.42jaiio.org.ar/proceedings/simposios/Trabajos/SSI/20.pdf>>. Acesso em: 09 jun. 2014.

VERÓN, E. Esquema para el análisis de la mediatización. **Diálogos de la Comunicación**, Lima, n. 38, 1997.

WELLER, M. **Social objects in education**. 2008. Disponível em: <[http://nogoodreason.typepad.co.uk/no\\_good\\_reason/2008/01/whats-a-social.html](http://nogoodreason.typepad.co.uk/no_good_reason/2008/01/whats-a-social.html)>. Acesso em: 18 jun. 2014.

WHEELER, S.; BOULOS, M. K. A. A cultura colaborativa e a criatividade destrutiva da Web 2.0: aplicativos para o ensino da medicina. **Revista Eletrônica de Comunicação, Informação & Inovação em Saúde**, Rio de Janeiro, v. 1, n. 1, p. 27-34, jan./jun. 2007. Disponível em: <<http://www.reciis.icict.fiocruz.br/index.php/recciis/article/view/49/50>>. Acesso em: 11 abr. 2014.

WILSON, T. D. The nonsense of 'knowledge management'. **Information Research**, v. 8, n. 1, Oct. 2002. Disponível em: <<http://www.informationr.net/ir/8-1/paper144.html>>. Acesso em: 12 set. 2014.

ZENG, M. L. Knowledge organization systems (KOS). **Knowledge Organization**, Frankfurt, v. 35, n. 2-3, p. 160-182, 2008.

## Desenvolvimento de simuladores na agropecuária

Luís Gustavo Barioni  
Adauto Luiz Mancini

### 1 Introdução

Simuladores computacionais são softwares que fazem uso de modelos matemáticos e algoritmos para representar um sistema (NUTARO, 2011). A simulação, por sua intrínseca associação com a modelagem matemática, é frequentemente designada “modelagem e simulação” (M&S). As finalidades da aplicação de M&S são diversas e incluem:

- a) Geração, formalização e avaliação do conhecimento científico, particularmente em sistemas complexos.
- b) Apoio ao processo decisório e à resolução de problemas gerenciais.
- c) Educação e treinamento.
- d) Transferência de tecnologia.

Apesar de suas origens na matemática aplicada e na computação, o estudo e aplicação da M&S tem se tornado fundamental nas ciências sociais, humanas, biológicas e da saúde.

Um dos grandes atrativos de simuladores está em prover um ambiente virtual que permite integrar com representações de sistemas, naturais ou artificiais sem as limitações do mundo real. Na realidade, estamos restritos por espaço e tempo e na instrumentação disponível para impor determinados tratamentos e mensurar respostas. Além disso, altos custos e riscos, questões éticas e ambientais, frequentemente, inviabilizam a experimentação em muitas situações. A utilidade da virtualização é evidenciada pelo funcionamento da nossa própria mente. De acordo com Pezzulo et al. (2013), nosso cérebro constantemente simula nossas interações com o meio antes de desencadear ações conscientes.

Do ponto de vista científico e tecnológico, a M&S também tem sido apontada como instrumento de excelência para estudar sistemas naturais e para apoiar a elaboração de projetos e a gestão de sistemas artificiais, por permitir considerar simultaneamente intrincadas interações entre componentes e processos em um sistema e formalizar a lógica associada à emergência de propriedade e fenômenos (JOHNSON et al., 2013; KEATING, 2008).

Na agropecuária há crescente aplicação de simuladores em diversas áreas, tais como: desenvolvimento e produção de plantas e de animais (JONES et al., 2003; OLTJEN, 1986), apoio à tomada de decisões gerenciais (BARIONI et al., 2010; FREER et al., 1997), dinâmica de pragas, doenças

e contaminantes (ZAGMUTT et al., 2013), avaliações de impacto ambiental (ITTERSUM et al., 2008), dinâmica de uso da terra (VERBURG et al., 2004), manejo da água (ASCOUGH et al., 2012) e avaliação de tecnologias em geral (WOODWARD et al., 2008).

É grande a diversidade de ferramentas e paradigmas para implementação de simuladores na agropecuária, o que reflete a busca por um balanço entre facilidade de aprendizado e uso, padronização, flexibilidade e desempenho, e de sua adequação às características da equipe de desenvolvimento (em fase de elaboração)<sup>1</sup>. Com a profusão de iniciativas de desenvolvimento de simuladores na agropecuária, algumas das maiores empresas de pesquisa na área têm alocado equipes dedicadas a seus próprios arcabouços e infraestrutura para simulação (e.g., o OMS (DAVID, 2013), desenvolvido pelo Departamento de Agricultura dos Estados Unidos (USDA); o APSIM (MCCOWN et al., 1996), desenvolvido pelo Commonwealth Scientific and Industrial Research Organization (CSIRO); e os arcabouços VLE (QUESNEL et al., 2007) e Record (BERGEZ et al., 2013), desenvolvidos pelo Institut National de la Recherche Agronomique (Inra).

Nesse contexto, este capítulo apresenta conceitos e experiências relacionados ao desenvolvimento de simuladores na agropecuária, com foco nos trabalhos em andamento na Embrapa Informática Agropecuária, Unidade da Empresa Brasileira de Pesquisa Agropecuária (Embrapa), e também uma visão de futuro sobre esses trabalhos.

## 2 Conceitos básicos

Nutaro (2011) distingue três componentes básicos em softwares de simulação computacional:

- 1) O modelo matemático.
- 2) O simulador.
- 3) A interface de entrada e saída.

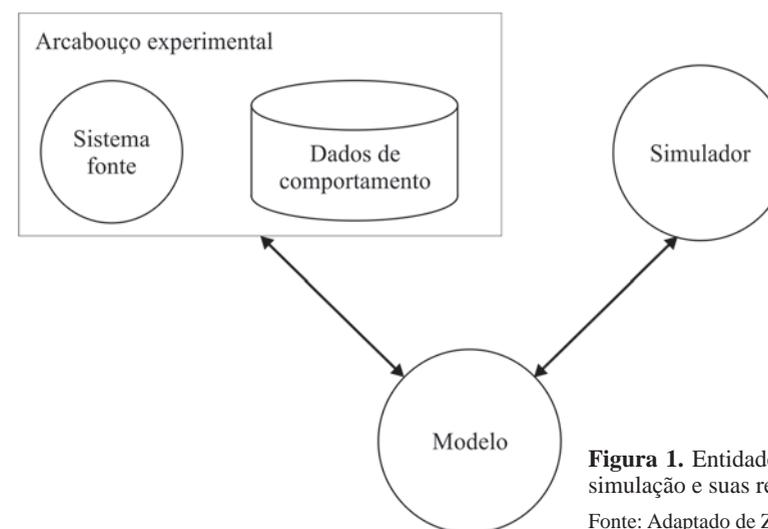
O modelo matemático é uma representação, utilizando linguagem e conceitos matemáticos, das relações e interações relevantes do sistema para o propósito para o qual o modelo se destina. Zeigler et al. (2000) utiliza o termo sistema-fonte para definir o sistema real que estamos modelando. O processo de desenvolvimento do modelo matemático a partir do sistema-fonte é normalmente denominado modelagem ou modelação.

O processo de modelagem se inicia pelo julgamento, de certo modo arbitrário ou empírico, de quais são a fronteira e os aspectos relevantes do sistema-fonte a serem reproduzidos. A fronteira do sistema define as variáveis, processos e inter-relações endógenos ao modelo. A fronteira é também o ponto de partida para identificarmos variáveis exógenas que são fatores determinantes do comportamento do sistema, também denominados em inglês *drivers*, os quais deverão fazer parte das entradas para a simulação, mas não serão afetados por ela.

O modelo conceitual é usado, sequencialmente, para definir, a partir do conhecimento teórico existente, como será representado o estado do sistema e quais os processos e fatores determinan-

tes (*drivers*) de sua dinâmica. Durante a modelagem conceitual deve haver, todavia, a escolha de um paradigma para a modelagem, que irá guiar os passos seguintes do processo de modelagem e de codificação do simulador. Vários paradigmas podem ser utilizados. Na agricultura há uma tendência para a aplicação do paradigma de dinâmica de sistemas contínuos, por meio de sistemas deterministas de equações diferenciais ou de diferenças. Entretanto, outros paradigmas, como simulação orientada a eventos discretos (BARIONI et al., 1999), simulações baseadas em indivíduos, simulação estocástica, e simulação-otimização são também utilizados. A escolha de um paradigma e a área de origem dos profissionais envolvidos no trabalho de modelagem e simulação direciona fortemente a metodologia de modelagem e codificação e os artefatos utilizados. O passo seguinte é a formalização matemática e algorítmica a partir do modelo conceitual e da codificação. Esses passos serão tratados mais detalhadamente ao longo do capítulo.

Zeigler et al. (2000) fazem uma útil distinção entre modelos e simuladores (Figura 1). Eles conceituam modelos como especificações (conjunto de instruções) e simuladores como os programas responsáveis por gerar o comportamento especificado em um computador. Modelos matemáticos podem, portanto, ser armazenados como arquivos contendo informações estáticas e independentes do programa que os executa. Um modelo pode, por essa abordagem, ser executado por mais de um simulador e o simulador pode executar diferentes modelos.



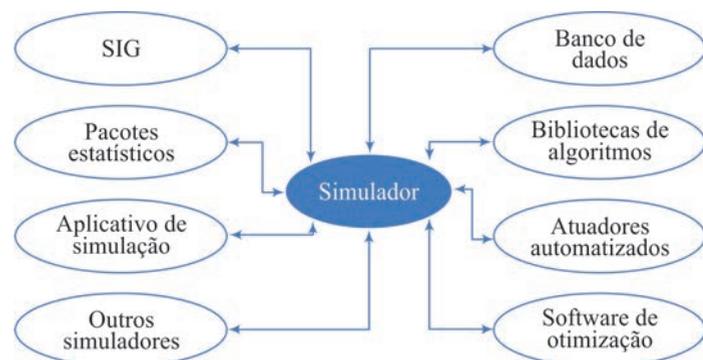
**Figura 1.** Entidades básicas em modelagem e simulação e suas relações.

Fonte: Adaptado de Zeigler et al. (2000).

Dados obtidos a partir do sistema-fonte são essenciais no processo de modelagem. Zeigler et al. (2000) denomina a estrutura envolvida na geração de dados sobre o sistema-fonte de arcabouço experimental. Dados observados são utilizados em diversas fases do processo, particularmente na definição da forma funcional das relações entre as variáveis, na determinação do valor de parâmetros (parametrização), na avaliação do modelo e em processos adaptativos envolvendo a contínua calibração do modelo a partir de dados observados. Dessa forma é equivocado pensar que M&S substitua ou enfraqueça o trabalho relacionado ao arcabouço experimental. Pelo contrário, espera-se que o arcabouço experimental seja ainda mais demandado e que essa demanda seja mais objetiva em função de lacunas de conhecimento identificadas no processo de representação do sistema para simulá-los.

<sup>1</sup> Comparison between three different approaches to implement a system dynamic model: an assessment by a multi-disciplinary team, de autoria de M. C. Freua, L. G. Barioni, R. Vilamiu, F. R. T. A. Dias a ser editado pela Embrapa Informática Agropecuária, 2014.

Além do modelo e do simulador, outros componentes de software ainda são necessários para que um software de simulação seja utilizado. Esses outros componentes estão relacionados ao pré-processamento, armazenamento e comunicação de dados de entrada e saída, bem como interfaces entre o simulador e o usuário. Existe grande variedade nas interfaces para simuladores (Figura 2).



**Figura 2.** Exemplos de possíveis interfaces de um simulador computacional com outros componentes em uma aplicação de simulação e de outros aplicativos em uma solução computacional para simulação de sistemas.

Muitos simuladores utilizam interfaces tradicionais, do tipo desktop, tais como a dos softwares DSSAT (JONES et al., 2003), Grazplan (FREER et al., 1997) e Invernada (BARIONI et al., 2010). Em outras oportunidades são utilizadas interfaces com sistemas de informação geográfica (PRIYA; SHIBASAKI, 2001), softwares estatísticos (BEN TOUHAMI et al., 2013), gerenciadores de processamento paralelo, sistemas de *workflow* (LUDÄSCHER, 2006), entre outros.

Em razão de simuladores serem mais precisos à medida que aumenta o número de iterações nos algoritmos de integração numérica, esses são frequentemente ávidos por capacidade de processamento. Frequentemente a aplicação de simuladores demanda computação de alto desempenho, envolvendo processamento paralelo em muitos processadores arranjados em *clusters*, grids ou nuvens.

Há ainda pouca exploração, na agricultura, de uso de interfaces avançadas como as encontradas em outras áreas de aplicação, tais como jogos educativos ou simuladores com atuadores mecânicos. Além disso, um simulador pode ser cliente de outro simulador na forma de pipeline. A Embrapa Informática Agropecuária também desenvolve um protótipo de jogo digital baseado em um simulador de sistemas de produção de bovinos (em fase de elaboração)<sup>2</sup>.

### 3 Métodos para o desenvolvimento de simuladores

Em muitas situações é possível que já exista um simulador adequado para o problema que se deseja. Adotar um simulador existente gera enorme economia e o uso de um determinado simulador

por várias equipes e em várias situações tende a contribuir para sua robustez e confiabilidade. É necessário, entretanto, verificar a adequação de um simulador ao problema e dados disponíveis. Algumas das principais verificações são:

1. O simulador possui as variáveis de saída (resultados) necessárias para a solução almejada?
2. Os dados requeridos para executar o simulador, i.e., os valores das variáveis de entrada, existem ou podem ser gerados com o custo e em tempo adequados?
3. A acurácia do simulador é satisfatória para o propósito do trabalho ou pode ser melhorada sem alterações estruturais?
4. A infraestrutura computacional disponível é suficiente para os resultados no tempo exigido?

Sempre que a resposta a uma dessas questões for negativa, justifica-se o trabalho de modelagem e desenvolvimento de um novo simulador.

O desenvolvimento de simuladores na agricultura, sob visão holística, envolve diversos passos ou subprocessos, que partem da estruturação do arcabouço experimental para a geração de dados, informação e conhecimento até o desenvolvimento das funcionalidades acessórias do software e, eventualmente, da aquisição de hardware necessário à simulação.

Podemos distinguir os trabalhos envolvendo simuladores na agropecuária em, pelo menos, cinco níveis com crescente complexidade:

1. **Aplicação:** O simulador é utilizado na sua forma original tendo como mudança apenas os dados de entrada para uma determinada situação.
2. **Avaliação:** O simulador é utilizado na sua forma original, mas os resultados são comparados a dados observados visando avaliar se o comportamento é adequado em determinadas situações.
3. **Calibração:** O simulador é utilizado na sua forma original, mas os parâmetros de seu modelo matemático são modificados de forma a gerar resultados os mais próximos possíveis das observações.
4. **Modelagem matemática:** O modelo do simulador é alterado estruturalmente nas suas formas funcionais, processos e componentes.
5. **Desenvolvimento de softwares de simulação:** Além do modelo, outros componentes do software de simulação são desenvolvidos desde sua concepção ou significativamente alterados.

O desenvolvimento de simuladores para a agropecuária é raro no Brasil, particularmente para aqueles que se mostram promissores para uso fora do ambiente do grupo de pesquisa no qual foi gerado. Assim, o mais frequente é termos trabalhos de pesquisa no Brasil visando ao aperfeiçoamento de simuladores desenvolvidos no exterior, com processos de M&S sendo desenvolvidos por parceiros internacionais. Embora se possa valer do conhecimento e do trabalho feito por equipes internacionais, essa dependência no processo de modelagem e codificação não pode ser considerada ideal. Essa estratégia é particularmente indesejável quando os simuladores se prestam à avaliação de temas sensíveis em nível nacional, tais como segurança alimentar e mitigação das emissões de gases de efeito estufa e nos casos nos quais se veja valor comercial na aplicação do simulador envolvendo, portanto, questões de propriedade intelectual e royalties.

Em virtude da complexidade, projetos de M&S em níveis mais abrangentes exigem eficiente gestão da equipe multidisciplinar e, normalmente, envolvem custos elevados, grande esforço e

<sup>2</sup> Desenvolvimento de um jogo educativo baseado em um simulador de sistema de produção de bovinos de corte (Resumos), de autoria de Jair Bortolucci Junior e Aduino Mancini a ser editado pela Embrapa Informática Agropecuária, 2014.

tempo. O reúso de produtos de trabalhos anteriores, tanto de modelagem quanto de geração e de organização de dados, torna-se muito importante para viabilizar muitos projetos de desenvolvimento. Simuladores agropecuários que se destacam no contexto internacional são tipicamente originários de atividades multidisciplinares, envolvendo profissionais de diferentes áreas, incluindo ciências agrárias, física, matemática, química e computação. Além disso, projetos bem-sucedidos e de longa vida útil, normalmente têm por característica eficiente feedback entre a geração de dados em campo e processos de M&S.

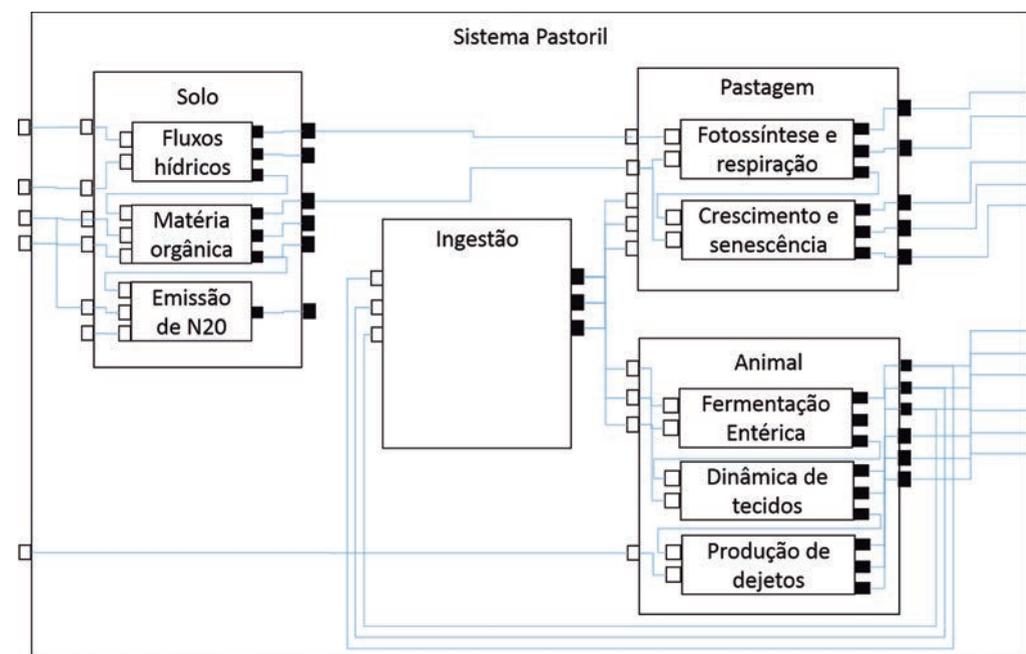
O desenvolvimento científico e tecnológico tem permitido a execução de simulações cada vez mais elaboradas. A obtenção de dados em larga escala, com aumento de resolução da grade dos dados de entrada e precisão dos valores obtidos, seja para processamento posterior ou em tempo real por meio de redes de sensores, associada ao contínuo aumento da capacidade de processamento dos computadores, está ampliando as possibilidades de monitoramento, reprodução e compreensão dos fenômenos observados. Os elementos dos sistemas de interesse podem variar em agregação, desde partículas subatômicas a galáxias. O uso de simuladores para estudar a relação entre processos e fenômenos, em diferentes níveis de agregação, para prever o comportamento de um sistema tem sido crescente, como também o número de níveis envolvidos.

Na maior parte dos casos, um ou mais modelos existem para um determinado processo a ser simulado. Uma vez que cada modelo de processo possa ser desenvolvido inicialmente de forma independente, é bastante conveniente que o modelo matemático de um simulador seja desenvolvido a partir do acoplamento de modelos componentes, facilitando a manutenção e o reúso de modelos componentes. Ademais, algumas abordagens e ferramentas de implementação de simuladores, particularmente a simulação orientada a objetos, permitem que os modelos componentes sejam agregados em estruturas hierárquicas (ZEIGLER et al., 2000), em analogia à agregação e decomposição de sistemas (WU; DAVID, 2002). Esse acoplamento de modelos desenvolvidos por diferentes equipes reimplementados para um determinado software também é realizado em importantes projetos de desenvolvimento de simuladores.

Na Figura 3, pode-se observar um exemplo de acoplamento de modelos componentes em uma estrutura hierárquica. Um modelo componente possui portas de entrada (representadas por pequenos quadrados vazios) e portas de saída (representadas por pequenos quadrados preenchidos em preto) que podem ser utilizadas para estabelecer conexões (linhas associando portas) com outros modelos. Modelos componentes podem abrigar outros componentes em tantos níveis quanto necessário. Portas de modelos componentes também podem ser associadas com portas no nível superior, permitindo que informações sejam transferidas entre níveis hierárquicos.

Olhando por essa perspectiva, Rotmans (2009) afirma que o maior desafio está, na atualidade, na combinação ou acoplamento dos modelos de forma a gerar soluções para as diferentes demandas por simulação. Seguindo filosofia similar, na Embrapa Informática Agropecuária tem-se trabalhado no desenvolvimento de componentes de software que facilitem a especificação de modelos componentes, a implementação do simulador, a execução de simulações e os processos de avaliação e calibração dos modelos.

O arcabouço (ou *framework*, do termo em inglês) de simulação MaCSim é o componente de software central para implementação de simuladores do Laboratório de Matemática Computacional da Embrapa Informática Agropecuária. É um arcabouço de simulação para implementação que



**Figura 3.** Exemplo do acoplamento em estruturas hierárquicas. Detalhes sobre portas, conexões e hierarquias são apresentados no corpo do texto.

permite o desenvolvimento de simuladores orientados a objetos e sua compilação como bibliotecas de vínculo dinâmico (MANCINI et al., 2013).

O paradigma de simulação orientada a objetos usado nessa ferramenta está alinhado com a maioria dos arcabouços de simulação utilizados atualmente como, por exemplo, JDEVS (FILIPPI; BISGAMBIGLIA, 2003); VLE (QUESNEL et al., 2007) e SimPy<sup>3</sup>. A linguagem escolhida foi C++, que oferece ampla versatilidade e eficiência do código e que possui compiladores para praticamente todos os sistemas operacionais atuais, além de ser suportada por diversos ambientes integrados de desenvolvimento gratuitos como Eclipse, NetBeans, Code::Blocks e Qt.

Os principais requisitos para a construção do arcabouço foram modularidade, suporte à hierarquia na estruturação de modelos, padronização de código e reúso de modelos componentes. Outros requisitos levantados pela equipe de desenvolvimento foram:

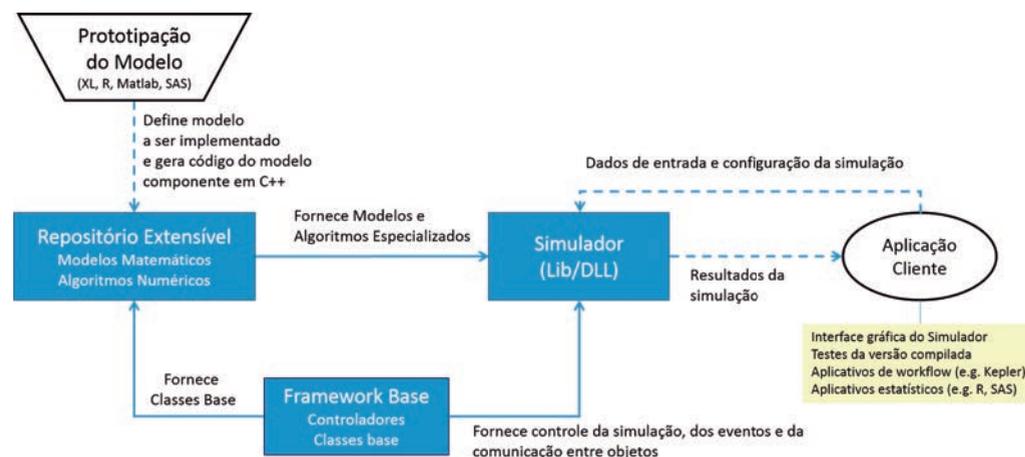
- Suporte à simulação contínua e híbrida.
- Especificação textual padronizada dos modelos para facilitar a interação com especialistas de domínio não programadores.
- Código de baixa complexidade para facilitar seu entendimento por alunos de graduação.
- Separação do código de controle da simulação do código para desenvolvimento de modelos.
- Interface simples para facilitar a comunicação entre a aplicação cliente e o simulador.
- Alto desempenho computacional via geração de arquivo binário executável.

<sup>3</sup> Disponível em: <[www.sourceforge.net/projects/simpy](http://www.sourceforge.net/projects/simpy)>.

A concepção do projeto do arcabouço de software permitiu o desenvolvimento independente dos modelos, de forma a gerar um repositório de modelos que possam ser acoplados conforme a necessidade de uma solução de simulação.

Além disso, o simulador é independente da aplicação cliente (por exemplo, uma interface gráfica específica do usuário ou um aplicativo genérico, como R). Isto é possibilitado compilando um simulador produzido no arcabouço (controle da simulação e o modelo específico do usuário) como uma biblioteca de vínculo dinâmico, que pode ser chamada pela aplicação cliente (Figura 4). Isto é útil porque:

- Um mesmo simulador pode ser usado por diferentes aplicações.
- O simulador e a aplicação cliente podem ser desenvolvidos por times diferentes usando possivelmente diferentes linguagens de programação.
- O simulador pode ser usado por desenvolvedores de software sem necessidade de conhecimento profundo sobre o código do simulador.



**Figura 4.** Fluxograma do processo de desenvolvimento e aplicação de simuladores desenvolvidos com as ferramentas desenvolvidas na suíte MaCSim.

O mesmo simulador pode ser usado para diferentes propósitos, atendendo aplicações clientes distintas (executadas independentemente) e dados da simulação podem ser armazenados de diferentes modos. Por exemplo, um simulador de uma fazenda pode ser usado para uma simulação determinística e pontual em um sistema de suporte à decisão, mas também para uma simulação em grade para uma área geográfica mais abrangente e cenários de mudança climática de longo prazo.

A separação do simulador dos outros componentes de software pode resultar em trabalho mais eficiente e especializado conduzido paralelamente por equipes dedicadas. Neste contexto, componentes de software, que não sejam o simulador propriamente, podem ser desenvolvidos por instituições parceiras. Um protocolo de comunicação padrão, definido pelo arcabouço, facilita o desenvolvimento e minimiza erros, uma vez que chamadas entre os componentes da aplicação cliente e o simulador não precisam ser redesenhadas para cada uso específico das aplicações clientes. A produtividade também pode ser aumentada, porque cada aplicação poderá usar o

mesmo protocolo com o simulador, de modo que desenvolvedores precisarão aprendê-lo apenas uma vez e não será necessário conhecimento detalhado sobre o código do simulador. Estimação de parâmetros, avaliação de modelo, simulações estocásticas, também são beneficiadas por se ter o simulador compilado como uma biblioteca e usado por uma aplicação cliente.

O arcabouço foi concebido para que a aplicação cliente se encarregue do armazenamento de dados. Esta abordagem difere daquela descrita por Bolte (1998), em que o armazenamento de dados é feito pelo software simulador. O desacoplamento dos componentes simuladores do serviço de armazenamento de dados provê flexibilidade, pois a aplicação cliente pode armazenar as saídas do modo mais adequado para atender sua demanda de análise (memória, arquivo texto ou banco de dados). Também melhora a eficiência porque evita armazenamento desnecessário dos dados de cada iteração gravando-os em disco a cada instante, por exemplo.

O MaCSim não possui interface gráfica para modelagem ou para a execução da simulação. Assim, para públicos sem treinamento em programação, seu uso para modelagem e simulação depende de ferramentas para prototipação e execução dos modelos desenvolvidos nesse framework.

Para facilitar a prototipação de modelos, desenvolveu-se uma planilha eletrônica, programada em Visual Basic para Aplicativos, que permite a especificação de modelos componentes (SILVA; BARIONI, 2013). A planilha suporta modelos estáticos e sistemas de equações diferenciais ordinárias. Essa ferramenta permite especificar, documentar, simular e gerar código dos modelos em C++ para o *framework* MaCSim e em linguagem R para prototipação mais avançada.

Conforme ilustrado na Figura 4, o *framework* de simulação necessita de uma interface para execução de simulações, estudo de cenários, calibração de modelos, entre outras funcionalidades. A Embrapa Informática Agropecuária desenvolveu, então, uma interface em linguagem R, denominada MaCSimR, para execução das bibliotecas de vínculo dinâmico contendo os simuladores. O MaCSimR permite:

- Acoplar os modelos, incluindo sua estrutura hierárquica, a partir de modelos componentes disponíveis na biblioteca do simulador.
- Carregar dados de arquivos em disco, de forma semi-automática ou utilizando funcionalidades nativas do R ou de suas extensões.
- Executar simulações.
- Recuperar trajetórias das variáveis de interesse (de entrada ou de saída) fornecendo somente o nome da variável.
- Plotar variáveis de interesse fornecendo seu nome.
- Executar experimentos com o modelo, alterando as variáveis de entrada.
- Utilizar algoritmos disponíveis em R para análises estatísticas, calibração e otimização em conjunto com simuladores compilados utilizando o MaCSim.

Experiências recentes com essas ferramentas em projetos de pesquisa da Embrapa têm demonstrado um grande potencial delas. Um exemplo de *script* no MaCSimR é apresentado na Figura 5.

```

source("MacSimR Simulation v2.R") # Runs the MacSimR script which should be in the working dir

Pecus <- Simulation(Simulator="DLLtoR.dll", Name = "Pecus")
ReadStdVal(GetInputs(Pecus), filename = "Input.txt", hasheader =F)
ReadColData(GetInputs(Pecus), "Clima_Piracicaba.csv", modelvar = "Tmin", filevar = "Tmin")
ReadColData(GetInputs(Pecus), "Clima_Piracicaba.csv", modelvar = "Tmed", filevar = "Tmed")
ReadColData(GetInputs(Pecus), "Clima_Piracicaba.csv", modelvar = "Rain", filevar = "chuva")

Run(Pecus, |
  t0 = 0,
  tf = 365,
  dt = 1,
  outtimes = seq(0, 360, by = 10),
  InitialDate = Sys.Date(),
  IntegrationMethod = "RK4",
  VarsToSave = c("Tmin", "Rain", "IME", "CH4"),
  AutoSave = T
)

```

Figura 5. Exemplo de *script* do MaCSimR para execução do modelo e plotagem de dados.

## 4 Perspectivas

Determinar as possíveis trajetórias de desenvolvimento e aplicação de simuladores na agropecuária brasileira para o futuro pode ser feito a partir de diferentes fontes de informação, dentre elas:

- Os simuladores, arcabouços de simulação e as formas de geração e organização de dados para M&S na agropecuária, já desenvolvidos ou em desenvolvimento, bem como as metodologias de desenvolvimento empregadas.
- A aplicação de simulação computacional em outras indústrias em que a simulação vem sendo aplicada há mais tempo, como na engenharia mecânica e mecatrônica, na climatologia, na neurociência e na indústria de entretenimento.
- As projeções de evolução de áreas relacionadas aos principais gargalos para o desenvolvimento de simuladores, tais como a formação de equipes multidisciplinares, a organização de dados, o desenvolvimento de software e a infraestrutura de hardware para suportar as simulações.

Temas como segurança alimentar, mitigação e adaptação às mudanças climáticas e comércio internacional têm sido os maiores demandantes de simuladores mais complexos. Nesses temas há tendência de trabalhar-se com maior número de níveis de agregação em simuladores na agropecuária, tornando frequente a busca por conexões entre observações na escala de parcela, processos biofísicos e da produção na escala da fazenda e consequências nos âmbitos regional, nacional ou global. Além disso, esses simuladores trabalham com abrangência temporal mais ampla e buscam prever os impactos de intervenções, tais como políticas públicas. Assim esses temas têm requerido simulações com abrangência de espaço e tempo muito mais amplas, gerando demanda por processamento milhares e, por vezes, milhões de vezes maiores que aquela demanda para a simulação de processos em simuladores de sistemas produtivos do início do milênio.

Em resposta às novas demandas e à expansão da aplicação da simulação na agropecuária, há atualmente forte tendência para o desenvolvimento de simuladores a partir de ferramentas que permitam alto desempenho computacional e reúso de modelos na forma de “soluções de simulação” a partir do acoplamento de modelos componentes (BERGEZ et al., 2013).

Embora simulações continuem sendo feitas a partir de populações, o uso de sensores que coletam informações fenotípicas em tempo real, aliados à parametrização de modelos com informação genômica, tem trazido à tona a possibilidade de simulação baseada em indivíduos e do uso de simuladores diretamente no melhoramento genético e na gestão de sistemas de produção de plantas e animais.

Em outras indústrias, aplicações importantes da simulação incluem o controle otimizado de sistemas, em intervenção automática por meios eletromecânicos e no treinamento de pessoas, áreas ainda incipientes na agropecuária. Além disso, em outras áreas há integração mais consistente com outras técnicas computacionais, tais como aquelas ligadas à inteligência computacional, otimização (i.e. simulação-otimização) e robótica. Observa-se que tanto a agricultura quanto a zootecnia de precisão caminham no sentido das demais indústrias mais intimamente ligadas à engenharia.

O uso de simuladores para treinamento de pessoas também parece ser bastante promissor nessa área que tem sido negligenciada na agropecuária. O treinamento por meio de jogos nos quais simuladores realistas são utilizados de forma iterativa pode promover experiência importante para gerenciar problemas de sistemas produtivos em situações de risco climático e de preços. Além disso, tal abordagem poderá também facilitar o ensino pela possibilidade de representar conceitos relacionados aos feedbacks que ocorrem em sistemas produtivos.

O desenvolvimento de simuladores agropecuários no Brasil parece ainda sofrer pela presença de vários gargalos, particularmente associados à falta de integração entre equipes de modelagem matemática, desenvolvimento de software e especialistas de domínio com pesquisa de campo em rede para o desenvolvimento de simuladores. Essa dificuldade na integração das equipes esbarra, via de regra, na falta de formação e treinamento de profissionais de ciências agrárias e pela carência de profissionais de ciências exatas com interesse em aplicações agropecuárias em detrimento de outras indústrias.

A Embrapa Informática Agropecuária tem investido, recentemente, na geração de ferramentas visando ao uso por profissionais com formação em ciências agrárias. O desenvolvimento dessas ferramentas tem ocorrido em conjunto com treinamento desses profissionais e a formação de redes de pesquisa. Tal iniciativa almeja dar suporte aos novos desafios de simulação na agricultura, por meio de uma visão holística e integrada, que parte da geração do dado até a solução de um problema de pesquisa auxiliado pelo desenvolvimento de um simulador.

## 5 Referências

- ASCOUGH, J. C.; DAVID, O.; KRAUSE, P.; HEATHMAN, G. C.; KRALISCH, S.; LAROSE, M.; AHUJA, L. R.; KIPKA, H. Development and application of a modular watershed-scale hydrologic model using the object modeling system: runoff response evaluation. *Transactions of the ASABE*, v. 55, n. 1, p. 117-135, 2012.
- BARIONI, G. L.; DAKE, C. K. G.; PARKER, W. J. Optimizing rotational grazing in sheep management systems. *Environment International*, New York, v. 25, n. 6-7, p. 819-825, Sept./Oct. 1999.
- BARIONI, L. G.; SILVA, O. D. D. da; OLIVEIRA, F. S. de. *Embrapa Invernada - Módulo de simulação de crescimento e composição de bovinos*. Versão 0.9. Campinas: Embrapa Informática Agropecuária, 2010. 1 CD-ROM.

BEN TOUHAMI, H.; LARDY, R.; KLUMPP, K.; BELLOCCHI, G. Bayesian calibration of the Pasture Simulation model (PaSim) to simulate emissions from long-term European grassland sites: a case study at Laqueuille (France). In: GREENHOUSE GASES AND ANIMAL AGRICULTURE CONFERENCE, 5., 2013, Dublin. **Proceedings...** Dublin: Cambridge University, 2013. p. 580.

BERGEZ, J. E.; CHABRIER, P.; GARY, C.; JEUFFROY, M. N.; MAKOWSKI, D.; QUESNEL, G.; RAMAT, E.; RAYNAL, H.; ROUSSE, N.; WALLACH, D.; DEBAEKE, P.; DURAND, P.; DURU, M.; DURY, J.; FAVERDIN, P.; GASCUEL-ODOUX, C.; GARCIA, F. An open platform to build, evaluate and simulate integrated models of farming and agro-ecosystems. **Environmental modelling & software**, Oxford, v. 39, p. 39-49, Jan. 2013.

BOLTE, J. Object-oriented programming for decision systems. In: PEART, R. M.; CURRY, R. B. (Ed.). **Agricultural systems modeling and simulation**. New York: Marcel Dekker, 1998. p. 629-650.

DAVID, O.; ASCOUGH, J. C.; LLOYD, W.; GREEN, T. R.; ROJAS, K. W.; LEAVESLEY, G. H.; AHUJA, L. R. A software engineering perspective on environmental modeling framework design: the object modeling system. **Environmental modelling & software**, v. 39, p. 201-213, Jan. 2013. DOI: 10.1016/j.envsoft.2012.03.006.

FILIPPI, J. B.; BISGAMBIGLIA, P. JDEVS: an implementation of a DEVS based formal framework for environmental modelling. **Environmental Modelling & Software**, Oxford, v. 19, n. 3, p. 261-274, Mar. 2003.

FREER, M.; MOORE, A. D.; DONNELLY, J. R. Grazplan: decision support systems for Australian grazing enterprises -II. The animal biology model for feed intake, production and reproduction and the GrazFeed DSS. **Agricultural Systems**, Essex, v. 54, n. 1, p. 77-126, May 1997. DOI: 10.1016/S0308-521X(96)00045-5.

ITTERSUM, M. K. van; EWERT, F.; HECKELEI, T.; WERY, J.; OLSSON, J. A.; ANDERSEN, E.; BEZLEPKINA, I.; BROUWER, F.; DONATELLI, M.; FLICHMAN, G.; OLSSON, L.; RIZZOLI, A. E.; WAL, T. van der; WIEN, J. E.; WOLF, J. Integrated assessment of agricultural systems – a component-based framework for the European Union (SEAMLESS). **Agricultural Systems**, Essex, v. 96, n. 1-3, p. 150-165, Mar. 2008. DOI: 10.1016/j.agry.2007.07.009.

JOHNSON, J. J.; TOLK, A.; SOUSA-POZA, A. A theory of emergence and entropy in systems of systems. **Procedia Computer Science**, Amsterdam, v. 20, p. 283-289, 2013. DOI: 10.1016/j.procs.2013.09.274.

JONES, J. W.; HOOGENBOOM, G.; PORTER, C. H.; BOOTE, K. J.; BATCHELOR, W. D.; HUNT, L. A.; WILKENS, P. W.; SINGH, U.; GIJSMAN, A. J.; RITCHIE, J. T. The DSSAT Cropping System Model. **European Journal of Agronomy**, Amsterdam, v. 18, n. 3-4, Jan. p. 235-265, 2003. DOI: 10.1016/S1161-0301(02)00107-7.

KEATING, C. B. Emergence in System of Systems In: Jamshidi, M. (Ed.) **System of systems engineering**: innovations for the 21st century. Hoboken: J. Wiley, 2008. p. 169-190. (Wiley series in systems engineering and management).

LUDÄSCHER, B.; ALTINTAS, I.; BERKLEY, C.; HIGGINS, D.; JAEGER-FRANK, E.; JONES, M.; LEE, E.; TAO, J.; ZHAO, Y. Scientific workflow management and the kepler system. special issue: workflow in grid systems. **Concurrency and Computation: Practice & Experience**, Chichester, v. 18, n 10, p.1039-1065, Aug. 2006. DOI: 10.1002/cpe.v18:10.

MANCINI, A. L.; BARIONI, L. G.; SANTOS, E. H. dos; DIAS, F. R. T.; SANTOS, J. W. dos; ABREU, L. L. B. de; TININI, L. V. S. **Arcabouço para desenvolvimento de simuladores de sistemas dinâmicos contínuos e hierárquicos**. Campinas: Embrapa Informática Agropecuária, 2013. 19 p. (Embrapa Informática Agropecuária. Boletim de pesquisa e desenvolvimento, 34). Disponível em: <<http://www.infoteca.cnptia.embrapa.br/bitstream/doc/981039/1/BolPesq34cnptia.pdf>>. Acesso em: 20 set. 2014.

MCCOWN, R. L.; HAMMER, G. L.; HARGREAVES, J. N. G.; HOLZWORTH, D. P.; FREEBAIRN, D. M. APSIM: a novel software system for model development, model testing and simulation in agricultural systems research. **Agricultural Systems**, Essex, v. 50, n. 3, p. 255-271, 1996. DOI: 10.1016/0308-521X(94)00055-V.

NUTARO, J. J. **Building software for simulation**: theory and algorithms, with applications in C++. Hoboken: J. Wiley, 2011. 347 p.

OLTJEN, J. W.; BYWATER, A. C.; BALDWIN, R. L.; GARRETT, W. N. Development of a dynamic model of beef cattle growth and composition. **Journal of Animal Science**, Champaign, v. 62, n 1, p. 86-97, Jan. 1986.

PEZZULO, G.; CANDIDI, M.; DINDO, D.; BARCA, L. Action simulation in human brain: twelve questions. **New Ideas in Psychology**, Oxford, v. 31, n. 3, p. 270-290, Dec. 2013. DOI: 10.1016/j.newideapsych.2013.01.004.

PRIYA, S.; SHIBASAKI, R. National spatial crop yield simulation using GIS-based crop production model. **Ecological Modelling**, Amsterdam, v. 136, n. 2-3, p. 113-129, Jan. 2001. DOI: 10.1016/S0304-3800(00)00364-1.

QUESNEL, G.; DUBOZ, R.; RAMAT, E.; TRAORÉ, M. K. VLE: a multimodeling and simulation environments. In: SUMMER COMPUTER SIMULATION CONFERENCE, 2007, San Diego. **Proceedings...** San Diego: Society for Computer Simulation International, 2007. p. 367-374. SCSC'07.

ROTMANS, J. Three decades of integrated assessment: the way forward. Introductory Keynote. In: INTEGRATED ASSESSMENT OF AGRICULTURE AND SUSTAINABLE DEVELOPMENT CONFERENCE, 2009, Wageningen. **Setting the agenda for science and policy**: proceedings. Wageningen: Wageningen University, 2009. p. 1-35. Paper. AgSAP 2009.

SILVA, R. D. R. da; BARIONI, L. G. **Prototipação de modelos e geração automática de código-fonte em planilha eletrônica para o framework de modelagem MaCSim**. Campinas: Embrapa Informática Agropecuária, 2013. 13 p. (Embrapa Informática Agropecuária. Boletim de pesquisa e desenvolvimento, 35).

VERBURG, P. H.; SCHOT, P. P.; DIJST, M. J.; VELDKAMP, A. Land use change modelling: current practice and research priorities. **GeoJournal**, Dordrecht, v. 61, n. 4, p. 309-324, 2004. DOI: 10.1007/s10708-004-4946-y.

WOODWARD, S. J. R.; ROMERA, A. J.; BESKOW, W. B.; LOVATT, S. J. Better simulation modelling to support farming systems innovation: review and synthesis. **New Zealand Journal of Agricultural Research**, Wellington, v. 51, n. 3, p. 235-252, 2008. DOI: 10.1080/00288230809510452.

WU, J.; DAVID, J. L. A spatially explicit hierarchical approach to modeling complex ecological systems: theory and applications. **Ecological Modelling**, Amsterdam, v. 153, n. 1-2, p. 7-26, July 2002. DOI: 10.1016/S0304-3800(01)00499-9.

ZAGMUTT, F. J.; SEMPIER, S. H.; HANSON, T. R. Disease spread models to estimate highly uncertain emerging diseases losses for animal agriculture insurance policies: an application to the US farm-raised catfish industry. **Risk Analysis**, New York, v. 33, n. 10, p. 1924-1937, Oct. 2013. DOI: 10.1111/risa.12038.

ZEIGLER, B. P.; PRAEHOFER, H.; KIM, T. G. 2nd ed. **Theory of modeling and simulation**: integrating discrete event and continuous complex dynamic systems. San Diego: Academic Press, 2000. 510 p. ill.



## Mineração de dados: conceitos e um estudo de caso sobre certificação racial de ovinos

Fábio Danilo Vieira  
Stanley Robson de Medeiros Oliveira

### 1 Introdução

#### 1.1 Descoberta de conhecimento em banco de dados

Nos últimos anos, observa-se que uma grande quantidade de dados cresce de forma acelerada em diversos campos de conhecimento, fato que dificulta a sua interpretação, pois o volume destes dados é maior que o poder de interpretá-los. Desta forma, surgiu a necessidade do desenvolvimento de ferramentas e técnicas automatizadas para minimizar esta situação, as quais pudessem auxiliar o analista a transformar os dados em conhecimento (HAN et al., 2011).

Grande parte dessas técnicas e ferramentas podem ser encontradas no processo de Descoberta de Conhecimento em Bases de Dados, da sigla em inglês *Knowledge Discovery in Databases* (KDD). Segundo Fayyad et al. (1996), a descoberta de conhecimento em bancos de dados é definida como um processo não trivial que busca identificar padrões novos, potencialmente úteis, válidos e compreensíveis, com o objetivo de melhorar o entendimento de um problema ou um procedimento de tomada de decisão.

O processo KDD se originou da intersecção de várias áreas de pesquisa, tais como aprendizado de máquina, reconhecimento de padrões, estatística, banco de dados, visualização de dados, inteligência artificial e computação de alto desempenho (FAYYAD et al., 1996). Por este motivo, as técnicas existentes no KDD não devem ser consideradas substitutas de outras formas de análise, como *Online analytical processing* (Olap), mas, sim, uma forma de se aperfeiçoar os resultados obtidos por meio das explorações realizadas pelas ferramentas atuais (REZENDE et al., 2003).

As aplicações das técnicas estão presentes em praticamente todos os setores do conhecimento humano. Na área de negócios, por exemplo, utilizam-se técnicas em detecção de fraudes em cartões, na criação de perfis de clientes de acordo com suas compras, entre outros. Na agricultura, podem ser utilizadas, em previsão de geadas, sistemas de alerta para a ferrugem do cafeeiro, sistemas de alerta para a ferrugem asiática da soja, etc. Na medicina, pode-se identificar terapias médicas de sucesso para diversas doenças. Na bioinformática, para se buscar padrões em sequências de DNA, entre muitas outras possibilidades.

Segundo Fayyad et al. (1996), o processo de KDD é interativo e iterativo, além de envolver vários passos, exibidos na Figura 1, com muitas decisões sendo feitas pelo especialista do domínio dos dados.

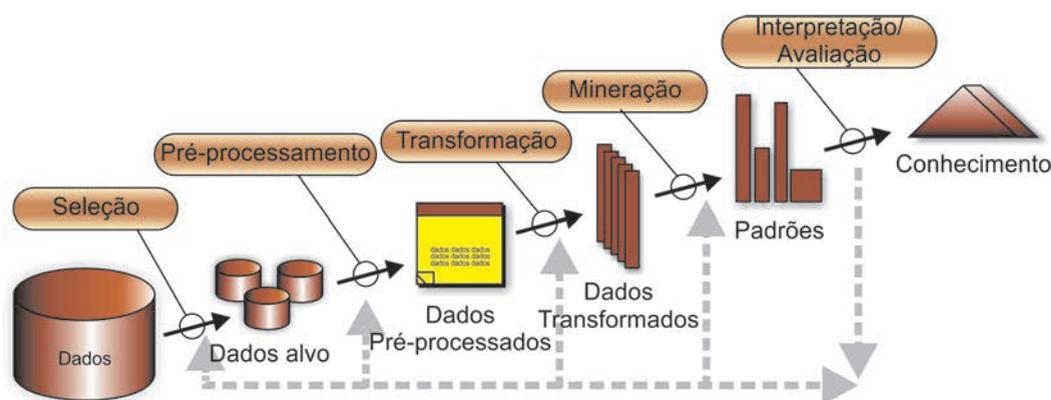


Figura 1. As fases do processo KDD.

Fonte: adaptado de Fayyad et al. (1996).

Os passos do processo KDD consistem em:

- 1) **Identificação do problema:** compreensão do domínio da aplicação e do tipo de conhecimento a ser procurado, além de se identificar o objetivo do processo KDD.
- 2) **Criação do conjunto de dados alvo (Seleção):** realizar a seleção de um conjunto de dados, ou se fixar num subconjunto de registros (instâncias), onde a descoberta deve ser feita.
- 3) **Limpeza de dados e pré-processamento (Pré-processamento):** neste passo estão operações básicas como remoção de ruídos quando necessário, coleta da informação necessária para modelar ou prever ruído, e decisão sobre quais estratégias se adotar para tratar atributos com valores faltantes.
- 4) **Redução de dados e projeção (Transformação):** busca por características úteis que possam representar os dados dependendo do objetivo da tarefa, visando à redução de dimensionalidade, ou seja, redução do número de atributos e/ou registros a serem considerados para o conjunto de dados.
- 5) **Mineração de dados (Mineração):** escolha do(s) algoritmo(s) de mineração de dados e de métodos a serem aplicados para a busca por padrões de interesse numa forma particular de representação ou conjunto de representações.
- 6) **Interpretação dos padrões descobertos (Interpretação/Avaliação):** realizam-se análises dos padrões descobertos com o objetivo de descobrir se estes apresentam conhecimento novo em aplicações práticas. Algumas vezes, há a necessidade de se retornar aos passos 1-6 para avaliação posterior.
- 7) **Implantação do conhecimento descoberto (Conhecimento):** incorporação deste conhecimento à performance do sistema ou, simplesmente, documentá-lo e reportá-lo às partes interessadas.

## 1.2 Tarefas e técnicas de mineração de dados

Uma tarefa de mineração de dados consiste na especificação do que se pretende buscar, ou que tipo de regularidade ou padrões interessa encontrar.

Na etapa de mineração de dados propriamente dita deve ser feita a escolha da tarefa a ser empregada, assim como a definição do algoritmo. Esta escolha deve ser baseada nos objetivos que se deseja atingir com a solução a ser encontrada. As possíveis tarefas de um algoritmo para se extrair padrões podem ser agrupadas em preditivas e descritivas (HAN et al., 2011), como ilustradas na Figura 2.

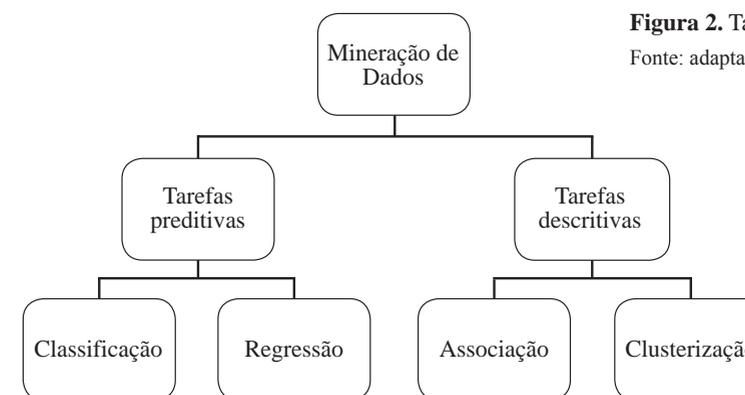


Figura 2. Tarefas de Mineração de Dados.

Fonte: adaptado de Rezende et al., 2003.

As tarefas preditivas têm como objetivo principal a construção de modelos que possam prever a classe de um novo evento a partir de exemplos ou experiências passadas com respostas já conhecidas. As tarefas descritivas procuram identificar padrões intrínsecos a um conjunto de dados que não possui uma classe determinada. A escolha de uma ou mais tarefas irá depender do problema a ser solucionado. As tarefas tradicionais de mineração de dados representadas na Figura 2 são brevemente descritas a seguir.

- **Classificação:** consiste na predição do valor de um atributo alvo do tipo discreto ou categórico por meio da construção de modelos e regras a partir de um conjunto de exemplos pré-classificados corretamente, para posterior classificação de exemplos novos e desconhecidos (HAN et al., 2011). O grande desafio para os algoritmos de classificação é gerar modelos que possuam boa capacidade de generalização, ou seja, que estejam aptos a prever, com alta taxa de acerto, os rótulos das classes para registros que não foram utilizados durante a construção do modelo (TAN et al., 2005).
- **Regressão:** técnica estatística muito empregada para se realizar predições (HILL et al., 2003). Essas predições procuram encontrar tendências de variações no conjunto de dados analisado em função dos atributos existentes. Possui um conceito semelhante à classificação, porém se aplica na predição de um valor alvo do tipo contínuo.
- **Associação:** determinam o quanto a presença de um certo conjunto de atributos nos exemplos de uma base de dados implica na presença de algum outro conjunto de atributos nos mesmos exemplos (AGRAWAL; SRIKANT, 1994). As regras de associação podem ser apresentadas no formato  $l \rightarrow r$ , onde  $l$  e  $r$  são, respectivamente, o lado esquerdo (*Left Hand Side*) e o lado direito (*Right Hand Side*), tal que  $l \cap r = \emptyset$ , de forma que representam conjuntos distintos de atributos. Basicamente, essas regras definem a relação existente entre  $l$  e  $r$ , demonstrando o quanto a presença de  $l$  implica a presença de  $r$ .
- **Agrupamento (clusterização):** é uma tarefa descritiva que procura identificar agrupamentos (*clusters*) finitos de objetos similares entre si e dissimilares entre os grupos no conjunto de

dados. De forma diferente da classificação, onde as denominações de classes são conhecidas, a clusterização analisa os dados onde as denominações de classes não estão definidas.

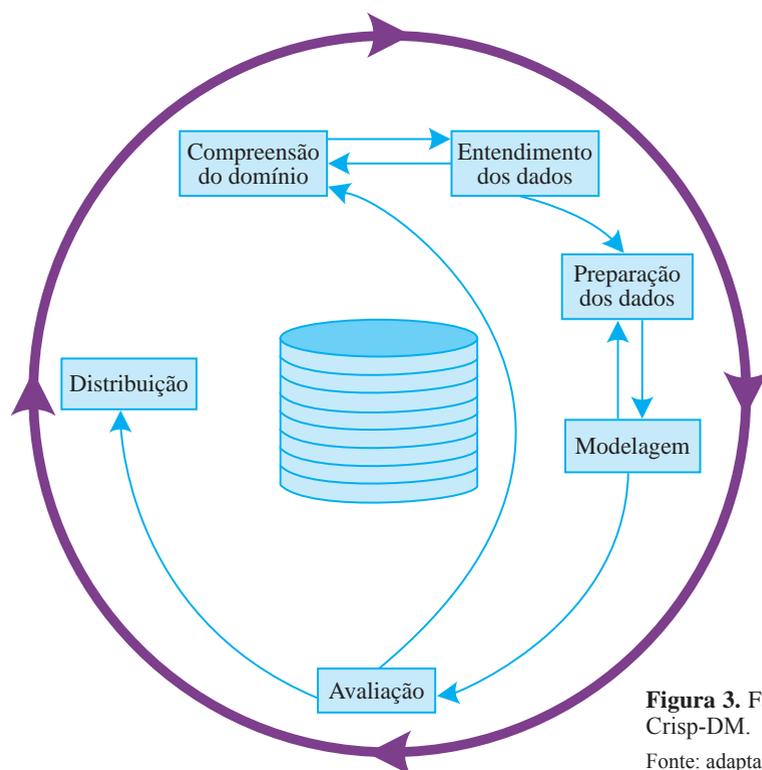
Cada tarefa de mineração de dados possui diferentes técnicas associadas. Dentre as mais populares estão (HAN et al., 2011): árvores de decisão, redes neurais, regressão linear ou não linear, k-vizinhos mais próximos. Existem também as abordagens híbridas, que utilizam duas ou mais técnicas em conjunto.

Não existe a técnica ideal, cada uma delas possui suas vantagens e desvantagens. Assim, ao se escolher uma técnica, deve ser realizada uma análise bem apurada do problema em questão, levando em consideração o formato dos dados e como o conhecimento descoberto pode ser representado. Se necessário, pode se aplicar mais de uma técnica para solucionar o mesmo problema e no final escolher o modelo que apresente os melhores resultados.

### 1.3 Modelo do processo de descoberta de conhecimento em bases de dados

Com o objetivo de padronizar o processo de descoberta de conhecimento, em 1996 foi criado o modelo de processo *CRoss-Industry Standart Process for Data Mining* (Crisp-DM), que divide o ciclo de vida de um projeto de mineração de dados em seis fases, a saber: compreensão do domínio, entendimento dos dados, preparação dos dados, modelagem, avaliação e distribuição (CHAPMAN et al., 2000).

As fases do modelo do processo estão ilustradas na Figura 3. O círculo externo traduz o aspecto cíclico de um projeto de mineração de dados, uma vez que após encontrar uma solução, o projeto não é necessariamente finalizado, e a partir de novos conhecimentos adquiridos podem ocorrer



**Figura 3.** Fases do modelo de processo Crisp-DM.

Fonte: adaptado de Chapman et al. (2000).

novos questionamentos que levam a ações mais específicas. As setas internas ilustram as relações entre as fases, que indicam que a sequência entre elas não é rígida, sendo comum haver a necessidade de voltar ou avançar entre as fases.

A seguir, encontra-se uma breve descrição de cada fase do processo:

- **Compreensão do domínio:** compreender os objetivos e requisitos do projeto e transformar esse conhecimento em um problema de mineração de dados. Definir um plano preliminar para atingir esses objetivos.
- **Entendimento dos dados:** inicia-se com uma coleção de dados inicial e prossegue com atividades de exploração de dados, para se familiarizar, identificar problemas de qualidade, fazer as primeiras hipóteses e identificar possíveis subconjuntos que possam abrigar informações ocultas sobre esses dados.
- **Preparação dos dados:** a fase de preparação dos dados contempla todas as atividades necessárias para a construção do conjunto de dados final, no qual serão aplicadas as técnicas de modelagem. As atividades incluem, por exemplo, limpeza de dados, seleção e transformação de atributos, entre outras.
- **Modelagem:** nessa fase são escolhidas e aplicadas as técnicas de mineração de dados, e seus parâmetros são calibrados. Diversas técnicas podem ser aplicadas ao mesmo problema, embora cada técnica necessite de formatos específicos e necessite voltar para a fase de preparação de dados.
- **Avaliação:** nesse estágio, tem-se o modelo (ou modelos) com boa qualidade. Os resultados são comparados e interpretados conforme a área de aplicação. É importante reavaliar todas as etapas do processo para se ter a certeza de que o modelo atende às necessidades e aos objetivos do projeto.
- **Distribuição:** a criação de modelos geralmente não finaliza um projeto. O conhecimento obtido deve ser documentado, organizado e apresentado para os usuários, para que estes possam saber quais ações devem ser realizadas para aproveitar os modelos criados.

## 2 Estudo de caso

### 2.1 Modelagem para certificação racial de ovinos

O Brasil possui diversas raças de ovinos que foram desenvolvidas a partir de raças trazidas pelos colonizadores portugueses, logo após o descobrimento. Ao longo desses quase cinco séculos, essas raças foram submetidas à seleção natural em diversos ambientes, a ponto de desenvolverem características de adaptação às diversas condições ambientais brasileiras. Essas raças passaram a ser conhecidas como crioulas ou localmente adaptadas. A maioria dessas raças encontra-se ameaçada de extinção, principalmente devido a cruzamentos indiscriminados com animais de raças exóticas que passaram a ser importadas a partir do final do século XIX (MARIANTE et al., 2009).

As raças localmente adaptadas, apesar de não possuírem o mesmo potencial produtivo das raças exóticas melhoradas, constituem uma importante fonte de informações que pode levar à descoberta de genes envolvidos com determinadas características adaptativas, tais como resistência a

diversas doenças e parasitas. Essas características permitem que os animais destas raças sejam mais adaptados que ovinos de outras raças (inclusive de raças exóticas melhoradas) a regiões de ambientes mais hostis. Essas informações fornecem um caminho muito interessante para futuras investigações, principalmente no entendimento da base genética envolvida na adaptação a estes ambientes (GOUVEIA, 2013).

Para evitar a perda deste importante e insubstituível material genético, a Empresa Brasileira de Pesquisa Agropecuária (Embrapa) decidiu incluir as raças localmente adaptadas no seu Programa de Pesquisa em Recursos Genéticos. Atualmente, a conservação dos recursos genéticos animais é realizada em bancos de germoplasma, que podem ser compostos de pequenos rebanhos de animais de uma raça que ficam submetidos à seleção natural (*in situ*), ou de material genético congelado, como sêmen, embriões e ovócitos (*ex situ*). Diversas raças localmente adaptadas estão presentes nestes bancos, sendo que as que possuem maior destaque nacional são as raças Crioula, Morada Nova e Santa Inês.

A seleção dos ovinos de uma determinada raça para compor estes bancos é realizada por meio de critérios tradicionais, tais como a avaliação de características morfológicas e produtivas. Entretanto, essa avaliação está sujeita a falhas, pois alguns animais cruzados mantêm características semelhantes àquelas dos animais locais. Com isto, identificar se os animais depositados no banco são ou não pertencentes a uma raça é uma tarefa que exige muita cautela.

Para auxiliar na busca de soluções para este tipo de problema, o emprego de tecnologias advindas das áreas da genética e da computação é fundamental para atingir resultados mais precisos e confiáveis. Nos últimos anos, houve um aumento na utilização de tecnologias que empregam análise do *deoxyribonucleic acid* (DNA) na área animal, sendo que as que fazem uso de marcadores moleculares baseados em polimorfismos de DNA se destacam entre as mais importantes.

Dentre os tipos de marcadores moleculares existentes, os do tipo *Single Nucleotide Polymorphism* (SNP) mostraram ser mais efetivos no auxílio da certificação racial de animais domésticos (PANT et al., 2012; SASAZAKI et al., 2011; SUEKAWA et al., 2010). Os marcadores SNP constituem uma variação que ocorre em apenas um único nucleotídeo da cadeia de bases nitrogenadas (Adenina, Citosina, Timina e Guanina) do DNA, afetando ou não o fenótipo alvo entre os membros de uma espécie em estudo. Atualmente, as novas tecnologias para geração destes dados moleculares fornecem metodologias que são capazes de genotipar de dezenas até centenas de milhares de marcadores SNP em microarranjos (*microarrays*) de DNA de alta densidade em um único ensaio.

Desta forma, selecionar os marcadores mais informativos para a identificação racial de um ovino torna-se um problema desafiador. Uma das formas de se realizar esta seleção é por meio de um processo de mineração de dados, cujo objetivo é encontrar padrões e tendências em grandes volumes de dados (HAN et al., 2011). Esse processo permite identificar e estudar o conjunto dos principais marcadores SNP. Para tanto, deve-se utilizar técnicas específicas que combinem seleção de atributos (ou variáveis) e geração de modelos preditivos. Estas técnicas devem ser capazes de gerar modelos que classifiquem novos exemplos a partir de experiências acumuladas em problemas anteriores e de lidar com problemas em que o número de atributos ( $p$ ) é muito maior que o número de observações ( $n$ ). De acordo com James et al. (2013), a combinação dessas técnicas contribuem para eliminação de atributos redundantes e não-informativos, simplificam o

modelo preditivo e reduzem o custo de processamento do algoritmo de aprendizado de máquina para construção do modelo.

Os modelos obtidos pelo processo de mineração de dados poderão ser utilizados na certificação racial dos animais já depositados nos bancos de germoplasma, e de novos animais a serem incluídos, assim como poderão ser utilizados por diversos segmentos ligados à ovinocultura, como por exemplo, por associações de criadores interessadas em certificar seus animais, e pelo Ministério Agricultura Pecuária e Abastecimento (Mapa), no controle de animais registrados que apresentam alelos de outras raças, possibilitando a reclassificação ou mesmo a revogação desses animais registrados.

Além disso, os marcadores SNP selecionados pelos modelos poderão ser empregados na construção de ferramentas de genotipagem de marcadores SNP de baixa densidade, como os microarranjos, por exemplo (KIM; MISRA, 2007; ROORKIWAL et al., 2013). Cabe ressaltar que, quanto menor o número de marcadores selecionados, menor o custo total de construção destas ferramentas de genotipagem, pois a preparação de cada SNP no arranjo custa um determinado valor (CAETANO, 2009).

## 2.2 Metodologia

O exemplo utilizado para demonstrar o processo de aplicação de técnicas de mineração de dados aborda o desenvolvimento de modelos para selecionar os principais marcadores SNP na identificação racial de animais pertencentes às raças Crioula, Morada Nova e Santa Inês. Para dar suporte aos procedimentos realizados no exemplo, optou-se por seguir o modelo de processo Crisp-DM, já explicado anteriormente. Cada uma das seis fases do processo de análise do Crisp-DM, para o estudo de caso analisado, estão descritas a seguir.

**Fase 1 - Compreensão do domínio:** Considerando que o objetivo principal do exemplo utilizado é o desenvolvimento de modelos baseados em técnicas de mineração de dados para selecionar os principais marcadores SNP para as raças Crioula, Morada Nova e Santa Inês, foi realizada uma pesquisa contínua em busca de conhecimentos sobre as características destas raças, bem como a situação atual da ovinocultura e suas projeções no cenário nacional e internacional. Também foram estudados alguns conceitos sobre os assuntos concernentes a marcadores moleculares SNP e genética populacional, procurando compreender suas aplicações dentro do campo da genômica animal. Além disso, buscou-se entender alguns conceitos relacionados às técnicas de mineração de dados apropriadas para o problema da pesquisa.

**Fase 2 - Entendimento dos dados:** O conjunto de dados analisado no exemplo foi obtido do Consórcio Internacional do Genoma Ovino (ARCHIBALD et al., 2010) por meio da Rede Genômica Animal, projeto da Embrapa. Este conjunto era composto por 72 animais das raças estudadas (23 animais da raça Crioula, 22 da Morada Nova e 27 da Santa Inês), sendo que cada animal possuía 49.034 marcadores SNP. Observa-se, então, que o conjunto de dados é uma matriz em que o número de marcadores ( $p$ ) é muito maior que o número de instâncias ( $n$ ), isto é,  $p \gg n$ . Cada um desses marcadores SNP possui um valor de genótipo, que é composto por dois alelos, sendo que cada alelo pode conter uma Adenina (A) ou uma Timina (T) ou uma Citosina (C) ou uma Guanina (G). A Figura 4 ilustra o formato do conjunto de dados de ovinos em estudo.

	SNP 1	SNP 2	SNP 3	SNP 4	SNP 5	SNP 6	...	Raça
72 animais	AA	AG	AG	AG	AG	CC	AC	Crioula
	GA	AG	AG	GG	GG	AC	AC	Morada Nova
	GA	GG	AG	GG	AG	CC	CC	Santa Inês
	...	...	...	...	...	...	...	...

49.034 SNP

**Figura 4.** Formato do conjunto de dados de marcadores SNP das três raças em estudo.

**Fase 3 - Preparação dos dados:** Na etapa de preparação dos dados, realizou-se uma verificação quanto à existência de amostras idênticas dentro do conjunto de dados e de marcadores SNP que tivessem um valor único de genótipo para todas as raças. Após a verificação, constatou-se que não existiam amostras idênticas. Entretanto, existiam 384 marcadores SNP com valor único para todas as raças, os quais foram removidos do conjunto de dados final.

**Fase 4 - Modelagem:** Na etapa da modelagem do exemplo utilizado, foram aplicadas técnicas que combinam seleção de atributos e desenvolvimento de modelos preditivos para identificar os marcadores SNP mais relevantes para três raças de ovinos. Entretanto, devido ao elevado número de atributos (marcadores SNP) e o baixo número de registros (animais), técnicas capazes de lidar com esta situação foram empregadas, a saber: *Least Absolute Shrinkage and Selection Operator* (Lasso), Random Forest e Boosting.

Lasso é um método de regressão penalizada utilizado para reduzir os efeitos dos atributos que não contribuem para identificação da classe (atributo-meta ou variável resposta), encolhendo seus coeficientes para zero e excluindo-os do modelo (TIBSHIRANI, 1997). O método é usado normalmente para estimar os parâmetros de regressão  $\beta = (\beta_1, \beta_2, \dots, \beta_p)^T$  no modelo da Equação 1:

$$y_i = \mu + \sum_{j=1}^p x_{ij} \beta_j + e_i = \mu + X_i \beta + e_i \quad (1)$$

onde,  $y_i$  é a raça do  $i$ -ésimo animal ( $i = 1, 2, \dots, n$ );  $\mu$  é o coeficiente denominado intercepto, cujo valor é comum a todos os registros;  $x_{ij}$  é o valor do genótipo do marcador  $j$  ( $j = 1, 2, \dots, p$ ) do animal  $i$ ; o coeficiente  $\beta_j$  representa o efeito do marcador  $j$  na raça;  $e_i$  é o erro residual. Em problemas de classificação, Lasso estima os coeficientes  $\beta_j$  do modelo por meio da maximização do logaritmo da função de verossimilhança, impondo a restrição de que a soma dos valores dos coeficientes absolutos seja limitada por uma constante (HASTIE et al., 2011).

Sendo  $\hat{\beta} = (\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_p)^T$ , a estimativa Lasso para problemas de classificação é definida pela função de máxima verossimilhança penalizada descrita na Equação 2:

$$l(\hat{\mu}, \hat{\beta}) = \frac{1}{n} \sum_{i=1}^n [y_i ( + \sum_{j=1}^p x_{ij} \beta_j ) - \log ( 1 + e^{\mu + \sum_{j=1}^p x_{ij} \beta_j } )] \quad (2)$$

$$\text{Sujeito à restrição } \sum_{j=1}^p |\beta_j| \leq t \text{ para } t \geq 0,$$

onde  $t$  é um parâmetro de penalização, também representado pela letra grega  $\lambda$  (lambda) em outra formulação da equação, e que deve ser determinado separadamente. Normalmente, os algoritmos de implementação do Lasso fornecem o valor ótimo para tal parâmetro, utilizando uma análise por validação cruzada de um intervalo de  $n$  possíveis valores.

Random Forest é uma técnica de classificação e regressão desenvolvida por Breiman (2001), que consiste num conjunto de árvores de decisão combinadas para solucionar problemas de classificação. Cada árvore de decisão é construída utilizando uma amostra aleatória inicial dos dados e, a cada divisão desses dados, um subconjunto aleatório de  $m$  atributos é utilizado para escolha dos atributos mais informativos. No final, Random Forest gera uma lista dos atributos mais importantes no desenvolvimento da floresta, que são determinados pela importância acumulada do atributo nas divisões dos nós de cada árvore da floresta (JAMES et al., 2013). Os principais passos do algoritmo Random Forest podem ser vistos na Figura 5.

Dado um conjunto de dados  $X = x_1, x_2, \dots, x_j$  e  $Y = y_1, y_2, \dots, y_k$ .

Para  $b = 1, 2, 3, \dots, B$ , repita:

- Cria uma amostra *bootstrap*  $(X_b, Y_b)$  com  $n$  exemplos de  $(X, Y)$ .
- Ajusta uma árvore de decisão  $f^b$  para o conjunto de treinamento  $(X_b, Y_b)$ , utilizando  $m$  atributos para a escolha de cada nó.

Fim de repetição.

Gera o modelo final:  $\hat{f}(x) = \sum_{b=1}^B f^b(x)$ , que calcula os votos obtidos pelos modelos  $f^b$ , resultando uma classificação final de acordo com a votação majoritária.

**Figura 5.** Algoritmo básico da técnica Random Forest.

Fonte: Breiman (2001).

De uma forma geral, uma árvore de decisão é um modelo gráfico representado por nós e ramos, onde os nós intermediários, ou decisórios, representam os testes de atributos (variáveis independentes), enquanto que os ramos representam os resultados desses testes. O nó localizado no topo da árvore representa seu início e é denominado nó-raiz. Já o nó externo, que não possui um nó descendente, localizado na extremidade inferior, é denominado folha ou terminal, e representa o valor de predição do atributo-meta ou classe (HAN et al., 2011). Para evitar *overfitting*, foi utilizada a abordagem Random Forest que, em geral, lida melhor com o problema de sobreajuste nos modelos (MEGETO et al., 2014).

A ideia principal da técnica Boosting é transformar múltiplos classificadores ruins em um único muito bom (FREUND; SCHAPIRE, 1999). Os métodos desta abordagem funcionam aplicando-se sequencialmente um algoritmo de classificação a versões reponderadas do conjunto de dados de treinamento, dando maior peso aos registros classificados erroneamente no passo anterior. O algoritmo que mostra a execução básica da técnica Boosting é descrito na Figura 6.

Para aplicação das técnicas de modelagem, escolheu-se o software R (versão 3.0.1). O pacote instalado para o algoritmo Lasso foi o glmnet (FRIEDMAN et al., 2010), para Random Forest foi instalado o pacote randomForest (LIAW; WIENER, 2002) e, para Boosting, foi instalado o algoritmo gbm (RIDGEWAY, 2013). Além destes, instalou-se o pacote caret (KUHN, 2013), utilizado para a escolha dos melhores valores para alguns parâmetros de cada técnica aplicada.

Dado um conjunto de dados de treinamento  $X = x_1, x_2, \dots, x_j$  e  $Y = y_1, y_2, \dots, y_k$ .

Define  $\hat{f}(x) = 0$  e  $resíduos_i = y_i$  para todos os registros do treinamento.

Para  $b = 1, 2, 3, \dots, B$ , repita:

- Ajusta um modelo  $f^b$  para o conjunto de treinamento  $(X, resíduos)$ .
- Atualiza  $\hat{f}$  com o novo modelo:  

$$\hat{f}(x) = \hat{f}(x) + f^b(x).$$
- Atualiza os resíduos (erros na classificação):  

$$resíduos_i = resíduos_i - f^b(x_i).$$

Fim de repetição.

Gera o modelo final:  $\hat{f}(x) = \sum_{b=1}^B f^b(x)$ , que calcula os votos obtidos pelos modelos  $f^b$ , resultando uma classificação final de acordo com a votação majoritária.

Figura 6. Algoritmo básico do algoritmo Boosting.

Fonte: James et al. (2013).

Foram realizados vários experimentos utilizando cada uma das técnicas, procurando obter modelos que fornecessem os melhores resultados em termos de acurácia e menor número de marcadores selecionados. Para tanto, antes da utilização do pacote caret, os principais parâmetros de cada uma das técnicas foram ajustados diversas vezes para atingir tal objetivo.

Lasso foi a primeira técnica a ser aplicada, e o único parâmetro testado foi o intervalo de possíveis valores para o coeficiente de penalização  $\lambda$  (lambda). O número padrão deste intervalo é de 100 valores possíveis (FRIEDMAN et al., 2010; JAMES et al., 2013), obtidos separadamente pelo algoritmo Lasso, via validação cruzada, sobre os dados analisados. Após a aplicação da técnica Lasso, utilizou-se Random Forest para a busca dos marcadores SNP mais relevantes, associados a cada uma das raças. Os parâmetros avaliados para Random Forest foram o número de árvores a serem construídas e o número de atributos selecionados para determinar o *split* em cada nó das árvores. Com a construção desta floresta, foi possível determinar os marcadores mais importantes para o modelo (do atributo mais importante ao menos relevante). Assim como Random Forest, Boosting foi utilizado para fornecer um modelo com a listagem dos marcadores mais importantes na identificação das raças. O único parâmetro testado para Boosting foi o número de classificadores a serem desenvolvidos para o modelo final. Os classificadores construídos pela técnica Boosting foram baseados em árvores de decisão, as quais foram construídas em distribuições ponderadas dos dados.

Após a obtenção dos modelos e dos conjuntos de marcadores mais importantes para identificação das raças, foi realizada uma análise da frequência alélica de cada um desses marcadores, a fim de verificar o quanto um alelo estava presente em uma raça e ausente em outras duas. Por fim, foi selecionado um subconjunto menor de marcadores SNP com maior potencial de identificação das três raças pesquisadas.

**Fase 5 - Avaliação:** nesta etapa do processo, os modelos já foram desenvolvidos. Antes de passar a fase final de desenvolvimento dos modelos finais, é importante revisar todos os passos executados, para verificar se os objetivos foram alcançados. No fim dessa etapa, uma decisão a respeito do uso dos resultados da análise deve ser tomada.

Para avaliar o desempenho dos modelos, dividiu-se o conjunto de dados inicial em duas partes disjuntas, sendo que uma parte constitui o conjunto de treinamento e outra o conjunto de teste. As técnicas utilizaram dois tipos de particionamento dos dados: validação cruzada e *bootstrap*. Na validação cruzada, os dados são particionados em  $k$  sub-conjuntos de tamanhos aproximadamente iguais, e o indutor é treinado e testado  $k$  vezes. Para cada uma das vezes, o indutor é testado com uma das partições e treinado com o restante. O *bootstrap* consiste em gerar os conjuntos de treinamento e teste a partir de uma amostragem randômica dos dados, repetindo esse processo de classificação por várias vezes. A cada ciclo, as amostragens são selecionadas com reposição, isto é, um mesmo exemplo poderá aparecer mais de uma vez no mesmo sub-conjunto.

Os modelos foram analisados por meio dos valores da acurácia e do coeficiente Kappa. A acurácia, ou taxa de acerto, fornece a porcentagem de observações que foram classificadas corretamente pelo classificador, enquanto o Kappa (COHEN, 1960) mede o grau de concordância entre as classes preditas e observadas, deduzindo o número esperado de acertos (utilizando uma classificação ao acaso) do número real de acertos do classificador (WITTEN et al., 2011).

**Fase 6 - Distribuição:** a construção dos modelos geralmente não é a fase final de um processo de mineração de dados. O conhecimento obtido deve ser organizado e apresentado para os usuários, para que estes possam saber quais ações devem ser realizadas para aproveitar os modelos criados. Dependendo do projeto, a distribuição pode ser simplesmente a geração de um relatório ou, em outras vezes, pode ser uma tarefa mais complexa. Em muitos casos, a tarefa da distribuição é responsabilidade do usuário não analista dos dados (não especialista), por isso a importância que estes tenham algum conhecimento da forma como podem utilizar os modelos.

A fase 5 será abordada com mais detalhes no decorrer da próxima seção (Resultados obtidos). Em relação à fase 6 do exemplo utilizado, esta será realizada em trabalhos futuros, quando os modelos obtidos serão repassados para especialistas em genômica animal, que poderão aplicá-los no desenvolvimento de ferramentas de genotipagem de baixa densidade, entre outras aplicações.

### 2.3 Resultados obtidos

Na aplicação do algoritmo Lasso, para obtenção do melhor valor de  $\lambda$ , avaliaram-se intervalos de 100 e de 1.000 valores possíveis. Entretanto, o número de marcadores selecionados e a acurácia permaneceram inalterados, mantendo-se, então, os 100 valores fornecidos por caret. Com o valor ótimo de  $\lambda$ , o algoritmo Lasso selecionou 29 marcadores relevantes, dos quais, cinco se destacaram para a raça Crioula, 12 para Morada Nova e 12 para Santa Inês. Os cinco marcadores que se destacaram para Crioula e suas respectivas informações estão descritas na Tabela 1. Cabe ressaltar que, a nomenclatura de marcadores SNP segue um determinado padrão na maioria dos organismos. Por exemplo, o marcador OAR2\_55861669.1, encontrado em ovinos, indica que o marcador está presente na espécie *Ovis Aries*, dentro do cromossomo 2 (OAR2), e sua posição dentro do cromossomo é 55.861.669. O número 1, no final do nome, indica a versão daquele marcador encontrado.

De forma geral, todos os marcadores mostraram alto potencial de identificação da raça Crioula, destacando-se, entre outros, quatro marcadores (OARX\_121724022.1, s56924.1, OARX\_78903642.1 e OARX\_29830880.1) pertencentes ao cromossomo X. Foi observado que todos os marcadores da raça Crioula possuem altas diferenças de frequências em relação às outras raças, o que se deve, provavelmente, ao fato de ela possuir as características físicas mais distintas entre elas, como possuir tamanho diminuto e ser lanada (PAIVA, 2005).

**Tabela 1.** Frequências alélicas dos marcadores SNP selecionados pelo algoritmo LASSO para a raça Crioula.

SNP	Cromossomo	Posição	Alelos*	Frequência alélica**		
				Crioula	Morada Nova	Santa Inês
OARX_121724022.1	X	121724022	[C/A]	0.98	0.02	0.05
OARX_29830880.1	X	29830880	[A/G]	0.80	0.00	0.05
OARX_78903642.1	X	78903642	[A/G]	0.95	0.07	0.09
s56924.1	X	53358543	[A/G]	0.98	0.13	0.15
OAR1_268303279_X.1	1	268303280	[G/A]	0.78	0.07	0.09

\* Alelo específico para a raça Crioula do lado esquerdo.

\*\* Frequência do alelo específico na população Crioula e nas raças Morada Nova e Santa Inês.

Para a raça Morada Nova, LASSO identificou os 12 marcadores listados na Tabela 2.

**Tabela 2.** Frequências alélicas dos marcadores SNP, selecionados pelo algoritmo LASSO para a raça Morada Nova.

SNP	Cromossomo	Posição	Alelos*	Frequência alélica**		
				Morada Nova	Crioula	Santa Inês
s05480.1	X	52592630	[G/A]	0.93	0.15	0.22
OAR1_187375309_X.1	1	187375310	[A/G]	0.86	0.02	0.31
OAR1_194627962.1	1	194627962	[G/A]	0.73	0.00	0.02
DU373896_534.1	3	139464759	[A/C]	0.82	0.35	0.15
s32131.1	4	22382506	[A/G]	0.98	0.32	0.42
s06182.1	5	30787155	[A/G]	0.93	0.15	0.31
OAR6_39029427.1	6	39029427	[A/G]	0.84	0.17	0.11
OAR9_39924477.1	9	39924477	[A/C]	0.95	0.17	0.33
OAR10_33338187.1	10	33338187	[A/G]	0.90	0.22	0.28
OAR17_22334380.1	17	22334380	[G/A]	0.79	0.19	0.13
OAR17_8472049.1	17	8472049	[A/G]	0.95	0.22	0.37
OAR20_45964534.1	20	45964534	[G/A]	0.75	0.00	0.15

\* Alelo específico para a raça Morada Nova do lado esquerdo.

\*\* Frequência do alelo específico na população Morada Nova e nas raças Crioula e Santa Inês.

Os destaques para a raça Morada Nova são dois SNP (OAR1\_187375309\_X.1 e OAR1\_194627962.1) no cromossomo um e dois SNP (OAR17\_8472049.1 e OAR17\_22334380.1) no cromossomo 17, além do total de seis marcadores com frequência acima de 90%. Foi observado ainda que há uma frequência relativamente maior dos alelos dos animais Morada Nova na raça Santa Inês. Isto talvez seja explicado pelo fato de os animais Santa Inês serem originários do cruzamento entre Morada Nova e outros ovinos sem raça definida do nordeste brasileiro, fazendo com que muitos ovinos Santa Inês preservem características genóticas do Morada Nova (PAIVA, 2005).

Para a Santa Inês, foram selecionados os 12 marcadores apresentados na Tabela 3.

Dentre os marcadores selecionados para a raça Santa Inês, três pertencem ao cromossomo dois (OAR2\_145195113.1, OAR2\_242658985.1 e s20468.1), três ao cromossomo três (OAR3\_153703374.1, OAR3\_165050963.1 e s16949.1) e três ao cromossomo sete

**Tabela 3.** Frequências alélicas dos marcadores SNP, selecionados pelo algoritmo Lasso para a raça Santa Inês.

SNP	Cromossomo	Posição	Alelos*	Frequência alélica**		
				Santa Inês	Crioula	Morada Nova
OARX_53305527.1	X	53305527	[A/G]	0.72	0.00	0.09
OAR2_145195113.1	2	145195113	[A/G]	0.74	0.04	0.38
OAR2_242658985.1	2	242658985	[A/G]	0.85	0.17	0.29
s20468.1	2	56248983	[A/G]	0.76	0.15	0.00
OAR3_153703374.1	3	153703374	[A/G]	0.76	0.41	0.13
OAR3_165050963.1	3	165050963	[A/G]	0.80	0.02	0.07
s16949.1	3	164901721	[G/A]	0.89	0.15	0.18
OAR5_93120389.1	5	93120389	[G/A]	0.89	0.19	0.38
OAR7_21409209.1	7	21409209	[G/A]	0.61	0.02	0.11
OAR7_94733688.1	7	94733688	[G/A]	0.98	0.37	0.59
s11241.1	7	30741909	[C/A]	0.81	0.35	0.34
s59000.1	18	45393237	[A/G]	0.87	0.30	0.38

\* Alelo específico para a raça Santa Inês do lado esquerdo.

\*\* Frequência do alelo específico na população Santa Inês e nas raças Crioula e Morada Nova.

(OAR7\_21409209.1, OAR7\_94733688.1 e s11241.1). Uma observação importante vem do fato de que dos três marcadores do cromossomo, três estão em posições muito próximas. De maneira geral, os marcadores para a raça Santa Inês têm altas diferenças de frequência alélica em relação às outras raças, tendo como destaques os marcadores OARX\_53305527.1 e s20468.1.

A acurácia atingida com o conjunto de 29 marcadores SNP selecionados pelo algoritmo Lasso foi de 100% na predição de novas raças, e o índice Kappa foi igual a 1. O algoritmo Lasso teve ótimo desempenho, tanto em termos de acurácia quanto computacionalmente (tempo de execução baixo, em torno de uma hora e 57 minutos), como demonstrado em Ayers e Cordell (2010), cujos resultados também confirmaram uma boa performance de outras técnicas de regressão penalizada, como Ridge Regression e Elastic-net.

Random Forest gerou uma listagem dos marcadores mais importantes para o modelo de identificação das raças ovinas. Experimentou-se modelos combinando de 1.000 a 5.000 árvores, e conjuntos aleatórios de atributos variando de 20 a 49.033 atributos para divisão (split) dos nós. Após esses experimentos, o melhor resultado obtido foi utilizando os parâmetros fornecidos pelo pacote caret, que resultou em 1.000 árvores e 313 marcadores para *split*. Selecionou-se, então, os 27 melhores SNP classificados, pois, a partir desta posição os SNP restantes contribuíam com menos que 2% para o modelo. Em Mokry et al. (2013), utilizou-se um critério de seleção diferente, no qual, primeiramente selecionou-se 1% dos SNP mais relevantes de cada cromossomo e, em seguida, foi selecionado 1% dos SNP mais importantes do subconjunto anterior, sendo selecionados 70 marcadores SNP pela técnica Random Forest, utilizando tal critério.

Do conjunto total de 27 marcadores, nove marcadores também foram selecionados pelo algoritmo Lasso. Agrupando-se os marcadores fornecidos pelo modelo Random Forest de acordo com a raça, desenvolveu-se três tabelas para análise da frequência do alelo específico de cada uma delas em relação às outras. A Tabela 4 mostra os marcadores predominantes na raça Crioula e as frequências dos alelos específicos desta raça em relação à Morada Nova e Santa Inês.

**Tabela 4.** Frequências alélicas dos marcadores SNP, selecionados pelo algoritmo Random Forest para a raça Crioula.

SNP	Cromossomo	Posição	Alelos*	Frequência alélica**		
				Crioula	Morada Nova	Santa Inês
OARX_121724022.1	X	121724022	[C/A]	0.98	0.02	0.05
OARX_29830880.1	X	29830880	[A/G]	0.80	0.00	0.05
OARX_78903642.1	X	78903642	[A/G]	0.95	0.07	0.09
s56924.1	X	53358543	[A/G]	0.98	0.13	0.15
OAR1_23724877.1	1	23724877	[G/A]	0.50	0.00	0.04
OAR2_212548956.1	2	212548956	[G/A]	0.80	0.04	0.18
OAR2_55853730.1	2	55853730	[A/C]	0.85	0.00	0.07
OAR11_18815864.1	11	18815864	[A/G]	0.93	0.34	0.22
s71482.1	14	41937578	[G/A]	0.91	0.18	0.50
OAR15_45152619.1	15	45152619	[G/A]	0.76	0.02	0.02
OAR16_39888776.1	16	39888776	[A/G]	0.89	0.11	0.15
s25195.1	25	7203123	[G/A]	0.93	0.02	0.30
s30024.1	25	7165805	[C/A]	0.91	0.02	0.28

\* Alelo específico para a raça Crioula do lado esquerdo.

\*\* Frequência do alelo específico na população Crioula e nas raças Morada Nova e Santa Inês.

Do conjunto de 13 marcadores identificados por Random Forest para a raça Crioula, quatro também foram identificados por LASSO (OARX\_121724022.1, OARX\_29830880.1 e OARX\_78903642.1, s56924.1). Os dois SNP do cromossomo 25 estão em posições próximas e com frequência acima de 90% dentro da raça, surgindo como bons separadores raciais. De forma geral, os SNP fornecidos por Random Forest se mostraram importantes na identificação da raça Crioula.

Na Tabela 5, os SNP com predominância na raça Morada Nova são listados.

**Tabela 5.** Frequências alélicas dos marcadores SNP, selecionados pelo algoritmo Random Forest para a raça Morada Nova.

SNP	Cromossomo	Posição	Alelos*	Frequência alélica**		
				Morada Nova	Crioula	Santa Inês
OAR1_194627962.1	1	194627962	[G/A]	0.73	0.00	0.02
OAR2_54691204.1	2	54691204	[G/A]	0.57	0.04	0.00
OAR18_65638912.1	18	65638912	[G/A]	1.00	0.56	0.41

\* Alelo específico para a raça Morada Nova do lado esquerdo.

\*\* \*\* Frequência do alelo específico na população Morada Nova e nas raças Crioula e Santa Inês.

O algoritmo Random Forest indicou três marcadores importantes para a raça Morada Nova. Como destaque, observa-se os marcadores OAR1\_194627962.1, indicado também pelo modelo Lasso, e OAR2\_54691204.1, com frequência acima de 50% na Morada Nova e praticamente ausente nas outras duas raças. O marcador OAR18\_65638912.1 se destaca com frequência de 100% na raça Morada Nova, apesar de sua frequência em outras duas raças ter ficado entre 40% e 60%.

Na Tabela 6 pode-se observar os SNP com alta frequência na raça Santa Inês.

**Tabela 6.** Frequências alélicas dos marcadores SNP, selecionados pelo algoritmo Random Forest para a raça Santa Inês.

SNP	Cromossomo	Posição	Alelos*	Frequência alélica**		
				Santa Inês	Crioula	Morada Nova
OARX_53305527.1	X	53305527	[A/G]	0.72	0.00	0.09
s61697.1	-	-	[C/A]	0.68	0.06	0.04
OAR1_175474366.1	1	175474366	[G/A]	0.55	0.24	0.00
s03528.1	1	28583773	[A/G]	0.92	0.43	0.23
s20468.1	2	56248983	[A/G]	0.76	0.15	0.00
OAR3_164788310.1	3	164788310	[G/A]	0.89	0.22	0.18
OAR3_165050963.1	3	165050963	[A/G]	0.80	0.02	0.07
OAR3_195698523.1	3	195698523	[A/G]	0.66	0.15	0.04
s16949.1	3	164901721	[G/A]	0.89	0.15	0.18
s69653.1	3	164951744	[G/A]	0.90	0.08	0.36
OAR9_76802154.1	9	76802154	[A/G]	0.96	0.32	0.50

\* Alelo específico para a raça Santa Inês do lado esquerdo.

\*\* Frequência do alelo específico na população Santa Inês e nas raças Crioula e Morada Nova.

Para a raça Santa Inês, 11 marcadores foram selecionados com altas frequências alélicas. Destes, quatro estavam presentes no modelo fornecido pelo algoritmo LASSO (OARX\_53305527.1, s20468.1, OAR3\_165050963.1 e s16949.1). Um dado interessante é que cinco marcadores são originados do cromossomo três (OAR3\_164788310.1, OAR3\_165050963.1, OAR3\_195698523.1, s16949.1 e s69653.1). O marcador s61697.1 também se destaca com alta frequência na raça Santa Inês e com frequências abaixo de 7% na raça Crioula e Morada Nova.

Para treinamento e teste, foram desenvolvidas e combinadas 1.000 árvores utilizando as amostras bootstrap. O comitê de classificadores que formaram a floresta obteve uma acurácia de 99% e Kappa de 0,98. Em relação ao desempenho, o modelo Random Forest foi gerado em duas horas e 44 minutos, ou seja, um tempo que pode ser considerado aceitável levando-se em consideração o elevado número de atributos do conjunto de dados.

Na aplicação da técnica Boosting, o único parâmetro testado foi o número de classificadores (neste caso, árvores de decisão) a serem construídos. Avaliou-se modelos desenvolvidos com totais entre 1.000 e 10.000 árvores, sendo que o melhor resultado, em termos de acurácia e Kappa, ocorreu com 1.000 árvores, número fornecido pelo pacote caret. Selecionou-se os 20 melhores marcadores, pois os SNP a partir desta posição pouco contribuíam (menos que 1%) para o modelo. Entre os 20 marcadores ordenados por Boosting, seis estavam presentes nos modelos Lasso e Random Forest, dois estavam somente em Lasso e sete somente no modelo Random Forest. Com isto, Boosting selecionou apenas cinco marcadores diferentes das técnicas anteriores. Na Tabela 7 estão descritos os SNP predominantes na raça Crioula e suas frequências.

Na lista de marcadores importantes para a raça Crioula, dois deles (OARX\_121724022.1 e s56924.1) foram indicados nos dois modelos anteriores, e outros dois (OAR2\_55853730.1 e OAR15\_45152619.1) foram selecionados no modelo Random Forest, demonstrando o alto potencial destes marcadores. Os marcadores indicados apenas no modelo Boosting (OAR4\_51441757.1, OAR6\_110447914.1 e s30024.1) também mostraram ser potenciais discriminantes de raças.

**Tabela 7.** Frequências alélicas dos marcadores SNP, selecionados pelo algoritmo Boosting para a raça Crioula.

SNP	Cromossomo	Posição	Alelos*	Frequência alélica**		
				Crioula	Morada Nova	Santa Inês
OARX_121724022.1	X	121724022	[C/A]	0.98	0.02	0.05
s56924.1	X	53358543	[A/G]	0.98	0.13	0.15
OAR2_55853730.1	2	55853730	[A/C]	0.85	0.00	0.07
OAR4_51441757.1	4	51441757	[A/G]	0.91	0.25	0.16
OAR6_110447914.1	6	110447914	[G/A]	0.67	0.04	0.02
OAR15_45152619.1	15	45152619	[G/A]	0.76	0.02	0.02
s30024.1	25	7165805	[C/A]	0.91	0.02	0.28

\* Alelo específico para a raça Crioula do lado esquerdo.

\*\* Frequência do alelo específico na população Crioula e nas raças Morada Nova e Santa Inês.

A Tabela 8 traz uma listagem dos marcadores com predominância na raça Morada Nova.

**Tabela 8.** Frequências alélicas dos marcadores SNP, selecionados pelo algoritmo Boosting para a raça Morada Nova.

SNP	Cromossomo	Posição	Alelos*	Frequência alélica**		
				Morada Nova	Crioula	Santa Inês
OAR1_194627962.1	1	194627962	[G/A]	0.73	0.00	0.02
s32131.1	4	22382506	[A/G]	0.98	0.32	0.42
s06182.1	5	30787155	[A/G]	0.93	0.15	0.31
s10365.1	10	21720029	[G/A]	0.45	0.00	0.00

\* Alelo específico para a raça Morada Nova do lado esquerdo.

\*\* \*\* Frequência do alelo específico na população Morada Nova e nas raças Crioula e Santa Inês.

O algoritmo Boosting separou cinco marcadores com maior frequência em Morada Nova, sendo um deles (OAR1\_194627962.1) presente nos dois modelos anteriores e dois (s32131.1, s06182.1) no modelo LASSO. O marcador OAR1\_194627962.1 possui frequência de apenas 2% na Santa Inês e ausente na Crioula, resultado que o confirma como um bom discriminante de raças. Os marcadores s32131.1 e s06182.1 surgem com frequência acima de 90% nos animais Morada Nova, o que também demonstra o bom potencial destes SNP.

A Tabela 9 apresenta os marcadores associados a raça Santa Inês.

Dentre os marcadores fornecidos pelo modelo Boosting para a raça Santa Inês, destacam-se três deles (OARX\_53305527.1, s20468.1, OAR3\_165050963.1) também selecionados pelas técnicas Lasso e Random Forest. Além disso, dois SNP (s39114.1, OAR9\_40217510.1) foram selecionados exclusivamente por Boosting. De forma geral, a maioria dos marcadores selecionados para a raça Santa Inês apresenta alta frequência de alelo. Destaque para os três SNP também indicados pelos dois modelos anteriores, atestando seu potencial de identificação da raça Santa Inês.

Para realização de treinamento e teste, o algoritmo Boosting foi executado por meio de validação cruzada em 10 subconjuntos de dados, sendo que o modelo final foi obtido por meio da média dos 10 subconjuntos. A acurácia e o Kappa obtidos pelo modelo, com a combinação dos classificadores ajustados, foi de 100% e 1, respectivamente. Observando esses resultados, pode-

**Tabela 9:** Frequências alélicas dos marcadores SNP, selecionados pelo algoritmo Boosting para a raça Santa Inês.

SNP	Cromossomo	Posição	Alelos*	Frequência alélica**		
				Santa Inês	Crioula	Morada Nova
OARX_53305527.1	X	53305527	[A/G]	0.72	0.00	0.09
s61697.1	-	-	[C/A]	0.68	0.06	0.04
s03528.1	1	28583773	[A/G]	0.92	0.43	0.23
s20468.1	2	56248983	[A/G]	0.76	0.15	0.00
OAR3_164788310.1	3	164788310	[G/A]	0.89	0.22	0.18
OAR3_165050963.1	3	165050963	[A/G]	0.80	0.02	0.07
s39114.1	3	232410568	[A/G]	0.59	0.08	0.07
s69653.1	3	164951744	[G/A]	0.90	0.08	0.36
OAR9_40217510.1	9	40217510	[C/A]	0.54	0.08	0.02

\* Alelo específico para a raça Santa Inês do lado esquerdo.

\*\* Frequência do alelo específico na população Santa Inês e nas raças Crioula e Morada Nova.

-se acreditar que há indícios de *overfitting*, porém os parâmetros ajustados para a execução do algoritmo foram obtidos pelo caret de forma a evitar um super-ajuste do modelo. O tempo de execução do algoritmo Boosting foi o menor entre os três modelos, sendo finalizado em uma hora e nove minutos. Este bom desempenho também foi obtido no trabalho de González-Recio et al. (2010), que utilizou o algoritmo L2-Boosting em dois conjuntos de marcadores SNP (de touros e frangos), obtendo alta precisão nas predições com um tempo computacional relativamente curto.

Com o desenvolvimento dos três modelos e a seleção dos principais marcadores para identificação das raças, foi realizada uma análise daqueles SNP que convergiam em dois ou três modelos. A intersecção dos modelos envolvendo a raça Crioula mostra que os marcadores OARX\_121724022.1 e o s56924.1 foram selecionados nos três modelos, demonstrando alta relevância na identificação da raça Crioula. O marcador OARX\_121724022.1, em especial, possui uma frequência de 98%, ou seja, demonstra ser um SNP com alto potencial de identificação da raça.

A intersecção presente nos três modelos, relativa à raça Morada Nova, exhibe o marcador OAR1\_194627962.1 com frequência de 73% para a raça Morada Nova e frequências praticamente nulas nas outras raças, o que caracteriza esse SNP como bom discriminante da raça. Os modelos Lasso e Boosting selecionaram os SNP s32131.1 e o s06182.1, os quais possuem frequências acima de 90% na raça Morada Nova, colocando-os também como altamente relevantes para a raça.

Em relação à raça Santa Inês, a intersecção mostra que, nos três modelos, há a presença de três SNP (OARX\_53305527.1, s20468.1 e OAR3\_165050963.1) que apresentam frequências acima de 70% em ovinos Santa Inês e abaixo de 10% em outras raças, confirmando alta capacidade na discriminação racial. Entre os marcadores obtidos por Random Forest e Boosting, destaca-se o s61697.1, com frequência de 68%, posicionando-o como um potencial identificador da raça.

Considerando apenas os marcadores selecionados por dois e três modelos, um total de 18 marcadores demonstra ter grande potencial na identificação das raças estudadas. Esse número de marcadores é próximo aos resultados de trabalhos relacionados à identificação racial em bovinos,

como em Suekawa et al. (2010), onde foram encontrados cinco marcadores por meio de análise de frequência alélica capaz de distinguir gados japoneses e americanos. Por sua vez, Sasazaki et al. (2011) desenvolveram um modelo no qual foram selecionados 11 SNP importantes para gados provenientes de rebanhos dos Estados Unidos.

#### 2.4 Considerações finais

A avaliação dos modelos com aplicação das três técnicas de mineração de dados escolhidas revelou resultados promissores para a seleção dos marcadores SNP mais informativos, que identificam as raças estudadas. Em particular, os modelos gerados pelas técnicas Lasso e Boosting obtiveram resultados melhores, em termos de acurácia e Kappa, em comparação com o modelo Random Forest. Considerando que o conjunto de dados utilizado possui um elevado número de atributos, as técnicas utilizadas reduziram o número de SNP para menos de 0,2%. Na intersecção dos marcadores que compõem os modelos, foram encontrados 18 SNP com maior potencial de identificação das raças, indicando que realmente os marcadores selecionados possuem alta correlação com a raça associada. Os modelos desenvolvidos podem ser utilizados na certificação racial de animais já depositados em bancos de germoplasma e de novos animais a serem inclusos nestes bancos, assim como poderão ser utilizados por diversos segmentos ligados à ovinocultura, como por exemplo, associações de criadores interessadas em certificar seus animais, e pelo Ministério da Agricultura, Pecuária e Abastecimento (Mapa), no controle de animais registrados que apresentam alelos de outras raças, possibilitando a reclassificação desses animais. Adicionalmente, a metodologia proposta poderá ser estendida para toda e qualquer espécie animal de produção.

### 3 Referências

- AGRAWAL, R.; SRIKANT, R. Fast algorithms for mining association rules. In: INTERNATIONAL CONFERENCE ON VERY LARGE DATABASES, 1994, Santiago. **Proceedings...** Santiago: Morgan Kaufmann, 1994. p. 1-32.
- ARCHIBALD, A. L.; COCKETT, N. E.; DALRYMPLE, B. P.; FARAUT, T.; KIJAS, J. W.; MADDOX, J. F.; MCEWAN, J. C.; HUTTON ODDY, V.; RAADSMA, H. W.; WADE, C.; WANG, J.; WANG, W.; XUN, X. The sheep genome reference sequence: a work in progress. **Animal Genetics**, Malden, n. 41, n. 5, p. 449-453, Oct. 2010. DOI: 10.1111/j.1365-2052.2010.02100.x.
- AYERS, K. L.; CORDELL, H. J. SNP selection in genome-wide and candidate gene studies via penalized logistic regression. **Genetic Epidemiology**, New York, v. 34, n. 8, p. 879-91, 2010.
- BREIMAN, L. Random forests. **Machine Learning**, Boston, v. 45, n. 1, p. 5-32, Oct. 2001. DOI: 10.1023/A:1010933404324.
- CAETANO, A. R. Marcadores SNP: conceitos básicos, aplicações no manejo e no melhoramento animal e perspectiva para o futuro. **Revista Brasileira de Zootecnia**, Viçosa, v. 38, p. 64-71, 2009. Número especial.
- CHAPMAN, P.; CLINTON, J.; KERBER, R.; KHABAZA, T.; REINARTZ, T.; SHEARER, C.; WIRTH, R. **CRISP-DM 1.0**: step-by-step data mining guide. Illinois: SPSS, 2000. 78 p.
- COHEN, J. A. A coefficient of agreement of nominal scales. **Educational and Psychological Measurement**, Durham, v. 20, p. 37-46, 1960.
- FAYYAD, U.; PIATETSKY-SHAPIRO, G.; SMYTH, P. From data mining to knowledge discovery: an overview. In: **Advances in knowledge discovery & data mining**. Menlo Park: American Association for Artificial Intelligence, 1996. p. 1-34.
- FREUND, Y.; SCHAPIRE, R. A short introduction to boosting. **Journal of Japanese Society for Artificial Intelligence**, Tokyo, v. 14, n. 5, p. 771-780, 1999.
- FRIEDMAN, J.; HASTIE, T.; TIBSHIRANI, R. Regularization paths for generalized linear models via coordinate descent. **Journal of Statistical Software**, Los Angeles, v. 33, n. 1, p. 1-22, 2010.
- GONZÁLEZ-RECIO, O.; WEIGEL K.A.; GIANOLA D.; NAYA H.; ROSA, G. J. M. L2-Boosting algorithm applied to high-dimensional problems in genomic selection. **Genetics Research**, New York, v. 92, n. 3, p. 227-237, 2010.
- GOUVEIA, J. J. de S. **A utilização da genômica de populações na análise das principais raças de ovinos brasileiros**. 2013. 98 p. Tese (Doutorado) - Universidade Federal do Ceará, Fortaleza.
- HAN, J.; KAMBER, M.; PEI, J. **Data mining**: concepts and techniques. 2nd. ed. San Francisco: Morgan Kaufmann, 2011. 529 p.
- HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. **The elements of statistical learning**: data mining, inference, and prediction. London: Springer, 2011. 745 p.
- HILL, C. M.; MALONE, L. C.; TROCINE, L. Data Mining and Traditional Regression. In: BOZDOGAN, Hamparsum. **Statistical data mining and knowledge discovery**. Knoxville: Chapman & Hall, 2003. p. 17.
- JAMES, G.; HASTIE, T.; TIBSHIRANI, R. **An introduction to statistical learning**: with applications in R. London: Springer, 2013. 429 p.
- KIM, S.; MISRA, A. SNP genotyping: technologies and biomedical applications. **Annual Review of Biomedical Engineering**, Palo alto, v. 9, p. 289-320, 2007.
- KUHN, M. **Caret**: classification and regression training. R package version 5.16-24, 2013.
- LIAW, A.; WIENER, M. Classification and regression by random forest. **R News**, Austria, v. 2, n. 3, p. 18-22, 2002.
- MARIANTE, A. S.; ALBUQUERQUE, M. S. M.; EGITO, A. A.; MCMANUS, C.; LOPES, M. A.; PAIVA, S. R. Present status of the conservation of livestock genetic resources in Brazil. **Livestock Science**, Amsterdam, v. 120, n. 3, p. 204-212, 2009.
- MEGETO, G. A. S.; OLIVEIRA, S. R. de M.; PONTE, E. D.; MEIRA, C. A. A. Árvore de decisão para classificação de ocorrências de ferrugem asiática em lavouras comerciais com base em variáveis meteorológicas. **Engenharia Agrícola**, Jaboticabal, v. 34, n. 3, jun. 2014.
- MOKRY, F. B.; HIGA, R. H.; MUDADU, M. A.; LIMA, A. O.; MEIRELLES, S. L. C.; SILVA, M. V. G. B.; CARDOSO, F. F.; OLIVEIRA, M. M. O.; URBINATI, I.; NICIURA, S. C. M.; TULLIO, R. R.; ALENCAR, M. M.; REGITANO, L. C. Genome-wide association study for backfat thickness in Canchim beef cattle using Random Forest approach. **BMC Genetics**, London, v. 14, n. 47, p. 1-11, 2013. DOI:10.1186/1471-2156-14-47.
- PAIVA, S. R. **Caracterização da diversidade genética de ovinos no Brasil com quatro técnicas moleculares**. 2005. 108 p. Tese (Doutorado)- Universidade Federal de Viçosa, Viçosa, MG.
- PANT, S. D.; SCHENKEL, F. S.; VERSCHOOR, C. P.; KARROW, N.A. Use of breed-specific single nucleotide polymorphisms to discriminate between holstein and jersey dairy cattle breeds. **Animal Biotechnology**, New York, v. 23, n. 1, p. 1-10, 2012.
- REZENDE, S. O.; PUGLIESI, J. B.; MELANDA, E. A.; PAULA, M. F. de. Mineração de Dados. In: REZENDE, S. O. **Sistemas inteligentes**: fundamentos e aplicações. São Paulo: Manole, 2003. p. 307-336.
- RIDGEWAY, G. **Generalized boosted regression models**. Version 2.1. [S.l.: s.n], 2013. 34 p.
- ROORKIWAL, M.; SAWARGAONKAR, S. L.; CHITIKINENI, A.; THUDI, M.; SAXENA, R. K.; UPADHYAYA, H. D.; VALES, M. I.; RIERA-LIZARAZU, O.; VARSHNEY, R. K. Single nucleotide polymorphism genotyping for breeding and genetics applications in chickpea and pigeonpea using the BeadXpress platform. **The Plant Genome**, Madison, v. 6, n. 2, p. 1-10, Aug. 2013.

SASAZAKI, S.; HOSOKAWA, D.; ISHIHARA, R.; AIHARA, H.; OYAMA, K.; MANNEN, H. Development of discrimination markers between Japanese domestic and imported beef. **Animal Science Journal**, Tokio, v. 82, n. 1, p. 67-72, 2011.

SUEKAWA, Y.; AIHARA, H.; ARAKI, M.; HOSOKAWA, D.; MANNEN, H.; SASAZAKI, S. Development of breed identification markers based on a bovine 50K SNP array. **Meat science**, Barking, v. 85, n. 2, p. 285-8, June 2010.

TAN, P.-N.; STEINBACH, M.; KUMAR, V. **Introduction to data mining**. Boston: Addison Wesley, 2006. 769 p.

TIBSHIRANI, R. Regression shrinkage and selection via the lasso, **Statistics in Medicine**, New York, v. 16, p. 385-395, 1997.

WITTEN, I. H.; FRANK, E.; HALL, M. A. **Data mining: practical machine learning tools and techniques**. 3rd. ed. San Francisco: Morgan Kaufmann , 2011. 664 p.

## Métodos, procedimentos e técnicas utilizadas na construção de AgroTIC

Leandro Carrijo Cintra  
Alan Massaru Nakai  
Jorge Luiz Correa

### 1 Introdução

Com o afloramento do método científico no século 16 ficou evidente que a humanidade teria enormes benefícios se tratasse os problemas agrícolas com um procedimento formal e replicável. Iniciou-se então um ciclo que perdura até o presente, no qual muito se avançou no conhecimento sobre cultivo de plantas e criação de animais, além do domínio dos mais variados microrganismos em diversos processos relacionados à agricultura e à alimentação.

A ciência desempenhou um papel fundamental e extremamente relevante para que a humanidade alcançasse o estágio de desenvolvimento em que se encontra nas ciências agrícolas; e atualmente há projetos de pesquisa que permeiam absolutamente todos os pontos de interesse no que diz respeito às questões agropecuárias. Apenas para citar alguns exemplos, sem a mínima intenção de exaurir toda a lista de áreas de atuação, pode-se mencionar projetos no estudo de solos; fisiologia de plantas e animais; nutrição animal e vegetal; genética e melhoramento de microrganismos, animal e vegetal; seleção genômica; rastreabilidade; monitoramento por satélite e veículo aéreo não tripulado (VANTS) aplicados à agricultura; modelagem agroambiental, impactos da agricultura no clima e impactos da mudança climática na agricultura.

O que todos estes projetos têm em comum é o fato de se basearem sempre no método científico para estabelecer os seus resultados. Método este que compreende um conjunto de passos que ajudam a formular, corrigir e evoluir teorias sobre observações e problemas de interesse.

Atualmente, o método científico tem sofrido uma revolução em função do uso massivo da tecnologia da informação (TI) nas mais diversas etapas. A etapa de experimentação, por exemplo, tem usado intensivamente a TI para a geração de dados e informações acerca de eventos sob estudo em várias circunstâncias. Com a criação de equipamentos eletrônicos e softwares especializados em capturar dados, têm-se gerado vultuosos volumes de dados para posterior análise e obtenção de resultados em várias áreas do conhecimento. Em diversas situações estes repositórios são referenciados como *big data*. A quantidade de dados gerados é tamanha que torna impossível a análise dos mesmos sem o uso de processamento computacional intenso, e novamente componentes da tecnologia da informação aparecem associados a outra etapa do processo científico, tornando a TI essencial para o avanço de uma ampla gama de áreas do conhecimento humano.

A relação entre o processo científico e a tecnologia da informação tem se tornado tão estreita que alguns especialistas advogam que se está vivenciando uma nova era nas ciências, na qual a exploração de dado é colocada como um novo paradigma na ciência moderna (HEY et al., 2011). Assim, há aproximadamente mil anos, a ciência era puramente empírica, baseada na descrição de fenômenos naturais; há poucos séculos, tornou-se teórica, baseada em modelos e generalizações; nas últimas décadas ficou explícito o paradigma computacional da ciência, baseado em simulações de fenômenos complexos; e atualmente, fala-se no paradigma baseado na exploração de dados, comumente referenciado como e-Science.

Independentemente das discussões sobre a atual configuração dos paradigmas científicos, é evidente que na atualidade as ciências não poderiam abdicar do uso da computação para alcançar as suas metas. Isto tem ocorrido também com uma ampla gama de estudos relativos à agricultura.

É neste ponto que se torna indispensável uma infraestrutura computacional adequada para fazer frente aos diversos grandes desafios que se apresentam atualmente no âmbito das pesquisas que permeiam o campo da agricultura. Para evidenciar esta demanda, pode-se citar dois casos, dentre vários outros, em atividades de projetos na Embrapa Informática Agropecuária. No primeiro, um projeto para sequenciamento do *Bos Indicus* (popularmente conhecido como Nelore), uma espécie com alto interesse comercial, gerando um volume de dados inicial da ordem de 1,5TB (um e meio terabytes). Durante a execução do projeto atingiu-se um patamar de 15TB de dados intermediários, tendo utilizado na fase crucial de montagem do genoma 500GB de memória RAM e 96 threads de processamento por um período de 3 dias consecutivos. O segundo caso está relacionado com o processamento e análise de séries temporais para avaliação de modelos em mudanças climáticas. Um projeto nesta linha de ação estima a necessidade de armazenamento temporário de 100TB para a execução de suas atividades.

Estes exemplos evidenciam a necessidade efetiva do uso de soluções computacionais de alto-desempenho nas atividades de pesquisas agrícolas atualmente. Sendo assim, segue-se uma discussão sobre as principais infraestruturas computacionais utilizadas para suporte aos diversos projetos de pesquisa realizados na Embrapa Informática Agropecuária.

## 2 Arquiteturas computacionais aplicáveis a problemas científicos

A principal característica exigida dos sistemas computacionais que apoiam os projetos científicos na Embrapa Informática Agropecuária é a escalabilidade. Esta característica diz respeito à capacidade que os sistemas têm de crescerem em relação a alguma de suas propriedades, tais como, capacidade de processamento, tamanho de memória RAM e capacidade de armazenamento. Quando se trata de escalabilidade, os sistemas podem encaixar-se em duas categorias: escaláveis verticalmente ou horizontalmente. Na primeira categoria estão as máquinas que individualmente podem atingir milhares de processadores, possuir dezenas de terabytes<sup>1</sup> de memória RAM e

sistemas de armazenamento especializados que podem atingir petabytes<sup>2</sup> de capacidade. Na segunda categoria encontram-se os *clusters* e *grids*<sup>3</sup>, que permitem a ampliação dos seus recursos com a adição de novos “nós”, ou seja, novas máquinas. O custo por processador ou por unidade de memória é mais expressivo nas máquinas escaláveis verticalmente, porém existem problemas que somente são bem equacionados em arquiteturas deste tipo. Na Embrapa Informática Agropecuária tem-se uma infraestrutura baseada nas duas arquiteturas.

Outro ponto vital relacionado ao tema infraestrutura computacional diz respeito à forma como os recursos são alocados aos usuários. Na atualidade, a computação em nuvem apresenta uma abordagem bastante prática para esta questão, o que tem motivado as organizações a investirem esforços e recursos financeiros nesta área. Um dos projetos da Embrapa Informática Agropecuária tem investigado o uso desta tecnologia aplicada aos problemas relacionados à infraestrutura que surgem comumente nas pesquisas agropecuárias.

### 2.1 Arquitetura altamente escalável

O Laboratório Multiusuário de Bioinformática da Embrapa (LMB), que possui sua infraestrutura computacional hospedada na Embrapa Informática Agropecuária, é um exemplo de arquitetura escalável. Sua missão é propiciar uma infraestrutura computacional moderna e atualizada que faça frente aos desafios relacionados à bioinformática na Embrapa. A bioinformática, por sua vez, trabalha com o armazenamento, a organização e o processamento de dados biológicos, os quais, atualmente advém de várias fontes. A principal demanda dos problemas desta área é por máquinas com grande capacidade de processamento e que tenham memória RAM considerável. Em virtude destas demandas, o laboratório utiliza uma infraestrutura escalável nas duas frentes, ou seja, tanto vertical quanto horizontalmente. A Figura 1 ilustra o ambiente computacional destinado ao processamento de dados biológicos na Embrapa.

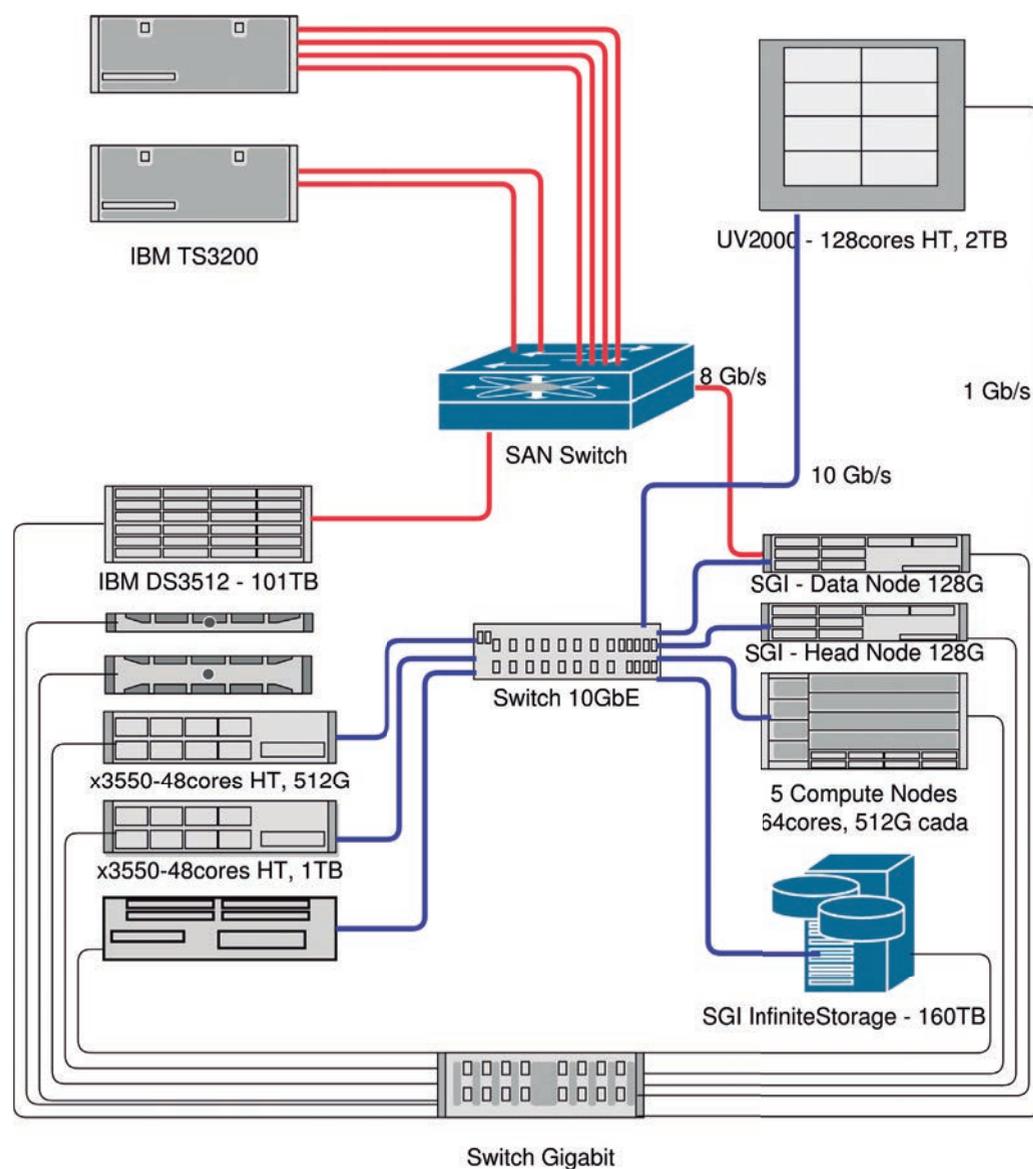
O ambiente computacional é constituído por dois storages para armazenamento. O primeiro é um *Storage Area Network* (SAN) com capacidade para armazenar 100TB de dados. O segundo é um *Network-attached Storage* (NAS) com capacidade para armazenar 220TB, sendo expansível a até 1,3PB (petabytes). Este é um exemplo de máquina com grande escalabilidade vertical.

Para atender às demandas por processamento com boa capacidade de memória, tem-se um grid computacional com a maioria dos nós de processamento possuindo quantidade mediana de memória e alguns poucos nós especiais, com grande volume. Assim, a capacidade de processamento pode ser expandida horizontalmente com a adição de novos nós simples. Já a capacidade de memória disponível para a solução de um mesmo problema pode ser expandida com a adição de nós especiais com alta capacidade de memória. Atualmente, o grid computacional em questão conta com oito nós para processamento e dois nós para controle e gerência do ambiente. Dos nós exclusivos para processamento, dois têm 1TB e 2TB de memória RAM e, respectivamente, 48 e 128 núcleos *hyper-threading* para processamento. Esta última máquina pode escalar até 2048 núcleos e 64TB de memória RAM, constituindo um exemplo expressivo de escalabilidade vertical. Os demais nós têm 512GB de memória RAM e 64 núcleos para processamento cada.

<sup>2</sup> Petabyte equipavale a  $10^{15}$  bytes.

<sup>3</sup> *Cluster* e *Grids* se referem a conjunto de computadores que trabalham juntos na solução de um problema. Usa-se o termo *cluster* para conjuntos fisicamente próximos e normalmente homogêneos e *grids* para conjunto de *clusters*.

<sup>1</sup> Terabytes: unidade de medida de volume de dados em um sistema computacional. Segundo o sistema internacional de medidas equivale a  $10^{12}$  bytes (caracteres).



**Figura 1.** Arquitetura com escalabilidade vertical e horizontal na Embrapa Informática Agropecuária. O grid ilustrado pode ter sua capacidade computacional incrementada adicionando-se novos nós (escalabilidade horizontal), ou expandindo os recursos de alguns nós específicos (escalabilidade vertical).

O ambiente é integrado por duas redes, uma baseada na tecnologia 10 GbE para o tráfego de dados a serem processados e outra, baseada na tecnologia *Gigabit Ethernet*<sup>4</sup>, para a gerência do ambiente, por onde trafegam apenas dados de controle dos sistemas. Este grau de conectividade é importante para que o sistema tenha escalabilidade horizontal. Todos os servidores rodam o

sistema operacional Linux, uma vez que a maioria das ferramentas de bioinformática disponíveis executam nesta plataforma.

O acesso do usuário ocorre via login remoto (*ssh - secure shell*), quando o mesmo necessita executar comandos *shell* no ambiente; ou via interface web, quando utiliza a ferramenta *Galaxy* para executar pipelines de análises. Independentemente dos meios de interação do usuário com o sistema, é vital que a mesma ocorra de forma transparente, sem a obrigatoriedade de se saber muitos detalhes para se utilizar adequadamente o ambiente. Neste sentido, é muito importante em ambientes de processamento científico o uso de sistemas de gerenciamento de recursos (*Distributed Resource Manager System -DRMS*), os quais controlam o escalonamento automático das tarefas (*jobs*) nas máquinas e gerenciam sua execução e resultados. Estes sistemas criam uma abstração do ambiente baseada no conceito de filas de execução. Desta forma, os usuários passam a trabalhar com filas com características específicas, nas quais podem submeter seus trabalhos. A seleção das máquinas, e o instante em que estes trabalhos iniciarão sua execução, ficará a cargo do DRMS determinar.

Em resumo, a infraestrutura apresentada atende atualmente às necessidades de processamento de dados biológicos da Embrapa, e tem a possibilidade de expandir-se de forma bastante efetiva. A capacidade de armazenamento pode ser ampliada, tanto adicionando-se novos módulos ao *storage* já existente quanto adicionando-se novos *storages* ao ambiente. A capacidade de processamento pode ser ampliada com a adição de novos nós, e a capacidade de tratar problemas que exijam grandes quantidades de memória RAM pode ser ampliada com a adição de nós especiais, que tenham a memória necessária disponível.

## 2.2 Clusters Hadoop

O *Hadoop* (BORTHAKUR, 2007; LAM, 2011) é um arcabouço livre de software, mantido pela Apache, voltado para processamento distribuído de grandes quantidades de dados. Além de ser altamente escalável, permitindo a computação distribuída em milhares de computadores, o arcabouço implementa tolerância a falhas no nível da aplicação, provendo serviços de alta disponibilidade. Dentre os diversos componentes do arcabouço, destaca-se sua implementação do modelo de programação *MapReduce* (DEAN, 2008), originalmente proposto pela Google, para processamento e geração de grandes quantidades de dados.

O *MapReduce* aplica a técnica de dividir para conquistar, no qual o problema é dividido em problemas menores e processados separadamente, de forma distribuída, nas máquinas do *cluster*. Neste modelo de programação, a computação é realizada em termos de dois tipos de tarefas *maps* (mapeadores) e *reducers* (redutores). As tarefas de mapeamento são responsáveis por executar a computação sobre frações dos dados de entradas (*splits*) e geram resultados intermediários. Cada *split* é processado por uma instância do mapeador. Os resultados dos mapeadores são consolidados pelos redutores, que agregam os resultados intermediários e geram o resultado final.

O formato dos *splits* variam conforme a natureza dos dados de entrada. Um *split* pode ser, por exemplo, uma quantidade específica de linhas de um arquivo muito grande, uma quantidade fixa de bytes ou um arquivo inteiro, quando os dados de entrada são uma grande coleção de arquivos.

O software que implementa o *MapReduce* (ex. *Hadoop*) cuida do escalonamento das tarefas, ou seja, paraleliza automaticamente a computação dos mapeadores e redutores nas máquinas do *cluster*. As tarefas são monitoradas e, no caso de falhas, são reexecutadas de forma transparente

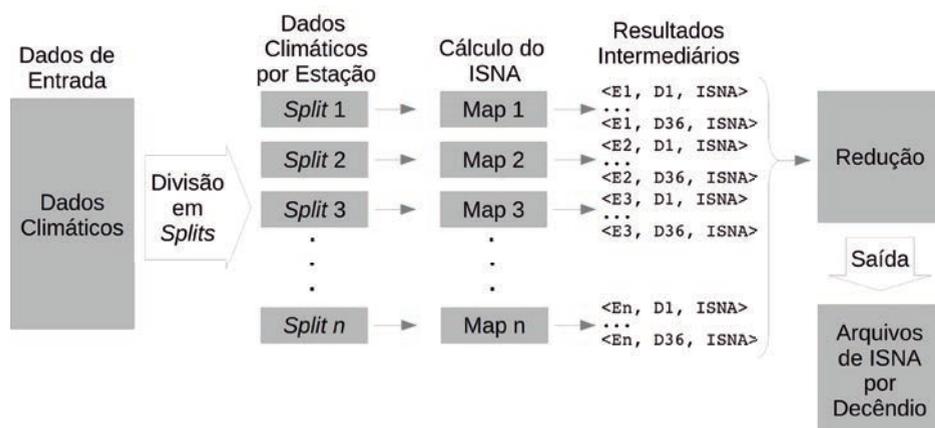
<sup>4</sup> 10 GbE e Gigabit Ethernet são ambas tecnologias para comunicação em rede de computadores. A diferença principal entre ambas está no fato da primeira possibilitar velocidades de 10 Gbps e a segunda, 1 Gbps.

ao usuário. Estas características facilitam o desenvolvimento das aplicações, pois os programadores não precisam se preocupar com aspectos de escalonamento e tolerância a falhas.

No Laboratório de Modelagem Agroambiental (LMA) da Embrapa Informática Agropecuária, o *Hadoop MapReduce* é utilizado na simulação de cenários agrícolas futuros (NAKAI, 2013). Estas simulações são realizadas por meio de uma metodologia de zoneamento de risco climático utilizando dados de projeções climáticas.

A metodologia de zoneamento utilizada é baseada no cálculo do balanço hídrico, que requer parâmetros, como: coeficiente das culturas, capacidade de armazenamento de água do solo, evapotranspiração e séries temporais de chuva. Utiliza-se um modelo baseado em balanço hídrico chamado Bipzon (ASSAD, 1986; FOSTER, 1984; VAKSMANN, 2000) que calcula o Índice de Satisfação da Necessidade de Água (Isna) para cada coordenada desejada. Este índice é utilizado pelos especialistas para determinar os cenários agrícolas com base no risco climático. Atualmente, os cenários agrícolas futuros são simulados com base em dados históricos de milhares de estações climáticas e dados de modelos de projeção climática.

A Figura 2 ilustra como o *MapReduce* é utilizado na simulação de cenário agrícolas (NAKAI, 2013). Em um primeiro momento, dados climáticos são divididos em *splits* de forma que cada split contenha os dados de uma estação climatológica. O *Hadoop* cria uma instância de mapeador para cada *split*. O mapeador implementa o modelo Bipzon para calcular o valor do Isna correspondente a uma estação para todos os decêndios do ano. Todos esses Isnas são agregados pelo redutor, que gera um arquivo contendo os Isnas de todas as estações por decêndio. Posteriormente, esses arquivos serão espacializados para criação dos mapas dos cenários agrícolas decendiais.



**Figura 2.** Esquema da simulação de cenários agrícolas utilizando *MapReduce*;  $n$  é o número de estações.  $\langle En, Di, Isna \rangle$  corresponde ao valor do Isna da estação  $n$ , no decêndio  $i$ .

O uso do *Hadoop MapReduce* na simulação de cenários agrícolas futuros é apenas um exemplo de uso desta tecnologia na pesquisa agropecuária. Neste caso, especificamente, o *Hadoop* tem possibilitado a realização de estudos envolvendo um grande número de cenários com mais agilidade.

### 2.3 Cloud para infraestrutura computacional

A computação em nuvem (*cloud computing*) tem se tornado um grande foco de pesquisa nas últimas décadas, sendo até considerada como um novo modelo de se fazer computação. Embora este tipo de desenvolvimento seja impulsionado majoritariamente por questões comerciais, essa nova tecnologia pode ser aplicada em benefício de diversas outras áreas. Exemplo disso é sua aplicação neste novo modelo de processo científico intimamente relacionado com a tecnologia da informação.

A utilização de tecnologia da informação em métodos científicos pode ser, de uma forma simples, entendida em duas partes: a parte lógica que utiliza a implementação de um método, um algoritmo ou um modelo, para executar uma ou mais tarefas, e uma parte física, responsável pela execução da parte lógica. O *National Institute of Standards and Technology* (THE NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY, 2010) define a computação em nuvem como sendo:

“Um modelo para permitir acesso via rede, conveniente e sob demanda, a um conjunto de recursos computacionais configuráveis e compartilhados (como redes, servidores, armazenamento, aplicações e serviços) que podem ser rapidamente provisionados e liberados com o mínimo de esforço de gerenciamento e interação com o provedor do serviço” (tradução nossa).

Ou seja, a computação em nuvem pode ser entendida como uma nova maneira de provisionar recursos computacionais. Seu uso no gerenciamento de infraestrutura altera, diretamente, o modo até então utilizado para realização de atividades de gerência de tecnologia da informação. O controle no uso de ativos computacionais, provisionamento de infraestrutura de hardware (processamento e armazenamento), criação e liberação de serviços, gerenciamento de máquinas virtuais e capacidade de aumentar ou diminuir recursos computacionais, tudo sob demanda, são algumas dessas atividades cuja gerência torna-se extremamente versátil. Esta versatilidade significa otimização no tempo do processo científico.

Além disso, existem outras variáveis envolvidas no uso de nuvens computacionais, principalmente quando utilizada no suporte ao desenvolvimento de pesquisa científica. Atividades de pesquisa envolvem detalhes como propriedade intelectual, sigilo de informações, armazenamento de dados sensíveis, dentre outros relacionados. A maior parte da utilização de nuvens computacionais é dentro de um mercado de venda de serviços, por terceiros, impactando diretamente nesses detalhes. Neste contexto, torna-se mais compreensível as duas vertentes de nuvens computacionais mais utilizadas: nuvens públicas e nuvens privadas. Nuvens públicas são ofertas de serviços sob demanda por empresas, capazes de hospedar infraestrutura computacional, dados e até mesmo software. Já as nuvens privadas são estruturas semelhantes, com os mesmos objetivos, porém estabelecida dentro das próprias instituições que farão uso das mesmas. Além da versatilidade, tem-se o controle total sobre os ativos computacionais, e, principalmente, sobre os dados trabalhados, o que é muito importante em processos científicos.

Existem atualmente alguns pacotes de software que permitem a implementação de ambientes de nuvem, tanto para armazenamento quanto para processamento. O *OpenStack* é o mais difundido, sendo um conjunto de softwares, de código aberto, que permite a criação de uma nuvem para processamento e armazenamento. Uma vez estabelecida, esta nuvem cria um ambiente virtual para pesquisadores executarem seus processos que necessitem de recursos computacionais. Neste contexto, as nuvens têm desempenhado o papel de aproximar a tecnologia da informação

de pesquisadores das diversas áreas do conhecimento. A obtenção de recursos computacionais para o desenvolvimento de suas atividades passa a ser independente de administradores de redes e infraestrutura. As nuvens tem contribuído para tornar a tecnologia da informação um recurso mais imediato no desenvolvimento de métodos científicos. Em um exemplo prático, uma infraestrutura de nuvem permite a um pesquisador alocar uma determinada quantidade de recursos de armazenamento, inserir seus dados, realizar o processamento e armazenar os resultados de forma rápida e independente.

Utilizando o antigo paradigma de compra de equipamentos de TI para a hospedagem de sistemas e execução de tarefas, o custo para o desenvolvimento de certos processos científicos seriam ainda bem altos. As nuvens computacionais, além de todas as características citadas, inovam também em relação a tarifação. Em tempos passados um contrato de prestação de serviços de tecnologia da informação quase sempre reservava os recursos contratados, mesmo se eles não fossem usados. As nuvens permitem o que se chama de elasticidade, de modo que as capacidades dos recursos podem ser aumentadas ou diminuídas facilmente. Esta característica permite um tipo de cobrança sob demanda, ou seja, o usuário paga pelo que for usado. Despesas antes consideradas como Capex (*CAPitalEXpenditure*) passam a ser consideradas Opex (*OPerationalEXpenditure*), otimizando o uso de recursos financeiros.

Em virtude destas características positivas, a Embrapa Informática Agropecuária vem trabalhando com uma nuvem privada destinada especificamente para aplicações em problemas que surgem em projetos de pesquisa voltados para a agricultura.

### 3 Considerações finais

Não restam dúvidas de que a tecnologia da informação tem impactado e impactará positivamente o processo científico. Sendo assim, o papel da infraestrutura computacional empregada nos projetos de pesquisa torna-se relevante, o que exige uma atenção especial das organizações com relação à infraestrutura disponibilizada. As discussões de aplicações, métodos e técnicas apresentados neste livro mostram a necessidade de uma infraestrutura computacional moderna e adaptável aos mais diversos cenários.

### 4 Referências

ASSAD, E. D. **Simulation de l'irrigation et du drainage pour les cultures pluviales de riz et de mas en sols de bas-fonds brasilia**. Montpellier: IRAT, 1986. 10 p. IRAT. Memories et Travaux, 13.

BORTHAKUR, D. **The Hadoop Distributed File System**: architecture and design. [S. l.]: The Apache Software Foundation, 2007. 14 p. Disponível em: <[http://hadoop.apache.org/common/docs/r0.18.0/hdfs\\_design.pdf](http://hadoop.apache.org/common/docs/r0.18.0/hdfs_design.pdf)>. Acesso em: 18 out 2013.

DEAN, J.; GHEMAWAT S.; Mapreduce: simplified data processing on large clusters, **Communications of the ACM**, v. 51, n. 1, p. 107-113, 2008.

FOSTER, F. **Simulation du bilan hydrique des cultures pluviales. prsentation et utilisation du logiciel BIP**. Montpellier: IRAT-CIRAD, 1984. 63 p.

HEY, T.; TANSLEY, S.; TOLLE, K. (Org.). **O quarto paradigma**: descobertas científicas na era da eScience. São Paulo: Oficina de textos, 2011. 263 p. il.

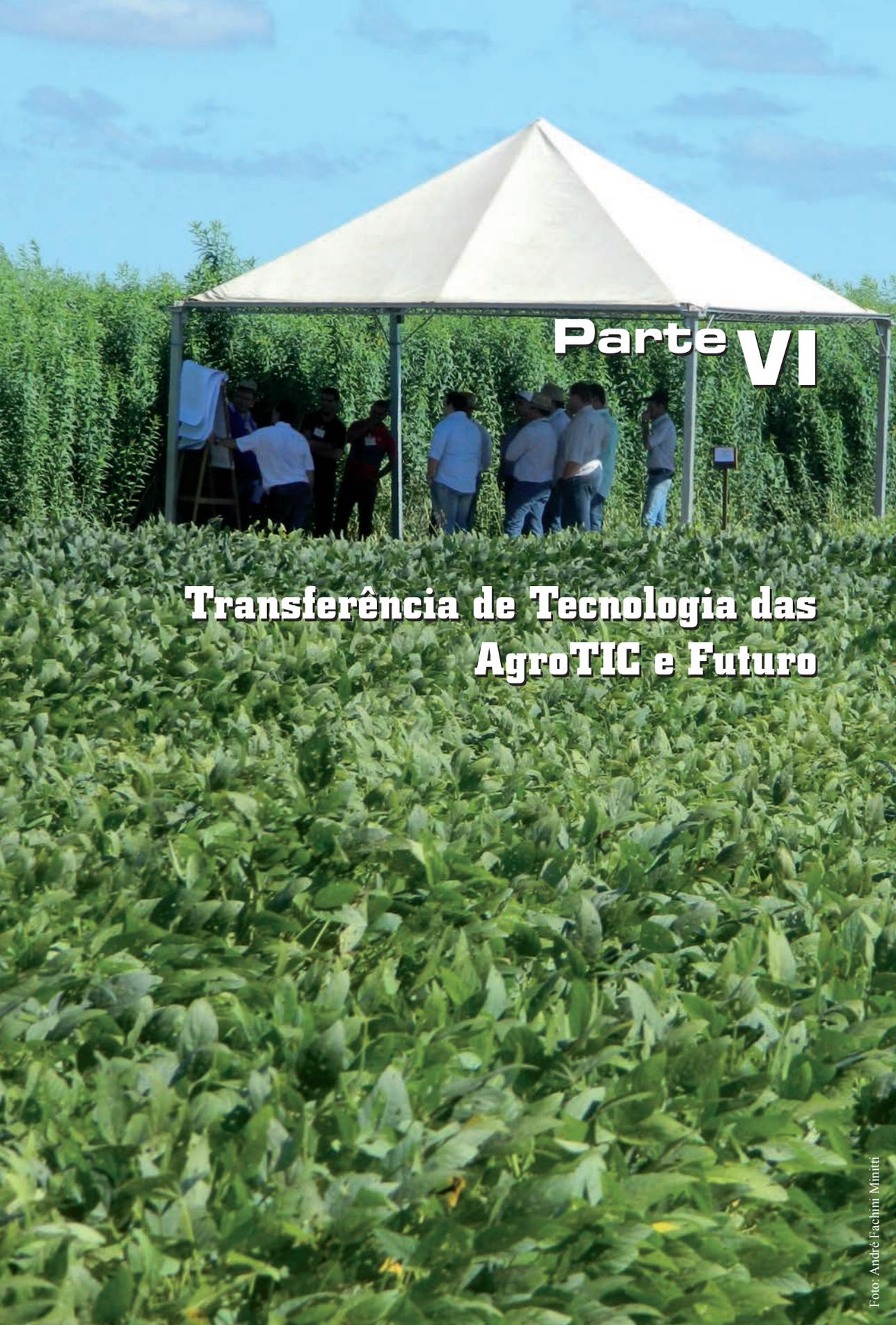
LAM, C. **Hadoop in action**. Stamford: Manning, 2011. 312 p. il.

NAKAI, A. M. **Otimizando o Hadoop MapReduce para tarefas pequenas**: um estudo envolvendo simulações de cenários agrícolas. Campinas: Embrapa Informática Agropecuária, 2013. 5 p. (Embrapa Informática Agropecuária. Comunicado técnico, 115).

THE NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY. **NIST Cloud Computing Program**. 2010. Disponível em: <<http://www.nist.gov/itl/cloud/>>. Acesso em: 1 out. 2014.

VAKSMANN, M. **Le modle bipode**: Logiciel. Bamako: IRAT, 1990.



A group of about ten people are gathered under a white pyramid-shaped canopy in a field of green plants. The canopy is supported by metal poles. The people are dressed in casual attire, including white shirts and hats. The background shows a clear blue sky with some light clouds. The foreground is filled with dense green foliage.

# Parte VI

## Transferência de Tecnologia das AgroTIC e Futuro



## Mercado de AgroTIC e transferência de tecnologia

Martha Delphino Bambini  
Luciana Alvim Santos Romani

### 1 Introdução

A chamada revolução das tecnologias de informação e comunicação (TIC), ocorrida na década 90 e descrita por diversos autores sob várias nomenclaturas (ATKINSON, 2004; PEREZ, 2009), têm levado a uma maciça disseminação destas tecnologias na maioria dos setores econômicos e na vida privada dos indivíduos, representando uma importante mudança de paradigma que afetou muitos processos de cunho social e econômico.

Apesar do potencial de contribuição que as inovações em TIC podem ter para aumentar a eficiência de processos, bem como a quantidade, a qualidade e a inserção dos produtos agropecuários no mercado, alguns aspectos vêm limitando sua adoção no setor, como: limitações de crédito e investimento, falta de infraestrutura de telecomunicações em áreas com baixa densidade demográfica, bem como de informação ou habilidades e competências para uso, e baixo nível de escolaridade de pequenos agricultores e da agricultura familiar.

Mesmo com estas limitações, as TIC estão cada vez mais presentes no setor agropecuário, representadas pelo segmento de AgroTIC (descrito na próxima seção). Estas tecnologias têm sido cada vez mais encontradas no campo, representadas por aplicações baseadas em: software embarcado, redes de sensores para monitoramento de diversos aspectos da produção, sistemas de alerta de eventos climáticos ou de manifestações de doenças, sistemas de informação geográfica e GPS (*Global Positioning Systems*, em inglês), aplicações de sensoriamento remoto, modelos de simulação e sistemas de suporte à decisão, sistemas web e aplicativos para dispositivos móveis, entre outras.

Mesmo que a inserção destas tecnologias no setor agropecuário tenha sido mais lenta, se comparada a outros segmentos da economia, como a indústria e o setor financeiro e bancário, verifica-se um movimento de desenvolvimento e adoção de aplicações de TIC motivado pela necessidade de aumentar a eficiência das atividades e processos desenvolvidos por organizações do setor agropecuário e pela - ainda lenta - disponibilização de infraestrutura de redes e Internet no setor rural do Brasil, mas que vem aumentando dia a dia.

Identificando tal necessidade e o potencial de aplicação destas tecnologias a diversas áreas da agropecuária, empresas atuando no setor de TIC passaram a interessar-se por este mercado, considerando seu potencial de expansão (ZAMBALDE et al., 2011). Os centros de pesquisa e uni-

versidades também vêm ampliando o desenvolvimento de pesquisas e tecnologias baseadas em TIC para aplicação na agropecuária a fim de atender a várias demandas deste segmento, buscando redução de perdas, minimização de riscos e aumento de produtividade e lucratividade do setor.

Este capítulo busca analisar o mercado de TIC aplicadas ao setor agropecuário no Brasil e os modelos e formas de disponibilização de tecnologias geradas pela Embrapa Informática Agropecuária com foco em diferentes públicos de interesse. Unidade temática de pesquisa da Empresa Brasileira de Pesquisa Agropecuária (Embrapa), a Embrapa Informática Agropecuária vem desde 1985 desenvolvendo soluções de tecnologia de informação para o setor agropecuário.

A primeira seção deste trabalho apresenta um panorama dos segmentos do mercado de TIC, com especial ênfase às suas aplicações na agropecuária, considerando o contexto brasileiro e avaliando indicadores relacionados à Indústria Brasileira de Software e Serviços (IBSS). Na sequência, o estudo de caso aborda conceitos relacionados ao processo de transferência de tecnologias geradas por Institutos Públicos de Pesquisa (IPP). No item seguinte são descritos exemplos de tecnologias desenvolvidas, transferidas e disponibilizadas pela Embrapa Informática Agropecuária, mediante diferentes modelos de negócio, visando atingir públicos de interesse distintos, entre eles: agentes da Extensão Rural, técnicos de Governos Estaduais e Federal, produtores, empresas privadas, pesquisadores e estudantes.

## 2 Aplicações AgroTIC e seus mercados

As TIC podem ser entendidas como um conjunto de componentes tecnológicos individuais, normalmente organizados em sistemas de informação baseados em computador (o que se convencionou chamar de tecnologias de informação ou TI), combinados aos avanços no campo das telecomunicações (que sustentam o componente Comunicação) (BAMBINI et al., 2013).

Vários estudos apontam as TIC como ferramentas potenciais para o desenvolvimento agrícola, especialmente na América Latina e Caribe (GOYAL; GONZÁLES-VELOSA, 2013). Entre os benefícios esperados estão: a redução dos custos de comunicação entre os agentes da agropecuária (especialmente relacionados ao uso de telefones celulares e ferramentas de internet); a redução dos custos de acesso a serviços (como seguro e crédito) e informação (de mercado ou tecnológica), bem como sua contribuição para o atendimento às demandas por distribuição e controle de qualidade; incrementos de renda; redução de riscos relacionados a eventos climáticos, pragas e desastres naturais.

Inovações em tecnologias de informação aplicadas ao meio rural podem contribuir para os sistemas agrícolas a fim de aumentar a quantidade, a qualidade e a inserção dos produtos agropecuários no mercado, pelo acesso a informações sobre preços, quantidade e qualidade demandadas por parte de potenciais compradores. Sua aplicação pode contribuir para a melhoria do gerenciamento, eficiência e sustentabilidade para o uso de recursos e insumos como terra, nutrientes do solo, água, energia, trabalho e acima de tudo informação e conhecimento (BALLANTYNE et al., 2010).

AgroTIC podem ser definidas, de maneira geral, como tecnologias de informação e comunicação aplicadas ao setor agropecuário. Dentre suas mais variadas aplicações temos: biopredição; biossimulação; agricultura de precisão envolvendo a aplicação de inteligência embarcada, automação

e rede de sensores locais para monitoramento da produção e controle de processos como mapeamento de solos, monitoramento de doenças e de variáveis meteorológicas e irrigação inteligente; atividades de sensoriamento remoto (a fim de obter mais dados sobre a produção e aspectos ambientais e climáticos), mineração de dados e big data (a fim de organizar e armazenar estes dados, visando lidar com a complexidade e geração de informações e conhecimentos para uso do setor); desenvolvimento de modelos de inteligência computacional e simulação para emissão de alertas e suporte à decisão agropecuária; uso de Sistemas de Informação Geográfica e GPS para apoio à decisão agropecuária, bem como sistemas de informação e gestão do conhecimento implementados em aplicações web e mobile.

### 2.1 Categorias de TIC e seus mercados no Brasil

A Organização de Cooperação e Desenvolvimento Econômicos (OCDE) (ORGANIZAÇÃO DE COOPERAÇÃO E DESENVOLVIMENTO ECONÔMICOS, 2005) divide os produtos de TIC em quatro grandes categorias: equipamentos de informática (hardware); equipamentos de telecomunicação; componentes eletrônicos e produtos de software. Nesta perspectiva, o mercado de TIC refere-se ao conjunto de valores pagos por empresas, residências, órgãos públicos, instituições educacionais, entre outras, aos fornecedores dessas quatro grandes categorias de produtos de TIC.

A seguir são detalhados os segmentos do mercado de TIC por tipos de produto gerado:

- **Equipamentos de informática** (hardware): servidores, computadores pessoais (PCs, no acrônimo em inglês), estações de trabalho, equipamentos de comunicação de dados e periféricos, tais como unidades centrais de processamento (CPUs, no acrônimo em inglês), dispositivos de armazenagem, impressoras, bem como sistemas operacionais vendidos com o equipamento e equipamentos de comunicação de dados.
- **Software**: programas de computador e suas customizações; sistemas e aplicativos.
- **Serviços de TI**: tanto os prestados por empresa contratada como por equipe interna divididos nas categorias: consultorias, capacitações, serviços de implementação, processamento e suporte de informática. Inclui também orçamento interno para operação de estrutura de TI, atividades internas de customização de software e outras despesas internas envolvendo TI.
- **Telecomunicações**: equipamentos para redes públicas e privadas e prestação de serviços, como telefonia fixa e móvel e serviços de televisão a cabo.

De acordo com estudo da Associação Brasileira das Empresas de Software (2014), o setor de TIC, que inclui hardware, software e serviços, representa 2,74 % sobre o Produto Interno Bruto (PIB) brasileiro e 3% do total de investimentos em TI no mundo, movimentando em 2013 US\$ 61,6 bilhões de dólares.

Deste valor, US\$ 10,7 bilhões vieram do mercado de software e US\$ 14,4 bilhões do mercado de serviços, e a soma destes dois segmentos já superou 40% do mercado total de TI, indicando que o Brasil integra o grupo de economia que privilegia o desenvolvimento de soluções e sistemas.

Ainda segundo estudos da Associação Brasileira das Empresas de Software (2014), os segmentos de software e de serviços, conjuntamente, tiveram, em 2013, um crescimento de 10,1%, superando os demais setores da economia brasileira, também acima do crescimento do PIB que foi em torno de 2,3%.

A Indústria Brasileira de Software e Serviços (IBSS), caracterizada por Duarte (2012), é formada por empresas de natureza pública ou privada divididas em:

- Indústria de software, que atua em desenvolvimento e licenciamento de software (sob encomenda, customizável ou não).
- Indústria de serviços de TI, responsável por atividades de consultoria, suporte técnico, manutenção e outros serviços em TI. Esses serviços dividem-se em tratamento de dados, provedores de serviços de aplicação e de hospedagem na internet, portais, provedores de conteúdo e outros serviços de informação na internet e reparação e manutenção de computadores e de equipamentos periféricos e de comunicação.

Esta definição do escopo da IBSS foi delineada pela Associação para Promoção da Excelência do Software Brasileiro (Softex) com base na Classificação Nacional de Atividades Econômicas (CNAE) versão 2.0, desenvolvida sob a coordenação do Instituto Brasileiro de Geografia e Estatística (IBGE). Por meio dos critérios da CNAE, as empresas da IBSS foram classificadas pela principal atividade, geradora da maior fonte de receita.

A pesquisa da Softex, descrita por Duarte (2012), classificou as empresas da IBSS pela atividade correspondente a sua maior fonte de receitas, e os resultados são apresentados na Tabela 1.

**Tabela 1.** Número de empresas da IBSS com 20 ou mais pessoas ocupadas em 2009, classificadas por sua fonte principal de receitas.

Atividade principal	Qtde	%
Desenvolvimento e licenciamento de software customizável	489	21,18
Reparação e manutenção de computadores e de equipamentos periféricos	410	17,76
Tratamento de dados, provedores de serviços de aplicação e de hospedagem na Internet	341	14,77
Desenvolvimento de software sob encomenda	295	12,78
Suporte técnico, manutenção e outros serviços em TI	270	11,69
Consultoria em TI	202	8,75
Desenvolvimento e licenciamento de software não customizável	174	7,54
Reparação e manutenção de equipamentos de comunicação	87	3,77
Portais, provedores de conteúdo e outros serviços de informação na Internet	41	1,78
Total de empresas com mais de 20 pessoas ocupadas em 2009	2.309	

Fonte: adaptado de Duarte (2012).

Estima-se, pela pesquisa da Softex, que existam, em 2012, 73 mil empresas atuando na IBSS. Se o crescimento verificado no período apurado for mantido, em 2014 o número deve passar para a ordem de 80 mil empresas. Em relação à distribuição das empresas segundo o seu porte, temos que 95% delas possuem até 19 pessoas ocupadas e menos de 1% contam com 100 ou mais pessoas ocupadas.

A pesquisa indica que grande parte do contingente de empresas da IBSS possui até quatro funcionários, sendo que muitas destas eram formadas por apenas um sócio, sem necessidade de uma estrutura empresarial formal para seu funcionamento. As análises efetuadas na pesquisa indicam que isto ocorre em virtude da predominância de uma forma de contratação de mão de obra fora do regime de trabalho da Consolidação das Leis Trabalhistas (CLT), por meio do que o mercado considera como “contratação de pessoa jurídica”.

Segundo Duarte (2012), estima-se que existam 11 mil empresas da IBSS na faixa de 5 ou mais pessoas ocupadas. Da mesma forma, a pesquisa da Associação Brasileira das Empresas de Software (2014) identificou 11.232 empresas atuando no setor de software e serviços em 2013, sendo que quase a metade dedicadas à distribuição e comercialização. Desse total, 24,1% atuam no desenvolvimento e produção, 49,8% dedicam-se à comercialização e distribuição e 26,1% prestando serviços. Analisando o porte das empresas dedicadas ao Desenvolvimento e Produção, tem-se que 43,9% são microempresas, 49,6% são categorizadas como pequenas empresas, 5,2% podem ser consideradas médias empresas e apenas 1,3% são grandes empresas.

Verifica-se que 54% das empresas com mais de 20 pessoas ocupadas estão classificadas em três atividades principais (DUARTE, 2012): desenvolvimento e licenciamento de software customizável; reparação e manutenção de computadores e de equipamentos periféricos e tratamento de dados, provedores de serviços de aplicação e de hospedagem na internet.

Em 2013, o faturamento do setor de software e serviços associados atingiu o patamar de US\$ 25,9 bilhões, representando um crescimento de 10,1 % em relação a 2012. Deste total, o faturamento do segmento de software foi de US\$ 10,9 bilhões, uma elevação de 13,5% em relação ao ano anterior. Evidencia-se uma tendência de crescimento do mercado brasileiro de software e serviços entre 2004 e 2014, em geral (ASSOCIAÇÃO BRASILEIRA DAS EMPRESAS DE SOFTWARE, 2014).

No caso da Indústria de Software, destacam-se os segmentos de produção local e produção local para exportação, somando 23,3% da receita. No que se refere aos serviços de TI, destacam-se as receitas de consultoria em sistemas e processos, os serviços de outsourcing (terceirização, em português) e as atividades de suporte e manutenção de software, totalizando 67,6% da receita de serviços.

Apesar da tendência de crescimento do mercado de TI e em especial da indústria de software e serviços, verifica-se que o setor agropecuário é um tímido consumidor de software no Brasil, como pode ser visto na Tabela 2, com valores do mercado comprador por setor econômico.

Os setores financeiro, de telecomunicações e a indústria representam quase 71% do mercado comprador de software no Brasil. A agroindústria é responsável por apenas 2,2 % das compras deste mercado (ASSOCIAÇÃO BRASILEIRA DAS EMPRESAS DE SOFTWARE, 2014).

**Tabela 2.** Segmentação do mercado doméstico comprador de software.

Setor econômico comprador doméstico	Volume (Milhões de US dólares)	Participação percentual
Finanças	3204	26,40
Serviços e Telecomunicações	2970	24,40
Indústria	2451	20,20
Governo	1061	8,70
Comércio	973	8,00
Outros	703	5,80
Óleo e gás	530	4,30
Agroindústria	265	2,20
Total	12.157	

Fonte: adaptado de Associação Brasileira das Empresas de Software (2014).

De maneira geral, a agroindústria caracteriza-se pelas empresas que transformam processos ou produtos de origem agropecuária bem como aquelas voltadas para a distribuição destes produtos. Se no ramo industrial do setor agropecuário a adoção de software é percentualmente baixa, nos empreendimentos agrícolas (fazendas e propriedades rurais) a taxa de adoção de Tecnologias de Informação e Comunicação, de maneira geral, é ainda mais tímida.

Mendes et al. (2011) destacam que, historicamente, a introdução da TI em estabelecimentos agrícolas ocorre mais lentamente quando comparada a outros setores econômicos, em vários países do mundo. Mesmo nos países desenvolvidos, à exceção de algumas propriedades mais tecnificadas que aplicam sensores, aplicativos e robótica nas atividades de produção agrícola, de maneira geral, avalia-se que a adoção de TIC na agropecuária está bastante atrasada em relação a outros setores da economia (GELB et al., 2008).

Entre algumas limitações ao uso de AgroTIC temos: a falta de infraestrutura de telecomunicações no meio rural, bem como de informação ou habilidades e competências para uso destas tecnologias, como alfabetização e conhecimentos específicos, bem como a falta de crédito e não-priorização de investimentos neste tipo de tecnologia.

Francisco e Pino (2004) identificaram algumas variáveis associadas à adoção da internet em unidades de produção agrícola no Estado de São Paulo, mostrando que adotam mais facilmente a internet as propriedades: de caráter empresarial e de maior porte; que já possuem um nível tecnológico mais elevado tanto em atividades produtivas (adotando práticas de inseminação artificial, sementes melhoradas e conservação do solo), quanto em atividades administrativas (usando procedimentos de contabilidade e escrituração agrícola); que possuem proprietários menos idosos e com nível superior de instrução; cujos proprietários mantêm atividades, especialmente econômicas, fora da propriedade; e aquelas cujos proprietários fazem parte de cooperativas/associações de produtores.

Consideramos que estes indicadores são fatores de influência relacionados à adoção de AgroTIC em âmbito geral. Observa-se que, em geral, o emprego de tecnologias de informação e comunicação no âmbito da propriedade agrícola, ocorre primeiramente nas atividades de administração do negócio e depois nas técnicas de cultivo e criação.

## 2.2 Aplicações AgroTIC: software para o setor agropecuário

Considerando as categorias do mercado de TIC mencionadas nesta seção, apresentamos alguns tipos e exemplos de aplicações de AgroTIC.

No caso de equipamento de hardware, temos como exemplo o emprego de equipamentos eletrônicos associados a sensores e software embarcado para monitoramento das atividades agropecuárias.

O software pode ser definido, em linhas gerais, como conjunto de programas de computador, procedimentos, documentações correlatas e dados associados.

Em termos de AgroTIC implementadas por software existem diversos sistemas de informação e de suporte à decisão que podem ser empregados para apoiar o desenvolvimento de ações de gestão da propriedade, de ações técnicas visando ao controle de processos técnicos como ações de fitossanidade e alimentação do rebanho, entre outras.

Outros sistemas de base web podem apoiar a tomada de decisão tanto de produtores, quanto de gestores e técnicos de governo fornecendo informações relativas a condições do tempo, sobre

indicadores sociais e agrícolas de um município, bem como possibilitando o acesso a publicações e conteúdo qualificado sobre diversos sistemas produtivos e criações.

O foco deste trabalho será a análise das aplicações de software aplicado à agropecuária, desenvolvido pela Embrapa Informática Agropecuária, considerando diferentes tipos de produtos: serviços web, software executável para desktop, software livre de código aberto, aplicativos para dispositivos móveis.

### 2.2.1 Aplicativos móveis: estado atual e potencial de expansão

Tendo em vista o potencial do mercado de mobilidade brasileiro, esta seção apresenta com maior detalhamento as questões mercadológicas associadas à aplicação de dispositivos e aplicativos móveis ao setor agropecuário.

Estima-se que o mercado de mobilidade corporativa deve movimentar cerca de US\$ 40 milhões em 2014, segundo levantamento apresentado pela Tata Consultancy Services, em parceria com a Perfecto Mobile, de acordo com Aplicações (APLICATIVOS..., 2013).

Diniz (2010) ressalta, em relação ao conceito de mobilidade, uma grande revolução que ocorreu nessa primeira década do século XXI na qual se destaca o papel central dos aparelhos celulares na vida dos indivíduos. Os chamados smartphones foram criados, no início dos anos 2000, com uma interessante proposta de integrar tarefas comuns aos PDAs (*Personal Digital Assistant*, em inglês), aos celulares e aos computadores pessoais. Inicialmente, os custos dos smartphones eram muito elevados e a tecnologia existente bastante simples, se comparada aos equipamentos atuais. Nesta época, algumas grandes empresas começaram a atuar neste segmento, focando nichos específicos como executivos e entusiastas da tecnologia.

Já em 2007, houve uma mudança no público-alvo dos aplicativos móveis, tendo em vista que as telas táteis (touch) começaram a ser usadas nos telefones. Algumas características dos aparelhos que começaram a ser comercializados eram o multi-toque e o fato de não necessitar do uso de uma caneta stylus (especializada em telas táteis), dando mais possibilidades e opções ao usuário.

Nesse ano foi lançado o iPhone, a primeira geração de smartphones projetada pela empresa Apple (2014) e comercializada em lojas próprias. Este utiliza o sistema operacional da Apple, o iOS. Atualmente, este produto está em sua oitava geração com o lançamento dos iPhones 5C, 6 e 6 Plus em 2014. Características como a sua interface de usuário, baseada em tela *touch* e teclado virtual; a possibilidade de conexão via rede Wi-Fi e sua capacidade de processamento, atraíram um grande número de usuários. Este grande número de vendas acabou por criar um novo paradigma no mercado de smartphones, tornando a Apple uma das empresas mais valiosas do mercado em 2011/2012 (em termos de venda de suas ações). O iPhone é o smartphone mais vendido em vários países como os EUA e o Japão.

Até 2007, os celulares basicamente faziam ligações, tiravam fotos, enviavam mensagens de texto (*Small Message Service*, em inglês) e armazenavam uma agenda de contatos. O iPhone transformou o celular em um computador móvel, acoplando maior capacidade computacional e um maior número de sensores. Criou-se um cenário inovador para a criação de milhões de aplicações, sofisticadas e contextuais.

Em 2007, foi lançado o Android, um sistema operacional para smartphones desenvolvido pela Google e distribuído sob o modelo *open-source* (software livre de código aberto). Este lançamento envolveu a criação da Open Handset Alliance, um consórcio de empresas de hardware,

software e telecomunicações, voltadas ao avanço do modelo aberto para equipamentos móveis. Este sistema é popular entre empresas de tecnologia que necessitam de um produto disponível, de baixo custo e customizável para equipamentos de alta tecnologia.

Desta forma, estabeleceram-se dois modelos de negócios principais para o desenvolvimento e comercialização de aplicativos, chamados Apps. O modelo de negócios aberto do Android contrasta fortemente com o modelo da Apple, que tem por diretriz produzir equipamentos e tecnologias que não permitem intercambiamento ou interoperabilidade com tecnologias de outros fornecedores, ao contrário do Android.

As informações do iPhone são gerenciadas em website comercial próprio da Apple: o iTunes. Nele, o usuário pode comprar aplicativos ou atualizar seu sistema operacional. Atualmente foi lançado o iTunes 7.6, para incluir versões de 64 bits do Windows XP e Vista.

A natureza aberta do Android encorajou uma grande comunidade de desenvolvedores e entusiastas a usar esta plataforma livre como uma base para projetos comunitários, que criam novas funcionalidades para usuários avançados. O sucesso deste sistema operacional fez com que este fosse alvo de ações de litígio de patentes no que se chamou de guerra dos smartphones (QUEM INVENTOU..., 2011).

Dados de 2011 evidenciam as cinco principais empresas atuando no mercado de smartphones: Nokia (24,3%), Apple (18,7%), RIM (14%), Samsung (10,8%) e HTC (8,9%) - outras empresas, somadas, representam 23,2% (MERCADO..., 2011). A Apple, mesmo comercializando apenas dois aparelhos, vendeu quase 19 milhões de iPhones e cresceu 114,1%. Já a Samsung, que apostou em aparelhos Android de primeira linha e campanhas publicitárias monstruosas, cresceu nada menos do que 350%.

Existe uma intensa competição pela posição de líder de mercado, antes estagnado por opções limitadas. Com isso, empresas de telefonia viram-se obrigadas a investir em informática ou buscarem novos parceiros, e as empresas da área de tecnologia entraram nesse mercado, seja como proprietárias de hardware e software ou somente de um deles.

Os custos de acesso à internet móvel estão bem mais baratos, assim como os telefones, que se tornam mais acessíveis às pessoas comuns, tornando-se um objeto de desejo de grande parte da população. As várias funções que os telefones podem realizar (como fotografar, acessar a internet, fazer ligações, enviar mensagens e servir de organizador pessoal com cadernos de contatos e agenda pessoal) bem como a possibilidade de instalar aplicativos diversos são a chave do sucesso dos smartphones atualmente. Em 2011, o número de Apps disponíveis no mercado chegava à notável marca de 1 milhão. Hoje, apenas pouco mais de dois anos depois, o número quase dobrou: em todo o mundo são 1,8 milhão de Apps. Os Apps e suas funcionalidades tornam a experiência de se utilizar um smartphone muito mais divertida e produtiva. Pequenos Apps podem informar, entre outras coisas, o horário de trens ou indicar um bom restaurante (APLICATIVOS..., 2014).

Informações identificadas em pesquisa pela Bitkom, a associação das empresas alemãs de tecnologia da informação, mostrou que 38% dos usuários de smartphones estão dispostos a pagar pelos Apps e quase metade dos entrevistados (45%) quer usar apenas aplicativos gratuitos. A pesquisa também mostrou que um em cada seis usuários não baixou nenhum App para o seu smartphone (APLICATIVOS..., 2014).

No caso do Brasil, segundo Pellanda (2010), o tópico da comunicação móvel vem crescendo em complexidade à medida que penetra em diferentes culturas e classes sociais, com grande impacto em diferentes camadas econômicas. Mesmo caracterizando-se por um país com extremas diferenças sociais, o Brasil é uma das nações a adotar, de forma mais intensa, novas tecnologias e a cultura digital.

No campo da telefonia celular, dados de 2010 apontam a existência de 140 milhões de aparelhos ativos, sendo 81% deste número comercializado em planos pré-pagos, cujo modelo de pagamento é responsável pela grande popularização da comunicação wireless no país. A comunicação móvel está transformando atividades econômicas e sociais ao permitir que várias funções da economia formal e informal sejam realizadas por celular. Atividades de profissionais freelancers, por exemplo, como manicures, encanadores, vendedores passam a ter escritórios móveis. Adicionalmente, uma grande parcela da população brasileira não possui ainda acesso a telefones fixos devido à inexistência de infraestrutura em áreas rurais ou de favelas, por exemplo. Neste aspecto, a tecnologia sem fio transpõe esse problema, caracterizando-se como um instrumento de inclusão digital e social. No setor agropecuário, os aparelhos celulares vem ajudando a levar informação e conhecimento para produtores de países em desenvolvimento como a Índia e alguns países da África. Os equipamentos móveis, mais baratos do que computadores de mesa, são mais fáceis de usar e podem ser transportados para outros locais, conforme a conveniência, tornando-se mais acessíveis para as populações locais e podendo se tornar um importante instrumento para tomada de decisão. Os celulares possuem um grande potencial de atuar no sentido de aumentar a qualidade de vida de populações rurais ao prover acesso à informação, serviços e produtos. Entre as informações mais buscadas estão a previsão e o estado do tempo.

### 2.3 Geração de software para o setor agropecuário: atores e histórico

Em relação às empresas privadas atuando no segmento de software agropecuário, um estudo realizado entre 2008 e 2010 sobre o mercado brasileiro de software para o agronegócio identificou 180 empresas privadas atuando neste setor (MENDES et al., 2011).

Destas, 162 empresas concordaram em participar da pesquisa acima mencionada, que apontou uma concentração geográfica (85,8%) no eixo Sul-Sudeste. Estas empresas estão distribuídas em 65 municípios brasileiros. Utilizando as classes definidas pelo Sebrae para o porte das empresas, foi identificado que 95,8% das empresas são microempresas (57,4%) ou de pequeno porte (38,3%). Das 162 empresas, 70% foram fundadas antes de 2000. A pesquisa indicou que elas trabalham com o desenvolvimento de 337 softwares próprios e 65 desenvolvidos por outras empresas. Estes softwares enquadram-se nas categorias de administração/gerenciamento, manejo animal, cultivo vegetal e controle de processo e/ou de atividades rurais (MENDES et al., 2011).

No que tange à atuação e papel dos Institutos Públicos de Pesquisa e Universidades, de acordo com Zambalde et al. (2011), os estudos para a aplicação de Tecnologia de Informação (TI) no agronegócio brasileiro iniciaram-se com: a criação do Centro de Informática (Ciagri) em 1984, vinculado à Escola Superior de Agricultura “Luiz de Queiroz” (Esalq) da Universidade de São Paulo (USP), em Piracicaba, SP; a criação da Embrapa Instrumentação Agropecuária em 1984, em São Carlos, SP e a criação da Embrapa Informática Agropecuária em 1985, em Campinas, SP.

Nota-se que o Ciagri, embora tenha atribuições relativas à execução de projetos e da política de TI na Universidade, buscou capacitar sua equipe em desenvolvimento de aplicações e uso de microcomputadores na agricultura.

A Embrapa Instrumentação Agropecuária, unidade de pesquisa temática da Embrapa, foi criada com a missão de atuar no desenvolvimento de tecnologias de instrumentação voltadas para o agronegócio, como máquinas, equipamentos, sensores e automação de processos.

O Núcleo Tecnológico para Informática Agropecuária (NTIA), da Embrapa, foi criado no modelo “fábrica de software”, tendo como resultados iniciais o desenvolvimento de um software para gerenciamento e análise dos dados e operações matemáticas. Em 1993, o NTIA passa a ser chamado Centro Nacional de Pesquisa Tecnológica em Informática para a Agricultura (Cnptia), sendo designado por Embrapa Informática Agropecuária, em 1996, e consolidando-se como uma referência em TI aplicada à agropecuária. Possui grupos atuando em modelagem agroambiental, modelagem matemática, bioinformática, visão computacional, processamento de imagens, organização da informação, sistemas de informação e aplicativos móveis para a agropecuária e geoprocessamento.

Em 1989, foi fundado o Laboratório de Automação Agrícola (LAA) da Politécnica USP com o objetivo de desenvolver e aplicar TI no agronegócio, mais especificamente em eletrônica embarcada em máquinas e equipamentos, agricultura e zootecnia de precisão, controle de ambientes, robótica e processamento pós-colheita.

De forma geral, as universidades têm buscado maior participação no contexto da TI no agronegócio (MENDES et al., 2011) com destaque para as universidades federais de Lavras (UFLA), de Juiz de Fora (UFJF) e Viçosa (UFV), Pontifícia Universidade Católica do Paraná (PUC-PR) e universidades estaduais de Ponta Grossa (UEPG) e Londrina (UEL).

Nota-se que tanto os atores privados quanto públicos localizam-se no eixo Sul-Sudeste do Brasil com destaque para os estados de São Paulo e Minas Gerais.

### 3 Transferência de tecnologia na Embrapa

Esta seção apresenta os processos de transferência de tecnologia desenvolvidos por Institutos Públicos de Pesquisa (IPPs) e universidades, no papel de transferidores, e várias categorias de clientes, como empresas privadas, indivíduos, institutos públicos, centros de pesquisa, entre outros.

Assafim (2010) define “transferência de tecnologia” como um processo que envolve aquisição e disponibilização de operações e implica na transmissão ou troca de conhecimento e/ou tecnologias entre duas ou mais partes. O conceito de transferência assume a existência de um possuidor da tecnologia e de outro ator que dela necessita.

Manimala e Thomas (2012) identificaram seis elementos essenciais sobre o processo de transferência de tecnologia internacional, que podem ser aplicadas a processos de transferência em geral:

- Um cedente, a entidade que possui um ativo de conhecimento e quer transferi-la para outra entidade.
- Um cessionário (receptor da tecnologia), a entidade que deseja esse ativo, a fim de empregá-lo e convertê-lo em benefícios.
- O conteúdo, a natureza do ativo de conhecimento a ser transferido.

- O mecanismo pelo qual o ativo será transferido, a fim de permitir a transferência e usá-lo de forma adequada e eficaz.
- O objetivo da transferência, em relação ao cedente, bem como ao cessionário (que podem ter diferentes agendas).
- Os canais, instrumentos e modos de transferência, como o licenciamento, acordos de cooperação, joint ventures, start-ups etc.

O processo de transferência de tecnologia refere-se, em geral, a um ato comercial formalizado juridicamente através de dispositivos contratuais que visam:

- Transmitir bens imateriais (criações, segredos de negócio ou software) protegidas por instrumentos de propriedade intelectual (ex. patentes e registro de software), e / ou;
- Transferir conhecimento tecnológico (know-how).

Institutos de Pesquisa Públicos (IPPs) podem gerar vários tipos de resultados, a partir de atividades de pesquisas desenvolvidas conforme descrito por Drucker e Goldstein (2007): novos conhecimentos; desenvolvimento do capital humano; know-how, tecnologias e protótipos; ativos de propriedade intelectual (patentes, marcas, programas de computador registrados); infraestrutura de conhecimentos que estimula a formação de clusters ou economias de aglomeração; e contribuições para o desenvolvimento econômico, social e cultural regional. Entre estes resultados destacam-se as tecnologias e protótipos baseados em tecnologias de informação e comunicação (TIC).

Este capítulo analisa o contexto de processos de transferência de tecnologias geradas por IPPs brasileiros, considerando o caso da Empresa Brasileira de Pesquisa Agropecuária (Embrapa), com foco para a Embrapa Informática Agropecuária e a geração de AgroTIC.

A Embrapa é uma empresa pública de pesquisa agropecuária vinculada ao Governo Federal do Brasil. A Embrapa tem a missão de viabilizar soluções de pesquisa, desenvolvimento e inovação para a sustentabilidade da agricultura, em benefício da sociedade brasileira (2014). A empresa, fundada em 1973, é uma organização em rede, formada por 47 centros de pesquisa e de serviços distribuídos por todo o País. A Empresa tem também vários laboratórios e escritórios comerciais no exterior, como Estados Unidos, Panamá, Venezuela, e também na África, Ásia e Europa.

O foco deste trabalho é a análise dos processos de transferência de AgroTIC promovidos pela Embrapa Informática Agropecuária, unidade de pesquisa temática da Embrapa. Sua missão é viabilizar soluções de pesquisa, desenvolvimento e inovação em tecnologia da informação para a sustentabilidade da agricultura em benefício da sociedade brasileira.

A Unidade possui 110 empregados, sendo 65% deles vinculados a nove grupos de pesquisa, atuando nas áreas de: bioinformática aplicada, biologia computacional, organização da informação, software livre, geotecnologias, inteligência computacional, novas tecnologias, modelagem agroambiental e matemática computacional.

O estudo de caso enfoca vários tipos de tecnologias e modelos de transferência visando atender a diferentes segmentos de clientes.

A Embrapa, enquanto IPP e Instituição de Ciência e Tecnologia (conforme designação da Lei de Inovação Brasileira), tem como atividade-fim a pesquisa e desenvolvimento com foco na inovação, concretizada por meio de processos de transferência e disponibilização de tecnologias

que promovem a adoção e utilização de seus resultados de pesquisa por vários segmentos da sociedade brasileira.

Dentre vários públicos de interesse da Embrapa, destacam-se: agentes da extensão rural; cooperativas e associações de produtores; agricultores e pecuaristas de diversos portes, com destaque para a agricultura familiar; empresas privadas atuando junto ao agronegócio, como fornecedores de insumos e sementes, processadores de alimentos, entre outras; institutos de pesquisa, universidades e estudantes.

Buscando fortalecer o processo de transferência de tecnologias, a Embrapa, corporativamente, definiu três grandes categorias de resultados de pesquisa a fim de organizar e gerir seu portfólio de soluções tecnológicas. Conforme descrito em Embrapa (2014a,b) as categorias são:

- **Processos** - são metodologias, processos agroindustriais e práticas agropecuárias.
- **Serviços** - são soluções tecnológicas não materiais, como análises e levantamentos, consultorias, mapeamentos, zoneamentos, monitoramento, serviços Web, sistemas de produção, treinamentos e capacitações.
- **Produtos** - são as soluções tecnológicas estruturadas fisicamente que a Embrapa oferece ao mercado ou à sociedade em geral. As categorias de produtos são: agente de controle biológico, agrotóxico, alimento, animal, bebida, bioproduto/formulações/congêneres, corante, cultivar, embrião, estirpe/cepa fertilizante/corretivo, implemento, inoculante, linhagem, máquina/equipamento, medicamento, microrganismo, multimídia, produto biotecnológico, ração, sêmen, software, vacina.

No que se refere às AgroTIC geradas pela Embrapa, temos como produtos os equipamentos eletrônicos, instrumentos e os softwares e, como serviços, os serviços web. Estes constam de websites que disponibilizam, na internet, resultados de pesquisas na forma de relatórios, publicações, informações e dados para uso gratuito de vários públicos de interesse da Empresa.

A partir do movimento de aceleração da adoção de TIC no meio rural nos anos 2000, a Embrapa vem intensificando a utilização deste tipo de tecnologia como instrumento para transferência de tecnologia e de conhecimento, seja pela divulgação de resultados de pesquisa, documentos e publicações via internet, seja pela disponibilização de serviços web e pelo licenciamento de software para apoio a diversas atividades agropecuárias.

Considera-se que, com a grande disseminação de smartphones entre a população em geral e também entre os produtores agrícolas, os aplicativos móveis tem um grande potencial de consolidação como instrumentos efetivos de apoio à gestão agrícola, oferecendo funcionalidades simples para ajudar o produtor a gerenciar suas atividades. Ao oferecer informações meteorológicas, calculadores para suplementação de dieta bovina ou regulagem de equipamentos via reconhecimento de imagens e padrões, entende-se que este tipo de instrumento pode apoiar a decisão do produtor de uma forma simples, ágil e móvel (sem a necessidade de acessar um computador desktop, por exemplo) em qualquer local da propriedade, algumas vezes sem necessidade de conexão com a Internet.

Este trabalho apresenta, a seguir, alguns casos de transferência e disponibilização de produtos de software e serviços web gerados pela Embrapa Informática Agropecuária e seus parceiros.

Uma relação completa das tecnologias disponibilizadas pela Embrapa Informática Agropecuária pode ser encontrada no Anexo I desta obra.

### 3.1 Modelos de transferência de produtos Agrotic: múltiplos casos

#### 3.1.1 Serviços Web

Conforme já foi mencionado, serviços web envolvem, em geral, a disponibilização de websites na internet, oferecendo resultados de pesquisas na forma de relatórios, publicações, informações e dados para acesso e uso gratuito por vários públicos de interesse da Empresa.

##### 3.1.1.1 Sistema Agritempo

O Sistema de Monitoramento Agrometeorológico (Agritempo) é um serviço disponibilizado via internet desde 2002 e compreende um website com informações e dados agrometeorológicos gratuitos com cobertura para todo o território brasileiro. O sistema foi gerado e mantido em parceria com o Centro de Pesquisas Meteorológicas e Climáticas Aplicadas à Agricultura (Cepagri) vinculado à Universidade Estadual de Campinas (Unicamp).

Desenvolvido inicialmente para atender a uma demanda do Ministério da Agricultura, Pecuária e Abastecimento (Mapa), o sistema Agritempo oferece informações e dados agrometeorológicos qualificados para apoiar a tomada de decisão de gestores e produtores do setor agropecuário, bem como técnicos do Governo Federal.

A Figura 1 apresenta a interface de entrada no sistema Agritempo.

O modelo de disponibilização escolhido para esta tecnologia foi a disponibilização gratuita, via internet, buscando aumentar a disseminação e o uso das informações agrometeorológicas para tomada de decisão agrícola, tendo em vista o amplo impacto do seu uso na minimização dos riscos associados ao clima, reduzindo perdas e aumentando a eficiência das práticas agrícolas.



**Figura 1.** Website do sistema Agritempo.

Fonte: Agritempo (2014).

Este sistema concentra uma rede em três dimensões: uma rede física de sensores meteorológicos; uma rede de tecnologia de informação e comunicação (TIC) que permite captar os dados dos sensores e armazená-los em uma base de dados, acessível via internet; e uma rede formada por instituições e indivíduos que compartilham dados e conhecimentos visando fortalecer a base de dados do Agritempo.

O sistema Agritempo gerencia dados e informações de uma rede de mais de 1.400 estações meteorológicas, pertencentes a várias instituições parceiras. Além disso, dados e informações obtidas a partir de imagens de satélites também integram a base de dados do sistema. Em abril de 2014 foi lançada a segunda versão do Agritempo, que se utiliza de uma plataforma computacional mais atual, alinhada com o paradigma atual da internet.

A nova funcionalidade WebGIS (Geographic Information Systems), disponível na versão 2.0 do Agritempo, permite a consulta a diferentes tipos de mapas e informações georreferenciadas para todo o Brasil. Por meio de diferentes níveis, denominados camadas, que podem ser adicionadas e removidas de acordo com a necessidade da pesquisa, é possível ainda obter informações detalha-

das por ponto (localização no mapa) e realizar medições (comprimento e área no mapa). Por meio desta funcionalidade, o usuário tem acesso a todos os mapas temáticos gerados pelo Agritempo diariamente em formato geotiff, que permite que sejam usados localmente em outros softwares GIS. Além disso, o sistema disponibiliza shapes com contornos regionais, estaduais, municipais, por biomas, dentre outros. O WebGIS também permite a impressão de um relatório com o mapa gerado após as definições de diferentes camadas e alterações realizadas pelos usuários.

O sistema Agritempo consiste em uma importante inovação de produto e de processo no campo da agrometeorologia. A principal inovação de processo oferecida por esta tecnologia refere-se à automação de tarefas para a geração das informações agrometeorológicas, permitida pelo uso de TIC, tornando o sistema totalmente automático e independente de ação humana. Todo o processo de recebimento de dados, sua incorporação na base de dados, a construção de mapas, ocorre automaticamente, utilizando-se de várias ferramentas construídas pela equipe e por softwares livres específicos, sendo realizado pelo sistema sem a intervenção humana. Isto proporciona maior rapidez e precisão e oferece mais qualidade à própria base de dados uma vez que o sistema efetua automaticamente alguns testes nas variáveis coletadas.

A gratuidade configura-se como uma inovação de mercado tendo em vista que muitas instituições disponibilizam dados e informações agrometeorológicas mediante pagamento.

A customização dos produtos meteorológicos para agricultura é outra novidade, tendo em vista que antes de 2002 havia poucos websites que ofereciam produtos específicos para este público (como mapas de previsão e monitoramento do tempo como informações relativas a risco de geadas por cultura, deficiência hídrica, necessidade de irrigação, condições de colheita, entre outras), e atualmente poucos o fazem gratuitamente.

Uma ação que vem sendo associada ao processo de disseminação do uso de sistemas web na Embrapa Informática Agropecuária é a capacitação de multiplicadores, entendidos como indivíduos formadores de opinião, que tem potencial para repassar os conhecimentos aprendidos sobre o uso da ferramenta.

No caso do sistema Agritempo 2.0 estão previstas atividades de capacitação de multiplicadores nas cinco regiões do País com a promoção de programas de treinamento sobre a tecnologia, além de ações comunicação mercadológica com a divulgação em feiras e disponibilização de vídeos na internet.

#### 3.1.1.2 Ageitec

A Agência Embrapa de Informação Tecnológica (Ageitec) é um repositório de informações tecnológicas validadas da Embrapa e de seus parceiros, organizadas de forma hierárquica numa estrutura ramificada denominada “árvore do conhecimento” (EMBRAPA, 2014b).

As árvores do conhecimento contêm informações validadas sobre todas as etapas da cadeia produtiva dos produtos (cultivo e criação) e sobre os temas diversos. Nos primeiros níveis da hierarquia estão os conhecimentos mais genéricos e, nos níveis mais profundos, os mais específicos.

A Ageitec apresenta diferentes tipos de árvore do conhecimento, dependendo da natureza do conteúdo. As categorias dividem-se em: cadeias produtivas agropecuárias, cadeias produtivas florestais, de temas agroecológicos e de temas básicos.

No caso das cadeias produtivas agropecuárias, cada árvore do conhecimento estrutura-se em três eixos: pré-produção, produção e pós-produção.

O usuário pode acessar a informação de três maneiras: em árvore hiperbólica (apresentação gráfica da árvore do conhecimento), navegação em hipertexto ou pela utilização de serviço de busca (SILVA; OLIVEIRA, 2012).

A Figura 2 apresenta a interface de navegação em árvore hiperbólica, representando visualmente a árvore do conhecimento.

A motivação para o desenvolvimento desta tecnologia foi o oferecimento de informações tecnológicas qualificadas, apresentadas em linguagem mais adequada ao contexto do produtor rural, do agente de extensão rural e dos estudantes. Buscou-se traduzir os resultados de pesquisa apresentados em publicações científicas para linguagem mais simples, utilizando recursos audiovisuais como fotos e vídeos, e oferecendo informações complementares (como documentos e outros arquivos para download).

A Embrapa Informática Agropecuária e a Embrapa Informação Tecnológica, Unidade de serviço da Embrapa (localizada em Brasília, DF) desenvolveram uma metodologia para organizar, tratar, armazenar e divulgar as informações tecnológicas geradas pela Embrapa e suas instituições parceiras. Esta metodologia foi implementada a partir de um sistema gestor de conteúdo, possibilitando a validação e qualificação das informações nele inseridas por intermédio da figura de perfis de edição e revisão.

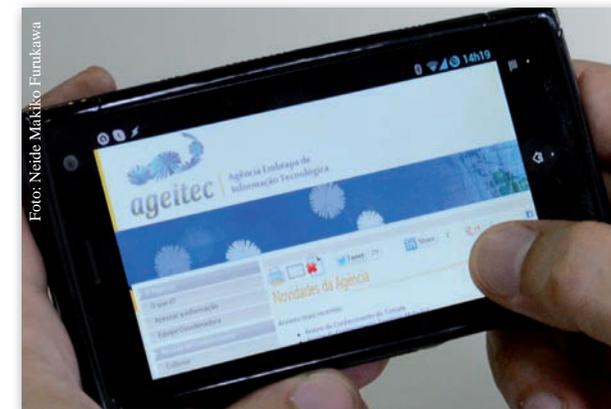
Considerando a motivação em aumentar o acesso dos públicos de interesse da Embrapa aos resultados de suas pesquisas, o modelo de disponibilização selecionado foi gratuito, via internet.

Com média de 100 mil visitas por mês, o website da Agência de Informação Embrapa passou por uma atualização de leiaute e pela disponibilização de novas ferramentas, com o objetivo de aumentar a interatividade com os visitantes e aprimorar a coleta de informações sobre os acessos (SILVA; OLIVEIRA, 2012).

A ferramenta “Logômetro” é um sistema de análise de visitas a websites que provê informações qualificadas sobre o tráfego e o acesso aos conteúdos e páginas disponibilizadas. Foi construída a partir do software livre Piwik e disponibiliza informações como números de acessos, localização georreferenciada das visitas, palavras-chave utilizadas, páginas mais acessadas, tempo médio de visita. Integrado ao website da Ageitec, o logômetro permite descobrir e minerar dados relativos às visitas de maneira geral ou em relação a uma árvore do conhecimento específica.

#### 3.1.1.3 WebAgritec

O WebAgritec é um sistema computacional de acesso e utilização via web que tem por objetivo auxiliar a tomada de decisão de profissionais ligados ao setor agropecuário. O sistema busca



**Figura 2.** Ageitec - Interface em forma de árvore hiperbólica.  
Fonte: Embrapa (2014b).

auxiliar o produtor em diversas etapas do plantio da cultura, atuando desde a escolha da semente até a colheita.

O WebAgritec é constituído por um sistema web de planejamento, previsão e monitoramento da produção agrícola que reúne informações sobre as seguintes culturas: arroz, feijão, milho, soja e trigo.

Ao utilizar o sistema, o usuário do WebAgritec poderá obter informações sobre:

- A época mais adequada para efetuar o plantio.
- A cultivar mais apropriada para seus propósitos.
- A indicação de calagem e adubação para cada cultura.
- As previsões e tendências das condições climáticas antes, durante e depois da safra.
- As doenças e deficiências nutricionais que porventura possam aparecer no decorrer da safra.
- O monitoramento da safra com estimativa de produtividade.
- O acompanhamento da safra via agenda da propriedade.

Em função do tipo de conteúdo e da linguagem utilizada, para acessar a plataforma é necessário possuir formação técnica ou experiência prática no setor agropecuário.

Assim, dentre os perfis de públicos pretendidos para o sistema têm-se: engenheiros agrônomos e técnicos agropecuários que trabalham com assistência técnica ou extensão rural, seja como consultores e autônomos; empresas de extensão rural públicas e privadas; técnicos vinculados a associações de produtores, cooperativas ou sindicatos rurais; empresas privadas, empresas de crédito ou seguro agrícola e produtores rurais.

A plataforma computacional do WebAgritec organiza conteúdos e informações sobre várias áreas do conhecimento com base nos resultados de projetos de Pesquisa e Desenvolvimento (P&D) desenvolvidos pela Embrapa e a partir de demandas de vários setores da agropecuária, conforme apresentado na Figura 3.

O sistema possui oito módulos que orientam o usuário desde o planejamento da cultura até a colheita. São eles: Cultivares, Zoneamento, Adubação, Previsão, Monitoramento, Diagnóstico e Multimídia e a funcionalidade Agenda.

O modelo de transferência do sistema WebAgritec está sendo definido. Atualmente está em fase de negociação junto a atores do setor privado.

### 3.1.2 Software executável para desktop

#### 3.1.2.1 Invernada

O Invernada é um sistema de apoio ao planejamento de produção de bovinos de corte (INVERNADA, 2014). Foi desenvolvido pela Embrapa Informática Agropecuária em parceria com empresas privadas do agronegócio.

A tecnologia é capaz de prever, por meio de modelos estatísticos, o balanço hídrico do solo, crescimento e composição fracional da pastagem, pastejo, seletivo, ingestão de matéria seca, além de configurar dietas a partir da biblioteca de alimentos do sistema Invernada para melhorar o desempenho do animal e otimizar a formulação de dietas, seja por meio de custo mínimo da matéria seca, por lucro máximo e por mínimo custo diário de alimentação.



Figura 3. Organização da informação e conhecimento implementada pelo WebAgritec.

Crédito: Sílvia Maria Fonseca Silveira Massruhá.

O sistema incorpora uma série de bibliotecas de dados, modelos matemáticos de processos biológicos, otimizadores numéricos lineares e não-lineares, e ferramentas auxiliares que permitem análises do sistema de produção. A ferramenta permite a realização de análises adicionais, permitindo a exportação de dados para planilhas eletrônicas no formato Excel e csv (valores separados por vírgulas) que podem ser importados por outras planilhas eletrônicas, pacotes estatísticos e outros softwares.

As análises efetuadas pelo sistema são: simulação e prognóstico de desempenho; comparação do efeito de mudanças gerenciais e adoção de tecnologias e otimização.

As funcionalidades de simulação e prognóstico de desempenho incorporam o conhecimento científico em modelos estatísticos e modelos de processos dinâmicos (dinâmica de sistemas) que permitem representar o resultado dos processos envolvidos na produção animal, gerando tecnologias de apoio à tomada de decisões. Os modelos utilizados são: balanço hídrico do solo, crescimento e composição fracional da pastagem (folha, haste e material morto), pastejo seletivo, ingestão de matéria seca (considerando efeito substitutivo de pastagem-suplemento), crescimento e composição corporal dos animais. Vários prognósticos podem ser realizados, como: crescimento dos animais, crescimento da pastagem, estoques de forragem, fluxo de caixa e retorno econômico.

A funcionalidade de comparação do efeito de mudanças gerenciais e adoção de tecnologias estima, por meio de modelos matemáticos, o efeito de mudanças realizadas em diversas variáveis do controle gerencial sobre o desempenho bioeconômico do sistema de produção, podendo considerar inclusive efeitos do clima.

A função otimização incorpora diversos otimizadores para apoio à tomada de decisão em sistemas de produção de bovinos de corte, entre eles: formulação de dietas de custo mínimo da matéria seca; formulação de dietas de lucro máximo (mínimo custo de produção) e formulação de dietas de mínimo custo diário de alimentação.

O programa consiste em um software executável que pode ser utilizado em computadores pessoais como desktops e notebooks. É disponibilizado gratuitamente para download mediante um cadastro mínimo de dados (nome, e-mail, profissão, estado, cidade).

Os pré-requisitos de uso são conhecimentos básicos em informática e um computador que possua um bom processador. O sistema é disponibilizado com um manual de usuário detalhado para facilitar o uso da ferramenta.

### 3.1.3 Software de código aberto

#### 3.1.3.1 Potion

A bioinformática pode ser definida como uma área de estudo marcada por forte interdisciplinaridade, envolvendo a aplicação de conceitos biológicos e ferramentas práticas de computação para entender, gerar, processar, organizar e propagar informação biológica (Spengler, 2000). Este campo de estudo caracteriza-se pela necessidade de recursos computacionais com grande capacidade de processamento e armazenamento, e por uma intensa geração de dados e informações (configurando-se como um contexto de *big data* e de *e-science*). Uma equipe de pesquisas em bioinformática necessita de pessoal altamente qualificado e interdisciplinar, transitando por diferentes campos de estudo como a biologia, a tecnologia de informação, a matemática, a agronomia, a zootecnia, entre outros. Estes profissionais são geralmente encontrados em ambientes acadêmicos onde cursam pós-graduação em programas de biologia, genômica ou bioinformática.

Fernald et al. (2013) destacam que a bioinformática, entre outros campos científicos e tecnológicos, como a genômica e a farmacogenética, podem ainda ser considerados emergentes e vem crescendo em termos de aplicação, não tendo ainda chegado ao seu limite de potencial inovativo.

No que se refere às aplicações de genômica e bioinformática, pode-se dizer que ainda não se conhece a função biológica de parte considerável dos genes encontrados em projetos genoma (HONGO; LOBO, 2012). Esse vasto universo de genes desconhecidos constitui um campo fértil para a busca de genes interessantes, visando a aplicações de biotecnologia. No que tange às espécies de interesse agropecuário, pode-se dizer que estes genes desconhecidos formam um vasto campo de buscas para localização de genes para ganhos de produção. O desenvolvimento de métodos computacionais possibilita a automatização de processos de detecção de genes desconhecidos que apresentem potencial de contribuir para traços fenotípicos interessantes em espécies animais e vegetais estudadas pela Embrapa.

Uma estratégia ainda não explorada para a detecção de genes potencialmente interessantes é a busca por grupos de genes homólogos - grupos de genes encontrados em espécies diferentes - sob evidência de seleção positiva.

A maioria dos genes homólogos é conservada (HONGO; LOBO, 2012). Entretanto, alguns poucos grupos de genes homólogos evoluem apresentando uma forte pressão seletiva para a variação, ao invés da conservação.

Considerando que as espécies estudadas pela Embrapa têm sido alvo de seleção artificial para alguns poucos fenótipos de interesse visando ganho de produtividade, pode-se supor que os genes sob evidência de seleção positiva nessas espécies serão, possivelmente, associados a fenótipos de produtividade.

Nesse contexto, a busca por genes sob evidência de seleção positiva em genomas de espécies de interesse da Embrapa constitui uma importante ferramenta para indicar possíveis genes associados a um maior ganho de produção nessas espécies. No entanto, muitos passos para a detecção de seleção positiva são computacionalmente custosos. Uma estratégia para contornar este problema é a implementação de programas paralelizados, uma vez que a detecção de seleção positiva em cada grupo de homólogos independe das buscas realizadas em outros grupos.

O software POTION (POSitive selecTION) é um software paralelizado para a detecção de grupos de genes homólogos sob evidência de seleção positiva em escala genômica, tendo sido idealmente concebido para ser executado em servidores que possuam vários processadores, embora também funcione em desktops.

A ferramenta POTION é um software modular e facilmente expansível que utiliza diversos programas que são o estado da arte em seus respectivos campos, tais como OrthoMCL para a detecção dos grupos de homólogos, MUSCLE para o alinhamento dos grupos de proteínas homólogas, phylip para a construção de árvores filogenéticas e PAML para a detecção de seleção positiva. O POTION é capaz de adequar os arquivos de saída de cada um dos software listados acima para o próximo software do pipeline. O programa final produzido possui aproximadamente 1500 linhas de código e utiliza diversos módulos sofisticados de bioinformática previamente desenvolvidos para perl (bioperl). O usuário pode controlar o comportamento de todos os softwares de terceiros por parâmetros globais definidos no início da execução do pipeline.

A ferramenta foi desenvolvida por intermédio de um projeto financiado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e, considerando sua especificidade e potencial para contribuir para o avanço do conhecimento e pesquisas em bioinformática, foi disponibilizada gratuitamente na plataforma GoogleCode como um software livre de código aberto.

### 3.1.4 Aplicativo móvel

#### 3.1.4.1 Gotas

O sistema Gotas, disponível para desktops e também na versão mobile, é um programa de computador que auxilia a calibrar a deposição de pulverizações dos produtos fitossanitários visando tornar este processo mais eficiente e evitando o desperdício.

O processo de produção de gotas ou pulverização tem na agricultura um papel fundamental para a produção de qualquer cultura vegetal. Sua aplicação consiste na colocação do produto fitossanitário (calda) que se encontra dentro da gota, na superfície da planta desejada (alvo). Visando proteger a cultura contra os prejuízos que possam ser causados por um agente externo, como pragas e ervas daninhas, a classe de produtos mais utilizada nas aplicações é a dos agrotóxicos ou defensivos agrícolas.

A aplicação é efetuada através de bicos de pulverização presentes no implemento agrícola, sendo a análise das gotas produzida por estes uma das principais maneiras de quantificar a eficiência da aplicação. A distribuição, o tamanho e o espectro das gotas, por exemplo, são fatores comumente utilizados para a avaliação de um processo de pulverização.

O software Gotas foi desenvolvido neste contexto, objetivando o auxílio aos agricultores para que estes possam calibrar devidamente os bicos de pulverização e obter parâmetros adequados de disposição de agrotóxicos nos alvos desejados.

Desenvolvido pela Embrapa Meio Ambiente (Jaguariúna, SP) e pela Embrapa Informática Agropecuária (Campinas, SP), o Gotas é de um sistema disponibilizado gratuitamente e busca atingir o produtor de forma mais direta, contribuindo para uma atividade agrícola mais sustentável.

Entre as funcionalidades incorporadas nesta versão, destacam-se uma ferramenta para recortar determinada área da amostra selecionada; possibilidade de salvar o experimento em formato compatível para uso dos resultados em planilha de cálculo; recursos para salvar e recuperar o experimento inteiro, com todas as imagens das amostras analisadas; e uma ferramenta para eliminar as amostras indesejadas.

Para alvo de amostragem, é recomendada a utilização de cartão comercial sensível à água, disponível no mercado. Esse cartão, com imagem digitalizada, é processado pelo programa que oferece vários parâmetros de deposição. Com isso, o agricultor poderá decidir sobre a melhor combinação de bicos de pulverização, consumo de calda, velocidade de aplicação etc., que deem o máximo de deposição no alvo desejado. A Figura 4 apresenta a interface do sistema Gotas.



Figura 4. Sistema Gotas.

A calibração de deposição de gotas de pulverização é importante tanto para a aplicação de produtos químicos como para produtos biológicos.

A versão *mobile* do sistema foi desenvolvida para o sistema operacional Android, sistema livre utilizado pela plataforma Google Play (GOOGLE PLAY, 2014).

### 3.1.5 Sistemas de apoio à gestão da pesquisa agropecuária

#### 3.1.5.1 Sistema Ainfo

O sistema Ainfo, desenvolvido em 1991 pela Embrapa Informática Agropecuária, possui funcionalidades para captura, gerenciamento e recuperação da informação, disponibilizando toda produção científica e o acervo documental técnico-científico adquirido pela Embrapa. O Menu principal do sistema Ainfo está apresentado na Figura 5.

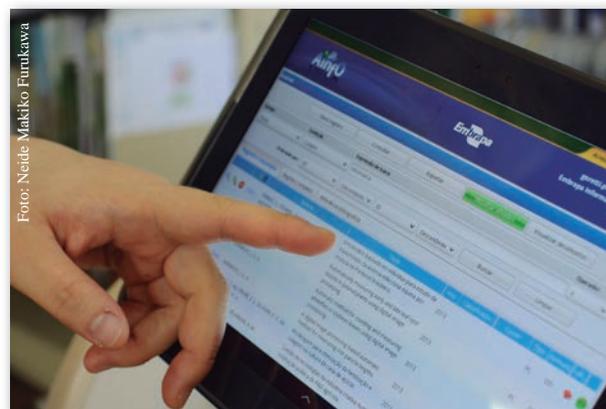


Figura 5. Menu principal do sistema Ainfo.

A versão web do Ainfo foi lançada em 2009 e consolidou todas as bases locais das bibliotecas em uma plataforma única e integrada, que possui mais de 97 mil publicações digitais de 44 Unidades. A tecnologia, atualmente na versão 6, vem sendo utilizada por diversas instituições brasileiras como a Empresa de Pesquisa Agropecuária e Extensão Rural de Santa Catarina (Epagri), Empresa de Pesquisa Agropecuária de Minas Gerais (Epamig), Empresa Baiana de Desenvolvimento Agrícola (EBDA), dentre outras.

Em âmbito internacional, o Instituto Nacional de Investigación Agropecuaria (Inia), do Uruguai, firmou um acordo de cooperação técnica internacional com a Embrapa Informática Agropecuária para implementar o software e disponibilizar o serviço de gestão do acervo documental internamente ao Inia.

Este acordo tem por objetivo fortalecer o intercâmbio técnico-científico entre os dois países por meio do intercâmbio de tecnologia, conhecimentos e dados. Além da implantação do Ainfo, o acordo prevê o apoio na integração da base de dados do Inia com outros sistemas, além de capacitação técnica na implementação e uso do sistema Ainfo. Além disso, o Inia Uruguai disponibilizará o acesso aos metadados de publicações do seu acervo documental relativo à pesquisa agropecuária do Uruguai, estimado em 40 mil itens, para a plataforma computacional do Sistema Aberto e Integrado de Informação em Agricultura - Sabiia, mantido pela Embrapa.

O modelo de transferência da tecnologia Ainfo ao Inia Uruguai foi como software proprietário, disponibilizado mediante um contrato de licença de uso. O contrato previu o fornecimento do código-fonte aberto ao Inia, mediante cláusula de confidencialidade, a fim de possibilitar ações de customização e adaptação da tecnologia.

## 4 Considerações finais

Este trabalho apresenta um panorama do mercado brasileiro de AgroTIC, com ênfase no segmento de software agropecuário. Diversos atores participam como ofertantes deste mercado, como empresas privadas, Institutos Públicos de Pesquisa (IPPs) e universidades.

O capítulo apresenta um estudo de caso da Embrapa Informática Agropecuária enquanto geradora de produtos de software destinados ao mercado agropecuário. Foram apresentados elementos conceituais associados aos processos de transferência de tecnologia em IPPs bem como a categorização de resultados de pesquisa empregada pela Embrapa e a descrição de algumas tecnologias e modelos de disponibilização.

No que se refere aos modelos de transferência de produtos de software gerados por IPPs, ressalta-se a necessidade de buscar maior variedade de alternativas de disponibilização e modelos de licenciamento, considerando as necessidades de diferentes públicos-alvo, de forma a propiciar sua utilização por agentes do mercado, desde empresas privadas até setores governamentais, como ministérios e empresas de extensão rural.

Quanto às tendências tecnológicas das AgroTIC, destaca-se o potencial de crescimento dos aplicativos móveis como instrumento de disseminação de conhecimentos e tecnologias agropecuárias, considerando a grande disseminação dos smartphones e sua ampla utilização pelos indivíduos, seja em âmbito pessoal ou profissional.

## 5 Referências

AGRITEMPO - **Sistema de Monitoramento Agrometeorológico**. Disponível em: [www.agritempo.gov.br](http://www.agritempo.gov.br). Acesso em: 26 set. 2014.

AGROPEDIA BRASILIS. **Ainfo**: tela menu principal. [2014]. Disponível em: <https://www.agropediabrasilis.cnptia.embrapa.br/web/ainfo/administracao-geral-do-sistema>. Acesso em: 26 set. 2014.

APLICATIVOS móveis corporativos devem faturar US\$ 40 milhões em 2014. **Computer World**, São Paulo, set. 2013. Disponível em: <http://computerworld.com.br/negocios/2013/09/20/aplicativos-moveis-corporativos-devem-faturar-us-40-milhoes-em-2014/>. Acesso em: 11 set. 2014.

APLICATIVOS para celulares movem mercado bilionário. **Carta Capital**, jul. 2014. Disponível em: <http://www.cartacapital.com.br/tecnologia/aplicativos-para-celulares-movem-mercado-bilionario-8851.html>. Acesso em: 11 set. 2014.

APPLE. 2014. Disponível em: <https://www.apple.com/>. Acesso em: 11 set. 2014.

APPS contextuais são o novo modelo dos aplicativos móveis. 2014. Disponível em: <http://cio.com.br/tecnologia/2014/09/01/apps-contextuais-sao-o-novo-modelo-dos-aplicativos-moveis/>. Acesso em: 11 set. 2014.

ASSAFIM, J. M. L. **A Transferência de tecnologia no Brasil**: spectos contratuais e concorrenciais da Propriedade Industrial). Rio de Janeiro: Lumen Juris, 2010. 333 p.

ASSOCIAÇÃO BRASILEIRA DAS EMPRESAS DE SOFTWARE. **Mercado brasileiro de software**: panorama e tendências. São Paulo, 2014. 24 p.

ATKINSON, R. D. **The past and future of America's economy**. Washington, D. C.: Edward Elgar, 2004. p. 92-140.

BALLANTYNE, P.; MARU, A.; PORCARI, E. M. Information and communication technologies - opportunities to mobilize agricultural science for development. **Crop Science**, Madison, v. 50, Mar./Apr. 2010. DOI: 10.2135/cropsci2009.09.052.

BAMBINI, M. D.; MENDES, C. I. C. M.; MOURA, M. F.; OLIVEIRA, S. R. M. Software para agropecuária: panorama do mercado brasileiro. **Parcerias Estratégicas**, Brasília, DF, v. 18, n. 36, p. 175-198, jan./jun. 2013.

DINIZ, E. **Pesquisa sobre o uso das tecnologias da informação e da comunicação no Brasil 2005-2009**. Tradução Karen Brito Sexton. São Paulo: Centro de Estudos sobre as Tecnologias da Informação e da Comunicação, 2010. p. 67-70. Edição Especial comemorativa 5 anos.

DRUCKER, J.; GOLDSTEIN, H. Assessing regional economic development impacts of universities: a review of current approaches. **International Regional Science Review**. Philadelphia, v. 30, n. 1, p. 20-46, 2007. DOI: 10.1177/0160017606296731.

DUARTE, V. Caracterização da IBSS. **Software e serviços de TI**: a indústria brasileira em Perspectiva, Campinas, v. 2, p. 108-131, 2012.

EMBRAPA. **Sistema de Gestão de Soluções Tecnológicas da Embrapa**: manual do usuário. Brasília, DF, 2014a. 43 p.

EMBRAPA. **Ageitec - Agência Embrapa de Informação Tecnológica**. 2014b. Brasília, DF. Disponível em: [www.embrapa.br/ageitec](http://www.embrapa.br/ageitec). Acesso em: 26. set. 2014.

FERNALD, K. D. S.; WEENEN, T. C.; SIBLEY, K. J.; CLAASEN, E. Limits to biotechnological innovation. **Technology and Investment**, 4, p. 168-178, Aug. 2013. DOI: 10.4236/ti.2013.43020.

FRANCISCO, V. L. F. dos S.; PINO, F. A. Fatores que afetam o uso da internet no meio rural paulista. **Agricultura em São Paulo**, São Paulo, v. 51, n. 2, p. 27-36, jul./dez. 2004.

GELB, E.; MARU, A.; BRODGEN, J.; DODSWORTH, E.; SAMII, R.; PESCE, V. **Adoption of ICT enabled information systems for agricultural development and rural viability**. 2008. Disponível em: [http://www.fao.org/docs/eims/upload/258775/workshop\\_summary\\_final.pdf](http://www.fao.org/docs/eims/upload/258775/workshop_summary_final.pdf). Acesso em: 22 out. 2014.

GOOGLE PLAY. 2014. Disponível em: <https://play.google.com/>. Acesso em: 11 set. 2014.

GOYAL, A.; GONZÁLES-VELOSA, C. Improving agricultural productivity and market efficiency in Latin America and the Caribbean: how ICTs can make a difference? **Journal of Reviews on Global Economics**, Washington, D.C., v. 2, p. 172-182, 2013.

HONGO, J. A.; LOBO, F. P. POTION: um software paralelizado para a detecção de grupos de genes homólogos sob evidência de seleção positiva em escala genômica. In: MOSTRA DE ESTAGIÁRIOS E BOLSISTAS DA EMBRAPA INFORMÁTICA AGROPECUÁRIA, 8., 2012, Campinas. **Resumos...** Brasília, DF: Embrapa, 2012. p. 163-166.

INVERNADA. 2014 Disponível em: <http://www.invernada.cnptia.embrapa.br/>. Acesso em: 26 set. 2014.

MANIMALA, M. J.; THOMAS, K. R. Learning needs of technology transfer: Coping with Discontinuities and Disruptions. **Journal of the Knowledge Economy**, New York, v. 4, p. 511-539, May 2012. DOI 10.1007/s13132-012-0095-6.

MENDES, C. I. C.; OLIVEIRA, D. R. M. S.; SANTOS, A. R. (Org.). **Estudo do mercado brasileiro de software para o agronegócio**. Campinas: Embrapa Informática Agropecuária, 2011. 184 p.

MERCADO de smartphones já cresceu 80% em 2011: Apple e Samsung projetam disputa. 2011. Disponível em: <http://gizmodo.uol.com.br/mercado-de-smartphones-ja-cresceu-80-em-2011-apple-e-samsung-projetam-disputa/>. Acesso em: 11 set. 2014.

ORGANIZAÇÃO DE COOPERAÇÃO E DESENVOLVIMENTO ECONÔMICOS. **Perspectivas da tecnologia da informação**: as tecnologias da comunicação e da informação e a economia da informação. São Paulo: Senac, 2005. 494 p.

PELLANDA, E. C. Mobilidade: O crescimento das mídias móveis e o impacto nas relações sociais. In: DINIZ, E. **Pesquisa sobre o uso das tecnologias da informação e da comunicação no Brasil 2005-2009**. Tradução Karen Brito Sexton. São Paulo: Centro de Estudos sobre as Tecnologias da Informação e da Comunicação, 2010. p. 61-66. Edição Especial Comemorativa 5 anos.

PEREZ, C. **Technological revolutions and techno-economic paradigms**. 2009. Disponível em: [http://www.carlotaperez.org/papers/TOC\\_TUTWorkingPaper.html](http://www.carlotaperez.org/papers/TOC_TUTWorkingPaper.html) > Acesso em: 18 mar.2012.

QUEM INVENTOU essa guerra? **Época Negócios**, nov. 2011. Disponível em: <http://epocanegocios.globo.com/Revista/Common/0,ERT262473-16642,00.html>. Acesso em: 11 set. 2014.

SILVA, A. C. F. da; OLIVEIRA, L. H. M. de. Novas ferramentas de interatividade e coleta de informações disponibilizadas no website da Agência de Informação Embrapa. In: MOSTRA DE ESTAGIÁRIOS E BOLSISTAS DA EMBRAPA INFORMÁTICA AGROPECUÁRIA, 8., 2012, Campinas. **Resumos...** Brasília, DF: Embrapa, 2012. p. 143-146.

SPENGLER, S. Computers and Biology: Bioinformatics in the information age. **Science**, New York, v. 287, n. 5456, p. 1221-1223, Feb. 2000.

ZAMBALDE, A. L.; SCHNEIDER, H.; LOPES, M. A.; PAGLIS, C. M.; BAMBINI, M. D. Tecnologia da Informação no agronegócio. In: MENDES, C. I. C.; OLIVEIRA, D. R. M. dos S.; SANTOS, A. R. dos. (Org.). **Estudo do mercado brasileiro de software para o agronegócio**. Campinas: Embrapa Informática Agropecuária, 2011, v. 1. p. 39-72.

## ANEXO I - Portfólio de Tecnologias da Embrapa Informática Agropecuária

### Agência Embrapa de Informação Tecnológica (Ageitec)

[www.embrapa.br/agencia](http://www.embrapa.br/agencia)

Site com informações sobre os sistemas produtivos de diversas culturas (arroz, feijão, milho, cana etc) e sobre temas relacionados à agropecuária nacional (meio ambiente, agronegócio do leite, suínos etc).

### Agritempo

[www.agritempo.gov.br](http://www.agritempo.gov.br)

Sistema de Monitoramento Agrometeorológico. Site de informações agrometeorológicas com cobertura para todo o território brasileiro, atualizado diariamente.

### AgroLivre

<https://repositorio.agrolivre.gov.br/>

Repositório de softwares com temáticas diversas, cujo objetivo é aumentar a oferta de software para o domínio agropecuário por meio do incentivo e apoio a projetos de software no setor.

### Ainfo versão 6

[www.ainfo.cnptia.embrapa.br](http://www.ainfo.cnptia.embrapa.br)

O Ainfo é um sistema informatizado para gestão de acervos impressos e digitais de bibliotecas, incluindo todas as fases do fluxo de tratamento da informação.

### Alice

[www.embrapa.br/alice](http://www.embrapa.br/alice)

O Repositório Acesso Livre à Informação Científica da Embrapa (Alice) é um serviço coordenado pela Embrapa Informação Tecnológica que reúne, organiza, armazena, preserva e dissemina, na íntegra, informações científicas produzidas por pesquisadores da Embrapa.

### Árvore Hiperbólica

<https://repositorio.agrolivre.gov.br/>

Envolve ferramentas computacionais para criação, edição e navegação na web de árvores hiperbólicas, conhecidas como hipereditor e hipernavegador. Refere-se a uma ferramenta de estruturação e visualização de informações organizadas de forma hierárquica.

### BDPA

[www.embrapa.br/bdpa](http://www.embrapa.br/bdpa)

A Base de Dados da Pesquisa Agropecuária (BDPA) é um portal de busca para acesso ao acervo das bibliotecas de todos os centros de pesquisa da Embrapa e os documentos digitais gerados pela Empresa.

### Blue Star Sting

<http://sms.cbi.cnptia.embrapa.br/SMS/>

É a versão mais atual da suíte de programas e da base de dados STING, um conjunto de softwares de base web para visualização e análise abrangente de estruturas macromoleculares.

### Diagnose Virtual

[www.diagnose2.cnptia.embrapa.br/diagnose](http://www.diagnose2.cnptia.embrapa.br/diagnose)

Sistema de suporte à decisão que fornece diagnóstico de doenças para as culturas de feijão, trigo, arroz, milho e soja.

### e-Inventário

Sistema composto por um software e dispositivos eletrônicos (RFID e antena), conectados a um netbook, que permite o controle eficiente do patrimônio em ambientes internos, diminuindo a interferência humana para verificação e localização de bens.

### Gotas

<https://repositorio.agrolivre.gov.br/>

Sistema de Análise de Deposição de Agrotóxicos e um programa de computador que auxilia a calibrar a deposição de pulverizações dos produtos fitossanitários. O sistema possibilita a análise da distribuição de gotas no processo de pulverização e apresenta uma série de parâmetros para que o agricultor decida sobre a melhor combinação de bicos de pulverização, consumo de calda, velocidade de aplicação, entre outros fatores, que permitirão aplicar adequadamente a deposição no alvo desejado.

### Infoteca-e

[www.infoteca.cnptia.embrapa.br](http://www.infoteca.cnptia.embrapa.br)

Site que disponibiliza informações na íntegra de publicações, programas de TV e de rádio produzidos pela Embrapa.

### Invernada

[www.invernada.cnptia.embrapa.br](http://www.invernada.cnptia.embrapa.br)

Sistema que possibilita simular e comparar opções de manejo em pecuária de corte por meio de cenários, como: pastejo exclusivo, pastejo com suplementação ou confinamento. Além disso, formula dietas com várias opções de otimização.

### Planeja

[www.planeja.cnptia.embrapa.br](http://www.planeja.cnptia.embrapa.br)

A partir do levantamento de informações agrícolas e socioeconômicas, por meio do preenchimento de formulário de campo próprio para cada Unidade de Produção Agrícola (UPA) do município, possibilita a análise integrada dos dados, subsidiando a tomada de decisão de gestores municipais.

**Potion**

<https://code.google.com/p/potion-darwin>

Software para a busca de grupos de genes homólogos sob evidência de seleção positiva. O software está disponível na plataforma Google Code: <http://code.google.com>

**Navpro**

O Navpro é sistema para processamento e geração automática de produtos de imagens geradas por satélites da série NOAA (National Oceanic and Atmospheric Administration). Estes produtos são divulgados no site do sistema Agritempo por Estado da Federação.

**Sabiia**

[www.embrapa.br/sabiia](http://www.embrapa.br/sabiia)

O Sistema Aberto e Integrado de Informação em Agricultura (Sabiia) é um serviço, coordenado pela Embrapa Informação Tecnológica, que possibilita o acesso a textos completos de publicações científicas sobre agropecuária, disponíveis em diversas instituições nacionais e internacionais.

**Sisla**

[www.sisla.imasul.ms.gov.br](http://www.sisla.imasul.ms.gov.br)

O Sistema Interativo de Suporte ao Licenciamento Ambiental (Sisla) é um sistema de informação geográfica para a web com emissão de relatório do entorno de um determinado empreendimento sobre as variáveis limitantes, como: declividade, áreas protegidas e áreas indígenas.

**WebAgritec**

Sistema web que oferece serviços com aplicação regionalizada, como diagnósticos e alertas de pragas e doenças, dados sobre clima, condições de solo e estimativas de produção. As informações podem ser geradas em forma de relatórios, mapas, gráficos e tabelas, em linguagem acessível, apoiando a tomada de decisões por seus usuários. A tecnologia ainda não está disponível para acesso externo.

## Tecnologias emergentes - futuro e evolução tecnológica das AgrotIC

Maria Angelica de Andrade Leite  
Sílvia Maria Fonseca Silveira Massruhá  
Sílvio Roberto Medeiros Evangelista  
Kleber Xavier Sampaio de Souza

### 1 Introdução

O uso das Tecnologias da Informação e Comunicação (TIC) é uma realidade em todos os setores e ramos de atividades, tanto em nível operacional como estratégico. Na agricultura, como não poderia deixar de ser, as TIC vêm ganhando território cada vez maior. Com a popularização da internet, o uso das TIC foi potencializado em diversos campos. O entendimento desta evolução e suas tendências possibilitam a identificação das oportunidades e dos desafios que estão por vir na Era Digital. O setor agrícola tem buscado formas de assimilar a constante inovação destas tecnologias e garantir o aproveitamento das contribuições oferecidas por elas.

Na área tecnológica, as TIC, inicialmente, aplicadas à execução de procedimentos rotineiros, veem-se diante do crescente desafio de proporcionar, aos tomadores de decisão, a apresentação de informações confiáveis e atualizadas além de contribuir para automatização dos processos envolvidos no desenvolvimento agrícola. Por meio do trabalho colaborativo, profissionais distantes geograficamente trabalham em equipe. As TIC favorecem a tendência de que as fronteiras se tornem cada vez menos demarcadas, em relação ao seu meio ambiente, e as equipes trabalhem em parceria gerando novos conhecimentos e competências. Seu grande potencial reside na sua transversalidade, podendo agregar valor e benefício para as diversas áreas de negócios, mercado, agricultura e meio ambiente.

Algumas das inovações mais recentes prometem alavancar as pesquisas na agricultura. A convergência das áreas de Nanociência, Biotecnologia, Tecnologia da Informação e Ciência Cognitiva (NBIC), irá propiciar um grande salto qualitativo na forma como o mundo da agricultura pode ser transformado. A evolução da abordagem de sistemas, matemática e computação em conjunto com o trabalho em áreas NBIC permitirá, pela primeira vez, compreender o mundo natural e a cognição em termos de sistemas complexos e hierárquicos. Aplicado tanto para problemas específicos de pesquisa quanto para a organização geral da empresa de pesquisa, essa abordagem de sistemas complexos fornece consciência holística e oportunidades de integração, a fim de obter o máximo de sinergia ao longo das principais direções do progresso (KIM et al., 2012).

As próximas seções apresentam alguns campos que vêm recebendo a atenção da comunidade de pesquisa e prometem promover uma nova revolução no campo agrícola e na área de TIC. A seção 2 apresenta uma revisão dos trabalhos e conceitos que vêm sendo desenvolvidos na área da

agricultura de precisão, incluindo a robótica e o uso de robôs; a seção 3 apresenta os avanços da nanotecnologia em biotecnologia e em novos materiais que vão impulsionar os rumos da computação; a seção 4 discute a Internet das Coisas e como a computação pervasiva está se tornando uma realidade; a seção 5 apresenta como as TIC auxiliam na disponibilização e interpretação da informação e do conhecimento; a seção 6 trata de novos modelos de descoberta de conhecimento na era do big data e, finalmente, a seção 7 traz as considerações finais.

## 2 Robótica agrícola e uso de robôs

O crescimento da população mundial levou à necessidade de um nível crescente de padrão tecnológico em agricultura de precisão, tanto para otimização do uso de insumos e capital na preservação do ambiente quanto no aumento e intensificação da produção. Essa necessidade, por sua vez, criou uma exigência de novos métodos, ferramentas e estratégias para processos agrícolas. A robótica e as pesquisas de inteligência artificial podem oferecer novas soluções em agricultura de precisão para processos relacionados com a semeadura, a colheita e o controle de plantas daninhas além de aplicações de fertilizantes e pesticidas, visando melhorar a produtividade e a eficiência.

A aplicação de máquinas agrícolas na agricultura de precisão tem experimentado um aumento de investimento e de pesquisa, devido à utilização de aplicações de robótica na concepção de máquinas e execução de tarefas. Uma das áreas que tem ganhado mais força é a agricultura de precisão autônoma que consiste na operação, orientação e controle de máquinas autônomas para realizar tarefas agrícolas. Espera-se que, no futuro próximo, veículos autônomos estarão no coração de todas as aplicações de agricultura de precisão. O objetivo da robótica agrícola vai além da aplicação de tecnologias de robótica para a agricultura. Atualmente, a maioria dos veículos agrícolas automáticos utilizados para a detecção de plantas daninhas, a dispersão de agrotóxicos, terraplenagem, irrigação, e demais atividades agrícolas são tripulados. Um desempenho autônomo desses veículos vai permitir uma supervisão contínua do campo, desde o levantamento de informações sobre o meio ambiente, que podem ser adquiridos de forma autônoma, à execução adequada da tarefa a ser realizada pelo veículo. As qualidades mais importantes dos veículos agrícolas automáticos podem ser agrupadas em quatro categorias conforme mostrado a seguir:

- **Orientação:** a maneira como o veículo navega no ambiente agrícola.
- **Detecção:** a extração de características biológica do ambiente.
- **Ação:** a execução da tarefa para a qual o veículo foi projetado.
- **Mapeamento:** a construção de um mapa do campo agrícola com as suas características mais relevantes.

O processo de orientação necessita de informações sobre o meio ambiente (mapeamento) e os recursos atualmente detectados (detecção). Por exemplo, para semeadura ou colheita, a unidade de serviço deve estar ciente da presença de árvores ou obstáculos em movimento para sua navegação. Assim, um mapa do ambiente permitirá a navegação segura de uma unidade de serviço e as características detectadas permitirão um planejamento adequado para realizar ações (por exemplo, o nivelamento do terreno, aplicação de pesticidas, etc). Durante o mapeamento, um mapa do ambiente ao redor é construído e mantido para auxiliar o processo de navegação (orien-

tação). Esse mapa é composto pelos recursos ou medidas obtidas a partir do ambiente (detecção) e as informações sobre a localização da unidade de serviço para orientação e ação. A detecção é a aquisição de informações diretamente do ambiente agrícola. Esta informação é usada na fase de mapeamento do meio ambiente para orientar o processo de navegação (orientação) ou para executar uma determinada ação (por exemplo, a detecção de plantas daninhas ou aplicação de agrotóxicos). Finalmente, a fase de ação representa a forma como a unidade de serviço interage com o campo agrícola. Tal ação pode ser realizada na base de um processo de orientação (por exemplo, a colheita ou a semeadura), de detecção (por exemplo, remoção de ervas daninhas), ou de mapeamento (por exemplo, aplicação de agroquímicos com base nas informações anteriormente adquiridas).

Para incorporar a capacidade de agricultura de precisão autônoma, alguns problemas ainda devem ser contornados. As ferramentas desenvolvidas no campo robótico podem ser usadas como soluções propostas para estes problemas e para melhorar o processo agrícola. No entanto, o objetivo da robótica agrícola não é apenas a aplicação de tecnologias robóticas no campo da agricultura, mas também usar os desafios agrícolas para desenvolver novas técnicas e sistemas.

Algumas questões ainda permanecem em aberto nesta área e requerem um maior investimento em pesquisa. Uma delas se refere ao serviço de interação das máquinas com trabalhadores do campo, ou seja, como as máquinas irão atuar no campo, cumprindo sua tarefa e, ao mesmo tempo, detectando os trabalhadores presentes no campo, permitindo trabalhar cooperativamente com eles e mantendo-os em segurança. Outra questão se refere às manobras das máquinas autônomas nos campos. Neste caso, as habilidades de navegação, localização, orientação e manobras de giro exigem estratégias específicas que estão diretamente relacionadas com a disposição do ambiente e os recursos do veículo. Neste sentido, deve haver uma adaptação máquina-campo. Por fim, um outro ponto em aberto refere-se ao planejamento das tarefas, ou seja, a definição da sequência de execução das tarefas. Esta questão está intimamente relacionada com a versatilidade da unidade de serviço e consiste em ter uma arquitetura hierárquica com base nas prioridades e gerenciamento de tarefas. O sistema deve ser capaz de gerir os recursos disponíveis a fim de otimizar as tarefas agrícolas que devem ser executadas. (AUAT CHEEIN; CARELLI, 2013).

Para a solução destas questões, algumas áreas de pesquisa a serem aprofundadas se referem a estudos de interação homem-robô, o trabalho cooperativo e colaborativo, sistemas de controle, visão computacional, inteligência artificial, sistema de posicionamento global, sensoriamento remoto, entre outras. Além disso, existem os problemas relacionados com as características dos campos agrícolas pois, por exemplo, as habilidades necessárias para as máquinas autônomas executarem tarefas de colheita em uma plantação de café são diferentes daquelas referentes à colheita de soja. Por outro lado, existe uma preocupação relacionada ao uso excessivo de trânsito de maquinário nos campos agrícolas podendo provocar a compactação do solo, que impede a obtenção de oxigênio e de água que flui através dele (HARRIS, 2013).

Nesta linha de veículos autônomos, a empresa John Deere desenvolveu um protótipo de um trator automático que usa sinais de satélite para seguir rotas pré-programadas sem condutor humano. A empresa Kinze Manufacturing também utiliza uma abordagem similar para a sua solução de máquina autônoma. A condução autônoma no campo é considerada mais fácil do que na estrada, devido o ambiente ser mais previsível, não existindo pedestres em trânsito e nem outros veículos circulando na mesma via (BAUCKHAGE et al., 2012).

A robótica na agricultura de precisão foca mais nos campos, no cuidado e na manipulação da área das plantas e de seu entorno. Entretanto, existem pesquisas centradas nas raízes das plantas que, embora invisíveis, carregam grande inteligência, coletando informações sobre as propriedades físicas e composição química do solo, utilizando esta informação para decidir em que direção irão continuar crescendo. Aliado a isto, podem perfurar o solo empregando apenas uma fração da energia consumida pelas brocas artificiais além de serem consideradas sistemas altamente eficientes para exploração subterrânea. Olhando para este aspecto, já existem pesquisas para desenvolver dispositivos robóticos que se comportem como as raízes das plantas sendo seu objetivo construir robôs que possam monitorar a poluição do solo, detecção de minerais e, principalmente, a detecção de água possibilitando uma melhor gestão dos reservatórios subterrâneos (ROBOT PLANTS..., 2013).

Pelo que foi apresentado pode-se verificar que a área de agricultura de precisão tem dado bastante atenção na robótica agrícola no sentido de desenvolver máquinas e equipamentos agrícolas, com eletrônica embarcada e dispositivos robóticos, que permitam sua atuação da forma mais autônoma possível. Uma outra linha de pesquisa reside na construção de robôs de aspecto humanoide que possam desempenhar funções associadas aos seres humanos.

Robôs autônomos que podem atuar em locais perigosos para os humanos, pilotar veículos ou assumir o controle da direção em situações de emergência já são realidade. Universidades brasileiras já desenvolvem robôs que comandam dispositivos com o pensamento, especialmente para uso na medicina. Por meio de sensores de um eletroencefalograma – aparelho que mede a atividade cerebral – é possível captar os pensamentos dos usuários que são traduzidos por um software, o qual interpreta os dados obtidos e os transformam em comandos executados por robôs conectados à internet. Especialistas entendem que o aprendizado das máquinas apresenta características similares ao dos humanos. Entretanto, construir robôs tão inteligentes quanto pessoas é um dos principais desafios que motivam profissionais da robótica. Trata-se de algo extremamente complexo e envolve conhecimentos científicos sobre os fundamentos da cognição humana (RODRIGUES, 2014).

A agroindústria está defasada no uso de robôs em suas atividades e processos, se comparar à indústria, onde a utilização de robôs é presente em vários estágios de produção. Naturalmente, uma dificuldade é a adaptação de robôs para uso em ambientes mais rústicos e ao ar livre. A evolução tecnológica, assim como a evolução dos sensores, da inteligência artificial, dos agentes autônomos inteligentes, da precisão do *Global Positioning System* (GPS), dos equipamentos de telemática e da mecânica de movimentação de máquinas, permitiu que se vislumbre a utilização de robôs para atividades no campo (ZAPPA, 2014).

A Universidade de Illinois, por exemplo, desenvolveu uma geração de vários robôs autônomos que se movem nas linhas de plantio com o objetivo, em longo prazo, de assumir algumas das funções atualmente desempenhadas por equipamentos de grande porte. A ideia é colocar alguns robôs no campo, que se comunicam uns com os outros, para manejo e coleta de dados (PETERSON, 2014). Outra característica interessante é a futura capacidade de troca de informação entre os robôs que, espelhando-se no comportamento das abelhas que saem em busca do néctar e voltam para compartilhar a informação, podem encontrar plantas daninhas e comunicar esta localização a outros robôs para que eles possam atuar de forma conjunta.

Os robôs podem ser equipados com sensores ultrassônicos e sensores infravermelhos de detecção de movimento permitindo manobras dentro da linha de cultura, usando uma abordagem não

linear. O robô curva-se à esquerda se sente a presença de uma planta a sua frente e, em seguida, dirige-se para a direita, ao perceber uma outra planta. Como resultado, o robô acha o seu caminho entre as diversas linhas de plantio. Para fazer curvas no final de uma linha, sensores detectam quando ela acaba e sinalizam para que o robô vire. Quando estes robôs adquirirem uma grande capacidade de movimentação nos campos, eles podem ser programados e equipados para exercer funções como a detecção de doenças, de ervas daninhas, de insetos, de amostragem do solo ou mesmo aplicação de pesticidas ou fertilizantes, de forma precisa.

A viabilidade prática destes robôs, no auxílio das atividades agrícolas, dependerá da combinação de três fatores: robôs agrícolas, sistemas inteligentes e rede de sensores sem fio, espalhados no campo para monitorar a plantação e avaliar a sua vitalidade. Estes sensores devem ser capazes de coletar informações sobre o ambiente e a planta, a fim de tomar algum tipo de decisão em função dos parâmetros de entrada (ALISSON, 2014). Desta forma, vislumbra-se a possibilidade de uso de robôs em quase todas as fases de cultura, praticamente sem intervenção humana. Implementações de pequena escala já estão no horizonte. Este tipo de aplicação estará viável cientificamente em 2023 e financeiramente factível por volta de 2026 (ZAPPA, 2014).

Enquanto as experiências se concentram em atribuir capacidades humanas aos robôs, permitindo que sejam capazes de executar funções associadas a pessoas, seja no campo ou em outras áreas de atuação, por outro lado existem estudos e experimentos que incorporam elementos robóticos nos seres humanos. São os chamados cyborgs ou trans-humanos. O termo cyborg literalmente significa organismo cibernético, ou seja, um ser composto tanto da parte orgânica quanto da parte mecânica. Tradicionalmente, os cyborgs são tratados nas histórias de ficção científica. Atualmente, por meio da medicina moderna e, em particular, da prótese, o termo tem sido utilizado para se referir a alguns seres humanos que após procedimentos médicos complexos se submeteram a experimentos de implantes robóticos em função de necessidades médicas. Entretanto, alguns usam esta tecnologia para estender sua capacidade de percepção (LANXON, 2014; SOLON, 2014).

Implantes cibernéticos cerebrais, junto com Inteligência Artificial (IA) e Realidade Aumentada (RA) podem ser integrados no cotidiano das pessoas e alterar seu comportamento pessoal. Da mesma forma que a pesquisa do Google e Wikipedia mudaram a forma de buscar e de se lembrar da informação, a IA e a RA poderiam alterar a forma de pensar e interagir. Seguindo o modelo do Google Glass, a IA de um implante neural pode ajudar a analisar rostos, em uma festa, por exemplo, e determinar aqueles socialmente mais relevantes para o usuário. Usando RA projetada em um implante ótico, a IA destaca cada pessoa em sua linha de visão e, quando o usuário se aproximar, pode fornecer um dossiê de seus principais interesses e tipo de personalidade. Pode-se aplicar este nível de acesso à informação a qualquer atividade, quer se trate de grelhar um bife ou a realização de um transplante de coração. A IA com a sobreposição da RA podem melhorar radicalmente a habilidade e capacidade humana em exercer suas atividades (MUNKITTRICK, 2011). Este tipo de tecnologia permitirá que seres humanos recuperem capacidades perdidas, por meio de acidentes, ou habilidades inexistentes, em função de desordens ocorridas em sua concepção, tornando-os aptos a desenvolver funções em várias áreas de aplicação.

A humanização dos robôs, ou a robotização humana, trará inúmeras possibilidades futuras visando o aumento da habilidade de lidar com o mundo. Questões de ética certamente irão surgir para orientar os avanços tecnológicos e suas implicações nesta área.

### 3 Nanotecnologia

A natureza tem testemunhado a evolução dos sistemas adaptativos complexos e extremamente inteligentes na condução dos processos biológicos encontrados na vida cotidiana. Por exemplo, uma célula pode fundir processos genéticos, ricos em informação, com sensores em escala nanométrica e atuadores, tornando-se um sistema molecular autônomo eficiente. Estes processos básicos que ocorrem no nível molecular inspiram uma nova abordagem de engenharia: a fusão de biotecnologia (BT), nanotecnologia (NT) e tecnologia da informação (TI). A NT permitiu a produção de novos materiais e dispositivos em escala molecular. Avanços biotecnológicos têm permitido aos cientistas manipular fisicamente vias genéticas ou estirpes de engenharia de proteínas. A informática serviu como catalisador para organizar e compreender o vasto conhecimento do ponto de vista do sistema. A fusão de BT, NT e TI culminará em arquiteturas de sistemas que podem apresentar características dos sistemas biológicos (HO; CHEN, 2007).

O campo da nanotecnologia tem gerado grande interesse nos últimos anos por causa de seu impacto sobre diferentes áreas, como produtos químicos, eletrônicos, agricultura, medicina, farmacêutica e indústria espacial. As nanopartículas são grupos de átomos na gama de tamanho de 1-100 nm ( $1 \text{ nm} = 10^{-9} \text{ m}$ ). Estas nanopartículas possuem propriedades físicas, químicas, ópticas e mecânicas bem definidas. Um dos objetivos da nanotecnologia é ser capaz de projetar, construir e controlar nanosistemas, adaptando-os às necessidades especificadas.

Com relação às aplicações agrícolas, a Embrapa e seus parceiros possuem um extenso trabalho na área de nanotecnologia. Na edição número 6 da Revista XXI – Ciência para a Vida, várias áreas de estudo e aplicações de nanotecnologia na agricultura são descritas no artigo publicado por Reynol e Freire (2014).

A nanobiotecnologia é um novo campo de pesquisa que tem o potencial de revolucionar ambas, a nanotecnologia e a biotecnologia. Entre as possibilidades idealizadas estão as aplicações médicas e veterinárias. Um dos usos concebidos é a restauração de funções fisiológicas danificadas onde as funções naturais passam a ser exercidas por nanomáquinas. Por exemplo, as células pancreáticas podem ser substituídas por nanofábricas capazes de reiniciar a produção endógena de insulina. Um conjunto diferente de possibilidades seria a substituição de organismos vivos, como as bactérias, na produção de drogas e enzimas, por nanofábricas versáteis e eficientes (BUCCI et al., 2014).

Um aspecto importante da nanotecnologia está relacionado com a concepção de métodos experimentais para a síntese de nanopartículas de diferentes composições química, tamanho, forma e dispersividade. Biossíntese de nanopartículas metálicas, utilizando plantas, encontra-se como um dos ramos de pesquisa atualmente. A maioria dos organismos multicelulares possuem a capacidade de sintetizar nanopartículas quer intracelularmente, quer extracelularmente. No entanto, ao contrário dos métodos químicos e físicos, o percurso biológico produz nanopartículas de metais não-tóxicos, e também é eficaz em termos de custos. Assim, a síntese de nanopartículas metálicas utilizando plantas está surgindo como um ramo importante da nanobiotecnologia. Uma grande variedade de plantas e partes de plantas, incluindo folhas, caule, casca e enzimas de plantas demonstraram a síntese bem sucedida de nanopartículas metálicas. Mais importante ainda, a partir do ponto de vista da comercialização, a planta é um sistema biológico não patogênico o que a torna vantajosa para a síntese de nanopartículas metálicas. (RAI; YADAV, 2013).

O impacto da nanotecnologia na tecnologia da informação gira em torno da utilização de nanomateriais para criar componentes menores, mais rápidos, com mais memória e mais eficientes para uso em computadores e em uma escala cada vez menor. Nessa escala, os novos nanomateriais possuem propriedades não observadas no nível micro, que permitem o desenvolvimento de novos dispositivos e aplicações.

A International Business Machines (IBM), por exemplo, está investindo significativamente em áreas de pesquisa como a nanoeletrônica, fotônica de silício de carbono, novas tecnologias de memória e arquiteturas que suportam computação quântica e computação cognitiva. Nos computadores atuais utiliza-se os transistores de silício fabricados com a tecnologia CMOS (complementary metal-oxide-semiconductor). Estes transistores estão se tornando cada vez menores aproximando-se da nanoescala. Devido à natureza do silício, e às leis da física, eles estão se aproximando de um ponto de limitação física. Novas tecnologias serão necessárias para viabilizar o próximo passo na construção de computadores. Alternativas potenciais incluem novos materiais como grafeno, nanotubos de carbono e nanofotônicos de silício (IBM CORPORATION, 2014).

O grafeno é carbono puro, sob a forma de uma folha de espessura atômica. Trata-se de um excelente condutor de calor e eletricidade, e também é notavelmente forte e flexível. Os elétrons podem se mover no grafeno cerca de dez vezes mais rápido do que em materiais semicondutores comumente usados, tais como silício e germânio silício. Ele é um nanomaterial que tem sido recentemente referido como o silício do século 21. As propriedades ópticas e eletrônicas únicas deste nanomaterial vão permitir o desenvolvimento de uma nova geração de dispositivos eletrônicos, por exemplo, nanotransistores, nanomemórias, nanobaterias, nano transceptores e nano-antenas que vão abrir as portas para ultracomunicações de banda larga entre os nanodispositivos (JORNET; AKYILDIZ, 2012).

Os nanotubos de carbono (NTC) são cilindros tubulares de átomos de carbono que possuem propriedades mecânicas, elétricas, térmicas, ópticas e químicas extraordinárias. Estas estruturas exibem 200 vezes a força e 5 vezes a elasticidade do aço; 5 vezes a condutividade elétrica e 15 vezes a condutividade térmica do cobre; e, aproximadamente a metade da densidade do alumínio. Como um produto à base de carbono, os NTC não apresentam os problemas de degradação ambiental ou físicos, comuns aos metais de maneira geral, como a expansão e contração térmica, corrosão e sensibilidade à radiação (NANOCOMP TECHNOLOGIES, 2014). Os nanotubos de carbono podem formar o núcleo de um dispositivo de transistor que irá funcionar de um modo semelhante ao transistor de silício corrente, mas com melhor desempenho. Eles podem ser usados para substituir os transistores em chips de computadores de alto desempenho e telefones inteligentes ultrarrápidos. Transistores de nanotubos de carbono podem funcionar como excelentes interruptores em dimensões moleculares de menos de dez nanômetros. Os estudos de circuitos eletrônicos sugerem uma melhora de cinco a dez vezes em seu desempenho em comparação com os circuitos de silício tradicionais (IBM CORPORATION, 2014).

Os nanofotônicos de silício utilizam pulsos de luz para a comunicação, em vez dos fios de cobre tradicionais. Eles proporcionam uma super rodovia para transmissão de grandes volumes de dados entre chips de computador em servidores, em grandes datacenters e em supercomputadores, aliviando, assim, as limitações de tráfego de dados. A tecnologia de nanofotônico de silício fornece respostas para os desafios de Big Data por possibilitar a conexão de grandes sistemas numa

distância de alguns centímetros ou alguns quilômetros, permitindo mover terabytes de dados por meio de pulsos de luz através de fibras ópticas (IBM CORPORATION, 2014)

O estudo e a utilização destas nanotecnologias estão permitindo o desenvolvimento de uma nova geração de computadores com a quebra do paradigma da forma como se imagina a computação atualmente. Os avanços de pesquisa exploratória irão levar a descobertas de chips de computador menores, mais rápidos e mais poderosos, possibilitando a nova era da computação incluindo a computação pervasiva, computação quântica, a computação neurosináptica e a computação neuromórfica, conforme abordado nas demais seções deste capítulo.

## 4 Computação pervasiva ou úbiqua e internet das coisas

A computação úbiqua tem como objetivo tornar a interação homem-computador invisível, ou seja, integrar as TIC com as ações e comportamentos naturais das pessoas. O termo invisível é utilizado no sentido de que as pessoas nem percebam que estão dando comandos a um computador, mas como se participassem de uma conversa com outra pessoa. Além disso, os computadores teriam sistemas inteligentes que estariam conectados ou procurando conexão o tempo todo, tornando-se assim onipresentes.

O primeiro passo para conseguir chegar a essa interação mais facilmente ou de forma invisível, é a utilização de interfaces naturais tais como: fala, gestos, percepção de presença no ambiente ou até mesmo a movimentação dos olhos. O segundo passo seria a geração de uma computação sensível ao contexto, tornando possível que os dispositivos possam capturar o contexto automaticamente. O contexto neste caso é a presença de uma pessoa no espaço ou qualquer tipo de movimento corporal, movimentação dos braços, dedos, cabeça, olhos e até movimentos faciais.

A computação úbiqua requer computadores pequenos, baratos e tecnologias de comunicação com ou sem fios que permitam a conexão com computadores de maior dimensão. Por exemplo, uma casa controlada por dispositivos de computação úbiqua deverá ter controle remoto da iluminação da casa, sistema de extinção de incêndios, sistemas de entretenimento integrados, sistemas para monitorizar a saúde dos ocupantes da casa, uma geladeira que avise os ocupantes da casa sobre produtos estragados ou fora da validade, entre outras funcionalidades (SANTOS, 2011; WIKIPÉDIA, 2014b).

A computação úbiqua ou pervasiva, em seus vários desdobramentos e aplicações, é considerada por muitos como o novo paradigma da Computação para o século XXI, o qual permitirá o acoplamento do mundo físico ao mundo da informação e fornecerá uma abundância de serviços e aplicações, permitindo que usuários, máquinas, dados, aplicações e objetos do espaço físico interajam uns com os outros de forma autônoma e transparente, criando a chamada Internet das Coisas (Internet of Things - IoT). Para construir este cenário, são necessários esforços de pesquisa multidisciplinares, envolvendo praticamente todas as áreas da Computação: sistemas distribuídos, sistemas móveis, redes de computadores, engenharia de software, entre outras.

Algumas tecnologias formam a base para a IoT como as etiquetas Radio Frequency Identification (RFID), as Redes de Sensores Sem Fio (RSSF), o GPS e as redes móveis que estão sempre evoluindo e possibilitando internet de alta velocidade e oferecendo serviços tanto para as áreas

urbanas quanto para as áreas remotas e as áreas rurais. As etiquetas RFID podem ser facilmente incorporadas em todos os tipos de coisas, devido ao seu tamanho reduzido e sua operação sem bateria. Porém, as etiquetas RFID têm capacidade restrita de processamento, de detecção ou de armazenamento de dados. Neste sentido, as RSSFs podem fornecer à IoT a capacidade de computação, de armazenamento de dados, e detecção necessárias. (JORNET; AKYILDIZ, 2012). A oferta de dispositivos conectados à internet pode ser tanto móvel como fixa, como por exemplo, refrigeradores, equipamentos de transporte, controladores de estoque de silos e armazéns. Um exemplo do uso da IoT é um sistema de produção agrícola que, pela análise de correlação entre a informação estatística da cultura e informações sobre o ambiente agrícola, utilizando sensores, tem melhorado a capacidade dos agricultores, pesquisadores e autoridades governamentais para analisar as condições atuais de clima, solo e planta e prever colheita futura (LEE et al., 2013).

Pesquisadores nos Estados Unidos estão desenvolvendo uma tecnologia que permitirá aos projetistas utilizar “tintas” especiais para imprimir sensores em miniatura dentro de máquinas e em superfícies quentes, duras e difíceis de alcançar. Esta tecnologia, chamada de gravação direta, irá permitir colocar sensores em lugares antes impossíveis contribuindo para acelerar a transição de conexão de máquinas e aparelhos à IoT. (DRIVES CONTROLS, 2014)

O suporte de um conjunto infinito de dispositivos de baixo consumo de energia, o apoio contínuo e acrescido de serviços de computação em nuvem e tempos de resposta muito curtos dentro do link de comunicação móvel permitirão o surgimento das redes de comunicação onipresentes. Uma área que tirará proveito das redes onipresentes é o conjunto de serviços em nuvem implantados. Hoje existem as soluções em nuvem como uma parte essencial dos hábitos de comunicação, tais como: calendário Dropbox, Google, serviços em nuvem da Amazon, e muitos mais. A fim de reduzir o atraso a esses serviços, hoje e no futuro, a nuvem física precisa ser distribuída geograficamente. Esta exigência será contínua no futuro, sendo que serviços em nuvem geograficamente espalhados vão se tornar componentes da rede de acesso. Um conceito que está surgindo é o de nuvens móveis, no qual a nuvem, ou subsistemas de nuvem são localizados nos dispositivos do usuário final. Por meio desta mudança de paradigma, a nuvem se torna distribuída, ou geograficamente espalhada. Como exemplo, se alguém desejar ler uma notícia na web haverá uma alta probabilidade de que no ambiente de nuvem móvel seja possível obter dados em cache que esteja geograficamente perto de sua estação de base, ou ainda melhor, a partir do dispositivo móvel da pessoa que está ao seu lado. As redes de acesso, como atualmente conhecidas, irão passar por mudanças drásticas, a fim de acomodar esta mudança de paradigma. Estas redes tornar-se-ão autoconscientes ou inteligentes (LEHNER; FETTWEIS, 2012).

Uma das estratégias bem sucedidas na competição empresarial é a eficácia da gestão de logística, especialmente quando os produtos agrícolas, que têm um ciclo de vida curto, estão em causa. Aliando a conectividade dos equipamentos à internet com a tecnologia de RFID, onde cada produto vegetal ou animal pode ser etiquetado, vislumbram-se aplicações de controle de estoque, irrigação inteligente e distribuição controlada de produtos. Neste caso, a gestão do tempo afeta diretamente a frescura dos produtos. Será possível acompanhar os produtos nas diversas etapas da cadeia de distribuição e, caso ocorra algum tipo de contaminação, eles poderão ser rastreados, a fim de verificar sua origem, contribuindo, desta forma, para a segurança alimentar, nutrição e saúde, promovendo o desenvolvimento sustentável da agricultura moderna (RFID BEEFS UP, 2013).

Outro exemplo de uso das etiquetas RFID é que uma vez que elas sejam transformadas em sensores RFID, elas podem ser utilizadas para detecção não invasiva de qualidade dos alimentos. Um sensor sem fio de baixo custo para a detecção, em tempo real, da qualidade e da possibilidade de contaminação dos alimentos já está sendo desenvolvido. Nesta abordagem, as etiquetas RFID são transformadas em sensores que têm potencial para criar uma alternativa robusta e barata, podendo ser usadas por pequenas empresas e indivíduos na indústria de alimentos e agricultura. Um exemplo é o uso das etiquetas RFID para medir a qualidade da carne ou de queijo em um supermercado usando a métrica de permissividade. O conceito também pode ser estendido para a área biomédica. Como a permissividade do músculo e tecido humanos é muito similar à permissividade da carne, uma aplicação potencial é uma antena usável para detectar tumores cancerígenos sob o músculo (RFID BEEFS UP, 2013).

Uma nova onda de avanço na computação ubíqua e na IoT está relacionada com o uso da nanotecnologia. Estão sendo propostos nanomateriais para desenvolver uma nova geração de nanocâmeras, nanophones e nanotransdutores acústicos que podem ser usados para gerar o conteúdo multimídia em nanoescala. Esses nanodispositivos vão superar as limitações dos dispositivos sensores multimídia atuais, proporcionando maior qualidade e capacidade de detecção de áudio, maior capacidade de armazenamento computacional e de dados, maior eficiência energética e taxas de dados de comunicação sem fio mais elevadas. A interligação de forma generalizada dos nanodispositivos de multimídia com redes de comunicações e, finalmente, com a internet define um sistema físico verdadeiramente cibernético que está sendo chamado como a internet de multimídia NanoCoisas (IoMNT). A IoMNT abre um campo inimaginável de aplicações em todas as áreas onde as coisas, os seres vivos e o ambiente estarão conectados de forma altamente complexa (JORNET; AKYILDIZ, 2012).

## 5 Informação, conhecimento e cognição

Na área da gestão da informação e do conhecimento, um tópico importante é garantir a disponibilidade, o acesso aberto e interoperabilidade dos dados relacionados com a agricultura bem como sua geoespacialização. Uma vez que dados relacionados à agricultura, e suas condições, estejam disponíveis de forma geoespacializada, vislumbra-se o cenário em que um agricultor posicionado em sua propriedade, e de posse de seu celular, possa ser geolocalizado obtendo informações que indiquem a melhor cultivar, condições de mercado e de produção para plantio em suas terras. Este agricultor pode estar localizado no Brasil, na América Latina ou na África, por exemplo.

Outra aplicação é a TeleAgricultura onde um agricultor localizado em uma região de difícil acesso pode se beneficiar de aplicações online que simulem e ensinem técnicas relacionadas ao plantio e manejo. Aliando à visão computacional e ao processamento de imagens, este mesmo agricultor pode, a partir de uma foto de uma folha com doença tirada de seu celular, obter o diagnóstico da doença e formas para seu tratamento. Na Austrália, por exemplo, a tendência de envelhecimento dos agricultores, a diminuição da força de trabalho rural, a necessidade de atrair e reter jovens agricultores, o crescimento no tamanho das propriedades e o crescimento de oportunidades comerciais internacionais sugerem a necessidade de investir em tecnologias que ajudem os agricultores a obter conhecimento da situação quando eles não têm condições para resolver os problemas sozinhos. Assim, a agricultura eletrônica tem o potencial de criar oportu-

nidades de serviços virtuais em agronomia, pecuária, saúde, estoque e suporte de máquinas, onde os consultores de base regional fornecem serviços de análise e consultoria de gestão para os prestadores de serviço locais. Conectividade de banda larga entre as comunidades rurais e as principais cidades irá garantir que esses prestadores de serviços locais tenham acesso ao conhecimento. Tais serviços podem auxiliar as atividades de extensão rural que tradicionalmente prestam serviço de assistência no campo (TAYLOR et al., 2013). Além disto, um outro serviço importante está associado com a Telemedicina onde parte da atenção à saúde pode ser realizada remotamente por meio de informação e redes de comunicação. Graças ao progresso em comunicação e processamento de tecnologias de informação (sensores, biossensores, comunicação sem fio, web semântica, telefones inteligentes, consoles de jogos, entre outras), a saúde onipresente poderá em breve ser uma realidade, levando melhores condições de vida para a população rural (AGOULMINE et al., 2012).

Outra área que tem sido vista como aplicação futura é a de cognição, onde os computadores exibirão as capacidades sensoriais de tato, visão, audição e paladar e a capacidade de interpretá-los. Pela característica do tato pode-se utilizar o dispositivo móvel para sentir texturas pela superfície da tela, auxiliando, por exemplo, na escolha de materiais para embalagens. Pela visão, os sistemas não apenas conseguirão enxergar e reconhecer o conteúdo de imagens e dados visuais, como também transformarão os pixels em significado, sendo capazes de entendê-los. Pela foto de uma fruta pode-se interpretar se ela está pronta para o consumo. Um sistema distribuído de sensores inteligentes irá detectar elementos de som, como pressão, vibrações e ondas sonoras, em diferentes frequências. O computador interpretará esses dados para prever quando árvores cairão em uma floresta ou quando um deslizamento ou tempestade são iminentes. Este sistema analisará o ambiente e medirá movimentos para alertar de perigos futuros. Papilas gustativas digitais ajudarão as pessoas a comerem de forma mais inteligente, uma vez que ajudarão a detectar se a comida é saudável. Minúsculos sensores embutidos ao computador ou ao celular detectarão se um animal ou uma pessoa está prestes a desenvolver alguma doença. Ao analisar odores, biomarcadores e milhares de moléculas na respiração de um animal ou de pessoa poderá ser possível ajudar os sistemas cognitivos a diagnosticar e monitorar, desde o início, os problemas de saúde (CRIVELINI, 2013).

O volume de informação disponível está acelerando, uma vez que as atividades do mundo têm sido cada vez mais expressas de forma digital nesta era do Big Data. Não se trata apenas do aumento de volume, mas também da velocidade, da variedade e da incerteza. A maioria dos dados agora vêm em formas não estruturadas, como vídeos, imagens, símbolos e linguagem natural. Assim, um novo modelo de computação é necessário para processar esta informação e melhorar e ampliar o conhecimento dos seres humanos. Mais do que serem programados para antecipar cada possível resposta ou ação necessária para executar uma função ou um conjunto de tarefas, os sistemas de computação cognitivos são treinados, usando a inteligência artificial e algoritmos de aprendizado de máquina para detectar, prever, inferir e, de certa forma, pensar.

Rumo à computação cognitiva, a IBM, utilizando nanotecnologia, neurociência e supercomputação, anunciou o desenvolvimento do chip denominado SyNAPSE, inspirado na arquitetura do cérebro, visando à construção do computador neurosináptico. O computador neurosináptico é do tamanho de um selo e consome energia equivalente a uma bateria de um aparelho auditivo. Trata-se do primeiro chip de computador neurosináptico a atingir a escala de um milhão de neurônios programáveis, 256 milhões de sinapses programáveis e 46 bilhões de operações sinápticas por

segundo por watt. O chip de 5,4 bilhões de transistores é construído graças ao processo de nanotecnologia da Samsung numa escala de 28 nm. Estes sistemas podem processar eficientemente dados sensoriais de alta dimensão, mesmo com ruídos, em tempo real, enquanto consomem ordens de magnitude menos energia do que as arquiteturas de computadores convencionais (MODHA, 2014).

Ao contrário dos sistemas especialistas atuais que necessitam de um especialista humano para elaborar as regras a serem codificadas no sistema, os computadores cognitivos podem processar linguagem natural e dados não estruturados, e aprender pela experiência, da mesma forma que os humanos. Eles se tornarão os novos sistemas de apoio à decisão. Os sistemas de computação cognitivos irão utilizar imagem e reconhecimento de voz para compreender o mundo e interagir mais facilmente com os humanos. Usando análises visuais e técnicas de visualização de dados, computadores cognitivos podem exibir dados de uma forma visualmente atraente, para auxiliar a tomada de decisão, baseadas no enorme volume de dados (IBM RESEARCH, 2014). Suas aplicações incluem analisar dados referentes à modelagem de sistemas biológicos, sistemas complexos e toda a gama de informação que diariamente se torna online tornando praticamente impossível que os seres humanos interpretem todo este conhecimento sem auxílio da computação.

A transformação da informação codificada digitalmente em objetos sólidos, pelo uso das impressoras tridimensionais (3D), também promete revolucionar a área da manufatura, da medicina e das ciências da vida. Apesar de a maioria das impressoras 3D atuais serem usadas para prototipagem rápida de produtos, ou para a produção de moldes, o seu uso para fabricar produtos finais também já está ocorrendo. Este processo é conhecido como manufatura aditiva. Para uma pequena quantidade de produtos a serem manufaturados, a manufatura aditiva apresenta custo benefício melhor que montar uma linha de produção. Além disto, a impressão 3D permite a customização do produto em função das necessidades dos usuários (THREE-DIMENSIONAL PRINTING..., 2011).

Embora a manufatura aditiva em si já seja um avanço na área industrial, a tecnologia de bioimpressão 3D promete revolucionar ainda mais as áreas de medicina e biologia. As bioimpressoras 3D irão imprimir células e poderão ser utilizadas para impressão de órgãos. A empresa Organovo (<http://www.organovo.com/>) já tem criado tecidos de órgãos humanos para uso em testes laboratoriais. Para o futuro já se fala em bioimpressão 3D “in situ”, onde se espera desenvolver técnicas de impressão de tecidos direto no corpo humano. À medida que a bioimpressão 3D avance, poderá ser possível imprimir órgãos criados a partir das próprias células dos pacientes evitando, assim, o problema da rejeição. Junto com a evolução da nanotecnologia e da engenharia genética, a bioimpressão 3D é uma ferramenta poderosa para aqueles em busca de prolongamento da vida. Por meio destas ciências, os médicos, engenheiros e cientistas da computação estão, cada vez mais, aprendendo a manipular tecidos vivos em seu nível celular mais básico possibilitando aumentar cada vez mais o tempo e a qualidade de vida das pessoas (BARNATT, 2013).

## 6 Data science, computação quântica e neuromórfica

A Computação Ubíqua e Internet das Coisas, a miniaturização dos dispositivos até a escala nanométrica, os resultados dos sequenciamentos de genomas e as redes de sensores, dentre outras

aplicações da computação, têm gerado um volume cada vez maior de dados, com crescente variedade e velocidade de coleta, resultando no que se chama de Big Data.

Este acúmulo constante nos dados, resulta na possibilidade de construir ciência de uma outra forma: ao invés de construir teorias suportadas por exercício mental sobre teorias previamente estabelecidas, usando os dados simplesmente para validá-las, o aprendizado de máquina pode inverter este processo construtivo. Os dados em abundância e o processamento computacional massivo podem auxiliar na investigação científica (DHAR, 2013) dentro de um processo que tem sido chamado de “Data Science”, ou ciência dos dados.

No artigo escrito pelo editor da Wired Magazine, Chris Anderson (ANDERSON, 2008) discorre sobre o fim da teoria. Este artigo inicia com uma frase do estatístico inglês George Box que afirma que “todos os modelos estão errados, mas alguns são úteis”. Isto decorre do fato de que modelos são aproximações/simplificações de uma realidade que se deseja representar e, sendo simplificações, algum aspecto ficará de fora dos modelos teóricos.

Na era da inundação causada pela informação na escala do hexabite, os dados disponíveis para análise são tamanhos que é possível fazer inferências sobre eles antes mesmo de se possuir uma teoria que relacione causa-efeito. A análise matemática sobre os dados pode ser feita antes e a contextualização do que foi aprendido a posteriori. Exemplificando, o tradutor do Google, e o sistema computacional Watson da IBM, capaz de responder perguntas em linguagem natural, não entendem a semântica dos dados e o significado das perguntas do jogo de perguntas e respostas Jeopardy, transmitido pela televisão americana. Tudo é feito matematicamente por algoritmos de aprendizado de máquina que não compreendem a natureza dos dados que estão tratando. Apesar disto, as traduções estão melhores a cada dia e o Watson venceu seus concorrentes humanos no jogo Jeopardy.

Na construção científica aprende-se que não existe efeito sem causa e que a mera correlação matemática pode ser apenas coincidência. Entretanto, grande parte dos modelos sobre sistemas complexos como, por exemplo, os que envolvem interações sociais, economia e interações solo-planta-atmosfera são inerentemente incompletos. O que está por trás da escolha dos modelos resultantes do aprendizado de máquina pode ser sintetizado pela atualização que Peter Norvig, consagrado autor de livros sobre inteligência artificial, fez da frase de Box: “todos os modelos estão errados, e mais, você pode obter sucesso sem eles”.

Na física, o modelo newtoniano foi sucedido pelo einsteiniano que também não responde adequadamente às questões do mundo subatômico. Para tanto, estão sendo construídas teorias n-dimensionais para as quais são necessários aceleradores de partículas caríssimos capazes de comprová-las ou refutá-las à luz novamente de um modelo. Eis o porquê da preferência de um simples modelo preditivo, baseado na análise dos dados, em detrimento de uma teoria incompleta, já que este modelo também pode ser atualizado periodicamente.

A biologia também tem desafiado a teoria: as redes de interação gênicas entre o deoxyribonucleic acid (DNA) e o ambiente tem refutado o que parecia ser determinado apenas pelo DNA. Anderson comenta em seu artigo (ANDERSON, 2008) que Craig Venter, cientista pioneiro em biologia sintética, iniciou sequenciando organismos individuais e passou a sequenciar ecossistemas inteiros e, com isto, acabou por descobrir milhares de bactérias e outras formas de vida desconhecidas, tendo “avançado a biologia mais do que qualquer outro da sua geração”, e tudo o que ele usou foram algoritmos matemáticos.

Toda essa atividade científica, fortemente apoiada na computação, constitui uma enorme pressão no aumento do poder de processamento computacional que já começa a dar sinais de arrefecimento. Em 1965, o presidente da Intel, Gordon Moore, previu que a cada 18 meses a capacidade dos processadores duplicaria, o que se tornou conhecido como Lei de Moore. Ocorre que, ao se aproximar da escala atômica, previu-se também que esta ascensão não poderia ocorrer indefinidamente, tendo o seu fim previsto para 2020 (KURZWEIL, 2001; WIKIPEDIA, 2014a).

Várias soluções têm sido propostas para avançar neste limite. Em 2003, Phil Kuekes, cientista do laboratório de ciência quântica da HP obteve a patente de um chaveador “crossbar” que permite a construção de portas lógicas explorando junções em escala molecular entre nanofios que se cruzam perpendicularmente.

A computação quântica também pode auxiliar. Os computadores convencionais são construídos a partir de chips de silício que contêm milhões ou bilhões de transistores em miniatura. Cada um destes pode ser “ligado” ou “desligado” para representar um valor de qualquer um “1” ou “0”. Os computadores convencionais podem armazenar e processar dados usando “dígitos binários” ou “bits”. Em contraste, os computadores quânticos irão trabalhar com “bits quânticos” ou “qubits”. Estes são representados em hardware usando estados quânticos em vez de transistores que são transformados “on” ou “off”. Devido às leis peculiares da mecânica quântica, qubits individuais podem representar um valor de “1”, “0” ou ambos os números ao mesmo tempo. Isso ocorre porque as partículas subatômicas usadas como qubits podem existir em mais de um estado, ou uma superposição de estados, exatamente no mesmo ponto no tempo. Este fato permite que os computadores quânticos possam ser várias ordens de magnitude mais poderosos que os computadores digitais convencionais, possuindo o potencial para realizar processamento paralelo massivo. Isto significa que os computadores quânticos serão mais eficazes no desempenho de tarefas como reconhecimento de visão, diagnóstico médico e outras formas de processamento de inteligência artificial que dependam de atividades complexas de reconhecimento de padrões para além das capacidades de ambos os computadores tradicionais e seres humanos (BARNATT, 2012).

Outra possibilidade são os computadores neuromórficos (MONROE, 2014). Na computação tradicional, os dados são trazidos da memória, processados e armazenados novamente na memória, processo conhecido como arquitetura von Neumann. Nos computadores neuromórficos, dados e memória ficam juntos em um conjunto enorme de “neurônios primitivos”, cada um se comunicando com outros milhares de neurônios, imitando o que ocorre no cérebro. Como dados e processamento estão muito próximos, estes computadores são extremamente eficientes do ponto de vista energético. Correntemente, os computadores neuromórficos estão sendo desenvolvidos como parte do bilionário projeto europeu Human Brain Project primeiramente para compreender o funcionamento do cérebro e melhorar a tecnologia.

Além do aumento da demanda por processamento, tem-se também o aumento da capacidade de armazenamento, que também segue a Lei de Moore. Apenas para citar um exemplo, o Grande Colisor de Hádrons (LHC) do Centro Europeu de Pesquisas Nucleares (CERN) gera 30 petabytes de dados brutos por ano que necessitam ser analisados por um “grid” de computadores em rede que compreende 140 centros de computação distribuídos em 35 países.

A principal abordagem para a demanda maciça de processamento tem sido o processamento em grid e a computação paralela, porém a natureza das aplicações nem sempre permite a escolha destes modelos de processamento. Jacek Becla, do National Accelerator Laboratory nos Estados

Unidos comenta que, normalmente, na computação de alto desempenho tem-se poucos dados e pouca entrada e saída (I/O) destes dados, diferentemente do que ocorre na era do Big Data em que se tem que trabalhar com um trilhão de informações em memória, de uma só vez (WRIGHT, 2014).

A necessidade intensiva de transferir dados de entrada e saída para a memória tem-se revelado um desafio à construção de aplicações que explorem paralelismo, pois na maior parte do tempo o processador permanece em estado ocioso, à espera de que o dado seja obtido da memória, para que o processamento possa continuar. Isto significa que as aplicações devem ser construídas para transferir a maior quantidade de dados possível para trabalhar na memória e otimizar as aplicações para explorar o paralelismo neste cenário.

A capacitação necessária no currículo para os que trabalham com Big Data e Data Science é ampla, pois inclui computação distribuída, computação paralela, computação tolerante a falhas, estatística, especialmente a Bayesiana, conhecimento sobre correlação e causalidade, e habilidade para formulação de problemas que resultem em soluções efetivas (DHAR, 2013). O futuro para quem desenvolve aplicações voltadas para as ciências da vida e, em particular, para a agricultura é bastante promissor e desafiador. Certamente o que não ocorrerá será a monotonia.

## 7 Considerações finais

A ação sinérgica dos quatro campos científicos e tecnológicos, que apresentaram crescimento acelerado nas últimas décadas, como a NT, a BT as tecnologias de comunicação e informação e as ciências cognitivas (neurociência), tem sido intitulada Convergência Tecnológica. Algumas das possíveis aplicações práticas da nova Convergência Tecnológica residem na melhoria da saúde e da capacidade física humana considerando tópicos como nanobioprocessadores para a pesquisa e o desenvolvimento de novas estratégias terapêuticas, incluindo aquelas resultantes da bioinformática, da genômica e da proteômica; implantes de base nanotecnológica e biosistemas regenerativos para substituir órgãos humanos ou para monitoramento do bem-estar fisiológico; dispositivos em escala nanométrica para a intervenção médica, plataformas multimodais para aumentar a capacidade sensorial, em especial para os indivíduos com déficits visuais e auditivos; interfaces cérebro-cérebro ou cérebro-máquina; remoção das barreiras de comunicação determinadas por incapacidade física, pela diferença de línguas, pela distância geográfica e pelos diferentes níveis de conhecimento, levando, assim, a um aumento na efetividade e na eficiência da cooperação entre ambientes educacionais, corporativos, de governo e outros. Outras áreas de aplicação são a expansão da cognição e da comunicação humana, devendo-se atribuir alta prioridade aos esforços multidisciplinares que levam à compreensão da estrutura, das funções e do aprimoramento potencial da mente humana (CAVALHEIRO, 2007).

Dada a complexidade e heterogeneidade das tecnologias emergentes como as TIC, a nanotecnologia, a biotecnologia, a robótica e a agricultura de precisão, e suas aplicações na agroindústria, é natural que ainda não se tenha uma perspectiva sistemática, integrada e interdisciplinar entre elas. Em alguns campos bem instalados como a biotecnologia surgem, de tempos em tempos, novas técnicas que potencializam a capacidade de conhecer, modificar e conservar organismos de utilidade atual ou potencial. A Agro-Nanotecnologia, por sua vez, apresenta o desafio dobrado de ser um campo novo somado a várias questões de risco e de ética que começam a

ser discutidos. A expectativa é que estas áreas se integrem de forma multidisciplinar surgindo a ideia das NanoRoboTICs ou AgriNanoRoboTICs ou Metagenômica-Agricultura de Precisão (conforme discutido no âmbito do Programa Cooperativo para o Desenvolvimento Tecnológico Agroalimentar e Agroindustrial do Cone Sul - ProciSur). Estas tecnologias emergentes trabalhando de forma sinérgica vão possibilitar otimizar os recursos humanos, o conhecimento e a capacidade de inovação para garantir o avanço da produção agrícola e do agronegócio. Tem se falado em sustentabilidade computacional uma vez que o uso das tecnologias emergentes na agricultura devem colaborar para garantir a sustentabilidade ambiental, a segurança alimentar e, conseqüentemente, melhorar a qualidade de vida, bem estar e a saúde da humanidade.

A percepção do valor da inserção das TIC e NBIC para promover avanços na agricultura e na qualidade dos produtos ou serviços em oferta aos agricultores é interdependente. Sem agricultores que valorizam a inserção das TIC na agricultura, fazendo uso de seus serviços, não há como criar um mercado viável em torno delas. Sem um mercado viável, não tem como desenvolver um serviço de alta qualidade. E, na falta deste serviço, não há como facilitar a absorção das tecnologias. No “círculo dependência” da agricultura computacional, a demanda do mercado está intimamente ligada à qualidade dos produtos ou serviços oferecidos, que por sua vez influencia a demanda do mercado, bem como o nível de investimento no exterior, educação e infraestrutura, os quais alimentam o ciclo (TAYLOR et al., 2013). Além disto, os serviços disponíveis devem ter o potencial de possibilitar a comunicação em escala global considerando as diferenças de idiomas bem como a distribuição geográfica envolvendo todos os setores do agronegócio.

Por fim, uma questão que não pode deixar de ser abordada é que apesar das inúmeras contribuições do uso das TIC nas diversas áreas que atingem a população mundial, o seu uso de forma não racional tem preocupado as autoridades. As TIC, como indústria, também são responsáveis por impactos no meio ambiente. Atualmente, fala-se da TIC Verde onde a grande preocupação é a continuidade da oferta dos serviços de TIC respeitando o meio ambiente. Muitos provedores de processamento de dados e de hospedagem de serviços web, e de nuvem manipulam enormes quantidades de dados, softwares e servidores tanto para organizações quanto para indivíduos. Para prover este tipo de serviços, estas indústrias necessitam de um poderoso parque computacional. Um dos maiores impactos deste tipo de indústria está no consumo de energia e na dissipação de calor. Não é raro que os centros de dados, construídos atualmente, consumam mais energia do que a população da cidade onde eles estão localizados. Este fato tem forçado que os centros de dados se desloquem para locais onde a energia seja mais barata ou que os recursos naturais ajudem na dissipação de calor, como por exemplo, movendo-se para regiões geladas.

O impacto das TIC no clima terrestre e nos seus recursos em escassez é uma outra preocupação. Estudos recentes indicam que as emissões de dióxido de carbono dos centros de dados ultrapassam as emissões de muitas nações, individualmente. Adicionalmente, muitos equipamentos de Tecnologia da Informação (TI) contêm substâncias químicas tóxicas como mercúrio e chumbo muitos dos quais são lançados ao meio ambiente pelo descarte inadequado de equipamentos de TIC obsoletos.

Pelo cenário descrito, é necessária uma colaboração séria entre tecnólogos, desenvolvedores, pesquisadores, consumidores e políticos para alcançar uma TIC verde e autossustentável. A falha em atacar estes problemas pode ser devastadora para o meio ambiente e vai acarretar que as contas com os gastos de energia consumam cada vez mais os investimentos em TIC. O

esforço de apenas um ramo não é suficiente para diminuir os gastos de energia em TIC e a sua pegada no meio ambiente. É necessário um esforço conjunto incluindo áreas multidisciplinares da indústria, pesquisa e sociedade. Neste processo, algumas iniciativas incluem: melhoramento da eficiência da infraestrutura de TIC utilizando recursos e equipamentos de baixo consumo; alocação eficiente de recursos de TIC para execução das tarefas utilizando técnicas como virtualização e computação autônoma; desenvolvimento de novas tecnologias, materiais e equipamentos de baixa voltagem; uso de tecnologias avançadas de resfriamento para dissipação do calor juntamente com equipamento para geração de energia; estabelecimento de programas de reciclagem na base do reduzir-reciclar-reusar e disseminação de informação sobre a importância dos temas ligados à TIC verde. Muitos fabricantes têm estabelecido políticas para recolhimento de seus produtos obsoletos tentando diminuir sua pegada no ambiente. Uma outra iniciativa é a diminuição das substâncias tóxicas como, por exemplo, o chumbo e o mercúrio em seus produtos (YOUSIF, 2009).

Este capítulo apresentou um levantamento das tendências das TIC incluindo os avanços nas áreas de computação, nanotecnologia e robótica que, aplicadas, em conjunto com a biologia, a gestão da informação e do conhecimento, a agricultura de precisão e a engenharia genética prometem alavancar o desenvolvimento nas cadeias produtivas agrícolas e melhorar as condições de vida do meio rural. O futuro é promissor. A capacidade humana em gerar novos conhecimentos é cada vez mais impulsionada por toda esta evolução tecnológica e a expectativa é de um horizonte onde todos estes avanços levem ao aumento da sustentabilidade ambiental, social e econômica da agricultura na América do Sul.

## 8 Referências

- AGOULMINE, N.; RAY, P.; WU, T. Efficient and cost-effective communications in ubiquitous healthcare: wireless sensors, devices and solutions. **IEEE Communications Magazine**, New York, v. 50, n. 5, p. 90-91, May, 2012.
- ALISSON, E. **Sensor identifica insetos pela frequência do batimento das asas**. São Paulo: Agência Fapesp, 2014. Disponível em: <[http://agencia.fapesp.br/sensor\\_identifica\\_insetos\\_pela\\_frequencia\\_do\\_batimento\\_das\\_asas/1937/](http://agencia.fapesp.br/sensor_identifica_insetos_pela_frequencia_do_batimento_das_asas/1937/)>. Acesso em: 15 set. 2014.
- ANDERSON, C. The end of theory: the data deluge makes the scientific method obsolete. **Wired Magazine**, San Francisco, June 2008. Disponível em: <[http://archive.wired.com/science/discoveries/magazine/16-07/pb\\_theory](http://archive.wired.com/science/discoveries/magazine/16-07/pb_theory)>. Acesso em: 9 out. 2014.
- AUAT CHEEIN, F. A.; CARELLI, R. Agricultural Robotics: unmanned robotic service units in agricultural tasks. **Industrial Electronics Magazine**, New York, v. 7, n. 3, p.48-58, Sept. 2013. DOI: 10.1109/MIE.2013.2252957.
- BARNATT, C. **Bioprinting**: a future sharper's toolbox. 2013. Disponível em: <<http://www.explainingthefuture.com/bioprinting.html>>. Acesso em: 20 set. 2014.
- BARNATT, C. **Quantum computing**: a guide to computing. 2012. Disponível em: <<http://www.explainingcomputers.com/quantum.html>>. Acesso em: 19 set. 2014.
- BAUCKHAGE, C.; KERSTING, K.; SCHMIDT, A. Agriculture's technological makeover. **Pervasive Computing**, New York, v. 11, n. 2, p. 4-7, Apr./June 2012. DOI: 10.1109/MPRV.2012.25.
- BUCCI, E. M.; BUCCI, O. M.; SORRENTINO, R. Nanotechnology and life: an engineer's perspective [Point of View]. **Proceedings of the IEEE**, New York, v. 102, n. 6, p. 930-935, June 2014. DOI: 10.1109/JPROC.2014.2320315.

CAVALHEIRO, E. A. A nova convergência da ciência e da tecnologia. **Novos estudos CEBRAP**, São Paulo, n. 78, p. 23-30, jul. 2007. Disponível em: <<http://www.scielo.br/pdf/nec/n78/04>>. Acesso em: 21 set. 2014.

CRIVELINI, W. **Bancos de Dados & BI**: computação cognitiva explorando tato, olfato e paladar. [S.l.]: IBM DEVELOPERWORKS, 2013. Disponível em: <[https://www.ibm.com/developerworks/community/blogs/ibmacademiccell/entry/computa\\_c3\\_a7\\_c3\\_a3o\\_cognitiva\\_explorando\\_tato\\_olfato\\_e\\_paladar?lang=en](https://www.ibm.com/developerworks/community/blogs/ibmacademiccell/entry/computa_c3_a7_c3_a3o_cognitiva_explorando_tato_olfato_e_paladar?lang=en)>. Acesso em: 21 set. 2014.

DHAR, V. Data science and prediction. **Communications of the ACM**, New York, v. 56, n. 12, p. 64-73, Dec. 2013. Doi:10.1145/2500499.

DRIVES & CONTROLS. **Technology can 'print' 3D sensors inside machinery**. 2014. Disponível em: <[http://www.drivesncontrols.com/news/fullstory.php/aid/4471/Technology\\_can\\_\\_91print\\_92\\_3D\\_sensors\\_inside\\_machinery.html](http://www.drivesncontrols.com/news/fullstory.php/aid/4471/Technology_can__91print_92_3D_sensors_inside_machinery.html)>. Acesso em: 28 ago. 2014.

HARRIS, A. Agricultural technology to feed the world. **Engineering & Technology Magazine**, v. 8, n. 10, não paginado, Oct. 2013. Disponível em: <<http://eandt.theiet.org/magazine/2013/10/feed-the-world.cfm>>. Acesso em: 28 ago. 2014.

HO, C.; CHEN, J. M. When worlds collide. **IEEE Nanotechnology Magazine**, New York, v. 1, n. 1, p.18-21, Sept. 2007.

IBM CORPORATION. IBM Announces \$3 billion research initiative to tackle chip grand challenges for cloud and big data systems. 2014. Disponível em: <<https://www-03.ibm.com/press/us/en/pressrelease/44357.wss>>. Acesso em: 18 out. 2014.

IBM RESEARCH. Cognitive computing. Disponível em: <<http://www.research.ibm.com/cognitive-computing/index.shtml#fbid=njQjG3Ua-Iu>>. Acesso em: 20 out. 2014.

JORNET, J. M.; AKYILDIZ, I. F. The internet of multimedia nano-things in the Terahertz Band. In: EUROPEAN WIRELESS CONFERENCE, 18., Poznan, 2012. **Proceedings...** Poznań: Poznań University of Technology, 2012. p. 1-8, 18-20. <Disponível em: <http://www.ece.gatech.edu/research/labs/bwn/papers/2012/c4.pdf>>. Acesso em 20 out. 2014.

KIM, K.-H.; SHIM, W.; MOON, Y.-H.; KIM, K.-H.; SON, J.-K.; KWON, O.-J The structure of bio-information-nano technology convergence from firms' perspective. In PICMET, 2012, Vancouver. **Technology management for emerging technologies**: proceedings. New York: IEEE, 2012. p. 579-588. PICMET '12.

KURZWEIL, R. **The law of accelerating returns**. 2001. Disponível em: <<http://www.kurzweilai.net/the-law-of-accelerating-returns>>. Acesso em: 21 set. 2014.

LANXON, N. **Practical transhumanism**: five living cyborgs. 2014. Disponível em: <<http://www.wired.co.uk/news/archive/2012-09/04/cyborgs>>. Acesso em: 06 out. 2014.

LEE, M.; HWANG, J.; YOE, H. Agricultural production system based on IoT. In: IEEE INTERNATIONAL CONFERENCE ON COMPUTATIONAL SCIENCE AND ENGINEERING, 16., 2013, Sidney. **Proceedings...** New York: IEEE, p. 833-837, 2013. CSE 2013. Disponível em: <<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6755306>>. Acesso em: 20 out. 2014. DOI 10.1109/CSE.2013.126.

LEHNER, W.; FETTWEIS, G. Paving the path for the future technology developments [includes 9 white papers]. In: IEEE TECHNOLOGY TIME MACHINE SYMPOSIUM, 2012, Dresden. **Proceedings...** New York, 2012. p. 1-38. DOI: 10.1109/TTM.2012.6509029.

MODHA, D. **Introducing a Brain-inspired computer**: TrueNorth's neurons to revolutionize system architecture. 2014. Disponível em: <<http://www.research.ibm.com/articles/brain-chip.shtml>>. Acesso em: 16 out. 2014.

MONROE, D. Neuromorphic Computing Gets Ready for the (Really) Big Time. **Communications of the ACM**, New York, v. 57, n. 6, p.13-15, June 2014. DOI: 10.1145/2601069.

MUNKITTRICK, K. When will we be transhuman? seven conditions for attaining transhumanism. **Discover Magazine**, July 2011. Disponível em: <<http://blogs.discovermagazine.com/sciencenotfiction/2011/07/16/when-will-we-be-transhuman-seven-conditions-for-attaining-transhumanism/>>. Acesso em: 12 out. 2014.

NANOCOMP TECHNOLOGIES. What are carbon nanotubes? 2014. Disponível em: <<http://www.nanocomptech.com/what-are-carbon-nanotubes>>. Acesso em: 21 out. 2014.

PETERSON, D. **University of Illinois creates robot farmers**. 2014. Disponível em: <[http://www.eurekalert.org/pub\\_releases/2004-07/uoia-uo070604.php](http://www.eurekalert.org/pub_releases/2004-07/uoia-uo070604.php)>. Acesso em: 10 set. 2014.

RAI, M.; YADAV, A. Plants as potential synthesiser of precious metal nanoparticles: progress and prospects. **IET Nanobiotechnology**, Herts, v. 7, n. 3, p. 117-124, Sept. 2013. doi: 10.1049/iet-nbt.2012.0031.

REYNOL, F.; FREIRE, V. Nanotecnologia: no mundo das moléculas. **XXI Ciência para a Vida**, Brasília, DF, n. 6, p.12-27, 2014.

RFID BEEFS up. **Electronics Letters**, London, v. 49, n. 25, p.158, Dec. 2013. DOI: 10.1049/el.2013.3883. Disponível em: <<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6678438&isnumber=6678435>>. Acesso em: 20 out. 2014.

RODRIGUES, N. Inteligência artificial: da ficção para a realidade. **XXI Ciência para a Vida**, Brasília, DF, n. 6, p. 28-31, 2014.

ROBOT plants: putting down roots. New York: Alacra Store, 2013. Disponível em: <[http://news.alacrastore.com/economist/Robot-plants-Putting-down-roots-EN\\_EN\\_MAIN\\_20130727T000000\\_0065](http://news.alacrastore.com/economist/Robot-plants-Putting-down-roots-EN_EN_MAIN_20130727T000000_0065)>. Acesso em: 25 ago. 2014.

SANTOS, A. F. G. **Computação onipresente**: ficção ou realidade? 2011. Disponível em: <[https://www.ibm.com/developerworks/community/blogs/tlcb/entry/computacao\\_onipresente\\_ficcao\\_ou\\_realidade?lang=en](https://www.ibm.com/developerworks/community/blogs/tlcb/entry/computacao_onipresente_ficcao_ou_realidade?lang=en)>. Acesso em: 21 set. 2014.

SOLON, O. **The cyborg foundation**: we urge you to become part-machine. 2014. Disponível em: <<http://www.wired.co.uk/news/archive/2013-10/30/cyborg-foundation>>. Acesso em: 6 out. 2014.

TAYLOR, K.; GRIFFITH, C.; LEFORT, L.; GAIRE, R.; COMPTON, M.; WARK, T.; LAMB, D.; FALZON, G.; TROTTER, M. Farming the Web of Things. **IEEE Intelligent Systems**, Los Alamitos, v. 28, n. 6, p. 12-19, Nov./Dec. 2013. DOI: 10.1109/MIS.2013.102. Disponível em: <<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6617630>>. Acesso em: 20 out. 2014.

THREE-DIMENSIONAL printing from digital designs will transform manufacturing and allow more people to start making things. In: 3D PRINTING The printed world, 2011. Disponível em: <<http://www.economist.com/node/18114221>>. Acesso em: 20 set. 2014.

WIKIPEDIA. Moore's Law. Wikipédia, the free encyclopedia. 2014a. Disponível em: <[http://en.wikipedia.org/wiki/Moore%27s\\_law](http://en.wikipedia.org/wiki/Moore%27s_law)>. Acesso em: 10 out. 2014.

WIKIPEDIA. **Computação Ubíqua**. 2014b. Disponível em: <[http://pt.wikipedia.org/wiki/Computação\\_ubíqua](http://pt.wikipedia.org/wiki/Computação_ubíqua)>. Acesso em: 5 set. 2014.

WRIGHT, A. Big data meets big science. **Communications of the ACM**, New York, v. 57, n. 7, p. 13-15, July 2014. DOI: 10.1145/2617660. Disponível em: <<http://cacm.acm.org/magazines/2014/7/176202-big-data-meets-big-science/fulltext>>. Acesso em: 20 out. 2014.

YOUSIF, M. Towards green ICT - keynote. **Ercim News**, France, n. 72, p. 3, Oct. 2009. Disponível em: <<http://ercim-news.ercim.eu/en79/keynote/664-keynote-towards-green-ict>>. Acesso em: 20 set. 2014.

ZAPPA, M. **15 Emerging agriculture technologies that will change the world**. May 2014. Disponível em: <<http://www.businessinsider.com/15-emerging-agriculture-technologies-2014-4>>. Acesso em: 5 set. 2014.





# Parte VII



# Experiência de TIC na América do Sul





## **Tecnologias de Informação e Comunicação e sua relação com a agricultura - Chile**

Gustavo Alejandro Chacón Cruz

### **1 Introdução**

No escopo de um compromisso assumido pela Plataforma de Tecnologias Emergentes do Procisur<sup>1</sup>, o objetivo deste documento consiste em diagnosticar e analisar a situação em 2014, o impacto e a evolução das TIC na pesquisa agrícola, na transferência de conhecimento e no desenvolvimento agrícola no Chile.

A Plataforma de Tecnologias Emergentes do Procisur tem a missão de pesquisar o estado atual da aplicação e da demanda de tecnologias emergentes na agricultura do Cone Sul, tais como: TIC, biotecnologia avançada, agricultura de precisão e nanotecnologia.

A agricultura está incorporando novas tecnologias para enfrentar os desafios que levam à evolução dos padrões tecnológicos de maneira a produzir alimentos para uma população mundial estimada em 9,2 bilhões para 2050. (FAO, 2009). Neste contexto, a tecnologia da informação e da comunicação (TIC) já está sendo utilizada e tem um papel cada vez mais importante na gestão da cadeia de valor. As TIC estão evoluindo de modo contínuo e têm atraído um aumento dos investimentos públicos e privados na pesquisa, na transferência de tecnologia, na extensão e no desenvolvimento agrícola.

TIC é uma designação geral, definida pela Agência Norte-Americana para a Cooperação Internacional (Usaid), como a combinação de hardware, software e ferramentas de produção que possibilitem o intercâmbio, o tratamento e a gestão da informação e conhecimento. De acordo com a Usaid, as TIC são tecnologias e métodos para armazenar, gerir e processar a informação (por exemplo, computadores, software, livros, PDAS, tablets, bibliotecas digitais e não digitais) e para comunicar informação (por exemplo, correio, e-mail, rádio, televisão, telefones celulares, localizadores, internet, entre outros). É indispensável para o setor agrícola reduzir o seu preço, aumentar a acessibilidade, adaptabilidade e as novas capacidades de seu uso. Os produtores podem usar a Internet, o telefone e outras ferramentas digitais para a observação do clima, a gestão de frotas de veículos, a rastreabilidade dos produtos agrícolas, a busca de informação sobre preços dos insumos, a contratação de produtos ou serviços; o acesso a mercados, variedades,

<sup>1</sup> Disponível em: <<http://www.procisur.org.uy>>.

técnicas de produção, serviços de armazenamento ou processamento de matéria-prima, entre outras aplicações.

Felizmente, na atualidade, estão cada vez mais disponíveis os dispositivos móveis (telefones inteligentes, tablets etc), infraestrutura (redes de telecomunicações móveis e instalações de computação em nuvem) e, em especial, as aplicações em tempo real que auxiliam a resolução de problemas diversos. Em 2013, no Chile, os smartphones consolidaram-se como o padrão de telefones móveis, e os tablets superaram o mercado do PC em unidades vendidas (IDC, 2013).

O uso das TIC têm permitido o encontro entre agricultores, especialistas e outros agentes para selecionar as melhores soluções tecnológicas para uma determinada situação ou localização. Serviços especializados dos setores públicos e privados, tais como o uso de satélites e sensores remotos, o armazenamento e o processamento de grandes quantidades de dados e as aplicações móveis têm sido utilizados para planejar a produção agrícola, evitar perdas de colheitas, minimizar riscos e melhorar a gestão da cadeia financeira.

Estes exemplos representam somente um subconjunto da informação e da comunicação que pode ser proporcionado ao setor agrícola por meio dos serviços cada vez mais comuns e acessíveis das TIC na agricultura. E estes resultados têm promovido mudanças nas formas de produção de alimentos. Com este conjunto de técnicas é possível planejar a produção agrícola, florestal e animal, e utilizar, mais eficientemente, recursos naturais bióticos e abióticos, químicos e insumos biológicos, o capital humano e o conhecimento gerado por instituições científicas e tecnológicas.

O uso das TIC também possibilita reduzir os riscos associados às mudanças climáticas, como as doenças e pragas. As inovações relacionadas às TIC permitem produzir de modo eficiente, aumentando a quantidade e qualidade, cumprindo também as exigências do mercado. Hoje em dia a produção de alimentos, seja convencional, orgânica ou seguidora de outros protocolos, como, por exemplo, a produção integrada Global-GAP<sup>2</sup>, devem cumprir as normas de rastreabilidade e de qualidade de bem-estar animal e ambiental.

As novas formas de produção estão sendo beneficiadas pelo uso das TIC. A agricultura de precisão (AP) beneficiou-se da utilização de tecnologias da informação e da comunicação na agricultura. Neste contexto, as TIC são vistas aqui como a realização das seguintes funções (RUSTEN; RAMIREZ, 2003):

- a) Que o conhecimento tecnológico é um importante componente para o desenvolvimento do setor agrícola.
- b) Que as TIC aceleram o desenvolvimento do setor para organizar e facilitar a organização e a transferência de conhecimento entre os agentes da indústria.
- c) Que as organizações têm um papel fundamental na identificação das necessidades dos métodos de gestão adequados para a tomada de decisões e em novas necessidades tecnológicas para o uso das TIC em AP de modo mais eficaz, eficiente e fácil de usar.

É importante destacar que as TIC não são um fim em si, de forma isolada, para o desenvolvimento agrícola, e ainda é muito cedo para que tenhamos uma ideia clara, baseada por uma análise

rigorosa, de como as TIC podem apoiar o desenvolvimento agrícola. Se, por um lado, existem avaliações dos seus impactos, continua existindo dúvidas sobre como estes esforços são replicáveis, escaláveis e sustentáveis para uma população mais ampla e diversificada, em um contexto de agricultura globalizada, que requer alimentos seguros e rastreáveis, além de inovações produzidas de maneira sustentável.

## 2 TIC na construção das AgroTIC

### 2.1 Infraestrutura de alto rendimento

De acordo com um estudo encomendado pela Subsecretaria de Telecomunicações do Governo do Chile à Faculdade de Economia e Neg5ócios da Universidade do Chile, publicado em abril de 2014 (RIVERA et al., 2014), 66% dos chilenos declaram fazer uso permanente da internet (cerca de 12 milhões de pessoas), porcentagem que alcançaria 70% no fim deste ano, caso as tendências observadas nos últimos anos sejam mantidas. Da mesma forma, 62% dos lares chilenos contam com acesso próprio a internet. A pesquisa mantém a diferença estabelecida de acesso entre lares de zonas urbanas e lares de zonas rurais, uma vez que existe uma diferença de 25% em favor dos primeiros (65% contra 40%). A pouca relevância (42,6% em zonas urbanas e 37,6% em zonas rurais) e a pouca frequência de uso (18,5% e 19,2%) aparecem como as razões mais mencionadas pelos entrevistados para não acessar a rede em suas casas. Em terceiro lugar aparecem os custos de acesso e de equipamento (23,5% em zonas urbanas e 13,9% em zonas rurais). O estudo confirmou, de igual maneira, um aumento consistente nas conexões móveis tanto no espaço urbano como no rural. No âmbito rural, no entanto, 18% dos lares acessam através de banda larga fixa, 53% possuem banda larga móvel e 42,4% têm acesso através de telefones inteligentes (smartphones).

O Instituto de Pesquisas Agropecuárias (Inia<sup>3</sup>) do Chile, organização sem fins lucrativos, vinculada ao Ministério da Agricultura, é a principal instituição de pesquisa agrícola do país. Inia conta com uma estrutura de hardware e software que busca satisfazer as necessidades básicas para o desenvolvimento de projetos de TIC rural, ainda que, em algumas áreas esteja-se chegando a um nível de obsolescência que exigirá uma atualização no médio prazo. O Inia demanda o fortalecimento da sua equipe interna para enfrentar os novos desafios em matéria de informática agropecuária e desenvolver novas pesquisas e serviços para a agricultura nacional. Atualmente, a estratégia consiste em estabelecer parcerias com empresas privadas ou especialistas em informática que possuem conhecimento e disponibilidade para desempenhar atividades específicas.

O Inia Chile conta com uma rede IP/MPLS de cobertura nacional que cobre a totalidade de seus 10 Centros Regionais de Pesquisa. As conexões são, na maioria, de alta capacidade, e maiores ainda nas sedes, onde há cerca de 200 usuários. Contam com backbones de fibra ótica para vincular os edifícios principais de escritórios e laboratórios. Em geral a rede suporta as demandas de internet, conferências via web e outras aplicações. A rede possui uma divisão para o trânsito de dados, voz e videoconferência.

<sup>2</sup> Disponível em: <<http://www.globalgap.org/>>.

<sup>3</sup> Disponível em: <[www.inia.cl](http://www.inia.cl)>.

## 2.2 Gestão e análise de dados

Considera-se a grande necessidade no futuro de análise e processamento de dados (*Big data*) em todos os campos das TIC, particularmente nos dados meteorológicos adquiridos pelas estações meteorológicas automáticas (EMA) das redes agrometeorológicas (Inia<sup>4</sup> e Agromet Chile<sup>5</sup>).

## 2.3 Gestão da informação e do conhecimento

No âmbito do Inia existem diversos portais orientados à entrega de informação para a gestão do conhecimento. Um deles é o portal Trigos del Sur<sup>6</sup>, uma página web onde é publicada informação para agricultores, assistência técnica, extensionistas da IX Região de La Araucania no sul do Chile. O objetivo principal é apoiar os processos de tomada de decisão relacionados com o uso das variedades para contribuir com a melhoria dos índices de qualidade industrial e a competitividade do setor.

Na nova versão da página do Inia<sup>7</sup> web, com lançamento no fim de 2014, incluiu-se uma seção para todos os portais web dos projetos desenvolvidos ou em execução.

O Programa Nacional de Melhoramento Genético está implantando um sistema para classificar os rebanhos ovinos na base de dados, o que permitirá dispor, na rede, todos os materiais genéticos ovinos disponíveis.

Também a médio prazo se disponibilizará na web a informação de rebanhos vacunos compilada em projetos de pesquisa e desenvolvimento. Atualmente, a equipe encontra-se em processo de organização e padronização dos formatos de projetos, relatórios e outros.

## 2.4 Modelos e simulações

Inia Chile desenvolveu modelos para o controle de pragas. Por exemplo, um padrão para o voo da vespa “jaqueta amarela” na zona central do país<sup>8</sup>.

Assim também com o modelo de voo da mosca branca de tomate, que na sua primeira versão considera a entrada manual de temperaturas, mas na versão seguinte incluirá uma conexão à base de dados meteorológica.

Identifica-se um espaço para o desenvolvimento de modelos de balanços forrageiros, crescimento de prados em diversas áreas do país, sistemas de orçamento forrageiro, entre outros.

<sup>4</sup> Rede Agrometeorológica de Inia. Disponível em: <<http://agromet.inia.cl>>.

<sup>5</sup> Rede Agroclimática Nacional do Ministério da Agricultura. Disponível em: <[www.agromet.cl](http://www.agromet.cl)>.

<sup>6</sup> Disponível em: <[www.trigosdelsur.cl](http://www.trigosdelsur.cl)>.

<sup>7</sup> Disponível em: <[www.inia.cl](http://www.inia.cl)>.

<sup>8</sup> Sistema Avispa-t. Disponível em: <[www.avispa-t.cl](http://www.avispa-t.cl)>.

## 3 TIC em biotecnologia

### 3.1 Genômica aplicada

No Inia Chile a genômica é aplicada para o melhoramento genético e, para seu desenvolvimento, utiliza-se a bioinformática em várias etapas (bioinformática ou biologia computacional pode ser identificada como uma ferramenta da genômica).

A título de exemplo trabalhos em bioinformática no Inia Chile consiste em:

- Desenvolvimento de marcadores genéticos para o programa de melhoramento genético da uva de mesa.
- Desenvolvimento de sistemas de previsão para problemas fisiológicos de pós-colheita de maçãs.

Nas videiras, os marcadores genéticos são pontos de referência dos genomas que possibilitam diagnosticar o potencial genético para algumas características (calibre, presença de sementes, consistência etc). Desta forma podemos estimar o fenótipo<sup>9</sup> que a planta apresentará quando ainda está em estufa, em vez de ter que aguardar que o fenótipo se manifeste aos três anos de idade da planta no campo. Isto possibilita descartar antecipadamente os segregantes<sup>10</sup> que não têm o potencial genético para os requisitos do programa de melhoramento, focando-se nas linhas e nos fenótipos mais complexos para a análise.

Quanto às maçãs, um problema associado ao armazenamento prolongado sob frio é a escaldadura superficial. Problema que não ocorre em todas as colheitas e para o qual requer um sistema de predição baseado em um metabólito ou um conjunto de genes expressos que se associem ao problema sem ter, necessariamente, uma relação causal. A ideia é aplicar um sistema de predição à colheita que nos possibilite decidir se tais maçãs serão armazenadas por 180 dias ou se será preciso vendê-las antes dos 60 dias de armazenamento.

Para desenvolver os marcadores e os sistemas de predição, utilizam-se ferramentas de genômica e bioinformática.

Tanto as videiras quanto as maçãs são espécies para as quais já contamos com genoma sequenciado (de acesso público). Em videiras utiliza-se o genoma para desenhar marcadores para regiões ou genes particulares, e nas maçãs utiliza-se o genoma para caracterizar a expressão de alguns genes que acreditamos estar relacionados com o problema fisiológico. Quando falamos em desenhar marcadores, ou identificar genes para caracterizá-lo, referimo-nos a procedimentos bioinformáticos realizados em programas bioinformáticos (alguns livres instalados localmente, outros livres em servidores públicos e outros mediante pagamento instalados localmente) que analisam sequências, anotam o genoma, comparam as sequências, analisam as mutações e desenharam moléculas (para isolar fragmentos ou genes).

Outro exemplo aplicado às videiras é a construção de mapas genéticos de alta saturação para a uva de mesa. O mapa genético é a representação de seus cromossomos que serve de base para

<sup>9</sup> **Fenótipo:** é qualquer característica ou traço observável de um organismo, como a sua morfologia, desenvolvimento, propriedades bioquímicas, fisiologia e comportamento.

<sup>10</sup> **Segregantes:** grupo de indivíduos geneticamente variáveis que constituem a descendência de uma população híbrida que resulta de um cruzamento de duas linhas diferentes geneticamente uniformes.

estudar a associação entre variações de sequências e fenótipos. A construção do mapa foi feita pelo Inia Chile com ferramentas genéticas clássicas e com o apoio do genoma de referência para a espécie. Cada posição dos mapas (6.000 posições) foi corrigida com base no genoma de referência mediante análise bioinformática. Paralelo a isso, e considerando que o genoma de referência da espécie foi definido a partir de um genoma de uma videira de vinificação, o genoma da variedade “Sultanina”, considerada um ícone das variedades modernas de uva de mesa, foi sequenciado e montado, permitindo a identificação de milhões de variantes estruturais em relação ao genoma de referência. Esse trabalho foi uma colaboração entre Inia-Chile, Universidade do Chile e Universidade Andrés Bello..

## 4 TIC nas cadeias de fornecimento

### 4.1 TIC nos recursos naturais e mudanças climáticas

Existem no Chile diversos projetos, serviços públicos e privados orientados a informar sobre recursos naturais e mudança climática servindo-se das TIC. A seguir nomeamos alguns deles vigentes na atualidade.

AGROMET<sup>11</sup> é a rede nacional de estações meteorológicas automáticas de livre acesso, administrada pelo Ministério da Agricultura do Chile.

Uma equipe de Inia Tamel Aike, na cidade de Coyhaique, localizada no sul do Chile, está implementando uma plataforma web que entregará a caracterização de todos os vales produtivos da Região de Aysén, quanto à capacidade do uso dos solos, ecorregiões, uso atual e potencial, fertilidade dos solos, características físico-químicas dos solos, taxonomia dos solos e outros.

O **Observatório Agroclimático do Ministério da Agricultura do Governo do Chile**<sup>12</sup> contém muita informação para a gestão e mitigação do risco agropecuário. Uma das fontes de informação são mapas de estiagem baseadas em imagens por satélite. Este trabalho é desenvolvido em conjunto com Unesco, Minagri, Inia e publicado pela Unidade de Emergências Agrícolas do Ministério da Agricultura do Chile.

**DropControl**<sup>13</sup>, da empresa chilena Wiseconn, é um sistema sem fio de gestão integral de fertilização e irrigação e monitoramento das condições de cultivo mediante uma interface na internet. Permite o acionamento remoto de válvulas solenóides, bombas hidráulicas e outros dispositivos ON/OFF, que, além disso, realiza funções de registro de monitoramento de sensores de um campo

A empresa **AgroPrime**<sup>14</sup> oferece ferramentas de gestão (software agrícola) para melhorar os processos produtivos, minimizando custos e aumentando a entrada em hortas frutíferas e plantéis de *packing* no Chile e no Peru. Seu software agrícola se destaca quanto ao registro de atividades de produção e também por suas soluções móveis, inteligência de negócios e agricultura de precisão.

<sup>11</sup> Disponível em: <www.agromet.cl>.

<sup>12</sup> Disponível em: <www.agroclimatico.cl>.

<sup>13</sup> Disponível em: <http://www.dropcontrol.cl>.

<sup>14</sup> Disponível em: <http://www.agroprime.cl>.

**LemSystem**<sup>15</sup> entrega soluções tecnológicas de monitoramento e controle sem fio em tempo real para a otimização do uso de água e de energia. Seu sistema se destaca pelo monitoramento agrícola sem fio de umidade do solo e de variáveis ambientais.

### 4.2 TIC no monitoramento e uso da terra

O Ministério da Agricultura do Chile criou o projeto [Infraestrutura de Dados Espaciais] (**IDE Minagri**)<sup>16</sup> que consiste em um geportal onde todos os serviços do Ministério da Agricultura publicam suas informações disponíveis livremente. O IDE Minagri surge do anseio de contar com um sistema unificado que permitisse acessar toda a informação geoespacial do Ministério tem a finalidade de contribuir para a otimização da tomada de decisões em qualquer nível e em todos os serviços e instituições relacionadas. Um dos princípios fundamentais das IDE é a interoperabilidade, ou seja, a habilidade de organizações e sistemas díspares e diferentes de interagir com os objetivos convergentes e comuns, com a finalidade de obter benefícios mútuos.

A interação implica que as organizações participantes compartilhem informação e conhecimento através de seus processos de negócio por meio do intercâmbio de dados entre seus respectivos sistemas de tecnologia de informação e comunicações. Atualmente está em vigência a segunda versão do IDE Minagri, desenvolvida em conjunto com uma consultoria estrangeira, assumindo como referência os padrões internacionais. Está também em operação um Geportal onde os usuários podem acessar o geocatólogo, notícias, documentos diversos e o visualizador. Neste último é possível visualizar as diversas peças de informação originadas pelas diferentes instituições que participam do projeto, de acordo com os perfis de autorização estabelecidos caso a caso.

### 4.3 TIC e a segurança fitossanitária

**Modelo de determinação de risco de infecção da Bacteriose do Kiwi causada pelo *Pseudomonas Syringae pv Actinidiae***<sup>17</sup>. Mediante o financiamento do Fundo de Inovação Agrária (FIA) (Projeto PYT-2012-0213) desenvolveu-se um modelo de determinação de risco de infecção da Bacteriose do kiwi que possibilitará à indústria uma ferramenta descritiva para estabelecer de maneira rápida o risco potencial de infecção desta doença nos plantios de kiwi por meio de ferramentas geradas pelo modelo, de acordo com os dados por este expressados sobre as condições específica de regiões, agrozonas, localidades ou localização específica do prédio. O modelo é composto pelas três razões seguintes:

1. Mapa do Chile com zonas de risco agroclimático.
2. Curvas por localidade com Risco Agroclimático Mensal.
3. Calculadora de Determinação de Risco de Infecção de Psa.

O **Sistema Tizón Tardío (míldio)**<sup>18</sup> é uma aplicação para a detecção e controle a tempo do fungo que ataca os cultivos de batata. A doença pode levar à perda total do cultivo, afetando a compe-

<sup>15</sup> Disponível em: <http://www.lemssystem.com>.

<sup>16</sup> Disponível em: <http://ide.minagri.gob.cl>.

<sup>17</sup> Disponível em: <http://pdtpsa.comitedelkiwi.cl>.

<sup>18</sup> Disponível em: <http://tizon.inia.cl>.

titividade dos produtores de batata. O Instituto de Pesquisas Agropecuárias (Inia) há alguns anos trabalha na validação de modelos de prognóstico que possibilitem que os produtores tomem as melhores decisões para o controle da doença utilizando dados meteorológicos da Rede Agro Meteorológica do Inia junto ao desenvolvimento de plataformas para difundir a informação do prognóstico entre os produtores e assessores associados.

**Avispa-T**<sup>19</sup> é um pacote tecnológico criado pelo Inia La Platina como um complemento ao controle integrado da vespa “jaqueta amarela” (*Vespula germânica*). Esta aplicação informática integra conhecimentos, dados e interesses provenientes de diversas fontes. Primeiro: resultados do projeto de pesquisa FONDEF DO-31 1076 “Manejo integrado da vespa “jaqueta amarela” *Vespula germânica* (Fabricius) (Hymenoptera: Vespidae)”, liderado pela pesquisadora Patricia Estay do Centro Regional La Platina do Instituto de Pesquisas Agropecuárias, entre os anos 2004 e 2007. Segundo: dados em linha provenientes da rede nacional de Estações Meteorológicas Automáticas do Inia. E finalmente, a aplicação conta com o apoio da empresa ANASAC, que produz e comercializa o sebo “Vespugard”, necessário para controlar a vespa.

**Plataforma online de sensibilidade a botricidas na uva de mesa de exportação**<sup>20</sup>. Este projeto, liderado pelo Laboratório de Fitopatologia Vegetal da Faculdade de Ciências Agronômicas da Universidade do Chile, tem por objetivo criar uma plataforma de informação interativa em relação às mudanças nos níveis de incidência de resistência do tipo multidroga (MDR) e de sensibilidade às principais moléculas botryticidas em zonas produtoras de uva de mesa de exportação no Chile. A informação gerada permitirá ao setor produtor-exportador envolvido otimizar o uso de fungicidas mediante o desenho de programas mais eficazes com menor pressão fungicida (menor contaminação do meio ambiente, nível e número de resíduos na fruta), fornecendo, com isso, uma competitividade maior para a uva de mesa chilena nos mercados de exportação.

#### 4.4 TIC em processos de produção agrícola

**Epsilon Networks**<sup>21</sup> é uma empresa que oferece soluções de monitoramento remoto para a agricultura, por meio do uso de tecnologias de rede de sensores sem fio, sistemas de sensoriamento com cabo tradicionais e com equipamentos controladores ou *dataloggers* especializados na medição de variáveis de clima e solo à distância. No âmbito das estufas e câmaras de crescimento, fornece soluções para o monitoramento e controle de variáveis ambientais (temperatura, umidade, CO<sub>2</sub>, radiação) e também possui uma grande variedade de equipamentos de iluminação LED especializados no crescimento de plantas. Por último, a empresa está comprometida e participa ativamente de projetos de I+D, tanto no âmbito nacional como internacional.

A **Consultora Diestre Tecnología Ltda**<sup>22</sup>, é uma empresa dedicada a oferecer produtos e serviços tecnológicos orientados ao monitoramento de umidade no solo com a finalidade de entregar informação e automatização na agricultura. Com mais de 10 anos de experiência trabalhando com a tecnologia australiana da empresa Sentek, entrega ferramentas práticas que possibilitam oferecer informação objetiva para a tomada de decisões no status hídrico tempo e frequência de

irrigação, relacionando as variáveis solo-água-plantas e, desta forma, garantindo a produtividade dos cultivos. Procura-se, por meio de sensores de capacitância elétrica que meçam o conteúdo volumétrico da água no solo, alcançar um status de umidade ótimo no solo (máxima dinâmica da água), e, adicionalmente por meio do software IrriMAX, integrar a informação de umidade do solo, clima e plantas em forma gráfica, permitindo corrigir as estratégias de risco no tempo, aperfeiçoando assim nossa gestão de irrigação.

O software **FEDESOF**<sup>23</sup>, Gestão Produtiva, Administrativa e de Boas Práticas Agrícolas (BPA) é uma solução Cloud Computing à qual se acessa pela página ou aplicação web. FEDESOF se personaliza de acordo com as características próprias de cada empresa agropecuária. FEDESOF conta com os seguintes módulos:

- Orçamento.
- Vendas.
- Trabalhos de campo.
- Aplicação de agroquímicos.
- Inventário de insumo e materiais.
- Uso de maquinaria.
- Centros de custo.

**Eland**<sup>24</sup> oferece um sistema de informação web para diminuir ao máximo os problemas que a inserção manual dos dados que são trabalhados no campo representa. Esta aplicação permite o registro das atividades agrícolas, controle de inventário, gestão de orçamento de operações da temporada, entre outros. O sistema funciona com base em um mapa da propriedade, disponível em qualquer dispositivo que possa executar uma aplicação web.

**Agrosat**<sup>25</sup> é uma empresa de Temuco, no sul do Chile, que oferece diferentes soluções TIC para a agricultura. **Crosscheck**® é uma metodologia para o estudo de solos, plantas e fenômenos ambientais. **Fertimap** representa, por meio de mapas, as regiões com diferentes níveis de fertilidade e variabilidade existente. Baseado na análise da acidez ativa do solo, como informação primária, complementada e cruzada com informação topográfica; tem como objetivo principal entregar ao produtor um insumo que permita desenvolver um plano de fertilização básico, baseado em setores de fertilidade. O **ADN Suelo**® (Análise e Diagnóstico Nutricional do Solo) entrega mapas de fornecimento de nutrientes que permite conhecer a distribuição espacial dos elementos no solo. Desse modo, pode-se planejar melhor a distribuição de insumos, considerando que os terrenos de cultivo não são homogêneos. Além disso, permite otimizar o processo de amostragem, de recursos e maximizar a resposta econômica de cada área. O **ADN Foliar**® (Análise e Diagnóstico Nutricional Foliar) entrega, por meio de mapas, informação mais representativa do estado nutricional atual da planta. Esta análise complementa e potencializa um programa de fertilidade, permitindo corrigir oportunamente deficiências que afetem a qualidade ou o rendimento do produto colhido. **PreMonitor**® coloca à disposição dos agricultores uma visualização

<sup>19</sup> Disponível em: <<http://www.avispa-t.cl>>.

<sup>20</sup> Disponível em: <<http://www.botrytis.uchile.cl>>.

<sup>21</sup> Disponível em: <<http://www.epsilon.cl>>.

<sup>22</sup> Disponível em: <<http://cdtec.cl>>.

<sup>23</sup> Disponível em: <<http://cdtec.cl>>.

<sup>24</sup> Disponível em: <<http://www.eland.es>>.

<sup>25</sup> Disponível em: <<http://www.agrosatchile.cl>>.

georreferenciada e antecipada do rendimento do cultivo e/ou pomar em cada área produtiva, fornecendo detalhes sobre sua variação percentual e médias de produção ajustadas às respectivas variações. **Agrobook®** é uma solução que permite a captura de informação relevante no campo e em áreas produtivas, vinculada a plataformas on-line (web) que facilita a gestão, relatórios e visualização da informação gerada.

#### 4.4.1 Automatização e Agricultura de Precisão

**Reinsystem<sup>26</sup>**, é uma empresa de agricultura de precisão para o controle fitossanitário e para o controle e administração de irrigação. A solução de Reinsystem permite saber exatamente quando, quanto e onde regar eficientemente os pomares agrícolas. A aplicação analisa a disponibilidade de água e a velocidade de infiltração de água no solo mediante o uso de sensores de umidade em diferentes profundidades. **Controle fitossanitário**: este serviço permite capturar informação dos microclimas gerados nos pomares agrícolas para gerir de maneira eficiente os insumos agrícolas. Permite identificar as condições ambientais que afetam o desenvolvimento de eventos fitossanitários por meio de processamento de dados mediante a integração de algoritmos para o surgimento de doenças. A aplicação permite empregar de forma correta os insumos agrícolas e não de forma preventiva. **Controle e administração de irrigação**: este serviço possibilita saber exatamente quando, quando e onde regar eficientemente os pomares agrícolas por meio de aplicação de “Água Disponível”: a aplicação identifica o comportamento da umidade do solo, velocidade de infiltração e evapotranspiração que permite criar um eficiente manejo de irrigação.

O Programa de Agricultura de Precisão do Inia Quilimapu (Progap) enfatiza o uso da informação e das tecnologias emergentes para sintetizar e entregar ferramentas de decisão focadas a melhorar a rentabilidade do agricultor. A tecnologia utilizada baseia-se na interação de sensores, processamento de imagens, análise de modelos estatísticos, tecnologias da informação e comunicação e engenharia mecânica.

**Agrosuccess<sup>27</sup>** é uma empresa que fornece serviços de monitoramento de irrigações executadas, controle e monitoramento de equipamentos de irrigação, monitoramento de condições de solo, planta e ambiente. Além de estar diretamente relacionado com as plantas, o sistema pode comandar a abertura e fechamento de bombas de poço profundo ou da transposição de água entre tanques, registrando as aberturas e fechamentos das mesmas, além dos volumes de água bombeados. O sistema pode monitorar de forma contínua sensores de nível de poços e tanques, para saber, on-line, a disponibilidade de água nestes reservatórios. Também podem ser conectados consoles de estações meteorológicas automáticas com a finalidade de dispor de informação de clima, complementar à tomada de decisões sobre irrigação ou ao estudo de situações importantes para o desenvolvimento do cultivo .

A empresa **TSensor<sup>29</sup>** oferece soluções de telemetria para monitoramento on-line de variáveis como temperatura, umidade, pressão, fluxo, níveis ou qualquer outro processo ou variável que

necessite medir. Por exemplo, *Cold Chain Control System* é um sistema para controlar cadeias de frio on-line com alarmes que possibilitam atuar de forma imediata quando há algum problema. *Farm Management* serve para controlar as temperaturas e umidade de seus campos com alarmes para prevenção de geadas, cálculos de horas frio e de graus dia, além de alarmes fitossanitários configuráveis. *Cold Data Logger Systems* são sensores programáveis com memória que funcionam como termógrafos e podem ser baixados em um PC.

## 5 Aplicações e futuro de AgroTIC

### 5.1 Novas tecnologias - evolução tecnológica e o futuro de AgroTIC

Apesar de não existir medições que permitam quantificar exatamente o impacto das TIC na rentabilidade das empresas, fica claro que são um elemento definitivamente facilitador e indispensável na gestão moderna. Esse papel estratégico pode ser avaliado, sobretudo, quando comparamos empresas que incorporaram as TIC com aquelas que não o fizeram. Nesse caso, a ausência de acesso ao mundo digital converte-se em um fator de exclusão porque coloca a empresa à margem das modalidades de gestão modernas e à margem da informação oportuna e instantânea necessária para competir nos mercados globalizados (NAGEL; MARTINEZ, 2006).

É necessário entender as necessidades reais dos agricultores pequenos, médios e grandes para identificar os problemas e as oportunidades para o desenvolvimento de tecnologias que sinalizem a satisfazer essas necessidades.

Uma vez que os problemas são entendidos, é preciso desenvolver mecanismos para melhorar a transferência ou extensão de conhecimentos e acesso à informação. Por exemplo, construir e manter bases de dados de usuários, principalmente agricultores, clientes, pesquisadores, estudantes e tomadores de decisão de esfera privada e pública. Utilizar sistemas para a administração dessa informação do tipo Consumer Relationship Management.

Um mecanismo disponível via internet que instituições como o Inia poderiam explorar para melhorar o impacto da extensão do conhecimento são os seminários em linha ou webseminars. Isto consiste em fazer apresentações interativas multimídia por meio da internet, onde a audiência está geograficamente distribuída com um computador conectado à internet. O apresentador está no seu escritório, diante de seu PC, fazendo a apresentação, servindo-se de microfone e câmera, enquanto a audiência está escutando com a possibilidade de fazer perguntas por meio de um chat ou de um microfone. Finalmente, estes seminários podem ser gravados e armazenados em uma biblioteca que pode ser consultada quando necessário e a instituição que realiza os seminários fica de posse da base de dados de usuários e interesses de grande valor para futuros eventos ou atividades tanto de pesquisa como de difusão.

Outros exemplos de uso de TIC são:

- Acesso à informação de utilidade produtiva como boletins e informativos a partir de dispositivos móveis.
- Modelos de simulação de processos produtivos.
- Processamento de imagens para detecção de problemas fitossanitários ou para predição de colheitas.

<sup>26</sup> Disponível em: <<http://www.reinsystem.com>>.

<sup>27</sup> Disponível em: <<http://agrosuccess.cl>>.

<sup>28</sup> Disponível em: <<http://agrosuccess.cl>>.

<sup>29</sup> Disponível em: <<http://www.tsensor.cl>>.

Finalmente, de acordo com o relatório “ICT as a Catalyst to Enterprise Competitiveness” (DUCKER; PAINE, 2010) outro tema importante para a evolução das TIC tem relação com a força de trabalho. Tipicamente, os projetos estão centrados na criação de engenheiros altamente capacitados na aplicação de tecnologias e desenvolvimento de software, mas contar com usuários dentro da empresa, com habilidades avançadas em computação é igualmente importante. Muitas das soluções mais inovadoras provêm dos usuários e não dos engenheiros. As habilidades da força de trabalho devem melhorar para que as TIC realmente ajudem a melhorar a produtividade do negócio (PILAT, 2004). Frequentemente o trabalhador que entende bem do uso da tecnologia ajudará a encontrar maneiras em que a tecnologia terá um maior impacto na melhoria dos processos da empresa.

A Escola Virtual do PNUD<sup>30</sup> afirma que a importância da aplicação de tecnologias da informação e da comunicação e sua compreensão é fundamental para o desenvolvimento das pessoas em nossa sociedade atual. Neste sentido, o uso das TIC na agricultura pode aumentar as opções de um número maior de pessoas que não têm a possibilidade de obter informação relevante e de qualidade adequada.

## 5 Referências

- IDC. **Revisión de la actividad TI en Chile**. Santiago, Chile: ACTI, 2013. 18 p. il.
- DUCKER, M.; PAYNE, J. **Information communication technology as a catalyst to enterprise competitiveness**: research report. Washington, D. C.: Usaid, 2010. 37 p. il.
- FAO. **The state of food and agriculture, livestock in the balance**. Rome, 2009. 180 p. il.
- NAGEL, J.; MARTÍNEZ, C. **Chile**: agricultores y nuevas tecnologías de información. Santiago, Chile: CENDEC, 2006. 136 p.
- PILAT, D. The ICT Productivity paradox: insights from micro data. **OECD Economics Studies**, v. 38, n. 1, p. 37-65, 2004.
- RIVERA, J. C.; LIMA, J. L.; CASTILLO, E. **Estudio quinta encuesta sobre acceso, usos, usuarios y disposición de pago por internet en zonas urbanas y rurales de Chile**. Santiago: Facultad de Economía y Negocios, Universidad de Chile, 2014. 163 p. il.
- RUSTEN, E.; RAMIREZ, S. **Future direction agriculture and Information and Communication Technologies (ICTs) at Usaid**. Washington, D.C.: Academy For Educational Development: Winrock International, 2003. 80 p.

<sup>30</sup> Disponível em: <<http://www.escuelapnud.org/es/new-portal/en/87-sobre-la-escuela-virtual/cursos/460-information-and-communication-technologies-ict>>.

## **Tecnologias de Informação e Comunicação e sua relação com a agricultura - Argentina**

Miguel Angel Luengo

### **1 Introdução**

O propósito central deste documento é estabelecer um diagnóstico e analisar a situação atual, além do impacto e a evolução das TIC na pesquisa agrícola; a transferência de conhecimentos, o desenvolvimento agrícola e a redução da pobreza rural a partir dos trabalhos desenvolvidos pelos centros de pesquisa agrícola membros do Procisur, a saber, Inta, da Argentina, Embrapa, do Brasil, Inia, do Chile e Inia do Uruguai.

A agricultura está incorporando novas tecnologias para responder aos novos desafios que conduzem à evolução dos padrões tecnológicos e produzir alimentos para uma população mundial estimada em 9 bilhões em 2050. Isto acarretará um aumento de 70% da produção de alimentos de forma sustentável e segura<sup>1</sup>. Neste contexto, a tecnologia da informação e comunicação (TIC) já está sendo utilizada e tem um papel cada vez mais importante na gestão da cadeia de valor. As TIC estão evoluindo continuamente e têm atraído o aumento de investimentos públicos e privados para a pesquisa, a transferência de tecnologia, extensão e desenvolvimento agrícola.

TIC é um termo geral, definido pela Agência Norte Americana para a Cooperação Internacional (Usaid), como a combinação de hardware, software e ferramentas de produção que permitam o intercâmbio, o tratamento e a gestão de informação e de conhecimento. De acordo com a Usaid, as TIC são tecnologias e métodos para armazenar, gerir e processar a informação (por exemplo, computadores, softwares, livros, PDAs, tablets, bibliotecas digitais e não digitais) e para comunicar informação (por exemplo, correio, e-mail, rádio, televisão, telefones celulares, localizadores, internet, entre outros). Redução no preço, promoção da acessibilidade e adaptabilidade e de suas novas capacidades para seu uso são indispensáveis no setor agrícola. Os produtores podem utilizar a internet, o telefone e outras ferramentas digitais para: previsão do clima, a gestão de frotas de veículos, a rastreabilidade dos produtos agrícolas, para buscar informação sobre preços dos insumos, contratar serviços ou produtos, acessar mercados, variedades, técnicas de produção, serviços de armazenamento ou processamento de matéria-prima.

<sup>1</sup> Disponível em <[http://www.fao.org/fileadmin/templates/wfsfs/docs/synthesis\\_papers/C%C3%B3mo\\_alimentar\\_al\\_mundo\\_en\\_2050.pdf](http://www.fao.org/fileadmin/templates/wfsfs/docs/synthesis_papers/C%C3%B3mo_alimentar_al_mundo_en_2050.pdf)>

Felizmente, na atualidade, estão cada vez mais disponíveis os dispositivos móveis (telefones inteligentes, tablets etc), infraestrutura (redes de telecomunicações móveis e instalações de computação em nuvem) e, em especial, as aplicações em tempo real para a gestão dos sistemas de produção, a rastreabilidade dos produtos, informação de preços, maquinário e veículos, e gestão da cadeia de fornecimento. Por exemplo, perguntas feitas pelos agricultores sobre a forma de aumentar a produtividade e como reduzir ao mínimo os riscos derivados das condições climáticas podem ser respondidas com rapidez e precisão.

O uso das TIC permitiu o encontro entre agricultores, especialistas e outros agentes para selecionar as melhores soluções tecnológicas para uma determinada situação ou localidade. Serviços especializados, dos setores públicos e privados, como o uso de satélites e sensores remotos, o armazenamento e processamento de grandes quantidades de dados e as aplicações móveis foram utilizados para planejar a produção, evitar perdas nas colheitas, minimizar os riscos e melhorar a gestão da cadeia financeira.

Estes exemplos representam somente um subconjunto da informação e da comunicação que pode ser proporcionado ao setor agrícola por meio dos serviços cada vez mais comuns e acessíveis das TIC na agricultura. E estes resultados têm promovido mudanças nas formas de produção de alimentos. Com este conjunto de técnicas é possível planejar a produção agrícola, florestal e animal, e utilizar, mais eficientemente, recursos naturais bióticos e abióticos, químicos e insumos biológicos, o capital humano e o conhecimento gerado por instituições científicas e tecnológicas.

O uso das TIC também possibilita reduzir os riscos associados às mudanças climáticas, como as doenças e pragas. As inovações relacionadas às TIC permitem produzir de modo eficiente, aumentando a quantidade e qualidade, cumprindo também as exigências do mercado. Hoje em dia a produção de alimentos, seja convencional, orgânica ou seguidora de outros protocolos, como, por exemplo, a produção integrada Global-GAP, devem cumprir as normas de rastreabilidade e de qualidade de bem-estar animal e ambiental.

As novas formas de produção estão sendo beneficiadas pelo uso das TIC. A agricultura de precisão (AP) beneficiou-se da utilização de tecnologias da informação e da comunicação na agricultura. Neste contexto, as TIC são vistas aqui como a realização das seguintes funções (RUSTEN; RAMIREZ, 2003)<sup>2</sup>:

- 1) Que o conhecimento tecnológico é um importante componente para o desenvolvimento do setor agrícola.
- 2) Que as TIC aceleram o desenvolvimento do setor para organizar e facilitar a organização e a transferência de conhecimento entre os agentes da indústria.
- 3) Que as organizações têm um papel fundamental na identificação das necessidades dos métodos de gestão adequados para a tomada de decisões e em novas necessidades tecnológicas para o uso das TIC em AP de modo mais eficaz, eficiente e fácil de usar.

É importante destacar que as TIC não são um fim em si, de forma isolada, para o desenvolvimento agrícola, e ainda é muito cedo para que tenhamos uma ideia clara, baseada por uma análise

rigorosa, de como as TIC podem apoiar o desenvolvimento agrícola. Se, por um lado, existem avaliações dos seus impactos, continua existindo dúvidas sobre como estes esforços são replicáveis, escaláveis e sustentáveis para uma população mais ampla e diversificada, em um contexto de agricultura globalizada, que requer alimentos seguros e rastreáveis, além de inovações produzidas de maneira sustentável.

## 2 TIC na construção da AgroTIC

### 2.1 Infraestrutura de alto rendimento

Em âmbito regional, a infraestrutura de Alto Rendimento está disponível em diversas instituições para apoiar seus projetos de pesquisa, especialmente em universidades focadas em áreas como astronomia, bioinformática, química quântica, estudo de medicamentos, nanotecnologia, modelagem climática, entre outras.

Além disso, há o Sistema Nacional de Computação de Alto Desempenho (SNCAD), uma iniciativa conjunta entre o Ministério de Ciência, Tecnologia e Inovação Produtiva e o Conselho Interinstitucional de Ciência e Tecnologia (CICeT), inserido dentro de Programa de Grandes Instrumentos e Bases de Dados.

Seu propósito é consolidar uma rede nacional de centros de computação de alto desempenho pertencentes ao sistema científico e acadêmico interconectado para satisfazer a demanda crescente da comunidade científica e tecnológica nas áreas de armazenamento, computação em rede, de alto desempenho e de alto rendimento, de visualização e outras tecnologias emergentes.

Complementado, o Sistema Nacional de Redes Avançadas tem o objetivo de contribuir para a otimização e melhoramento da Rede Avançada Nacional Innovared, potencializar e apoiar as atividades da Innovared e, ao tempo, harmonizar as ações da mesma com as necessidades provenientes das instituições científico-tecnológicas onde o Inta atuar.

Internamente o Inta possui uma Rede MPLS que conecta aproximadamente 420 unidades em todo o país. É empregada para a gestão, mesmo que esta rede também seja utilizada como meio para a conexão da internet 2 através do Sistema Nacional de qualquer ponto da instituição. O projeto “Argentina Conectada” aparece como uma possibilidade de estender as prestações da Rede Federal de Fibra Ótica<sup>3</sup>, que possibilitará a conectividade em todo o país.

Além disso, o Inta conta com uma rede de equipamentos de videoconferência em cada um dos 15 centros regionais, nos centros de pesquisa e nos escritórios centrais, apoiado por um sistema colaborativo com mensagens instantâneas, de voz e e-mail.

Os serviços são prestados através de dois data centers onde estão alojados os servidores nos quais, por meio da virtualização, são configurados os servidores virtuais que fornecem as prestações aos usuários.

Um assunto sobre o qual se deve avançar é a integração dos serviços de TI às atividades de pesquisa e extensão. Os serviços principais consistem em apoio à gestão e atividades de suporte aos

<sup>2</sup> Disponível em <<http://www.globalgap.org>>

<sup>3</sup> Projeto Argentina Conectada. Disponível em <<http://www.arsat.com.ar/arsat-en-las-politicas-de-estado-argentino/argentina-conectada>>

pesquisadores e extensionistas, mas não estão integrados aos projetos de pesquisa que necessitam cálculos de alto desempenho, mineração de dados ou gestão de conhecimento.

### 2.3 Gestão e análise de dados

Na atualidade a geração de informação em grandes volumes possibilita a oportunidade de inteligência para o negócio. Maquinário, sensores e monitores geram constantemente informação de maneira tal que dificultam sua administração e armazenamento. Além disso, tem-se informação da produção intelectual gerada pelos pesquisadores e técnicos. Por isso é um desafio administrar e analisar esta informação.

O acesso à informação que a rede de sensores meteorológicos oferece, tanto do Inta como das redes privadas, permite sistematizar atividades como o monitoramento automático de insetos, prognósticos de chuvas, intensidade de chuvas, recálculo da refletividade, entre outros desafios e modelos apoiados pelas TIC.

### 2.4 Gestão da Informação e do Conhecimento

Esta atividade é incipiente dentro das AgroTIC. Não há projetos concretos neste escopo. No âmbito do Inta está sendo implementado um sistema de gestão da informação da produção intelectual da Instituição. A crescente geração de informação não estruturada demanda uma melhor organização dos conteúdos. A solução apoia-se em uma organização de metadados baseado no padrão da W3C, Resource Description Framework (RDF) e em motores de enriquecimento semânticos.

### 2.5 Modelos e simulação

Em relação às atividades de modelagem e simulação, as principais atividades que estão sendo realizadas vinculam-se aos prognósticos tanto sobre o clima quanto à estimativa de alguns cultivos, especialmente os de agricultura extensiva. O projeto “Sistema para a tomada de decisão estratégica sobre o estado e projeção dos cultivos em tempo operacional” tem como objetivo o desenho, desenvolvimento e implementação do software que possibilita contar com uma plataforma destinada a oferecer informação estratégica para a avaliação do estado dos cultivos e sua projeção futura.

Além disso, a calculadora de risco ecotoxicológico para aves permite avaliar rapidamente o dano às aves no campo argentino quando são aplicados pesticidas e avaliar o cumprimento das normativas vigentes como contribuição ao cuidado do meio ambiente.

O Instituto de Clima e Água<sup>4</sup> do Inta por meio de teledetecção e processamento de modelos calcula:

- Previsões agroclimáticas.
- Geadas.
- Inundações.
- Índices de vegetação.
- Evolução de cenários.

<sup>4</sup> Instituto de Clima e Água do Inta: <<http://climayagua.inta.gob.ar>> e <<http://sepa.inta.gob.ar>>

## 3 TIC em biotecnologia

O advento das novas metodologias de sequenciamento tem revolucionado, nos últimos anos, o desenvolvimento dos projetos que utilizem ferramentas de biologia molecular e genética aplicada a sistemas agropecuários. A característica saliente destas tecnologias, que em termos gerais se baseiam na paralelização da reação de sequenciamento por sínteses das moléculas de ADN, é a geração de grandes volumes de dados com arquivos resultantes no intervalo de 5 Gb a 600 Gb. O armazenamento e processamento desta grande quantidade de dados tem colocado a Biologia e a Bioinformática dentro do contexto do “Big Data problem”, que começa com a transferência dos arquivos entre servidores e se aprofunda com o processamento para transformar os dados obtidos, mediante tecnologias NGS, em informação e conhecimento útil.

A partir deste ano e com o financiamento do MinCeT, o Indear<sup>5</sup>, a Fundação Instituto Leloir e Inta, estão trabalhando de maneira integrada na criação de uma Plataforma de Genômica Nacional. O objetivo desta iniciativa é fornecer um serviço integral e acessível ao Sistema Nacional de Ciência e Tecnologia (SNCeT) na área de genômica de alta performance que inclui o desenho, a execução e o assessoramento em relação à estruturação de projetos genômicos e a análises dos dados produzidos, incluindo o armazenamento e o acesso remoto aos dados por parte dos usuários. Em articulação com a Plataforma de Bioinformática, que está sendo criada dentro do mesmo instrumento de Plataformas Tecnológicas (PPL) do MINCeT, está previsto também oferecer suporte e recursos de bioinformática para a análise dos dados NGS gerados e o desenvolvimento de atividades de educação relacionadas a estas disciplinas emergentes na Argentina.

Para favorecer o trabalho articulado dos nós que compõem a plataforma CATG, é necessário estabelecer uma rede de ponto a ponto de alta velocidade que permita o intercâmbio eficiente de dados entre os nós e destes com outros nós de outras plataformas, sobretudo a plataforma de Bioinformática. Também no caso do Nó de Sequenciamento e Genotipificação do Inta/CATG, que funciona em uma rede interna com a Unidade de Bioinformática (Instituto de Biotecnologia, CICVeA, Inta), está prevista a potencialização da conectividade intra e interinstitucionais por meio do sistema de redes avançadas. O salto qualitativo na conectividade permitirá acessar servidores de repositórios internacionais (NCBI<sup>6</sup> y EBI<sup>7</sup>) para o armazenamento de dados genômicos e a infraestrutura de computação HPC remota para cobrir as necessidades de cálculo intensivo que demandam algumas aplicações não paralelizáveis no momento, como a montagem de genomas complexos, entre outros.

O grau de integração que pressupõe o acesso a uma rede de alta velocidade é fundamental para a evolução das áreas de genômica e bioinformática em nosso país, considerando que é fundamental reunir capacidades e esforços para o desenvolvimento integral das mesmas, contribuindo, desta forma, para o ótimo aproveitamento dos resultados provenientes das áreas de Biologia Molecular e Biotecnologia, garantindo, conseqüentemente, o impacto tecnológico e econômico destas áreas.

<sup>5</sup> Instituto de Agrobiotecnologia Rosario, <<http://www.indear.com/>>

<sup>6</sup> National Center for Biotechnology Information, disponível em <<http://www.ncbi.nlm.nih.gov>>

<sup>7</sup> European Bioinformatics Institute, disponível em <<http://www.ebi.ac.uk>>

### 3.1 Bioinformática aplicada

- A área de genômica e bioinformática do Inta tem por objetivo desenvolver ferramentas de genômica assim como ferramentas de bioinformática para sua aplicação na exploração da diversidade genética, evolução, estrutura populacional, mecanismo de patogenicidade, mecanismos de resposta a estresses bióticos e abióticos, mapeamento genético e mapeamento por associação de caracteres de importância agrícola e florestal.
- Esta área é composta por pesquisadores de diferentes áreas, incluindo a biologia e a genética molecular, genética de populações, melhoramento molecular e a bioinformática, com vasta experiência na utilização e desenvolvimento de tecnologias de análise massiva de informação genômica, assim como conhecimentos genéticos para a interpretação biológica e aplicação em sistemas produtivos.
- Tem-se também a Plataforma Bioinformática Argentina<sup>8</sup> (BIA), destinada a fornecer serviços de bioinformática tanto para o Sistema Científico Nacional como para organismos e empresas públicas e privadas de diversos setores econômicos com projeção regional e internacional. O consórcio é composto por Conicet, Indear e pelas Universidades de Córdoba, Buenos Aires e San Martín. Os serviços consistem em sequenciamento ADN de alta performance e análise de grandes volumes de dados, desenvolvimento de bases de dados biológicas e de aplicações web para visualização e consulta, métodos estatísticos e computacionais para a análise de dados massivos em biologia, análise de estrutura e função de biomoléculas e desenho racional de medicamentos e proteínas, e computação de alta performance.

### 3.2 Fenotipagem e visualização por computador

A fenômica é reconhecida como um campo independente que auxilia o desenvolvimento de ferramentas para a fenotipagem em grande quantidade e alta precisão. O grupo composto por pesquisadores do Conicet, Inta, Universidade de Mar del Plata e setor privado trabalha em um protótipo industrial de uma plataforma automática para fenotipagem de plantas. Este desenvolvimento, na sua primeira versão, foi premiado em Innovar 2011 e ganhou uma menção outorgada pelo Instituto Balseiro por sua contribuição ao desenvolvimento da fenômica na Argentina. Foi desenvolvido pela PyME Adox S.A. e, por meio do uso de redes sem fio de sensores, permite o mapeamento da variabilidade espacial em ensaios a campo e o desenvolvimento de algoritmos para a automatização da análise de imagens.

### 3.3 Biologia computacional

A biologia computacional é uma ciência interdisciplinar originária de diferentes áreas como a matemática, química, estatística, física, biologia e ciências da computação. Na Argentina é praticada em grupos que desenvolvem em diferentes universidades atores do país. Existe um artigo que descreve o estado da arte na Argentina e os seus principais atores, “Biología computacional en Argentina”<sup>9</sup>.

<sup>8</sup> Plataforma Bioinformática Argentina (BIA), disponível em <<http://www.bioargentina.com.ar>>

<sup>9</sup> Biología computacional em Argentina, disponível em <<http://www.bioinformaticos.com.ar/biologia-computacional-en-argentina/>>

## 4 TIC nas cadeias de fornecimento

### 4.1 TIC nos recursos naturais e nas mudanças climáticas

O conhecimento e acompanhamento temporal da cobertura vegetal, a informação de previsões evolutivas da vegetação, das previsões meteorológicas e da zonificação de eventos relevantes resultam como fundamentais no processo de tomada de decisões para o manejo adequado e sustentável dos agrossistemas e dos recursos naturais. O Inta oferece serviços e realiza pesquisas por meio de seus grupos de Sistemas de Informação Geográfica e Teledeteção<sup>10</sup>.

A partir do processamento de imagens e processamento de dados meteorológicos é possível obter:

- Índices de vegetação normalizados.
- Evapotranspiração real.
- Anomalia ET.
- Temperatura de superfície.
- Albedo.
- Infravermelho térmico.
- Infravermelho a cores.
- Temperaturas limite nubladas.
- Animação.
- Cenários evolutivos.
- Eventos extremos.

O trabalho consiste em incorporar soluções de diferentes áreas: solos, regadio, produção agropecuária e recursos naturais, sendo representadas em diferentes bases de dados e em mapas temáticos que fazem parte dos Sistemas de Informação Geográficos. As imagens de satélite dos diferentes sensores Landsat, Spot, Radarsat são processadas, para fornecer apoio ao produtor nas diversas atividades que este realizar.

### 4.2 TIC e a segurança fitossanitária

O projeto FruTIC<sup>11</sup> fornece ferramentas tecnológicas para otimizar a gestão integral de empresas produtoras de cítricos vinculando o desenvolvimento do cultivo, as condições meteorológicas e a evolução de pragas e doenças. Estas ferramentas possibilitam que o produtor citrícola melhore a gestão integral da sua terra, reduza os custos e cumpra com requisitos de qualidade e preço para melhorar a sua rentabilidade, mantendo e aumentando seus mercados.

O FruTIC está desenvolvido para aqueles produtores que desejam melhorar sua rentabilidade e acessar normas de certificações de qualidade (Figura 1).

Um produtor integrado a FruTIC pode:

1. Conhecer o estado de brotação e floração em suas variedades.

<sup>10</sup> Instituto de Clima e Água, disponível em <[http://climayagua.inta.gob.ar/area\\_de\\_teledeccion\\_y\\_sig](http://climayagua.inta.gob.ar/area_de_teledeccion_y_sig)>

<sup>11</sup> FRUTIC, disponível em <<http://www.frutic.org.ar>>



Figura 1. Componentes do FruTic.

2. Conhecer a melhor data para a colheita da fruta.
3. Saber quanta água deve repor mediante a irrigação.
4. Receber alertas instantâneos de eventos meteorológicos.
5. Decidir com maior segurança se deve controlar ou não uma praga.
6. Conhecer a evolução das pragas e doenças de importância citrícola.
7. Conseguir um protocolo de produção, certificável por normas internacionais.

FruTIC oferece informação para aumentar a rentabilidade da empresa citrícola, possibilitando um uso mais eficiente de agroquímicos, uma maior qualidade da fruta e um acesso melhor a certificações internacionais. A informação fornecida por FruTIC é personalizada, segundo os requisitos dos produtores que integram o programa.

Na Argentina, o controle fitossanitário compete ao Serviço nacional de Saúde e Qualidade Agroalimentar, Senasa, que conta com um Sistema de Certificação Fitossanitária Eletrônica<sup>12</sup>.

### 4.3 TIC nos processos de produção agrícola

#### 4.3.1 Automatização e agricultura de precisão

A Agricultura de Precisão<sup>13</sup> continua caminhando a passos firmes, na Argentina (Figura 2), consolidando a tendência evidenciada ao longo da última década. O Inta centrou seus esforços

nos três elementos mais relevantes dos equipamentos aplicados à Agricultura de Precisão: monitores de rendimento; monitores de sementeira e sinalizadores via satélite para pulverizadores. Contudo, o grande desenvolvimento da atividade tem permitido um importante avanço das TIC, especialmente na área de telemetria e monitoramento de pulverização, sementeira e rendimento.

Os monitores de rendimento mantem-se como uma opção necessária para otimizar as tarefas de colheita. Em 2011 foram comercializados, aproximadamente, 1000 equipamentos, totalizando um plantel de 8.415 unidades no país, com um aumento de 13% em relação às vendas acumuladas até o ano anterior.

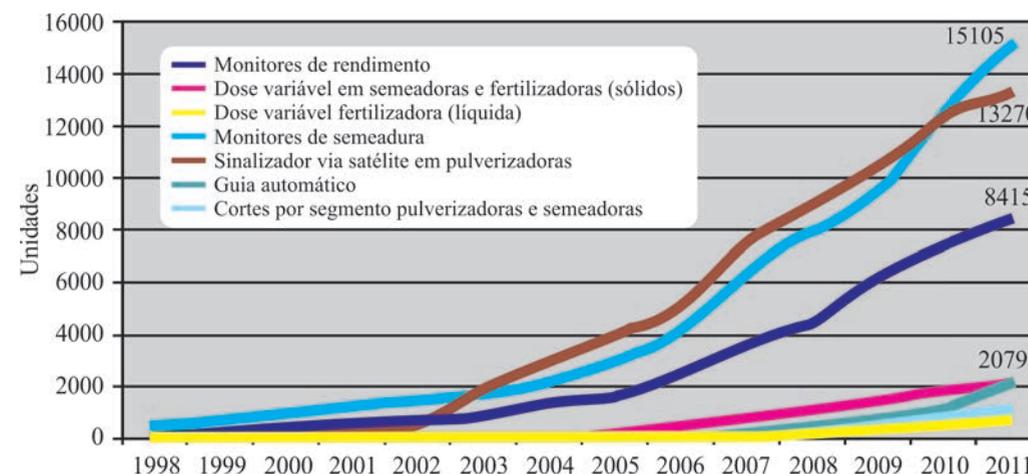


Figura 2. Evolução da adoção de ferramentas.

O salto foi ainda maior no segmento de monitores de sementeira. Em 2011 foram incorporados 2.545 aparelhos, com os quais as vendas acumuladas desde 1997 somaram 15.105 unidades. Desta forma, o parque disponível cresceu 20,2% em relação aos 12.560 equipamentos relatados até 2010.

Os sinalizadores por satélite continuam liderando as estatísticas de vendas acumuladas no país. Em 2011 foram incorporados 972 equipamentos para pulverizadores terrestres e 100 unidades especialmente adaptadas para aviões, totalizando um parque de 14.170 equipamentos comercializados desde 1997, representando um aumento de 6% em relação às unidades disponíveis até o exercício anterior.

Para completar o quarteto dos clássicos, os sistemas de dosificação variável somaram 425 novos equipamentos em 2011, incluindo 275 unidades para aplicações de sólidos (sementeira e fertilização) e 150 equipamentos preparados para aplicações variáveis de fertilizantes líquidos. Dessa maneira, o segmento acumula vendas em 2.829 unidades, com um aumento interanual de 17,6% em relação às existentes em 2010.

O Inta, por meio do grupo de agricultura de precisão, trabalhou no desenvolvimento de:

<sup>12</sup> Sistema de Certificação Eletrônica, disponível em <[http://www.senasa.gov.ar/requisitos\\_fitosanitarios/principal.html](http://www.senasa.gov.ar/requisitos_fitosanitarios/principal.html)>

<sup>13</sup> Site de Agricultura de Precisão, disponível em <<http://www.agriculturadeprecision.org/>>

- Monitor de Rendimiento Exacagro 128A, desenvolvido pela empresa IGB.
- Software GIS GeoAgro GIS.
- Sistemas de amortização para corpos de semeadura de modo a controlar a estabilidade dos mesmos e conseguir uma uniformidade na distribuição e profundidade da semeadura. Trabalho desenvolvido em conjunto com as empresas Agrometal e Baratec.

As áreas mais desenvolvidas em nosso país são:

- Telemetria geral de maquinaria (colheitadeira).
- Telemetria da informação gerada por equipamentos (monitor de rendimento, monitor de pulverização, monitor de semeadura, estações meteorológicas, sistemas de irrigação etc).
- Envio de informação georreferenciada para um servidor específico (gps, poketpe, Tablet etc). Isto é muito utilizado no monitoramento de pragas e ervas daninhas. A informação é enviada para uma plataforma web em tempo real e georreferenciada;
- Processamento da informação em plataformas web.

Em relação às ferramentas de Agricultura de Precisão de apoio tecnológico:

- Amostradores de solo pneumáticos com GPS: armazenam a amostra na memória de um Data Logger em forma georreferenciada para logo cruzar a informação obtida da análise do laboratório. É possível também fazer um trajeto programado para recolher amostras em lugares específicos do lote.
- Penetrômetros e sondas para medir compactação, profundidade do lençol freático e profundidade até a camada restritiva ou endurecida, entre outras coisas: esta informação é armazenada na memória de um Data Logger em forma georreferenciada para, depois, ser processada no escritório com software específico ou universal.
- Estações meteorológicas: armazenam a informação em algum dispositivo, como um Datta Logger para, posteriormente, baixa-lo em um PC ou levá-lo, em forma remota, via GRPS, para um servidor web ou telefone celular. Estas estações podem ser estáticas ou montadas nos pulverizadores terrestres ou aéreos.
- Monitores de semeadura: dentro deste grupo estão os monitores controladores de semeadora e os computadores de semeadura. Os primeiros controlam cada um dos sensores que equipam a semeadora para monitorar o adequado trabalho da mesma, velocidade de trabalho, densidade de semeadura, doses de fertilizante, monitoramento do nível das carretas agrícolas dos grãos e de fertilizantes, também conhecidas como “carretinhas”, obstrução de algum duto de queda de semente etc. Os computadores de semeadura cumprem a mesma função dos monitores controladores, mas têm, por sua vez, a capacidade de comandar motores que incidem sobre a semeadora para que esta realize a aplicação variável de insumos. Este comando se realiza através de software com dados georreferenciados para que a máquina responda em tempo real à variação de insumos em função do que foi prescrito para cada ambiente. Estes monitores também podem enviar a informação que está sendo considerada em forma remota para algum telefone celular e/ou página na web.
- Monitores de pulverização: com eles ocorre o mesmo processo que os monitores de semeadura, sendo que somente estes consideram a informação gerada pelo pulverizador, seja terrestre ou

aéreo, ou seja, cumprem a função de controlar e também, em alguns casos, atuar de maneira variável na aplicação dos insumos. Do mesmo modo, estes monitores enviam a informação de forma remota para um telefone celular ou servidor web. Por sua vez, tendem a estar sincronizados com estações meteorológicas montadas no mesmo pulverizador, o qual indica ao sistema as condições na qual está aplicando. Esta informação também pode estar chegando ao produtor ou dono da máquina, através do seu celular ou de uma página da web.

- Dentro deste escopo de aplicação também estão os monitores para fertilização sólida ou também para aplicar emendas em geral.
- Monitores de rendimento: esta ferramenta costuma ser equipada por série nas diferentes colheitadeiras. São encarregados de armazenar toda informação revelada pelos sensores que aparelham a colheitadeira e que são encarregados, junto ao GPS, pela confecção dos mapas de rendimento. Da mesma forma que as demais ferramentas, alguns destes monitores armazenam a informação em uma memória interna, em uma memória flash e também enviam os dados para um servidor web sob a forma remota. Por outro lado, alguns destes monitores não somente gerenciam dados de rendimento, mas também estão monitorando o estado geral da máquina e as condições nas quais está colhendo. Em muitas ocasiões esta informação é também enviada a um servidor que controla o estado geral das peças e de suas horas de trabalho para identificar alguma mudança ou reparo necessário.
- Estes monitores de rendimento também têm suas versões universais, as quais podem ser montadas em qualquer colheitadeira que não venha com o equipamento de precisão.
- Monitores em carretas-balanças: muitos desses monitores somente armazenam a informação em uma memória interna, mas, na atualidade, já existem desenvolvimentos nacionais que enviam de forma remota toda a informação que é gerada no mesmo momento para um telefone celular e também para um servidor web. Estes monitores podem enviar dados de localização da carreta, umidade do grão, destino do grão na colheita, horas de trabalho, entre outras.
- Nesta categoria das carretas, ou “carretinhas”, o mesmo está ocorrendo com os misturadores para alimentação do gado. Há desenvolvimentos nacionais que monitoram toda a atividade do misturador e enviam essa informação a um software de plataforma web, o qual gerencia de maneira remota o destino de cada alimento, a composição, horas de trabalho, estoque de produtos etc.
- Software para a análise da informação: grande porcentagem do que é utilizado é de desenvolvimento norte-americano, mas também há desenvolvimentos nacionais que permitem administrar e editar informação georreferenciada. Atualmente existe um software de plataforma web que possibilita baixar a informação de maneira remota, conectado à internet e com vários usuários trabalhando sobre o mesmo sistema. Nesta plataforma é possível carregar mapas de rendimento, imagens por satélite, amostras de solo, caracterização de ambientes, cartas de solo, prescrições de aplicação etc. Neste software podem trabalhar vários usuários ao mesmo tempo, otimizando a informação de todos.
- Monitoramento em estabelecimentos leiteiros: neste caso fala-se de “criação de gado de precisão” e faz referência ao monitoramento geral do tambo, do local de ordenha, a produção geral e individual, alimentação e até monitoramento georreferenciado do animal para avaliar seus hábitos comportamentais no lote, no pasto, nos bebedouros etc.

- Monitoramento de sistema de irrigação: neste aspecto ocorre algo similar aos já apresentados. Toda a informação que pode ser visualizada em um painel de controle ou em um monitor que permita administrar a irrigação, atualmente é possível receber essa informação em um telefone celular ou em uma página web e também enviar informação a partir de outros lugares ao equipamento para que, por exemplo, interrompa a irrigação ou comece a regar em dado momento específico do dia, ou seja, é preciso existir uma conexão bidirecional da informação.

Hoje em dia, tanto no agro quanto na criação de rebanhos, qualquer dispositivo eletrônico tende enviar toda informação de relevância para uma página web ou para o celular e também manejar a informação de forma bidirecional.

## 5 Aplicações e futuro das AgroTIC

### 5.1 1 Novas tecnologias - evolução tecnológica e o futuro das AgroTIC

As diferentes áreas da agricultura e a proliferação de sensores requerem ações interdisciplinares que permitam a geração de informação para monitoramento e prevenção de eventos e desenvolvimento de políticas públicas. Existe a demanda de produtos e serviços que requerem o uso de tecnologia para melhorar os sistemas produtivos.

O uso de tecnologia para o desenvolvimento destas áreas está vinculado com outros conceitos como “Internet das coisas” e “BIG Data”: tudo interconectado e gerando volumes de dados em níveis superiores aos esperados e com diferentes graus de validade.

Neste ensejo, é preciso contar com uma atividade permanente que possibilite a atualização, a geração de conhecimento e o desenvolvimento de produtos que contribuam para a melhoria da qualidade de vida dos cidadãos e dos sistemas de produção.

A aquisição de dados vê-se favorecida, mas a possibilidade de executar inteligência de negócio e mineração de dados requer a gestão e o armazenamento para, logo, ser possível realizar inferências.

Enfrentamos, também, situação onde muitas dessas informações não provêm de sistemas transacionais, mas sim de documentos de trabalho, publicações e redes sociais.

Concluindo, tudo está preparado para disponibilizar a captura de dados em todo lugar e em todo momento. O desafio consiste em como gerar conhecimento a partir dessa quantidade de dados que cresce em larga escala. O futuro está claramente na exploração e análise da informação que se gera.

## **Tecnologias de Informação e Comunicação e sua relação com a agricultura – Uruguai**

Juan Manuel Soares de Lima Lapetina

### **1 Introdução**

A agricultura do conhecimento é um desafio e uma exigência para os países da região e as TIC (Tecnologias de Informação e Comunicação) contribuem efetivamente para o desenvolvimento da agricultura onde são aplicáveis em quase todas as fases e âmbitos da atividade agropecuária. A adoção das TIC pelas empresas agropecuárias é um processo complexo que contempla diversas etapas e está condicionado pela heterogeneidade estrutural e pela estratificação dos agricultores (CEPAL, 2013).

Existem tendências globais que pressionam para que as TIC sejam adotadas na agricultura, entre as quais destaca-se: as exigências de competitividade dos mercados; o aumento da oferta de operações eletrônicas por parte das instituições; as pressões sociais e comunicacionais geradas; a oferta tecnológica com crescente componente de TIC, e as ações de fomento digital.

O desafio, além da inclusão dos pequenos produtores, consiste em criar condições para a plena utilização dos recursos digitais pelas empresas de ponta na agricultura da região.

Em virtude do maior desenvolvimento que alguns países apresentam e da existência de experiências relevantes é conveniente estabelecer mecanismos que favoreçam o intercâmbio e a transferência entre países em matéria de fomento de TIC para a agricultura.

O objetivo principal deste trabalho é realizar uma compilação de informação referente às TIC de maior importância que estão sendo utilizadas em nível nacional nos diferentes setores da produção agropecuária.

### **2 TIC na construção das AgroTIC**

#### **2.1 Curso de introdução à bioinformática**

Com a finalidade de transferir a experiência de bioinformática adquirida pelo grupo de pesquisa do Inia (Instituto Nacional de Pesquisa Agropecuária), foi apresentado o curso: “Introdução à Bioinformática: aplicações em projetos genômicos de melhoramento genético” à Unidade de Pós-Graduação e Educação Permanente da Faculdade de Agronomia, que aprovou para ser incluído na sua lista de cursos oferecidos pelo Mestrado em Ciências Agrárias.

O curso foi realizado na Estação Experimental Wilson Ferreira Aldunate - Las Brujas, de 17 a 21 de março de 2014 com a participação de 16 estudantes provenientes do Uruguai e da Argentina. Contou-se com um corpo docente interdisciplinar, composto por pesquisadores da Faculdade de Agronomia, Faculdade de Ciências, Instituto Pasteur de Montevideu e Inia.

O curso focou-se no uso de ferramentas bioinformáticas para a análise de dados genômicos a ser aplicados no melhoramento de plantas e animais. Foram proporcionadas as bases teóricas e práticas das técnicas mais usadas de sequenciamento massivo e suas principais aplicações; em particular as ferramentas de detecção de marcadores moleculares SNPs a partir de dados provenientes da técnica de genotipagem por sequenciamento (do inglês, *Genotyping by Sequencing*) e de sequências de genomas completas.

A realização deste curso permitiu começar a consolidar uma massa crítica em bioinformática aplica ao melhoramento que possa contribuir para a execução de projetos genômicos em diversos âmbitos de pesquisa nacional.

## 2.2 Infraestrutura de alta performance

### 2.2.1 Capacidade de processamento

Com o advento das novas tecnologias de sequenciamento (*Next Generation Sequencing*) e a integração paulatina das mesmas às áreas de estudo do Inia, coloca-se a necessidade de dispor de uma capacidade computacional que possibilite realizar e executar diversos programas e processos bioinformatizados e estatísticos. Neste sentido desenvolveu-se e instalou-se um servidor de tecnologia IBM Blade Center que atualmente está instalado fisicamente na estação experimental de Las Brujas em Canelones.

O equipamento (Figura 1) conta com 1 TB de memória RAM, 40 núcleos de processamento e 20 TB de armazenamento. Está sendo usado por pesquisadores do Inia de forma satisfatória e tem cumprido a necessidade motivada originalmente.

Hoje é possível trabalhar em projetos estatísticos de bioinformática com este servidor e, em paralelo, está sendo usado nos projetos de Inia por participantes externos como, por exemplo,

pesquisadores da Faculdade de Agronomia da Universidade da República, conseguindo assegurar vínculos e fortalecer a força de trabalho e pesquisa.

### 2.3 Gestão da informação e do conhecimento

Ainfo é um sistema informatizado para a gestão da informação técnica e científica. Criado em software livre, é desenvolvido e atualizado pela equipe técnica da Embrapa. Possui acesso on-line, plataforma web com recursos 2.0, integra bases de dados e funcionalidades para as Bibliotecas.

Este projeto de automatização das Bibliotecas começou a ser desenvolvido há algum tempo (2011) em conjunto com a Coordenação da Unidade de Comunicação e Transferência de Tecnologia e a equipe da Unidade de Informática.

No ano de 2013 foi acordado um novo Projeto de Cooperação Técnica (PCT) entre o Inia e a Embrapa para a "formação de recursos humanos e transferências de tecnologia com o objetivo de apoiar a organização do acervo documental e informação tecnológica que o Inia produz através de seus projetos e programas de pesquisa", contando com uma equipe interdisciplinar de bibliotecários e especialistas em informática da Embrapa e do Inia.

Ainfo consta de 3 Módulos:

1. Gestor: inclui todas as etapas do fluxo de tratamento da informação: registro e controle de acervo (aquisições, empréstimos, devoluções, reservas etc.) fortalecendo os processos de trabalho interno através da plataforma web integrada (administrador-usuário).
2. Digital: permite armazenar as publicações digitais e integrá-las ao módulo anterior tornando-as disponíveis ao usuário.
3. Consulta: por meio de uma avançada interface de busca permite aos usuários acessar on-line o Catálogo de Informação Agropecuária e todo o acervo bibliográfico e coleção de revistas disponíveis nas Bibliotecas Inia, mantendo atualizada a informação, em tempo real.

Este novo sistema permitirá:

- Melhorar a gestão da informação.
- Manter as coleções e revistas atualizadas no catálogo em tempo real.
- Reunir a produção científica gerada pelos pesquisadores do Inia, promovendo maior visibilidade aos trabalhos de pesquisa.
- Obter a referência imediata do documento para a referência bibliográfica.
- Acessar os documentos digitais.
- Emitir relatórios da produção técnico-científica.

No Inia, através do link <<http://ainfo.inia.org.uy/consulta/>> é possível acessar e conhecer o Catálogo, que em breve estará disponível no Portal Web, uma vez que ainda estão sendo feitos ajustes. (Figura 2).

Captura de imagem da página de apresentação:

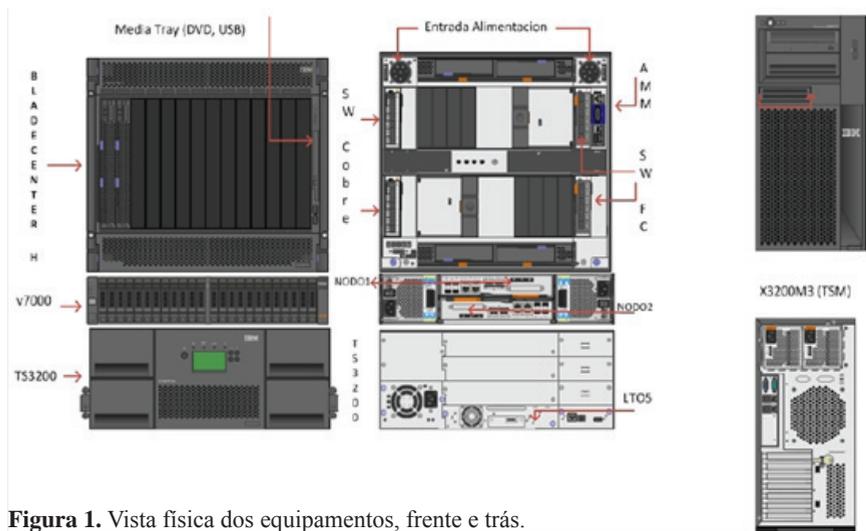


Figura 1. Vista física dos equipamentos, frente e trás.



Figura 2. Interface web do sistema Ainfo.

## 2.4 Modelos e simulação

### 2.4.1 MEGanE, Modelo de uma Exploração em Criação Extensiva de gado

Dieguez et al., 2012 - Instituto Planejamento Agropecuário (IPA)

MEGanE é um modelo dinâmico (o tempo é uma variável), empírico (reproduz as relações entre seus componentes a partir de observações experimentais e referências sem aprofundar-se nos processos) e determinístico (não está baseado em probabilidades). O modelo foi concebido para ser interativo, uma vez que permite que o usuário proponha situações do seu interesse. O conceito de extensivo refere-se àqueles modelos de exploração que não realizam melhoramentos forrageiros, tendo no Campo Natural (CN) a única fonte de alimento do rebanho. A Figura 3 apresenta o modelo MEGanE na forma de diagrama causal.

O peso vivo resultante determina o estágio de prenhez da vaca (output). O consumo também é uma resultante do modelo.

O modelo apresenta dois passos de tempo (*time-step*). Um diário que afeta o crescimento da pastagem, o consumo e a evolução de peso dos animais, e outro estacional, que afeta os coeficientes utilizados nos cálculos dos anteriores e permite ao usuário interagir com a simulação realizando operações.

Foi desenvolvido por uma equipe composta por especialistas nas áreas de informática, modelagem e simulação em conjunto com técnicos do Instituto Plan Agropecuario, do Uruguai. No processo, participaram técnicos e produtores da região do Basalto (norte do país) em um contexto de desenvolvimento participativo com a finalidade de resgatar o saber local. O MEGanE está definido para simular os resultados produtivos de uma exploração extensiva em CN sobre solos representativos do Basalto. Foi utilizado como ferramenta para suporte à tomada de decisões em

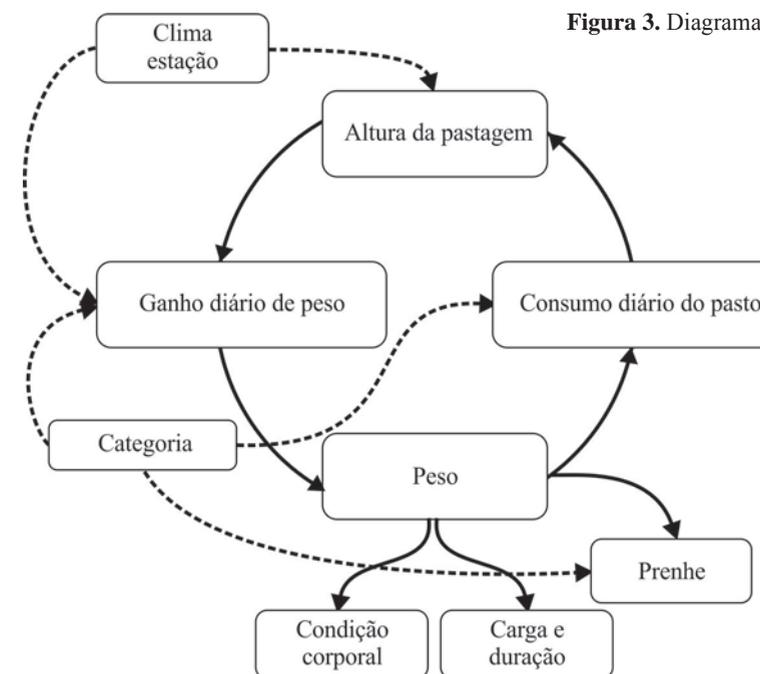


Figura 3. Diagrama causal do modelo MEGanE.

processos de aprendizagem com produtores, avaliando medidas de manejo e situações contrastantes de carga animal.

Foi idealizado para refletir sob forma amigável e intuitiva os efeitos das principais variáveis, como lotação e altura de pasto, em cenários prospectivos com um componente de variabilidade “climática” que afeta a taxa de crescimento do pasto.

Possibilita responder perguntas do tipo: como o ganho de peso afeta a lotação animal? Como o ganho de peso afeta a altura do pasto? Qual é a prenhez resultante ao trabalhar com diferentes lotações animais e/ou altura do pasto inicial? Qual é o efeito de uma redução ou aumento na taxa de crescimento do pasto no desempenho animal (produtivo e reprodutivo)?

### 2.4.2 Modelo de otimização agrícola-pecuário

Federação Uruguiaia de grupos Crea (Fucrea)

O modelo foi desenvolvido com base no Microsoft Excel como um conjunto de planilhas interconectadas que possibilitam completar uma matriz dois tipos de informação:

- As atividades: a ferramenta permite contemplar uma ampla gama de atividades da pecuária, rotações forrageiras, rotações agrícolas-forrageiras, rotações agrícolas, suplementação, transferência de forragem etc.
- As restrições: devem ser completadas as restrições do sistema, tais como, a área total melhorável e agrícola, balanço forrageiro etc.

Esta ferramenta determina qual é o sistema de produção (combinação de atividades) que maximiza o resultado econômico (margem bruta ou entrada de capital) dentro das opções possíveis no marco das restrições levantados (de natureza técnica-biológica e/ou aquelas restrições de caráter

empresarial). Também é possível determinar os custos de oportunidade dos diferentes recursos e sensibilizar o sistema diante de mudanças nas relações de preços.

Em função da complexidade dos sistemas bioeconômicos, com sua grande quantidade de componentes e inter-relações, é bastante difícil de modelar de maneira que esta ferramenta apenas pretende representar, simplificada, a realidade.

O resultado econômico é quantificado como a margem sobre alguns custos especificados (pastagens, sanidade, suplementação); ou como entrada de capital (margem anterior menos o resto dos custos da empresa).

A ferramenta desenvolvida permite adaptação às particularidades de diferentes regiões, definindo as diversas atividades (rotações agrícolas e forrageiras, atividades pecuaristas etc) e coeficientes técnicos de acordo com as características de cada região agroecológica a ser estudada.

Uma vez finalizada a entrada dos diferentes coeficientes e preços, esta ferramenta permite, além disso, determinar o sistema que maximiza o resultado, obter um menu de sistemas de produção alternativos que potencializam o resultado econômico para níveis variáveis de restrições.

O Modelo contempla a possibilidade de abordar o risco associado à variação de preços e rendimentos agrícolas por meio do software @Risk (Palisade Corporation).

Os preços que são considerados para avaliar o risco são aqueles dos produtos (pecuários e agrícolas) e de alguns insumos importantes. No que se refere aos aspectos físicos somente são consideradas as variações de rendimentos agrícolas.

Para realizar isso, naquelas variáveis de entrada que apresentam incerteza com relação ao valor futuro que será adotado, define-se uma determinada distribuição de probabilidade e não um único valor de preços ou rendimento.

#### 2.4.3 Modelo de simulação pecuarista

Juan Manuel Soares de Lima, 2009 - Instituto Nacional de Pesquisa Agropecuária (Inia)

Do ponto de vista da pesquisa, o objetivo do desenvolvimento deste modelo foi contar com uma ferramenta para analisar e avaliar os impactos produtivos e econômicos da implementação de determinadas técnicas, tecnologias, medidas de manejo ou cenários de preços sobre os resultados físico/econômicos de uma propriedade dedicada à pecuária, como ferramenta de apoio à transferência de tecnologia realizada pelo Inia.

O modelo simula o comportamento de um sistema de produção pecuarista em todas as suas etapas (cria, cria e engorda), considerando o desempenho animal, a dinâmica do rebanho, estratégias de manejo e o componente econômico.

Foi desenvolvido em Microsoft Excel, com componentes criados em Visual Basic para Aplicações (VBA), especificamente funções e procedimentos interativos.

É um modelo dinâmico e basicamente determinista, ainda que para a simulação de alguns processos como a engorda e venda para o frigorífico, utilizando critérios de estoque, tratando o grupo de animais como uma população cujo peso acompanha uma distribuição de probabilidade.

O intervalo de tempo utilizado (*time-step*) é de um mês, tendo sido considerado como o mínimo período necessário para simular certos eventos que ocorrem em alguns meses e/ou cuja duração é inferior ao ano.

O horizonte de simulação manejado é de 20 anos. Em função da duração dos ciclos biológico-produtivos envolvidos, verificou-se que este lapso é suficiente para garantir a estabilização dos processos que são simulados.

#### • Estrutura do modelo

Para facilitar a compreensão do funcionamento do modelo, dividiu-se em três blocos principais:

- a) Modelo animal: composto pelos mecanismos de crescimento e engorda que determinam o comportamento produtivo dos animais.
- b) Modelo populacional: compreende a dinâmica que determina a evolução do número de animais por categorias nos diferentes momentos do ano e as variáveis que os afetam.
- c) Modelo econômico: constituído por uma função de benefício que determinará o resultado do sistema produtivo em termos econômicos.

No Modelo animal, em função de uma disponibilidade e qualidade de forragem definidas, os animais são caracterizados pelo ganho (ou perda) diária de peso vivo. Deve-se esclarecer que a unidade não é o indivíduo, mas sim a categoria em cada mês.

O Modelo populacional estima a quantidade de animais em cada categoria ao longo do tempo em função de parâmetros inerentes à dinâmica populacional do modelo (taxas de nascimento e mortalidade por categoria, estrutura de idades, políticas de descarte dos animais, orientação produtiva do sistema - cria, engorda, cria + engorda - categorias de venda etc). O número de animais por categoria calculados em cada mês está relacionado com o peso correspondente, variável que, como se mencionou, constitui uma saída do modelo animal. Igual é o procedimento no tocante ao consumo por animal. Desta forma é possível estimar o volume total de quilos em estoque, mensal, e o volume de alimento consumido por mês. Por um procedimento de estoque, define-se, mês a mês, quais animais estão mais propensos à venda de acordo com o limite de peso definido para ser comercializado e a categoria. Este volume de peso a ser comercializado constitui uma das soluções deste sistema.

Finalmente, a produção estimada (representada pelas vendas se a estabilização do estoque foi alcançada), que constitui a saída do modelo populacional, é utilizada pelo Modelo Econômico para o cálculo de entrada e, ao descontar os custos, estimar a margem do sistema.

#### 2.4.4 Sistemas de Apoio à Gestão Florestal (SAG)

Hirigoyen e Scoz, 2014 - Instituto Nacional de Pesquisa Agropecuária (Inia)

Os Sistemas de Apoio à Gestão Florestal (SAG) são ferramentas informáticas baseadas em modelos de crescimento florestal que permitem simular diferentes opções de manejo, fornecendo alternativas para a tomada de decisões assertivas e de forma antecipada. Em geral os modelos são uma simplificação da realidade e, como tal, falham eles próprios em uma reprodução próxima ao real. Não obstante, são muito úteis para conhecer cenários de máximos e mínimos, por exemplo, o que torna possível a avaliação de risco. Se a base de dados sobre a qual se desenvolveu o modelo é suficientemente extensa, estes podem ser aplicados para a diversidade de condições de produção. Neste sentido, os SAG neles presentes permitirão realizar projeções de crescimento, produção e simulação de diferentes manejos.

O Programa Nacional Florestal do Inia colocou à disposição do setor alguns sistemas de apoio para a gestão de plantações para três espécies de importância florestal no Uruguai: *Eucalyptus*

*grandis* (METHOL, 2003), *E. globulus* (METHOL, 2006), *Eucalyptus* (METHOL, 2008) para plantações com destino à polpa e corte. Estes sistemas estão baseados em modelos empíricos de simulação de crescimento que permitem comparar manejos e locais alternativos desde o ponto de vista produtivo e econômico.

Atualmente encontra-se em sua etapa final a atualização dos modelos incluídos em SAG *globulus* e está sendo ampliada a base de dados para *E. dunnii*. Além disso, a atualização de modelos para *E. grandis* também está sendo realizada por meio de uma tese de doutorado.

A estes acrescentamos a necessidade de desenvolver sistemas de apoio para a tomada de decisões para outras espécies florestais de interesse, como é o caso dos pinheiros. Neste sentido, ao longo dos próximos meses estará também disponível o primeiro SAG para plantações de *Pinus taeda* no Uruguai, que representará a quarta espécie incluída em um SAG em 10 anos.

Finalmente é importante destacar o papel fundamental que o setor privado vem desempenhando nesta área de trabalho. Para o desenvolvimento e atualização destes modelos é indispensável contar com uma enorme quantidade de informação de qualidade sobre o crescimento das diferentes espécies e seus respectivos manejos em termos nacionais. Neste ensejo, muitas empresas contribuíram com seus dados de inventário, os quais, somados aos dados de ensaios do Programa Florestal, têm constituído a base destes modelos.

## 3 TIC na biotecnologia

### 3.1 Genômica aplicada

No Inia, com a criação da plataforma de Biotecnologia e do Banco de ADN genômico animal, surge a necessidade de contar com um software que possibilite que o Banco tenha um controle das amostras para a sua correta gestão, desde a entrada das mesmas até o processo de fornecimento de amostras para os projetos que as solicitam. Neste sentido, trabalhou-se em conjunto com uma equipe de Biotecnologia e desenvolveu-se um software que atende às características solicitadas. O mesmo foi desenvolvido em tecnologia web, que interage com uma base de dados desenhadas para surtir este efeito. (Figura 4).

#### Organização e gestão do Banco de ADN Genômico Animal

O Banco de ADN Genômico Animal é uma iniciativa estabelecida em conjunto com a Associação Rural do Uruguai (ARU) com o objetivo de conservar o material genético de ovinos e bovinos do nosso país para sua utilização em atuais e futuros projetos de pesquisa e seleção genômica (Tabela 1).

O Banco também inclui um laboratório de processamento de amostras para a extração do ADN com capacidade para conservar outros materiais biológicos, como órgãos e sêmen, também muito importantes para a aplicação tanto de técnicas de sequenciamento como de genotipagem para a I+D+i em genômica e seleção genômica. A gestão do mesmo é responsável tanto pela entrada e conservação do material genético como de sua administração e rastreabilidade desde o momento da sua chegada ao laboratório e durante a sua conservação.

### 3.2 Bioinformática aplicada



Figura 4. Interface do Banco de ADN Genômico Animal na web no Inia.

**Tabela 1.** Exemplos de tipo de populações consideradas no Banco de ADN Genômico Animal tanto ovinos como bovinos. Aparecem o número de amostras, projeto ao qual correspondem e as instituições participantes do mesmo.

Raça	Número de amostras	Tipo de população	Projeto e instituições envolvidas
Merino	2100	Rebanho	Geração de uma plataforma biológico-tecnológica de referência para estudos de seleção genômica aplicada ao melhoramento em ovinos no Uruguai: ênfase em resistência a parasitas (Inia, SUL, UDELAR, Universidade de Davis, Sociedades de Criadores, ARU).
Corriedale	700	Experimental	Linhas de seleção divergentes por resistência a parasitas gastrointestinais (SUL, Inia).
Hereford	400	Rebanho	Estudos de associação genômica para características da avaliação genética (Inia, SCHU, Iowa University).
Holando	1340	Plantéis comerciais (experimental)	Uso de polimorfismos genéticos para a melhoria da produção e fertilidade da vaca leiteira em condições pastoris (FPTA) (INML, UDELAR-FV, Inia).

#### Projeto de mapeamento associativo do arroz

Neste projeto têm-se trabalhado ativamente em vários aspectos bioinformáticos, desde a execução em conjunto do pipeline de Tassel (<http://www.maizegenetics.net/>) para descobrir SNPs passando pela criação de programas para processar a enorme quantidade de dados gerada até a criação de bases de dados que sirvam como base sólida para a informação gerada.

### 3.3 Biologia computacional

#### Base de dados de SNPs no arroz

Como parte do trabalho no projeto de Mapeamento Associativo do Arroz foi criada uma base de dados que permite identificar os diferentes SNPs e suas coordenadas no genoma do arroz para as linhas de pesquisa que trabalha o Inia segue (Figura 5).

type	start	end	score	strand	phase	attributes	id	chr	pos
gene	102680	102940	-	-	-	"D=LOC_045g0280.1Name=LOC_045g0280.1Pa=SI_181873	1		
mRNA	102680	102940	-	-	-	"D=LOC_045g0280.1Name=LOC_045g0280.1Pa=SI_181873	1		
exon	102680	102940	-	-	-	"D=LOC_045g0280.1Name=LOC_045g0280.1Pa=SI_181873	1		
intron	102680	102940	-	-	-	"D=LOC_045g0280.1Name=LOC_045g0280.1Pa=SI_181873	1		
gene	102680	103010	-	-	-	"D=LOC_045g0280.1Name=LOC_045g0280.1Pa=SI_181873	1		
mRNA	102680	103010	-	-	-	"D=LOC_045g0280.1Name=LOC_045g0280.1Pa=SI_181873	1		
exon	102680	103010	-	-	-	"D=LOC_045g0280.1Name=LOC_045g0280.1Pa=SI_181873	1		
intron	102680	103010	-	-	-	"D=LOC_045g0280.1Name=LOC_045g0280.1Pa=SI_181873	1		

Figura 5. Base de dados genômicos de arroz com localização de SNPs.

## 4 TIC nas cadeias produtivas

### 4.1 As cadeias produtivas

#### 4.1.1 Pecuária

##### 4.1.1.1 SNIG - Sistema Nacional de Informação Pecuária

O SNIG é um reflexo fiel e atualizado da situação atual e histórica dos atributos básicos, propriedade, localização e condições sanitárias do rebanho nacional e possui um conjunto de ferramentas adequadas para cada uma das áreas do Ministério de Pecuária, Agricultura e Pesca (MGAP) envolvidas, permitindo-lhes promover, regular e auditar a atividade do setor.

Este sistema de informação tem como objetivo principal assegurar a rastreabilidade do gado bovino desde o estabelecimento de origem do animal até o frigorífico, tanto individualmente como por grupos de animais, de acordo com as disposições e regulamentos do MGAP.

O SNIG teve duas estratégias fundamentais: melhorar o atual sistema de rastreabilidade grupal de Dicose (Controladoria da Divisão Pecuária), com a incorporação de novas tecnologias, enquadrando este processo nas normas vigentes e sem modificações no funcionamento atual, e a introdução gradual do Rastreamento Individual.

#### • Uma base de dados confiável e atualizada

Criou-se uma base de dados única e centralizada como forma de assegurar sua consistência e facilitar sua atualização. A integração, em um único sistema, dos dados de existências (provenientes da Declaração Juramentada Anual da DICOSE) e dos movimentos ou mudanças de propriedade (provenientes da Guia de Propriedade e Trânsito), permite, entre outras coisas, a manutenção de uma conta corrente atualizada por produtor e a adoção de medidas sanitárias mais efetivas em benefício do conjunto de produtores.

#### • Processamento ótico de formulários

Diversos formulários utilizados atualmente, e em particular a Guia de Propriedade e Trânsito, foram reformatados, adquirindo um aspecto já conhecido pelos produtores (similar à atual Declaração Juramentada, aos últimos Censos Agropecuários e ao último Censo de População e Residências). Assim, foi possível automatizar o processamento desses formulários para melhorar a qualidade dos dados, a eficiência dos processos e a consulta dos mesmos de qualquer ponto do país.

#### • Integração do sistema de informação geográfica

A distribuição da população animal no território nacional, assim como o detalhamento de seus movimentos, é um fator fundamental a considerar do ponto de vista sanitário. Por isso, foi montado um Sistema de Informação Geográfica, completamente integrado ao SNIG, que permite, a partir da informação recolhida nas Declarações Juramentadas anuais e das Guias de Propriedade e Trânsito, localizar no mapa cada um dos estabelecimentos registrados na Dicose e visualizar a origem e o destino de cada movimento de acordo com a data, espécie e categoria. Esta tecnologia começou a dar seus resultados tanto para a tomada de decisões operacionais como estratégicas.

A segunda estratégia do SNIG foi a introdução gradual do Programa de Rastreamento Individual. Para isso pôs em ação um Plano Piloto de caráter voluntário que culminou, em 1º de setembro de 2006, com a aprovação da lei de rastreamento obrigatório e a criação do Sistema de Identificação e Registro Animal, Sira.

O Uruguai passou por um caminho gradual até a incorporação de um sistema de rastreamento individual com o objetivo de satisfazer as crescentes exigências de nossos mercados compradores, incluindo nosso mercado interno.

**Rastreamento-definição** - É importante não confundir o rastreamento com a identificação dos animais. Esta última é apenas um meio para alcançá-lo. Algumas das definições de rastreabilidade mais comuns são:

- A capacidade técnica de identificação do animal desde seu nascimento até o final da cadeia de comercialização de seus distintos produtos.
- A habilidade para identificar a origem de um animal ou de seus produtos, tão distantes na sequência de produção quanto seja necessário, de acordo com o fim pelo qual a rastreabilidade tenha sido desenvolvida.

Quando se fala de rastreabilidade, pode-se referir a dois tipos: de produto ou de processo.

**Rastreabilidade de produto** - Trata-se do acompanhamento dos passos do animal desde o nascimento até o abate, fornecendo sua data de nascimento, lugar, proprietário, sexo e raça; e, por outro lado, seus movimentos, local de abate ou morte.

**Rastreabilidade do processo** - Além do mencionado, incorpora a informação de como foi produzido o animal e tudo que se refere aos aspectos sanitários.

Podemos concluir que a rastreabilidade do produto será, no futuro, um requisito adicional para entrar nos mercados mais exigentes; enquanto a rastreabilidade do processo de produção é o que traria um elemento diferenciador e permitiria um valor agregado ao produto final.

**Rastreabilidade do gado** - O conceito de rastreabilidade individual entrou com força na indústria de carne, especialmente a partir dos países da União Europeia, vinculado às crescentes exigências dos consumidores em relação à origem e à sanidade dos alimentos. O tema adquire

especial importância a partir de 1986 com o surgimento na Inglaterra do “mal da vaca louca” ou Encefalopatia Espongiforme do bovino (BSE).

O regulamento 820/97 da União Europeia exige de seus Estados-membros a implementação de um sistema de rastreabilidade individual para todos os animais e etiquetagem da carne assegurando a origem e outras condições de informação. O Uruguai, até esse momento, envia carne à EU por meio de seu sistema de rastreabilidade grupal implementado pela DICOSE, reconhecido por ela desde 1998.

Após a aparição da última epidemia de febre aftosa foi imposto controles mais rígidos que determinaram a necessidade de identificar, de forma individual, os animais com destino ao abate antes do embarque.

**Rastreabilidade individual-identificação** - O sistema de identificação individual do gado baseia-se em dispositivos permanentes com as seguintes características:

- Cada animal terá os dispositivos independentes e complementares (par de conjunto), um dos quais será um dispositivo visual, e o outro que conterà um dispositivo de radiofrequência (RFID) que armazena eletronicamente o mesmo número que está na identificação visual.
- Os dispositivos proporcionam uma identificação única e permanente durante toda a vida do animal, e que não pode ser alterada ou reutilizada.

O sistema de numeração individual será feito com 12 dígitos, entre os quais:

- Os 3 primeiros identificarão o país de origem do animal, neste caso a República Oriental do Uruguai, de acordo com as normas ISO 3166 ou normas internacionais equivalentes. O número ISO que identifica o Uruguai é o 858.

No caso da identificação visual esses três números são substituídos pela sigla UY.

- Os demais 9 dígitos identificarão o animal individualmente por meio de um número único e exclusivo.

O número do animal é o único dado armazenado no dispositivo eletrônico. O resto dos dados vinculados ao animal e seus movimentos será parte da base de dados do SNIG.

**Rastreabilidade individual - registro** - Os animais que possuam dispositivos de identificação individual deverão ser registrados no SNIG, incluindo os seguintes dados no momento da inscrição:

- Número do identificador.
- Dicose proprietário.
- Estação e ano de nascimento.
- Sexo.
- Raça e cruzamento.

Esses dados serão armazenados em uma base de dados, garantindo sua permanência no tempo e sua absoluta confidencialidade.

O SNIG focará os esforços iniciais na gestão da informação básica dos animais identificados e suas transações, e estabelecerá estratégias de integração com operadores privados especializados ou grupos de produtores que desejam maximizar o uso da informação do tipo produtivo.

### Rastreabilidade individual - movimento

- Operações de animais registrados no SNIG

Os produtores têm a responsabilidade de notificar ao MGAP toda mudança de propriedade, movimentos de entrada e/ou saída de animais identificados de seu estabelecimento.

Para isso deve recorrer aos serviços de um operador ou transportadora autorizada.

O adequado registro destas transações é o que permite manter a rastreabilidade dos animais.

- Transações de gados notificadas ao SNIG por meio de um operador autorizado

Os operadores são pessoas ou empresas devidamente registradas no Programa de Rastreabilidade Individual, os quais contam com equipamentos, software, capacitação e chaves de segurança que os permitem notificar em forma eletrônica o SNIG os movimentos e/ou mudanças de propriedade de animais com rastreabilidade.

- Transações de gado notificadas ao SNIG por meio de uma transportadora de fazenda

Um grupo de transportadoras tem sido registrado e conta com o equipamento necessário para a leitura e transmissão de dados ao SNIG.

### Rastreabilidade individual - história

- Registro de eventos

Os eventos notificados ao SNIG por meio de operadores de movimento, transportadoras, formulários, transações via Internet são registrados no sistema. Cada evento recebido é analisado em função da situação imediata anterior de cada animal, verificando sua consistência. Por exemplo: que o lugar de origem de um movimento coincida com o lugar de destino do último movimento registrado. Assim que eles são validados, a história do animal é atualizada ficando automaticamente disponível para a consulta, tanto pelo MGAP quanto pelo produtor proprietário do mesmo (Figura 6).

Eventos Animal						
CERRAR		AYUDA				
Animal		1213101		Estatus		Trazado
						Pag. 1/1
Imagen	Fecha	Tipo de evento	Propietario	Tenedor	Ubicación	
	27/10/04	REGISTRO - FORMULARIO DE REGISTRO DE ANIMALES AL PTI	030622706	030622706	030622706	
	30/04/05	GUIA DE PROPIEDAD Y TRANSITO	230622723	230622723	060613589	
						Pag. 1/1

Figura 6. Base de dados do site SNIG e detalhes de informações sobre diferentes eventos animas.

- Tipos de eventos

O primeiro evento de todo animal é seu registro no sistema.

Ao longo da sua vida poderão ser registrados: mudanças de propriedade, movimentos físicos, saídas do sistema, envio ao frigorífico etc.

No exemplo, mostra-se a história do animal 1213101. Ele mesmo foi registrado no SNIG em 27 de outubro de 2004 pelo produtor 030622706 que declara que o dito animal está em seu estabelecimento.

Em 30 de abril de 2005 o animal foi vendido ao produtor 230622723, que o transfere para o estabelecimento 060613589.

- Informação disponível

No caso de uma venda, o novo proprietário de um animal poderá acessar a história do animal desde o momento de seu registro, podendo consultar a partir do identificador:

- Os dados básicos com os quais foi registrado: sexo, idade, raça, cruzamento.
- Todos os proprietários anteriores.
- Os estabelecimentos por onde o animal esteve.
- A documentação relacionada a cada evento.
- Abate e destruição dos identificadores

Nos principais abatedouros de exportação do país técnicos da DIA supervisionam a retirada e destruição dos identificadores oficiais do SNIG.

Cada técnico do abatedouro conta com equipamento e capacitação adequados para a leitura e posterior transmissão ao SNIG dos números desses identificadores. Dessa forma, agrega-se como último evento de um animal rastreado a data e local de abate, ao mesmo tempo que se certifica oficialmente a destruição do dispositivo correspondente.

Fonte: site web MGAP, <www.snig.gub.uy>

#### 4.1.1.2 SEIIC - Sistema Eletrônico de Informação da Indústria de Carne

O Sistema Eletrônico de Informação da Indústria de Carne (SEIIC), também conhecido como Sistema “Caixas Pretas”, recebe e processa informação de quase todos os abates de gado bovino do país. Sua instalação e entrada em funcionamento supõe um passo enorme para nosso país e sua indústria mais tradicional de acordo com os desafios e oportunidades que estabelece a Sociedade da Informação.

Sendo uma das inovações mais importantes na incorporação de ferramentas tecnológicas chaves realizadas no Uruguai nas últimas décadas existe, no entanto, outro valor chave que faz a própria identidade deste Sistema.

De fato sua singularidade está, antes de mais nada, no fato de tratar-se de um Sistema de Informação em nível nacional que inclui atualmente 36 frigoríficos (o que significa mais de 95% do abate de bovinos em estabelecimentos industriais), resultado do trabalho conjunto de três administrações de Governo com os principais atores privados da indústria de carne bovina.

Este consenso ocorreu no âmbito do Instituto Nacional de Carnes (Inac) cujo Conselho é composto por representantes do Poder Executivo, da indústria frigorífica e dos produtores rurais, sendo o Inac o responsável da instalação e administração do Sistema.

Um aspecto que é desconhecido pela população uruguaia é que este sistema de controle do abate de bovinos é inédito na indústria de carne mundial e é o principal ativo em um mundo onde a segurança e qualidade dos alimentos se transformaram em uma prioridade estratégica para todos os países. Este é um ativo que o Uruguai deve conhecer e administrar com prudência.

#### Componentes do sistema

- Hardware: balanças, terminais, impressoras, scanners.
- Software: aplicativos, base de dados, informação eletrônica, integração de sistemas.
- Humano: entrada de informação, operação.

#### • Fase 1: as quatro balanças

A Fase 1 do SEIIC está instalada e operante em 36 indústrias desde 2007. Esta fase envolve quatro postos de pesagem instalados, controlados e mantidos pelo Inac. Os postos de pesagem verificam em balanças certificadas por organismos idôneos as seguintes medidas:

- a) O peso dos lotes de gado bovino que entram em cada indústria (Posto 1).
- b) O peso de cada animal depois de sangrado (Posto 2).
- c) O peso das duas meias-carcaças que obtém-se de cada animal antes da operação de acondicionamento das carcaças, conhecida como dressing (Posto 3).
- d) O peso das mesmas meias-carcaças após o processo de dressing. Além disso, a Fase 1 inclui uma balança (Posto 7 no fim da linha de produção) na qual pesam-se os produtos que saem para o mercado (normalmente o interno) sem terem sido colocadas em caixas (geralmente carne com osso).

As balanças citadas estão conectadas com terminais do sistema que coletam a informação (inviolável) das pesagens em cada posto e a armazenam em um servidor do SEIIC instalado em cada indústria, a partir do qual os dados são retransmitidos por redes de telecomunicações a um Servidor Central instalado na Inac.

O SEIIC abrange todo o país com um sistema uniforme de balanças calibradas e controladas periodicamente que proporcionam dados através de uma mesma estrutura para toda a indústria nacional, gerando múltiplas possibilidades de análises estatísticas, pelos quais várias já estão em uso.

A dimensão da complexidade e coordenação que supõe esse sistema realizado em nível nacional, fornecendo dados contínuos sobre o abate de bovinos, é dificilmente descrito. Uma imagem rápida é obtida só de pensar que seu funcionamento depende da interação diária, conjunta e contínua do Inac como administrador do Sistema com 36 Frigoríficos, 2 empresas de telecomunicações (sendo a principal a Antel - empresa pública de telecomunicações que fornece a comunicação *frame relay* às indústrias), várias empresas encarregadas de instalação e manutenção de hardware e software, órgãos públicos que fornecem pontos de apoio à rede de telecomunicações (Ministério do Interior), do Ministério de Pecuária, Agricultura e Pesca que, entre outras coisas, fornece informação sobre produtores rurais registrados na Direção correspondente (Dicose) e de

vários departamentos técnicos do Inac que operam de forma coordenada por meio de um Comitê Gerenciador do Sistema em que estão representados.

#### • Fase 2: a rastreabilidade

A Fase 2 do SEIIC inclui o funcionamento de três postos adicionais de obtenção de dados: a entrada da meia-carcaça no processo de desossa (posto 5), empacotamento dos cortes que resultam da desossa (posto 6) e envio final dos cortes de carne em caixas tanto para o mercado externo quanto interno (posto 7).

O objetivo estratégico central desta segunda fase será disponibilizar à indústria frigorífica nacional um sistema básico de rastreabilidade.

Seguramente isso gerará um novo valor agregado à carne natural uruguaia, dando ao setor privado mais ferramentas para que em seus planos de negócios incluam esses sistemas de garantia como parte da imagem da marca que constroem.

Finalmente, a ligação do SEIIC com o sistema de rastreabilidade do animal desde seu nascimento, utilizado pelo Ministério da Pecuária, permitirá, não apenas cumprir com os requisitos básicos que foram exigidos por alguns mercados chaves (ex: União Europeia em 2010), como também ampliará as possibilidades para que a indústria de carne nacional faça com que a carne rastreada do Uruguai cumpra os mais altos padrões mundiais.

#### Usos do sistema

Atualmente os dados gerados na Fase 1 são utilizados para gerar quatro conjuntos de informações básicas:

- Informação via web aos produtores rurais

Em primeiro lugar, um sistema de informação aos produtores rurais em duas variáveis: via web e via telefônica. Os produtores recebem do SEIIC uma senha que os permite acessar, por meio de uma página web, os dados de pesagem de seu lote de gado enviado aos frigoríficos, conhecendo, assim, os rendimentos de seus animais através das distintas etapas do processo até o momento posterior ao dressing. Além de ser um fator de transparência do mercado, essa também é uma ferramenta técnica muito útil para a tomada de decisões sobre os sistemas de produção e o tipo de gado a produzir.

- Informação via telefônica aos produtores rurais

A mesma informação, para torná-la mais acessível ao produtor rural, pode ser demandada por via telefônica através de uma chamada a um serviço gratuito, no qual as operadoras pedem simplesmente a senha do solicitante, verificam a autorização para acessar a seus próprios dados (simulando o sistema de controle de acesso por página web) e, caso se cumpra os requisitos de identidade verificada, fornecem a mesma informação disponível na web.

O produtor pode saber com precisão o rendimento conseguido pelo seu gado nos Postos 3 (pré-dressing) e 4 (pós-dressing).

Um aspecto complementar muito interessante, face a uma maior eficiência nas comunicações das indústrias com o Inac, é que o processo inclui a possibilidade de que a indústria frigorífica veja o conjunto de dados para o dia do abate coletados pelo SEIIC. Os dados que já estão no Servidor do Inac são confirmados posteriormente pela indústria e passam a ter, então, uma característica adicional: adquirem o status legal de declaração juramentada eletrônica. Este aspecto chave é o

que permitirá eliminar, paulatinamente, uma extensa lista de declarações e informação em papel que chegam hoje ao Inac com uma frequência menor (semanal ou mensal), substituindo-os por declarações eletrônicas que aumentarão a velocidade da informação, as quais serão coletadas e transmitidas diretamente por um único sistema para todo o país por meios automáticos e duplamente conferidos pelas indústrias e pelo Inac.

Desde 25 de setembro de 2007, as 36 indústrias incorporadas ao SEIIC estão publicando informação na web e desde o dia 1º de outubro fornecem, também, informação telefônica.

- Resumo Eletrônico de Cabeças Abatidas (“REC”)

O REC é um produto informativo interno para uso exclusivo do Inac com o objetivo de preparar as estatísticas que o instituto divulga semanalmente. É o primeiro programa de exploração do SEIIC que mostra a potência da informação “on-line”. Futuramente será uma ferramenta chave para o monitoramento do desempenho da indústria em uma base quase diária, o que colocará o Uruguai na liderança mundial nesse tipo de análise.

Este programa do SEIIC emite informação:

a) De cada indústria (de forma contínua) sobre as cabeças abatidas ao longo de cada dia dividida em 9 categorias de bovinos e indicando, inclusive, a terminação no caso de informação por cada indústria individualmente.

b) Informação dividida em quatro grandes categorias (vacas, novilhos, touros e vitelos) para o grupo de todas as indústrias do país incluídas no SEIIC.

Esta informação pode ser obtida para cada momento do dia (de manhã ou ao meio-dia, por exemplo) e, naturalmente, podem ser agregadas para construir estatísticas diárias, semanais, mensais ou anuais. Em um mercado tão dinâmico como o do abate bovino, contar com essa informação “on-line” oferece ao Inac uma força institucional muito destacada e uma capacidade de análise e formulação de políticas muito superior aos demais institutos que trabalham com o setor de carnes no mundo.

- Segmento eletrônico de carne natural

Outra função em uso é a colaboração com os serviços de Inac encarregados de certificar la Carne Natural Uruguaia. Neste caso o SEIIC permite fazer um rastreio chegando ao nível de detalhe de cada um dos bovinos enviados ao abate por produtores certificados pelo Inac às indústrias certificadas pelo Inac qualificadas dentro do sistema de produção de Carne Natural. Este é um primeiro passo para dar confiança a nossos compradores que se desenvolve com todas as suas possibilidades na Fase 2 com o programa de rastreabilidade que explicaremos mais adiante.

Novos usos do SEIIC estão em vias de se concretizar em breve, sendo, seguramente, o mais relevante a incorporação de informação econômica entre os dados recebidos pelo Sistema para seu processamento e divulgação pela Direção de Informação e Análises Econômicas do Inac.

Fonte: Inac <www.inac.gub.uy>

#### 4.1.2 Florestamento

##### 4.1.2.1 Produção florestal/madeira no Uruguai

- As empresas de florestais contam com sistemas de informação geográficas (SIG), nos quais se integra informação cartográfica que caracteriza espacialmente as plantações em relação aos

- mapas de solos, mapas topográficos, vias de drenagem, estradas etc, e que estão integradas a bases de dados do uso florestal (geodatabases), onde se registram as espécies, idade, manejos realizados e planejados, inventários, plano de colheita etc, para cada parcela florestal.
- Uma das fontes de informação dinâmica mais relevante que alimenta esses SIG, são as imagens aéreas, as quais podem ser de diferentes tipos, sistemas de captura e objetivos. Algumas das mais usadas nesse setor são:
    - Imagens de satélites: adquiridas de terceiros, ou seja, de empresas que as vendem. Pode ser de resoluções muito variadas, com tamanhos de pixel de 0,30 m a até 1 km de lado. Utilizam-se imagens com diferentes características espectrais, desde fotos RGB combinadas até imagens com mais de 10 faixas espectrais, as quais permitem realizar diversas composições e índices com o objetivo de destacar determinadas zonas, elementos ou atributos (biomassa florestal, água, profundidade de solos, força da plantação etc).
    - Imagens aéreas: podem ser capturadas através de veículos padrão (aviões pequenos) ou veículos não tripulados (UAV, drones). Estes serviços podem ser contratados localmente ou podem ser propriedade das empresas. Permitem capturar imagens visíveis de altíssima resolução espacial ou imagens térmicas no momento e com a frequência requerida pelas empresas, não dependendo das características dos sensores de satélites no que diz respeito à periodicidade e momento de aquisição. A alta resolução dessas imagens permite realizar tarefas de precisão com a delimitação da área de plantação efetiva e contagens com determinada frequência com o objetivo de reposição de plantas, quantificação de danos por geadas etc.
    - Lidar: Lidar (acrônimo do inglês *Light Detection and Ranging* ou *Laser Imaging Detection and Raging*) é uma tecnologia que permite determinar a distância de um emissor laser a um objeto ou superfície utilizando uma luz laser pulsada. Assim como ocorre com a tecnologia de radar, em que se usam ondas de rádio em vez de luz, a distância do objeto é determinada medindo o tempo de atraso entre a emissão do pulso e sua detecção através do sinal refletido. Em superfícies compactas sem vegetação têm-se um único retorno (toda a faixa de luz é refletida), em contraposição, em superfícies com vegetação o sensor é capaz de registrar vários retornos para um mesmo pulso de luz laser já que ele é parcialmente refletido pela vegetação. Esta capacidade multiretorno, fundamental para entender as aplicações florestais do Lidar, permite descrever com precisão a estrutura da vegetação, capturando informação tridimensional dos diferentes estratos do solo. Ao mesmo tempo o sensor registra as intensidades do sinal laser (visto que cada material reflete de maneira diferente o raio de luz laser) com as quais se podem produzir imagens de intensidades e tratá-las de forma similar às imagens capturadas com sensores passivos. Assim, o Lidar permite, além de gerar imagens da superfície, estabelecer um volume ou biomassa com o objetivo de realizar inventários florestais.
  - Muitas empresas florestais contam com software de otimização que permite desenvolver planos de colheita considerando espécies, idade e distância, buscando alcançar a máxima eficiência no uso de maquinaria, mão-de-obra e tempo.
  - Atualmente é recorrente o uso de aplicativos móveis que permitem visualizar grande parte da informação existente no SIG de forma que um operário possa contar, no campo, com a informação mais relevante quando se enfrenta determinado quadro florestal, o qual é atualizada

em tempo real. Igualmente, existem outros aplicativos móveis focados no monitoramento de operações, controle de colheita etc.

- Em algumas empresas realizam-se monitoramentos muito detalhados da maquinaria que se encontra em operação, os quais não se limitam apenas ao controle constante mediante tecnologia GPS, como também incorporam sensores na tomada de força, hidráulica, etc, realizando um diagnóstico preciso dos tempos de trabalho efetivo. Dessa forma é possível analisar mecanismos de melhoria da eficiência de trabalho modificando a operação de trabalhos, corrigindo os desenhos das plantações etc.
- Associado com a eficiência e com as medidas de segurança do trabalho, as frotas de empresas florestais contam com sistemas de rastreamento por satélite e sistemas de alarme frente a excessos de velocidade.

#### 4.1.2.2 Logística e transporte florestal

**Otimização logística** - Os grandes movimentos no setor são o transporte de matéria-prima desde o ponto de colheita até as indústrias de manufaturas e o traslado de produtos terminados até os mercados de destino (locais e no exterior). Empresas líderes em serviços de logística que trabalham com a indústria florestal utilizam TIC na programação e planejamento dos fretes, otimização de rotas e estradas a usar (estado das rotas, consumo de combustível etc.), acompanhamento dos envios por parte dos fornecedores e clientes e capacidade de fazer pedidos e reservas de forma eletrônica.

O objetivo do software é programar de forma eficiente o transporte de madeira em todos seus estados (florestas, serralheria, celulose, pátio de coleta) reduzindo ao máximo os gastos de transporte e respeitando as restrições técnicas, políticas e operacionais da empresa.

O sistema está fundamentado em duas noções básicas: um sistema administrativo centralizado que programa e controla todas as viagens, e um modelo de simulação que permite gerar as decisões.

Algumas empresas que exportam a madeira em fardos (toras) mencionaram que dispõem de sistemas de rastreabilidade, permitindo identificar cada fardo exportado com sua respectiva data de corte, parcela e campo correspondente, entre outras informações. Consultados se isso resultava em um preço maior ou se permitia o acesso a mercados mais exigentes, os entrevistados sinalizaram que não, mas que permitia um melhor controle de suas atividades produtivas e um serviço adicional para seus clientes. Por outro lado, com respeito à identificação das toras dentro do processo produtivo, algumas empresas disseram possuir sistemas de códigos de barras. No entanto, ainda que tenham a capacidade de acompanhar as toras na produção, não associam o produto terminado com a tora de origem, por isso, falta uma identificação completa do processo (floresta - tora - produto terminado). De acordo com os entrevistados, a identificação dos produtos terminados em relação à tora de origem não é um requisito para sua venda nos mercados internacionais. Nesse ponto, vale ressaltar que no mercado internacional existe uma forte tendência de exigir a identificação completa dos produtos terminados com a origem da madeira e as diferentes transformações e manufaturas sofridas por ela, especialmente no setor de móveis.

A União Europeia, por exemplo, estabeleceu requisitos especiais aos importadores de madeira que não demonstrem que os produtos provêm de florestas manejadas de forma sustentável ou

que não são resultados de derrubadas ilegais de florestas nativas. Esta situação, que ocorreu no Uruguai na exportação de painéis de madeira, mostra que há uma alta probabilidade de que este tipo de requisitos se generalizem.

#### 4.1.2.3 Fase secundária - elaboração da madeira

**Rastreabilidade** - A incorporação de códigos de barras, ou seja, a capacidade de identificar cada tora e/ou produto terminado e sua origem, permite agregar valor na produção por meio da possibilidade de garantir a origem da madeira (por exemplo se são florestas certificadas), reduzindo custos na cadeia e possibilitando o acesso a mercados internacionais. As vantagens podem ser analisadas em duas dimensões: por um lado, uma maior integração na cadeia de produção e, por outro, o manejo e monitoramento da produção nas indústrias de manufatura e remanufatura. Um exemplo da primeira dimensão é a produção de placas. A última etapa na serralheria é a classificação segundo as propriedades e qualidade. As placas são enviadas ao próximo elo, que continuará com o processo de manufatura, reclassificando-as segundo critérios e parâmetros distintos aos da primeira empresa. Se as placas fossem enviadas com suas respectivas informações de classificação, não haveria necessidade de reclassificá-las, diminuindo custos e tempo.

Em relação ao manejo e monitoramento da produção, os sistemas de rastreio facilitam a homogeneização das distintas sessões de produção, permitindo um melhor acompanhamento de cada tora ao longo da cadeia com maior controle sobre as etapas da produção e melhor acompanhamento entre a produção e os pedidos dos clientes.

**Sistemas de otimização, medição, processamento e análise de qualidade** - Os softwares de simulação são instrumentos cada vez mais utilizados na indústria. Estes instrumentos otimizam o uso da tora que entra na indústria, definindo os cortes de acordo com os produtos que demanda o mercado, qualidade de acordo com o uso final da madeira processada e, inclusive, permitem cumprir normas ambientais, reduzir o consumo de energia, melhorar as condições de produção e outras variáveis que se queira otimizar.

As soluções incluem desde simples modelos de programação linear até a instalação de sensores laser, de raios-x, luz, ultrassom, testes mecânicos, que recebem, processam e transmitem dados, até complexos sistemas integrados de decisão (DSS pela sigla em inglês).

Seu uso leva, também, à certificação de qualidade dos produtos, que é um fator chave para o acesso a mercados internacionais.

Fonte: Uruguay XXI, online, 2014.

#### 4.1.3 Laticínios

##### 4.1.3.1 Conexão com SAP, app off-line e múltiplos sistemas GeneXus para perfeito manejo de indústria láctea

Conaprole é uma cooperativa uruguaia, líder em processamento e distribuição de leite fresco, assim como também na produção de subprodutos do leite e sua comercialização, desde sorvetes e iogurtes até queijos e doces. A cooperativa, que trabalha com inúmeros processos diários – tanto para comercialização no interior do país como para a exportação – sempre está apostando na inovação tecnológica para maximizar e aperfeiçoar sua gestão.

É neste contexto que, desde o ano de 2000, a GeneXus (empresa uruguaia líder em desenvolvimento de software para múltiplas plataformas) e a SAP convivem na Conaprole; a SAP como

ferramenta de gestão administrativa e a GeneXus para o desenvolvimento de sistemas e funcionalidades unidas ao ERP (Planejamento de Recursos da Empresa).

Para implementar a comunicação entre GeneXus e SAP utiliza-se a tecnologia de BAPI's (Business Application Programming Interfaces), disponibilizadas como serviços que recebem os requerimentos do sistema desenvolvido na GeneXus (seja um pedido de dado da SAP ou um envio de dados para a SAP) e dispara a função remota que obterá os dados da SAP ou processará os dados recebidos.

Um dos primeiros sistemas GeneXus da Conaprole foi o GLF (Gestão de Leite Fresco) que maneja - em diferentes módulos - toda a operação de coleta do leite, a tomada de valores de qualidade, venda e pagamento aos produtores, entre outros processos em tempo real. GLF conecta-se com o ERP R/3 da SAP que administra outras áreas da empresa.

Por sua vez, um dos sistemas que integra o GLF, chamado de Info-Rut, interage com um sistema de Informação geográfica (GIS) e, através de algoritmos de base heurística, realiza cálculos para determinar, por exemplo, o melhor caminho que devem seguir os caminhões coletores. Esta ferramenta – desenvolvida em conjunto com a Faculdade de Engenharia – significou economia de US\$800 mil em seu primeiro ano de uso para o cálculo de distâncias, informou o gerente da área de informática da Conaprole.

Outros dos sistemas utilizados na Conaprole, desenvolvido na GeneXus, consiste no aplicativo na Internet que fornece aos produtores a possibilidade de acessar toda a informação tratada pelo GLF, ou seja: coleta de leite efetuada, análise de laboratório, compensações, datas de pagamentos e débitos correntes.

Até o ano de 2004, Conaprole utilizou um sistema de automatização de vendas para operar a partir de um Pocket PC que se comunicava com um sistema SAP via GPRS. Procurava-se, com isso, substituir o registro em papel e fax na hora de coletar os pedidos em supermercados e grandes clientes por meio de um aplicativo no Pocket PC. Diminuindo, conseqüentemente, os tempos de operações e as margens de erro. Assim mesmo a cooperativa conta com uma rede de subdistribuidores de produtos, os quais enviavam pedidos de grandes volumes através de formulários manuais, também; essa operação foi substituída por um sistema web desenvolvido na GeneXus, a qual permite enviar esses pedidos diretamente à SAP R/3, como pedidos de vendas, realizando os controles de estoque necessários.

Ao longo dos anos, as vantagens que a Conaprole encontrou com a GeneXus foram várias: “Com a GeneXus baixamos os custos de licenciamento, assim como os próprios custos de desenvolvimento” informou o Eng. Ricardo Scaich, gerente da Área de Tecnologia da Informação da Conaprole. Também se destaca a versatilidade, o menor tempo de desenvolvimento e a possibilidade de trabalhar com tecnologia de ponta com menor custo, em relação à integração com a SAP: “Este tipo de integração sendo tão natural no que faz a GeneXus nos abre uma grande gama de possibilidades de programação”, disse o engenheiro.

Atualmente a cooperativa está migrando seus sistemas mais importantes para a GeneXus X Evolution 2, visto que continuará com a estratégia de utilizar a GeneXus como ferramenta de desenvolvimento complementando a SAP, acompanhando os avanços que propõe a GeneXus em suas novas versões. Um exemplo é a utilização dos geradores para Smart Devices, nesse sentido, a Conaprole foi um dos primeiros testadores beta a utilizar a última versão GeneXus (Tilo) que

dispõe da funcionalidade de aplicativos off-line para Smart Devices para atualizar o aplicativo de pedidos (podendo incluir os mesmos ao dispositivo com conexão a Internet ou não) com o “Aplicativo Móvel AFV” (Aplicativo de vendas) para grandes áreas comerciais.

Fonte: Genexus, online, 2014

#### 4.1.4 Hortifruticultura

##### 4.1.4.1 Plan Senda

É um projeto chamado: “Democratização do acesso à informação no Mercado Modelo”, denominado também de Plan Senda. É executado de forma conjunta pela Comissão Administradora do Mercado Modelo (CAMM) e o escritório no Uruguai do Instituto Interamericano de Cooperação para a Agricultura (IICA), conta com o financiamento do Banco Interamericano de Desenvolvimento/Fundo Multilateral de Investimentos (BID/Fomin) e do International Development Research Centre (IDRC).

O projeto está dentro do programa ICT4BUSS do BID/Fomin.

A atual conjuntura nacional e internacional dos mercados e a reestrutura de logística e infraestrutura que está sendo impulsionada pela CAMM, constitui uma grande oportunidade para as 500 Pyme que operam no mercado.

Das empresas, 80% não contam com acesso a tecnologias de informação que lhes permita conhecer com precisão a realidade competitiva do setor, ficando em desigualdade de condições frente às 20% restantes.

Com este sistema busca-se melhorar a competitividade e produtividade das Pyme que operam no Mercado Modelo, baseando-se em:

- Gerar e fornecer informação de oferta e tendências do mercado às Pyme.
- Acesso à capacitação em tecnologias da informação.
- Assistência técnica e mecanismos de transferência tecnológica.
- Permite aproveitar o potencial de informação disponível como instrumento de apoio a tomada de decisões estratégicas.
- Para gerar informação de oferta e tendência de mercado para as Pyme produtoras e democratizar, assim, o acesso à informação mediante a criação e integração de uma comunidade virtual.

Os beneficiários são aproximadamente 2.200 Pyme que acessam o Mercado Modelo, 500 são beneficiários diretos do projeto. Estas Pymes são integradas por produtores, feirantes, armazéns, postos de frutas e verduras, autosserviços, supermercados não associados a cadeias, operadores e distribuidores do interior do país. Em sua maioria, são empresas familiares que contam com baixo nível de capacitação formal na área comercial e de negócios. Seu conhecimento do mercado foi incorporado através da vida cotidiana e da experiência.

##### • Objetivos e benefícios

- Capacitar as Pymes para a tomada de decisões estratégicas, disponibilizando informação atualizada sobre oferta e tendências do mercado e incrementar sua produtividade e rentabilidade.

- Gerar uma mudança cultural nas Pymes incorporando as TIC aos processos de trabalho e marcar diretrizes para melhorar sua gestão e acessar a novos mercados.
- Criar, consolidar e fortalecer os canais de comunicação entre a produção e o consumo por meio da comunidade virtual do Mercado Modelo.
- Democratizar o acesso à informação de oferta, preços e tendências incorporando transparência ao comércio de frutas e hortaliças.
- Criar uma comunidade virtual comprometida entre si e que sirva de base para projetos futuros de capacitação e de melhoria de gestão.
- Assentar as bases para replicar o projeto em mercados de abastecimento da América Latina, por meio da Federação Latino-americana de Mercados de Abastecimento, organização que apoia esse projeto.

##### • Soluções e serviços de TIC a serem desenvolvidos e testados

O projeto propõe a conformação de uma comunidade virtual, baseada em uma plataforma Web que integra:

- Um portal na Internet para a promoção das Pyme e seus produtos comercializáveis.
- Instalação de uma rede sem fio no Mercado para acesso à internet.
- Um sistema de suporte para a tomada de decisões baseado em indicadores de oferta, preços, e informação de tendências de mercado.
- Um campus virtual para a capacitação on-line.
- Uma ferramenta de comunicação e intercâmbio de informação, e-mail e fóruns.
- Uma Intranet com facilidades para o intercâmbio de informação e conhecimentos entre as Pyme (correio, fóruns, notícias, eventos, informação relacionada com a operação do Mercado Modelo).
- Financiamento conjunto para a compra por parte das Pymes de aproximadamente 60 laptops.

Instalação do Cyber Modelo, uma sala multimídia equipada, que permitirá capacitar aos interessados no uso da Internet e seus serviços.

Fonte: <<http://www4.mercadomodelo.net/tics-comunicacion.php>>

#### 4.2 TIC em recursos naturais e mudanças climáticas

Unidade de Agroclima e Sistemas de Informação (Gras) do Instituto Nacional de Pesquisa Agropecuária (Inia).

##### • Produtos disponíveis na web

##### 1. Disponibilidade de informação climática de estações automáticas on-line e em tempo real

##### 2. Análises do estado da vegetação a partir de informação de satélite

O índice de vegetação diferencia normal, IVDN ou NDVI, é uma variável que permite estimar o desenvolvimento de uma vegetação com base na medição, com sensores remotos, da intensidade da radiação de certas faixas do espectro eletromagnético que ela mesma reflete (Figura 7).

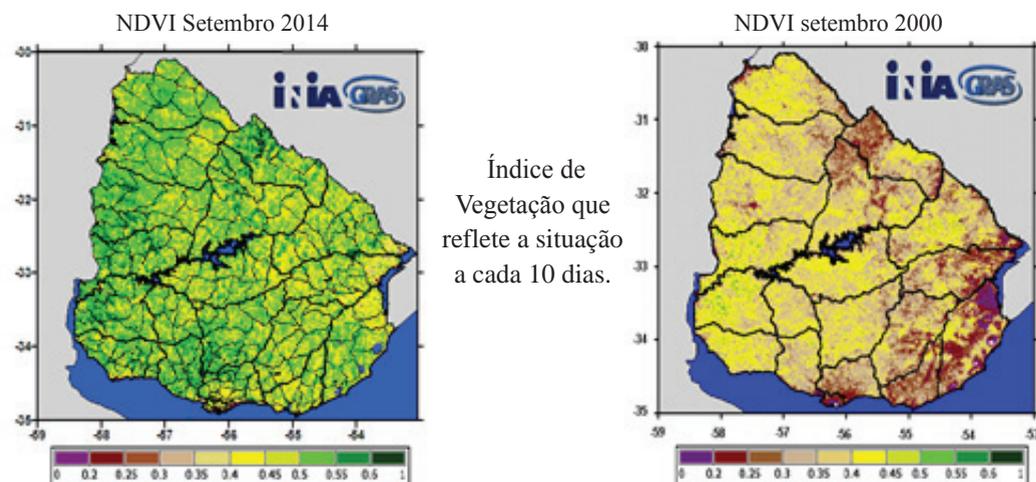


Figura 7. Índice de vegetação normalizado apresentado por Gras no site. Exemplos em situações constantes.

Estes são os resultados da interpretação das imagens produzidas a partir de informação captada pelo satélite NOAA-AVHRR através da Estação de Recepção do Inta Clima e Água Castelar em Buenos Aires - Argentina.

### 3. Análise do balanço hídrico dos solos a nível nacional mediante modelos

Nesta seção apresentam-se os resultados do modelo de balance hídrico para os solos do Uruguai. Este realiza uma estimativa do conteúdo de água disponível de uma região considerando o tipo de solo, a precipitação efetiva, a demanda potencial de água da atmosfera e a transpiração da vegetação.

Este modelo utiliza como variáveis de entrada:

- PRE: Precipitação Efetiva é calculada deduzindo da precipitação registrada em 84 estações meteorológicas (79 da DNM e 5 da Inia) um valor de escoamento superficial estimado em função da chuva antecedente (5 dias anteriores).
- ETP: Evapotranspiração Potencial ou demanda de água do solo por parte da pastagem é calculada em função de um modelo físico que estima a evapotranspiração potencial baseado em valores diários de: temperatura, umidade do ar, velocidade do vento e radiação solar. O método utilizado é chamado de Penman-Monteith.
- CR Água Solos: Capacidade de reter água do solo para cada uma das unidades de solo definidas segunda o Mapa de Reconhecimento de Solos do Uruguai escala 1:1.000.000 (Direção Nacional de Recursos Naturais Renováveis, Divisão de Solos e Águas, MGAP). O tipo de solo determina a capacidade de retenção máxima de água da zona de atividade das raízes. .

Para cada dia calcula-se a mudança de armazenagem mediante a aplicação de um modelo logarítmico que considera a retenção de água por parte do solo (adaptado de Thornthwaite C.W. e Mather J. R. "Instructions and tables for computing potential evapotranspiration and the water balance" Publ. In Climatology, 10:181-311, 1957).

O valor de transpiração diária corresponde à mudança de armazenagem calculada no passo anterior.

#### • Variáveis de saída do modelo

- IBH: Índice de Bem-Estar Hídrico que resulta da porcentagem da transpiração em função da demanda potencial diária (ETR/ETP).

Valores de IBH próximos a 1.0 indicam que a vegetação está em valores de transpiração próximos da demanda potencial. Por outro lado valores de IBH próximos de 0.0 indicam que a vegetação está em valores de transpiração muito abaixo da demanda potencial, indicando que climatologicamente a vegetação está sob estresse hídrico.

- ETR: Evapotranspiração Real.
- ANR: Água não retida. É a soma do escoamento superficial e excessos de água no solo (Água que excede o conteúdo de água do solo na capacidade do campo).
- ADI: Conteúdo de água disponível no solo.
- PAD: Porcentagem de água disponível, como  $ADI/CC*100$  (CC: capacidade de campo).

#### • Pressupostos

A metodologia considera que a vegetação de cobertura é:

- Uma pastagem (tipo gramínea).
- Intercepta 100% da radiação incidente.
- Encontra-se em ativo crescimento.

#### 4. Monitoramento ambiental por seção de variáveis como porcentagem de água disponível, produtividade primaria líquida aérea em pastagem, estado da vegetação através do índice de vegetação normalizada (NDVI) e água não retida

- Índice Verde (NDVI) proveniente do satélite MODIS: estimativa do valor médio por seção baseado em uma imagem composta a cada 16 dias durante o ano em curso e os valores médios por seção da média, o mínimo e o máximo de uma série histórica de imagens de 12 anos (2000-2011).
- Produção Primaria Líquida Aérea (PPNA) de pastagens: estimativas baseadas em imagens MODIS a cada 16 dias. A mesma está expressa em forma relativa como porcentagem (%) do valor atual em relação à média por seção (SP) dos valores médios, mínimos e máximos de uma série de estimativas de 12 anos (2000-2011). (INIA\_GRAS – UBA-LART).
- Porcentagem de água no solo (PAD): estimativa a cada 10 dias durante o ano em curso e pela média, percentual 10 e percentual 90 de uma série histórica de 25 anos (1985-2009) de estimativas do PAD.
- Água não Retida no Solo (ANR) como indicador de escoamento superficial: estimativa do acumulado (mm) cada 30 dias durante o ano em curso e pela média, percentual 10 e percentual 90 de uma série histórica de 25 anos (1985-2009) de estimativas do ANR.

#### 5. Monitoramento de cultivos baseados no estado da vegetação (NDVI)

Acompanhamento do estado geral de cultivos de verão em uma "amostra" estimada em 1.250.000 hectares com fazendas maiores que 25 hectares. Baseado no índice de estado da vegetação (NDVI\_MODIS) em 5 de março de 2014. Média de seções que contém uma área estimada maior que 1000 hectares

## 6. Previsão de DON para trigo durante o período de interesse (setembro a novembro)

### Desenvolvimento e Adaptação às Mudanças Climáticas (DACC)

Em 17 de janeiro de 2012 firmou-se entre o Banco Mundial e o Ministério da Pecuária, Agricultura e Pesca o Contrato de empréstimo 8099-UY Projeto de Desenvolvimento e Adaptação às Mudanças Climáticas DACC-MGAP. O objetivo geral é apoiar os produtores rurais para desenvolver um uso sustentável dos recursos naturais gerando uma maior adaptação à variabilidade e mudanças climáticas, promovendo uma modernização da gestão do MGAP na área da informação e serviços relacionados ao clima e aos recursos naturais.

#### • Estratégias

- Enfoque integral do desenvolvimento predial com opções de não retorno.
- Racionalização do acesso e gestão da água como base para o aumento da resistência.
- Apoios diferenciados por tipo de produtor segundo suas necessidades e capacidades.
- Construção de capacidades e modernização de serviços públicos em articulação com o setor privado e ao serviço do setor privado.

#### • O projeto está integrado pelos seguintes componentes

1º Desenvolvimento de um sistema de informação com camadas interoperáveis para a implementação de um Sistema Nacional de Informação Agropecuária, SNIA.

2º Mitigação e/ou adaptação aos efeitos das Mudanças Climáticas mediante à execução de subprojetos prediais integrais, assim como o aumento e/ou estabilidade da produção, entrada e desenvolvimento sustentável do mesmo, executado pela Direção Geral de Desenvolvimento Rural (DGDR)

3º Melhorar a gestão dos Recursos Naturais baseado na nova cartografia dos solos, implementação de planos de uso e manejo de solos, água e campo natural, executado pela Direção Geral de Recursos Naturais Renováveis (Renare).

4º Para o gerenciamento, articulação geral e relacionamento com a institucionalidade público-privada constituiu-se uma Unidade de Gestão de Projetos (UGP) que fará a administração, finanças, comunicação, difusão, capacitação, planejamento e monitoramento.

### 4.3 TIC no monitoramento e uso da terra

#### Uso de dados Georreferenciados do Uruguai-SIG

Um Sistema de Informação Geográfica (SIG) é uma integração de hardware, software e dados geográficos criada para capturar, armazenar, manipular, analisar e desenvolver em toda as suas formas a informação geograficamente referenciada.

Podemos encontrar nos SIG, com acesso web no Uruguai, vários tipos de dados para baixar, entre eles: Shapes, WMS (Web Map Service), KML (KMZ), Geo TIFF.

- **Shapes** - podem ser incorporados aos SIGs institucionais ou particulares. O acesso a informação faz-se por meio de um SIG de escritório (Kosmos, ArcGIS, gvSIG etc). O formato ESRI Shapefile (SHP) é um formato de arquivo informático proprietário de dados espaciais desenvolvido pela companhia ESRI, que cria e comercializa software para Sistemas de Informação

Geográfica como Arc/Info ou ArcGIS. Um Shapefile é um formato vetorial de armazenamento digital onde se guarda a localização dos elementos geográficos e dos atributos a eles associados. No entanto falta-lhe a capacidade para armazenar informação topológica. É um formato multi-arquivo, ou seja, é composto por vários arquivos informáticos.

- **Geo serviços Web** - a disponibilidade de dados é através de serviços WMS (e WFS) de forma que possam ser incorporados aos SIGs institucionais ou particulares. O acesso à informação realiza-se através de um SIG de escritório (Kosmos, ArcGIS, gvSIG etc). A partir do SIG de escritório pode-se realizar diversos tipos de análises, superposição com outras camadas e processamento de informação. Os mapas produzidos por WMS são gerados normalmente em um formato de imagem como PNG, GIF ou JPEG.
- **KLM** - estão disponíveis os dados de ordenamento territorial por meio arquivos KMZ (KML comprimido) de forma que possam ser visualizados no Google Earth.
- **GeoTIFF** - um arquivo TIFF é uma imagem com Informação Georreferenciado. GeoTIFF é um padrão de metadados de domínio público que permite que a informação georreferenciada seja encaixada em um arquivo de imagem de formato TIFF. A informação adicional inclui o tipo de projeção, sistema de coordenadas, elipsoide, datum e tudo que for necessário para que a imagem possa ser automaticamente posicionada em um sistema de referência espacial.
- **Metadados** - para uma correta utilização da Informação Geográfica é necessário considerar dados referidos a sua identificação, qualidade, tipo, referência espacial, entidades e atributos, distribuição, atualização, entre outros. Estas referências são obtidas por meio dos Metadados Geográficos.

Fontes de dados do Uruguai. Ministério da Pecuária, Agricultura e Pesca.

#### • SIG-RENARE (Dir. Nac. de Recursos Naturais Renováveis): alguns dos objetivos dessa direção são:

- Estabelecer, desenvolver e manter o Sistema de Informação Geográfica dos recursos naturais renováveis.
- Formular e coordenar ações com órgãos públicos nacionais e regionais no que se refere ao uso e manejo sustentável dos recursos naturais renováveis, assim como celebrar convênios com pessoas públicas e privadas, nacionais e estrangeiras para a execução dos objetivos, controlando seu cumprimento.
- Colaborar com a informação disponível em relação aos recursos naturais renováveis, ao regulamento ambiental do território.
- Difundir as normas que regulam o uso e manejo dos recursos naturais renováveis e as tecnologias que permitam o uso sustentável dos mesmos.
- **Visualizadores geográficos e web map service**

Através desse site são oferecidos serviços interativos de consulta sobre informação georreferenciada: mapas, imagens de satélite, fotos, tendo como objetivo difundir e distribuir informação sobre recursos naturais do Uruguai relacionados com a produção agropecuária. Estes serviços proporcionam ao usuário as ferramentas necessárias para interagir com a informação geográfica publicada tais como: selecionar camadas, “abrir ou fechar” mapas, identificar elementos, aproximar ou distanciar, medir, obter coordenadas, imprimir e baixar informação.

- Visualizadores geográficos
  1. Sistema de consulta predial às fotos aéreas. Serviço de consulta por padrão ao número de fotos aéreas 1:20.000 (Anos 1966-1968) de todo o país sobre os fotoíndices dos Serviço Geográfico Militar.
  2. Sistema de consulta a mapas interpretativos de solos (por exemplo: aptidão), assim como os tipos de coberturas e usos dos solos, imagens de satélites e cartografia básica.
  3. Aplicativo de Consulta Coneat.
  4. Aptidão para cultivo de verão, de seca.
  5. Mapas de armazenagem de Grãos.
  6. Obra de risco projeto Prenader.

#### - Serviço WMS

Estão disponíveis serviços WMS (Web Map Service = Serviços de Mapas Web) que permitem carregar, visualizar e consultas camadas de informação georreferenciadas em diversos programas de Sistema de Informação Geográfica (GvSiG, ArcGis etc). Também é possível a utilização destes serviços em outras ferramentas como Internet Explorer, Google Earth etc.

- Há dois serviços cujas URLs são:

- 1) <<http://www.renare.gub.uy/cgi-bin/mgap?>> contém mapas interpretativos de solos, assim como vários mapas de uso e coberturas do solo. Descarregar arquivo KML.
- 2) <<http://www.renare.gub.uy/cgi-bin/suelos>> corresponde aos mapas de solo publicados pela DGRNR

Fonte: Renare, MGAP, online, 2014

PRENADER- Coneat

Ministério de Indústria, Energia e Mineração

Visualizador geográfico de mineração (mapa 1:50.000 e informação hidrológica), permite buscas por padrão.

Serviços de Mapas Web (WMS) da Dinamige:

- Vigilância radiológica.
- Poços de hidrologia.
- Mapa geológico.
- Cadastro mineiro.

#### 4.4 TIC e a segurança fitossanitária

SCFFC - Sistema de Certificação Fitossanitária de Fruta Cítrica

O projeto inicial de rastreabilidade e certificação dos cítricos no Uruguai é um projeto quase terminado e em execução conjunta com a DGSA do MGAP há muito tempo. O mesmo consistiu em armar uma plataforma informática onde se pode fazer os levantamentos (GPS) dos quadros de plantação de cítricos (mais de 15000) em todo o país, e estes vetores georreferenciados estão

conectados a uma geodatabase com dados da citricultura. Esta base de dados é administrada pela DGSA, mas os dados são inseridos no sistema, via web, pelos próprios produtores ou por funcionários da DGSA.

Além disso, esse sistema conecta-se com a certificação fitossanitária e inspeções sanitárias para poder exportar para outros países. Como é dinâmico, as bases de dados e o georreferenciamento das bases de atualizam todos os anos. Os produtores ou empresas podem acessar seus croquis ou dados via internet.

O comércio internacional de produtos de origem vegetal, implica no risco de pragas, o que exige uma regularização fitossanitária. O MGAP por meio da Direção Geral de Serviços Agrícolas, que é a organização Nacional de Proteção Fitossanitária, tem a função de certificar os requisitos exigidos pelos países importadores.

Em 1998, novos requisitos da EU determinaram a implantação de um Sistema de Certificação Fitossanitária devendo modificar os procedimentos tradicionais de certificação fitossanitária no ponto de controle (chegada) à identificação de vários pontos de controle (lugar de produção, indústria de empacotamento, ponto de chegada). Foi necessário criar um Sistema de Rastreabilidade dos lotes de exportação ao longo da cadeia de produção-exportação.

As medidas fitossanitárias em cada **ponto crítico de controle** foram: no local de produção (inspeção de plantas), nas fábricas de empacotamento (inspeção da fruta e controle de tratamentos fitossanitários), nos pontos de chegada (emissão de certificado fitossanitário)

O **rastreamento** dos lotes de exportação é determinado com:

- Registro de operadores: nos locais de produção, fábricas de empacotamento, centros de armazenamento e exportadores.
- Identificação de Quinta-Quadro-Variada: por sistema de georreferenciamento identificam-se os locais de produção (quinta) com código alfanumérico, baseado nas coordenadas cadastrais de sua localização geográfica e um número correlativo e a unidade mínima de produção (quadro/variedade). Ex: O10B15- C 20 SOW.
- Identificação de pallets: registra-se os números de pallets empacotados de cada empresa.

Todo este processo de automatização realizou-se entre os anos de 2007 e 2010 por meio de um software desenvolvido para isso.

Os objetivos dos sistemas são:

- O cumprimento dos processos definidos e o das exigências dos mercados.
- Ter cada vez mais informação, apesar das dificuldades no manejo e na comunicação.
- A padronização da informação.
- Unificar os processos e os critérios da zona Norte e Sul do país.
- Cobrir toda a cadeia produtiva com todos os operadores.
- Escalabilidade todo o país.
- Alcançar uma análise histórica da informação.
- Informação georreferenciada.

## 4.5 TIC em processo de produção agrícola

### 4.5.1 Automatização e agricultura de precisão

A agricultura de precisão é um conceito agrônômico de gestão de parcelas agrícolas, baseada na existência de variabilidade no campo. Requer o uso das tecnologias de Sistemas de Posicionamento Global (GPS), sensores, satélites e imagens aéreas junto com Sistemas de Informação Geográfica (SIG) para estimar, avaliar e entender essas variações. A informação coletada pode ser usada para avaliar com maior precisão a melhor densidade de semeio, estimar fertilizantes e outros acréscimos necessários, e prever com mais exatidão a produção dos cultivos.

Alguns exemplos da utilização em agricultura de precisão:

- Determinação das necessidades da fertilização nitrogenada em trigo utilizando sensores ativos

Existem evidências claras de que a continuidade no tempo do sistema de agricultura contínua em semeio direto conduziu a uma redução da capacidade de aporte de N por meio de mineralização da matéria orgânica do solo (HOFFMAN, et al. 2011). Isto altera as doses e capacidade dos atuais métodos de recomendação de fertilização com N em diagnosticar doses ótimas em áreas com história agrícola em comparação áreas sem história agrícola (HOFFMAN, et al. 2013).

Por isso existe a necessidade de contar com ferramentas objetivas que permitam quantificar os efeitos antes mencionados.

Mediante técnicas de sensoriamento remoto é possível analisar diretamente, mediante espectrometria, o estado da vegetação (ou seja, conteúdo de clorofila, área foliar, nitrogênio absorvido), utilizando a refletância característica dos materiais a medir (vegetação, solo, restos secos) (HATFIELD et al. 2008). Os sensores passivos (câmeras ou sensores montados em satélites, aviões ou UAVs) medem a refletância ( $R = \text{luz refletida} / \text{luz incidente}$ ) da luz solar. Os sensores ativos medem a refletância da luz emitida pelo mesmo sensor, excluindo e, por tanto, em forma independente da intensidade da luz solar. Os sensores ativos são mais versáteis que os passivos e são mais simples de utilizar e operar.

Em ambos os casos, seja mediante sensoriamento remoto ativo ou passivo, não existe um modelo de diagnóstico e recomendação de fertilização nitrogenada para as condições nacionais.

- Monitoramento por satélite de crescimentos de pastagens

No ano de 2010 o Instituto Plano Agropecuário firmou um convenio com a Faculdade de Ciências e a Universidade de Buenos Aires para começar a trabalhar sobre a medição do crescimento das pastagens utilizando tecnologia por satélite. Desde 2011 estão trabalhando sobre distintos prédios localizados em todo o país. Dessa forma começaram a realizar medições mensais que permitem ao produtor ter outra fonte de informação para tomar decisões.

- Acompanhamento de forragens via teledetecção

O objetivo de utilizar este sistema é gerar e difundir informação atualizada sobre o crescimento das pastagens naturais das diferentes áreas agroecológicas.

O Sistema de Acompanhamento por teledetecção é um sistema de acompanhamentos da produtividade de forragens (também chamada de taxa de crescimento da forragem em âmbito da produção e produtividade primária líquida aérea, PPNA, no âmbito acadêmico). O sistema fornece

informação com o detalhe com o que mais frequentemente toma decisões de manejo do gado em um estabelecimento: em tempo real e por campo. O sistema utiliza informação via satélite que é traduzida à produtividade forrageira a partir dos croquis de cada campo, tipo de pastagem de cada campo, informação meteorológica e coeficiente ecofisiológico. A produtividade de forragem assim calculada é disposta em um site web que permite fazer consultas e ter um panorama do estado da forragem de uma área e momento e/ou descarregar a informação em planilhas de cálculo para uma análise posterior.

O sistema de acompanhamento da produtividade primária aérea do recurso se baseia em um modelo originalmente formulado por Monteith (1972). Ele estabelece que a produtividade forrageira (PF) em um período (um mês, por exemplo) está determinada pela quantidade de radiação fotossinteticamente ativa absorvida (RFAA) pelas plantas e a eficiência com que essa energia é transformada em matéria seca aérea. Por sua vez, a RFAA é o produto entre a radiação fotossinteticamente ativa incidentes (medida em estação meteorológica) e a fração desta que é absorvida pelas folhas verdes (fRFA) que depende da quantidade e disposição espacial da área foliar. A fRFA pode ser estimada com razoável precisão a partir da teledetecção e essa relação é o vínculo entre a produtividade da forragem e os satélites. A EUR pode ser calculada a partir de estimativas independentes da produtividade forrageira e a radiação absorvida ou ser tomada de estimativas da literatura ou a partir de modelos com base em variáveis ambientais. A EUR é muito menos variável que a RFAA para um recurso forrageiro determinado e, frente à falta de conhecimento mais detalhado, pode supor-se constante no modelo para calcular a produtividade forrageira desse recurso nesse ambiente. No entanto, espera-se no futuro contar com modelos que incluam variações por estação e variáveis ambientais. O índice de vegetação normalizado (IVN, também chamado de “índice verde”) é o mais comum entre os índices que se calculam a partir de observações via satélite para o acompanhamento da vegetação. O IVN está diretamente relacionado com a fRFA porque se baseia nas propriedades da vegetação verde de absorver intensamente a radiação visível (especialmente na longitude de onda vermelha), utilizada para a fotossíntese, e de refletir a maior parte da radiação na zona de infravermelho próxima. O acompanhamento forrageiro consta de um sistema de informação geográfica (SIG) que tem digitalizados os contornos e as subdivisões internas de um conjunto de estabelecimentos de gado que contribuam com informação e recursos econômicos ao projeto. Além disso, incorpora informação de radiação incidente (RFAI), índices de vegetação (IVN), coeficientes de eficiência no uso da radiação (EUR, calibrados com colheitas de biomassa especificamente para os recursos de cada zona), e uso da terra de cada lote previsto pelos agricultores. A partir dessa informação, uma série de cálculos internos gera estimativas de produtividade ou de radiação absorvida mensal (nos casos em que a EUR não foi caracterizada) para cada campo. O software especificamente desenvolvido maneja e armazena a informação em uma base de dados relacionados e implementa os procedimentos de carregamento de dados e cálculo de rotinas programadas (Grigera et al. 2007, *Agricultural Systems* 94: 637-648). As estimativas de produtividade forrageira são geradas por volta do dia 15 de cada mês, uma vez incorporada a informação meteorológica e de satélite do mês anterior.

A superfície predial monitorada em todo o país é de 41.662 hectares.

Fonte: Eng. Agr. Marcelo Pereira.

2009. IPA, LART e Fac. de Ciências. Implementação, difusão e transferência de um sistema de acompanhamento por satélite da produtividade ferragina em áreas de gado uruguayos.

Fonte: adaptado de Panario et al. 2011.

- Utilização de UAV's ou drones

No Uruguai, a soja provocou uma revolução tecnológica no campo, nas mãos de produtores que trabalham com modelos matemáticos, aviões sem pilotos (drones) e colheitadeiras de última geração.

A agricultura de precisão, entre outras coisas, refere-se à utilização de sensores remotos para escanear plantas e detectar o grau de hidratação, status sanitário e o ritmo de crescimento. Também estima um uso seletivo dos nutrientes e pesticidas que cada planta requer, reduzindo o custo e o impacto ambiental.

Os drones poderiam ser também a resposta à falta de mão-de-obra que sofre o campo e que com o tempo se agravará. Há poucos jovens interessados em trabalhar no campo e os trabalhadores que se dedicam a isso estão envelhecendo. A tecnologia pode dar resposta aos dois problemas, permitindo aos produtores incrementar as colheitas e reduzir custos. Em âmbito acadêmico existe o convencimento de que “os drones podem revolucionar a agricultura, reduzindo a necessidade de pesticidas e incrementando a produção”.

Estes aparatos são usados como um elemento estratégico para mapear o solo e determinar sua fertilidade, apostando na maior produtividade. Por meio de câmeras térmicas, estes pequenos aparelhos captam as necessidades de água ou de nitrogênio dos solos, indispensável para saber qual quantidade de fertilizantes será necessário.

Os drones serão cada vez mais precisos no exame e diagnóstico, indicando a dosagem de sementes e fertilizantes em tempo real.

Fontes: <[http://www.todoelcampo.com.uy/espanol/lo\\_nuevo\\_en\\_agricultura\\_los\\_drones\\_todo\\_una\\_revolucion-15?nid=7634#.VDbbmPlDX3Q](http://www.todoelcampo.com.uy/espanol/lo_nuevo_en_agricultura_los_drones_todo_una_revolucion-15?nid=7634#.VDbbmPlDX3Q)>, <[http://www.180.com.uy/articulo/35683\\_Drones-y-satelites-para-la-revolucion-de-la-soja-en-Uruguay](http://www.180.com.uy/articulo/35683_Drones-y-satelites-para-la-revolucion-de-la-soja-en-Uruguay)>

## 5 Aplicações e futuro de AgroTIC

### 5.1 Aplicações AgroTIC e transferência de tecnologia

#### • Uso de TIC em capacitação para a criação de gado

Em 2006 o Planejamento Agropecuário iniciou um projeto de validação de um Sistema de Transferência de Tecnologia para o setor agropecuário mediante o uso das Tecnologias e Informação e Comunicação (TIC), particularmente da capacitação à distância, que permita transferir conhecimentos, ferramentas de gestão e informação, aumentando a produtividade das empresas.

Cofinanciado pelo Banco Interamericano de Desenvolvimento, o projeto procurou provar que, sempre e quando não existirem limitadores de conectividade e acesso, as TIC têm um papel importante a desenvolver na capacitação do setor rural.

Como projeto piloto, definiu-se pela priorização da zona de criação e dentro dela foram escolhidas seis localidades, deixando-se de fora as capitais. Os lugares escolhidos contam com alguma associação que solicitou capacitação, localização estratégica, e sala pública ou privada

com computadores e conexão de internet. O público-alvo consiste em produtores, técnicos e assalariados e suas famílias.

Os cursos trataram de temas como rastreabilidade (para produtores e jovens), melhoramentos extensivos (para produtores), procriação de gado bovino (para produtores), manejo de vaca prenha (para jovens e produtores), doenças da reprodução (para produtores), pasto natural e ervas daninhas (para técnicos e produtores), gerenciamento da empresa de criação bovina (para técnicos), gestão (para produtores) e bem-estar animal (para técnicos).

Hoje os resultados são tangíveis através de cursos de educação à distância que realiza, anualmente, o planejamento agropecuário como parte da sua oferta de capacitação.

Fonte: <<http://www.plaagropecuario.org.uy/Contenido/53/Proyectos-ejecutados/>>

#### • Apoio e intercâmbio de experiências para o uso de tecnologias de informação na gestão de pequenos e médios produtores de gado

Trata-se de um “projeto de apoio e intercâmbio de experiências para o uso de tecnologias de informação na gestão dos pequenos e médios produtores de gado” que tem como principais objetivos “introduzir os produtores de menor desenvolvimento ao uso de novas Tecnologias de Informação e Comunicação (TIC) por meio da transferência de uma metodologia de fomento desenvolvida no Chile”. Da mesma forma, procura-se “transferir para o Chile a experiência dos produtores de gado do Uruguai em matéria de rastreabilidade e gestão de corte”, assim como “estabelecer laços estáveis e um mecanismo permanente de intercâmbio de experiências entre as instituições chilenas e uruguaias responsáveis pelo desenvolvimento produtivo para a agricultura de pequeno e médio porte”.

O projeto é financiado pela Agência de Cooperação Internacional do Governo do Chile (AGCI), executado pelo Planejamento Agropecuário e conta com a assessoria do Centro para o Desenvolvimento de Capital Humano do Chile (Cendec), que desempenhará o papel de articulador contribuindo com a sua experiência técnica na transferência da metodologia de incorporação de TIC. Teve uma duração de nove meses e começou a ser implantado em janeiro de 2007 por meio de uma primeira experiência piloto com 100 produtores da região leste do Uruguai. Este momento marcou o pontapé inicial da entrada dos produtores de gado uruguaios de baixos recursos escassos à sociedade da informação, elemento considerado fundamental para seu desenvolvimento e da produção pecuária do Uruguai. Os beneficiários foram pequenos criadores extensivos de gado, com renda anual por volta de US \$15 por ha. O projeto contou com uma doação de dinheiro de cerca de cem mil dólares oferecidos pela AGCI (66%) e pelo IPA (34%). (Projeto de Cooperação Horizontal Chile-Uruguai).

### 5.2 Novas tecnologias - evolução tecnológica e o futuro de AgroTICC

#### • Sistema Nacional de Informação Agropecuária (SNIA)

No mês de julho de 2012 concluiu-se o Plano Diretor do Sistema de Informação, Informática e Telecomunicações do Ministério da Agricultura, Pecuária e Pesca (MGAP). Nele foi validada uma série de metas gerais, entre as quais encontra o Desenvolvimento de Informação, permitindo a integração e acesso a dados de solos, cartografia, população animal e vegetal, entre outros. Esta estratégia interna de desenvolvimento de TIC ocorreu em paralelo ao aparecimento de novas leis de transparência, segurança e acesso à informação no país. Estas normas formam parte do esquema de construção do SNIA (BERTERRECTCHE, 2013).

É uma plataforma que procura compartilhar e analisar informação e gerar produtos e informação relevante e acionável, integrando:

- Dados de diferentes instituições.
- Bases de dados com diferentes formatos, projeções etc.
- Requer um enfoque multidisciplinar, dados de muitos diferentes “tipos” (clima, estatísticas, recursos naturais, Censos, amostragens etc).
- Grande dificuldade para unir bases de dados (interoperabilidade).

O SNIA é um conceito que está em permanente construção e evolução, focado em melhorar a gestão do conhecimento, por meio da integração da informação dos recursos naturais, do clima e dos diferentes setores produtivos até chegar à construção de um sistema de informação de produção integrado.

O processo de desenvolvimento do SNIA requer uma forte articulação institucional para permitir a interoperabilidade das bases de dados das organizações e a participação ativa e integrada das pessoas envolvidas nos processos de geração e análise de informação (BERTERRETCHE, 2013).

Com o objetivo de facilitar a interoperabilidade, instalou-se no Uruguai uma ferramenta informática desenvolvida pelo Instituto de Pesquisas sobre o Clima e Sociedade (IRI) da Universidade Columbia. O instrumento, conhecido como “Data Library”, permite organizar e integrar dados provenientes de diversas fontes, assim como processar, analisar e visualizar informação aplicando novas tecnologias e modelos matemáticos e visualizar os resultados em diferentes formatos gráficos.

No convênio celebrado entre o IRI e o Estado uruguaio busca-se desenvolver um conjunto de produtos de natureza diversa tais como o melhoramento das previsões estacionais e o monitoramento do clima e da vegetação, permitindo antecipar e coordenar ações diante da incidência de eventos prejudiciais, como, por exemplo, os déficits e excessos hídricos, realizar previsões de colheita ou avaliar a vulnerabilidade de uma determinada categoria produtiva.

A isto soma-se a contribuição da Universidade da República, do Inia e das demais organizações do setor agropecuário, o SINAE e os demais Ministérios, com o objetivo de fortalecer o relacionamento e o intercâmbio de informação de alto valor.

Devido à interoperabilidade das bases de dados que estão sendo construídas no âmbito do SNIA, atualmente trabalha-se na consolidação de vários projetos de curto prazo. Estes estão associados à melhoria da gestão do risco (alertas imediatos em pecuária, de caráter climático e sanitário, caracterização de risco para o desenvolvimento de seguros para a pecuária e para o setor de aves (frangos), controle de aplicação de agroquímicos, controle de desperdício de efluentes leiteiros e de engorda em confinamento), a integração de registros de produtores, a melhoria de diversos aspectos dos planos de uso e gestão do solo, a análise de dados de ensaios de cultivares e a análise de bacias hidrográficas para a promoção de irrigação. Foram formadas equipes interinstitucionais e interdisciplinares que trabalham na elaboração de tais produtos.

Em nosso país, o SNIA é um sistema de informação incipiente, que tem um enorme potencial em relação aos serviços que pode fornecer, os tipos de usuários que pode atingir e que, por sua modalidade de trabalho colaborativo, interdisciplinar e aberto, seu limite está unicamente condicionado à atitude e à imaginação com as quais o alimentamos.

Criação de mapas de risco agroclimático com base em diferentes fontes de informação. (Figura 8)..

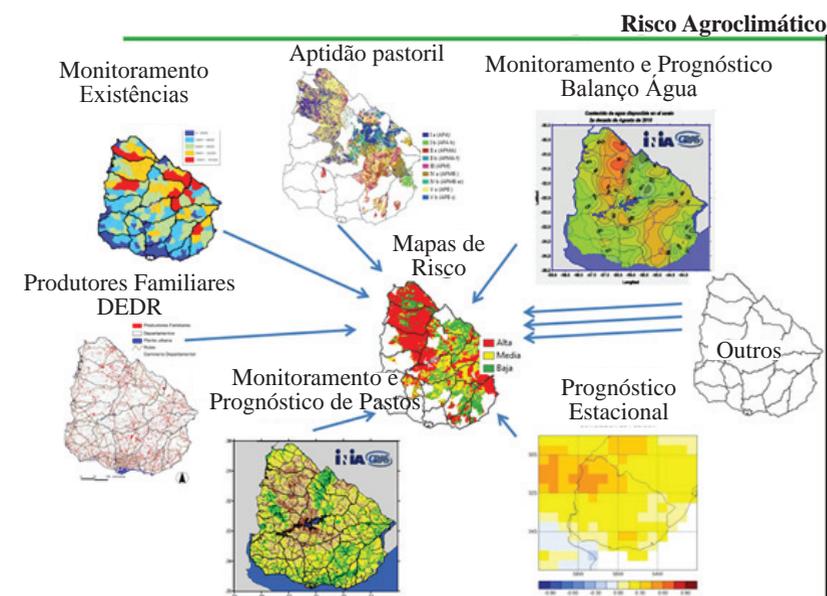


Figura 8. Esquema de geração de mapas de risco agroclimático no SNIA.

## Referencias

- BERTERRETCHE, M. 2013. Sistema Nacional de Información Agropecuaria (SNIA): enfoque, potencialidades y situación actual. En: Anuario OPYPA 2013. Montevideo: MGAP. Pág. 547-551
- DIEGUEZ et al. 2012. Modelización de una explotación ganadera extensiva criadora en basalto, *Agrociencia* 2012, 16 (2): 120-130.
- Genexus, 2014. Online: <http://www.genexus.com/Historias-de-exito/conaprole/conaprole?es>
- HIRIGOYEN A.; SCOZ, R. 2014. Online: <http://ainfo.inia.org.uy/digital/bitstream/item/3430/1/Jornada-28-2014-HIRIGOYEN-SCOZ.pdf>
- RENARE, MGAP. 2014. Online: <http://www.renare.gub.uy/>
- SOARES DE LIMA, J.M. 2009. Modelo bioeconómico para la evaluación del impacto de la genética y otras variables sobre la cadena cárnica vacuna en Uruguay. Tesis Doctorado, Universidad Politécnica de Valencia, Valencia, España. 240 pp.
- Uruguay XXI. 2014. Online: <http://www.uruguayxxi.gub.uy/wp-content/uploads/2012/07/Tecnolog%C3%ADa-de-la-Infomaci%C3%B3n-y-Comunicaciones.pdf>

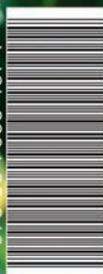




Ministério da  
Agricultura, Pecuária  
e Abastecimento

GOVERNO FEDERAL  
**BRASIL**  
PÁTRIA EDUCADORA

978385-7035-487-7



CGPE 11915