



OpenSLID: Uma ferramenta capaz de tornar mais eficientes os estudos sobre a dinâmica dos campos

César Huegel Richa¹, Leandro B. S. Volk^{2,3}, José Pedro P. Trindade^{2,3}, Leonardo B. Pinho³

¹ Programa de Pós-Graduação em Ciência da Computação, UFSC, Florianópolis-SC, Brasil,
cesar.huegel@posgrad.ufsc.br

²Embrapa Pecuária Sul, Bagé-RS, Brasil,
leandro.volk@embrapa.br, jose.pereira-trindade@embrapa.br

³Programa de Pós-Graduação em Computação Aplicada, UNIPAMPA/Embrapa, Bagé-RS, Brasil,
leonardopinho@unipampa.edu.br

RESUMO

A busca por práticas pecuárias mais eficientes tem impulsionado o desenvolvimento e uso de técnicas de sensoriamento remoto (SR), inclusive, para um melhor entendimento das relações solo-planta-animal, dependentes do conhecimento da dinâmica dos campos em regiões específicas. Apesar de eficazes, métodos convencionais de classificação de cobertura vegetal baseados em SR possuem potencial de incremento na automatização, mas também em relação a eficiência, em função da grande quantidade de imagens e metadados a serem avaliados. Este trabalho adota elementos de pesquisa descritiva para apresentar a ferramenta OpenSLID, desenvolvida em conjunto pela EMBRAPA e pela UNIPAMPA, capaz de agilizar a manipulação de metadados e de realizar, de forma mais eficiente, a busca por imagens de interesse. Após sintetizar os seus requisitos funcionais e os principais aspectos da sua arquitetura, são expostos resultados preliminares do seu uso, tendo como estudo de caso imagens de uma região do Rio Grande do Sul historicamente associada à pecuária, permitindo concluir que aumenta a eficiência do processo de classificação ao maximizar o conjunto de imagens que atendem ao padrão de interesse estabelecido, bem como por evitar que o usuário busque manualmente, de forma repetitiva e suscetível a erro, as imagens do Projeto Landsat.

PALAVRAS-CHAVE: Ciência da Informação, Pecuária de Precisão, Sensoriamento Remoto, Classificação de Cobertura Vegetal, Ferramentas de Software Aberto.

ABSTRACT

The search for more efficient livestock practices has promoted the development and use of remote sensing (RS) techniques, including a better understanding of soil-plant-animal relationships, dependent on the knowledge of the dynamics of the fields in specific regions. Although effective, conventional methods of vegetation coverage classification based on RS have potential of increment in the automation, but also in relation to the efficiency, due to the great amount of images and metadata to be evaluated. This work adopts elements of descriptive research to present the OpenSLID tool, developed jointly by EMBRAPA and UNIPAMPA, able to speed up the manipulation of metadata and to perform, more efficiently, the retrieval of images of interest. After synthesizing its functional requirements and the main aspects of its architecture, preliminary results of its use, taking as a case study images referring to a region of Rio Grande do Sul historically associated with livestock, allow us to conclude that OpenSLID increases the efficiency of the classification process by maximizing the set of images that meet the established interest pattern, as well as by preventing the user from manually retrieving Landsat Project images repetitively and susceptible to error.

KEYWORDS: Information Science, Precision Livestock Farming, Remote Sensing, Vegetal Coverage Classification, Open source Software Tools.

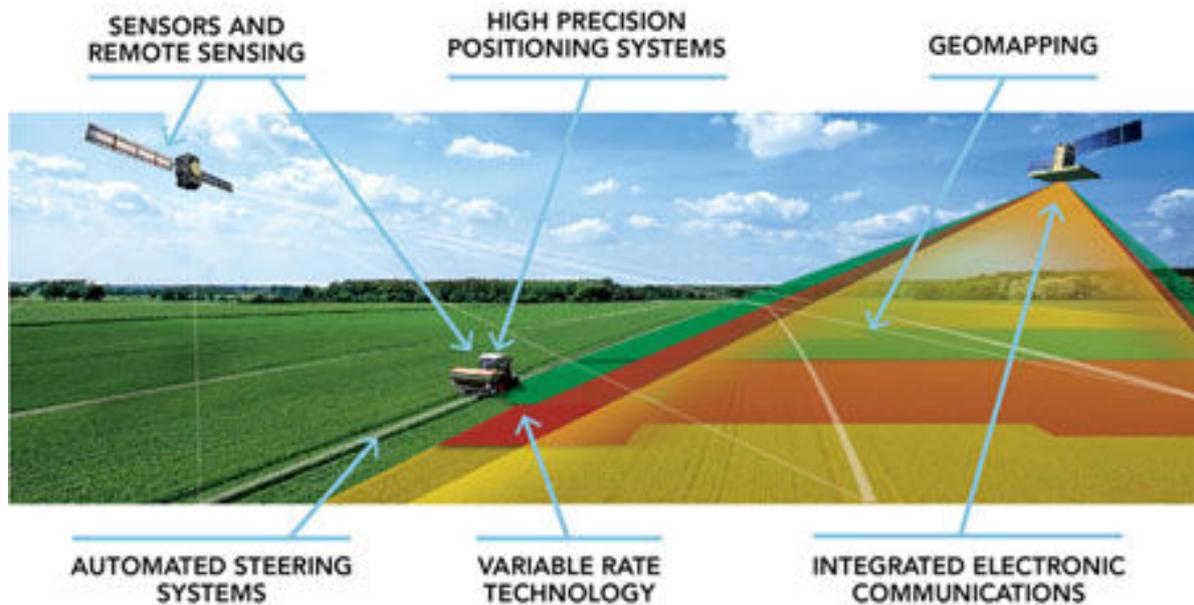
INTRODUÇÃO

Dados atuais apresentados pela CEMA (2017), Associação Europeia que representa a Indústria de Máquinas Agrícolas, apontam que a população mundial cresce em cerca de 240 mil pessoas por dia, podendo chegar a 9,6 bilhões em 2050, o que implica a necessidade de aumentar em 70%, de forma sustentável, a produção de alimentos. Neste contexto, aprimorar as atividades agrícolas é essencial para suprir este significativo crescimento na demanda por alimentos.

Em resposta a essa necessidade, surgiram, na comunidade científica e na indústria, diferentes ações de pesquisa, desenvolvimento e inovação envolvendo a aplicação de novas tecnologias em processos e produtos capazes de aumentar a eficiência dos sistemas de produção agrícola, alinhadas com o que se convencionou denominar por Agricultura de Precisão (AP). A Figura 1 ilustra um cenário de AP que combina conceitos e tecnologias de sensores e Sensoriamento Remoto (SR), sistemas de posicionamento de alta precisão, geoprocessamento, sistemas de colheita automatizada, tecnologias de taxa variável e de comunicações eletrônicas integradas.

Conforme a CEMA (2017), os principais desafios atuais da pecuária são os de (i) atender o crescimento da produção (estimulada pelo aumento no número de pessoas adotando dietas ricas em proteínas), (ii) aliviar a carga de trabalho dos envolvidos no manejo dos animais e assegurar a viabilidade econômica da propriedade rural (por meio de soluções escaláveis e custo-efetivas para o gerenciamento de propriedades rurais), e (iii) promover a sustentabilidade e conforto animal (não apenas aumentar a produção, mas também encontrar maneiras de minimizar o impacto ambiental da pecuária e de assegurar altos níveis de bem-estar e de saúde para os animais).

Figura 1 – Agricultura de Precisão: tecnologias e conceitos-chave.



Fonte: (CEMA, 2017)

Neste contexto, Berckmans (2008) destaca o impacto rápido e crescente da Ciência da Informação e a Tecnologia da Informação nos métodos usados na pecuária, fazendo surgir uma subárea de AP, denominada por Pecuária de Precisão, ou *Precision Livestock Farming* (PLF). Na PLF são aplicados princípios de engenharia de controle na otimização da produção e dos processos de gerenciamento amparados por transferência de informação eletrônica. Assim, técnicas potencialmente inovadoras, simples e de baixo custo podem ser empregadas para coleta de dados dos animais e do ambiente, seguidas pela avaliação destes por modelos computacionais baseados em conhecimento. Para Melo, Silva e Guimarães (2013), existem diversas ferramentas que podem ser utilizadas para AP, as quais combinam sensores, sistemas de informação, máquinas adaptadas e gestão de conhecimento. Entre elas, destaca-se o SR, utilizado para obter informações de objetos sem entrar em contato com os mesmos, sendo capaz de detectar e armazenar radiação eletromagnética em determinada faixa do espectro eletromagnético (intervalo com todas as possíveis frequências da radiação eletromagnética), e ainda gerar informações que possam ser interpretadas, seja na forma de imagem, gráfica ou tabelas. Em função das suas características, o SR se apresenta como possível solução para a gestão de um dos elementos principais das relações solo-planta-animal que se estabelecem em um sistema de produção pecuária: os campos.

Entender a dinâmica dos campos na paisagem pode ser estratégica para diferentes objetivos, podendo impactar na definição de políticas públicas, em ações duráveis de manejo para a conservação da pecuária de campo e/ou intensificação sustentável da produção animal, bem como na definição de práticas de manejo que possam inverter o atual processo de conversão dos Campos (OVERBECK et al., 2007). O conhecimento sobre a dinâmica dos campos em regiões específicas mostra-se fundamental para as práticas de extensão rural, permitindo que se

estabeleça um processo de educação informal que objetive o uso racional dos recursos naturais e a prestação da assistência técnica pertinente, para a consolidação das práticas conservacionistas que garantirão a continuidade e a sustentabilidade do processo produtivo da propriedade (DIAS et al., 1999). Como exemplo, um método baseado em SR foi usado por Trindade et al. (2016), para descrever, a partir de imagens de satélite disponibilizadas gratuitamente pelo Projeto Landsat, a dinâmica dos Campos – das regiões da Campanha, Fronteira Oeste e Missões do Rio Grande do Sul – com potencial de utilização pastoril na Produção Pecuária, independente de seu estado de conservação, desde a década de 1970 até o Outono de 2015. Com este método, foi possível evidenciar que a cobertura vegetal campestre nesta área poderá ser fortemente reduzida em se mantendo as mesmas condições que levaram ao padrão de substituição da cobertura vegetal campestre nas três regiões.

Embora seja eficaz, o método baseado em SR usado por Trindade et al. (2016) para o estudo da dinâmica dos campos envolve diversas etapas manuais e repetitivas, com potencial de aprimoramento. Considerando-se a grande quantidade de imagens a serem processadas e a necessidade de um estudo contínuo, a automatização do método de classificação, ou de certas etapas deste, foi identificada como uma característica desejável e relevante para tornar o método mais eficiente. Neste contexto, este trabalho apresentou a ferramenta *Open source Smart Landsat Imagery Downloader* (OpenSLID), desenvolvida como parte das ações de cooperação entre a EMBRAPA e a UNIPAMPA, capaz de agilizar a manipulação de metadados e de realizar, de forma mais eficiente, a busca por imagens de interesse.

O restante do trabalho está dividido em três seções: na seção de Material e Métodos é apresentado o material utilizado para a condução do trabalho, bem como a metodologia empregada; na seção de Resultados e Discussão são relatados os resultados obtidos, discutidos em relação a outros trabalhos e resultados associados; e na seção de Conclusões são sintetizadas as principais conclusões.

MATERIAL E MÉTODOS

Este trabalho faz uso de elementos de pesquisa descritiva para apresentar a ferramenta de *software* OpenSLID, resultante de um projeto amparado por atividades de pesquisa exploratória, realizadas por uma equipe interdisciplinar, sob a coordenação de pesquisadores das áreas de Agrárias da EMBRAPA e de Computação da UNIPAMPA. Esta foi desenvolvida tendo como principal requisito tornar mais eficiente, com base em ferramentas *open source*, o método de classificação da cobertura de vegetação baseado em SR proposto por Trindade et al. (2016). Após um estudo inicial sobre o método e as oportunidades para torná-lo mais eficiente, optou-se por concentrar o processo de desenvolvimento em requisitos relacionados à pesquisa, *download* e pré-processamento de imagens de satélite Landsat.

Com a ferramenta OpenSLID, a pesquisa por imagens pode ser realizada a partir de diferentes filtros, o que possibilita a recuperação de cenas que contemplem parâmetros específicos

que as classifiquem como de interesse. O *download* de cenas pode ser realizado, tanto por meio de filtros quanto por meio de uma lista contendo os identificadores das cenas de interesse. Já o pré-processamento de imagens deve ser realizado a partir de cenas obtidas previamente e que estejam disponíveis na árvore de subdiretórios da ferramenta.

A utilização desta ferramenta pode ocorrer de forma supervisionada ou automatizada. No caso de supervisionada, o utilizador interage com a ferramenta realizando a pesquisa, o *download* e o pré-processamento. No caso de automatizada os utilizadores podem definir um catálogo, através de um arquivo de entrada, contendo *datasets*, para os quais deverão constar filtros de pesquisa e parâmetros de *download* e pré-processamento. Além da automatização dos processos supracitados, existe o conceito de execução programada, que permite manter um banco de imagens de interesse atualizado de acordo com o catálogo de *datasets*. O sistema em questão possui como principais funcionalidades a pesquisa e o *download* de imagens de satélite Landsat. Adicionalmente, o sistema provê um módulo de pré-processamento, onde diferentes conjuntos de operações sobre imagens de satélite podem ser implementadas de acordo com a necessidade da aplicação. É válido ressaltar que o sistema não se propõe a realizar processamento de imagens de aplicações específicas, tão pouco a disponibilizar ferramentas de visualização e edição das imagens de satélite.

Dentre os sistemas relacionados estão: QGIS (antigo Quantum GIS), que provê funcionalidades, através de sua API (*Application Programming Interface*), para implementação do conjunto de operações de pré-processamento de imagens; *Landsat Bulk Metadata Service*, serviço de disponibilização dos metadados referentes ao inventário Landsat disponibilizado pela *United States Geological Survey* (USGS). Existem, pelo menos, outras duas ferramentas similares, Landsat Util (2017) e Landsat-Download (2017), ambas open source, desenvolvidas para busca e *download* de imagens de satélite, no entanto não possuem grande parte das características da OpenSLID, cujo desenvolvimento teve partes inspiradas nestas.

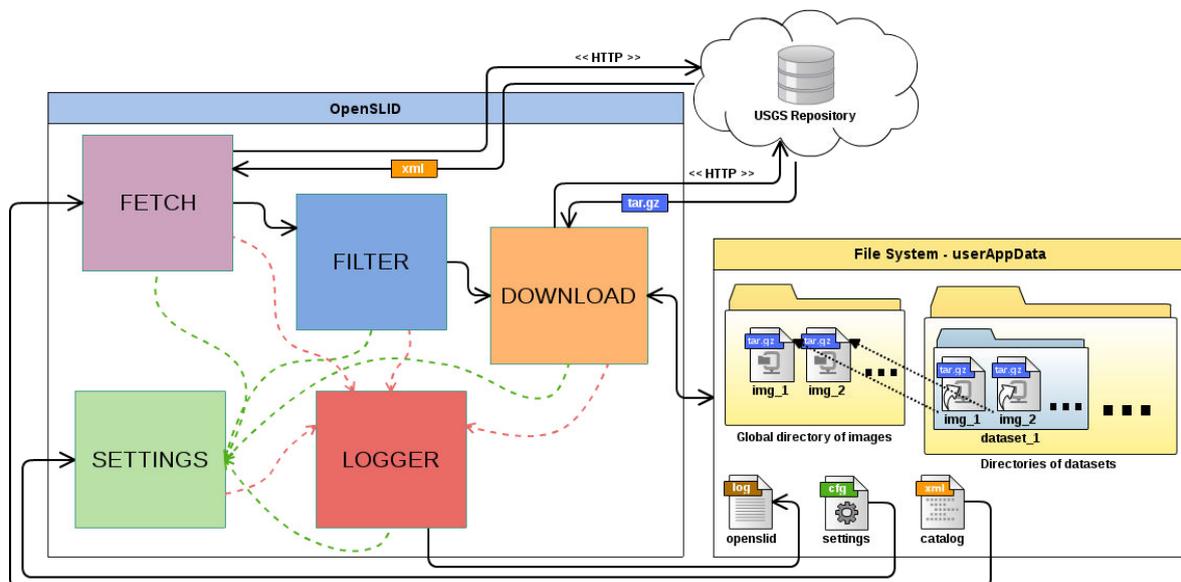
OpenSLID é destinada a usuários que utilizam massivamente imagens de satélite em seus projetos, principalmente quando existe a necessidade de análise temporal destas. Para estes usuários, em especial, a busca por imagens que satisfaçam determinados requisitos para utilização é uma tarefa custosa, pois demanda tempo, além da dificuldade de se avaliar os inúmeros metadados vinculados às inúmeras cenas de interesse e o tempo dispendido para *download* das imagens. Outro fator que intensifica a demanda por soluções que automatizem este processo é a necessidade de manter uma base de imagens de interesse atualizada, ou seja, a necessidade de realizar pesquisas periódicas ao repositório de imagens. A ferramenta viabiliza o uso compartilhado da base de imagens entre diferentes equipes e projetos por meio de um esquema de sincronização entre estações de trabalho e a base de imagens.

RESULTADOS E DISCUSSÃO

O protótipo da solução *open source* desenvolvida, OpenSLID, contempla atualmente a primeira etapa do processo de classificação de imagens que compreende a entrada (aquisição de imagens de satélite). Além de automatizar esta etapa, torna-a mais eficiente em termos de seus resultados (descoberta de imagens úteis para o estudo).

A arquitetura da solução desenvolvida, ilustrada na Figura 2, está dividida em cinco blocos de funcionalidades específicas: **FETCH**, responsável por realizar a leitura do arquivo “catalog.xml” (detalhado a seguir) e realizar a busca por metadados no repositório da USGS de acordo com os datasets e parâmetros de filtragem definidos no catálogo; **FILTER**, responsável por aplicar filtros específicos aos metadados resultantes da consulta ao repositório da USGS; **DOWNLOAD**, responsável por realizar o download das imagens que foram aceitas pelo filtro, salvá-las no diretório de imagens global e criar links simbólicos no diretório de imagens do dataset correspondente; **SETTINGS**, responsável por carregar as configurações setadas no arquivo “setting.cfg” e definir outras configurações padrões da ferramenta; **LOGGER**, responsável por gerenciar os níveis de logs (*debug, info, warning, error e critical*).

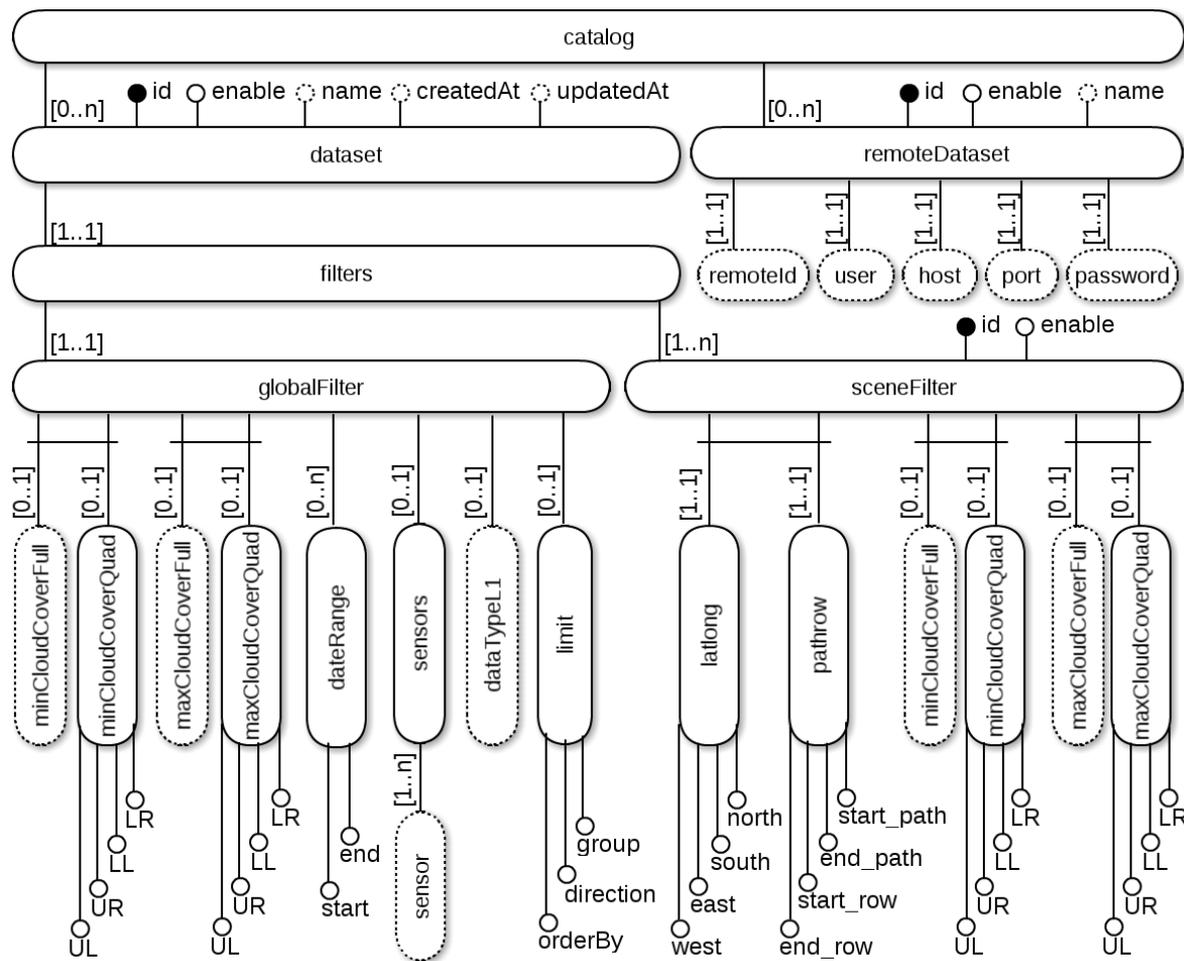
Figura 2 – Diagrama geral da arquitetura da solução.



Fonte: Próprios autores

A entrada da ferramenta consiste em um arquivo XML contendo um catálogo de datasets, onde estão definidos filtros específicos. Cada dataset corresponde a um agrupamento de cenas definidas por meio destes filtros, seguindo um modelo lógico hierárquico em XML que permite a representação de diferentes características referentes a estrutura e restrições. Dentre estas características estão: elementos simples e complexos; atributos simples; atributos identificadores; restrições de composição e cardinalidade entre elementos pai-filho; e ocorrência de disjunção entre elementos. A Figura 3 apresenta o modelo lógico XML destes dados e permite compreender seu nível de detalhamento.

Figura 3 – Modelo lógico XML.



Fonte: Próprios autores

Descreve-se a seguir os principais elementos do modelo XML:

- **catalog**: elemento raiz do modelo. É um elemento complexo que pode conter até n ocorrências do elemento “dataset”;
- **dataset**: representa um conjunto de imagens de trabalho definidas por meio de filtros. Também é um elemento complexo e é composto por um único elemento “filters”. Além disso, este elemento possui os seguintes atributos: “id”, atributo único e obrigatório que identifica o elemento em relação aos demais; “enable”, atributo obrigatório que indica se o *dataset* está ativo ou não; “name”, atributo opcional arbitrário para determinar um nome ao *dataset*; “createdAt”, atributo opcional que indica a data de criação do *dataset*; “updatedAt”, atributo opcional que indica a data da última atualização das definições do *dataset*;
- **filters**: elemento complexo que agrupa os filtros de um *dataset*. Este é composto por um único filtro global (*globalFilter*) e pelo menos um filtro de cena (*sceneFilter*). Foi criado pois o modelo será expandido e novas informações serão agregadas ao *dataset*;

- **globalFilter**: elemento complexo onde são definidos parâmetros relacionados à filtragem global das imagens de um *dataset*. Sua composição é dada em função das seguintes restrições: pode conter um único elemento de mínima cobertura de nuvens (minCloudCoverFull ou minCloudCoverQuad); pode conter um único elemento de máxima cobertura de nuvens (maxCloudCoverFull ou minCloudCoverQuad); pode conter um único elemento simples de data inicial (startDate); pode conter um único elemento simples de data final (endDate); pode conter um conjunto de filtro para sensores (sensors); e um elemento filtro simples para o tipo de dado (dataTypeL1);
- **sceneFilter**: elemento complexo onde são definidos parâmetros relacionados à filtragem de cena. Sua composição é dada em função das seguintes restrições: deve conter um elemento relacionado ao posicionamento geográfico das imagens (latlong ou pathrow); pode conter um único elemento para de mínima cobertura de nuvens (minCloudCoverFull ou minCloudCoverQuad); pode conter um único elemento de máxima cobertura de nuvens (maxCloudCoverFull ou minCloudCoverQuad);
- **latlong**: elemento complexo que indica, por meio de seus atributos obrigatórios (north, south, east, west), um retângulo cuja função é delimitar as imagens cobertas pelo respectivo filtro para compor as imagens de um *dataset*;
- **pathrow**: elemento complexo que indica, a partir de atributos obrigatórios (start_path, end_path, start_row, end_row), um conjunto de imagens por meio de suas referências de path e row.
- **min,maxCloudCoverFull**: elementos simples que contém, respectivamente, um valor percentual mínimo e máximo de cobertura de nuvens em toda a imagem;
- **min,maxCloudCoverQuad**: elementos complexos que possuem os seguintes atributos: UL (upper left), indica o quadrante superior esquerdo; UR (upper right), indica o quadrante superior direito; LL (lower left), indica o quadrante inferior esquerdo; e LR (lower right), indica o quadrante inferior direito. Estes dois elementos indicam, respectivamente, um valor percentual mínimo e máximo de cobertura de nuvens em cada quadrante da imagem;
- **remoteDataset**: representa um *dataset* remoto, localizado em outra máquina (e.g. base compartilhada de imagens). É um elemento complexo composto por um elemento que identifica o *dataset* remoto e outros cujos dados são utilizados para efetuar a sincronização via protocolo *Secure Copy* (SCP). Seus atributos seguem o mesmo padrão do elemento “dataset”.

Cabe destacar que o formato XML foi escolhido com base nas seguintes características: apresentar alto nível de padronização; ser baseado em texto simples, o que o torna de fácil leitura e edição; permitir validação da sua estrutura e de seus dados por meio de um XML Schema. Esta validação ocorre no início da execução do módulo FETCH, garantindo sua estrutura correta e a consistência dos dados contidos, já que foi definido em conformidade com o mesmo XML Schema dos metadados da USGS.

A principal contribuição da ferramenta está na precisão dos filtros (e.g. filtro de cobertura de nuvem por quadrante) aplicados aos metadados, maximizando o conjunto de imagens úteis resultante para um determinado estudo. Além disso, a ferramenta permite a definição de um *dataset* remoto, o que permite construir, dentro de uma organização, um servidor de imagens, contendo diferentes *datasets* para diferentes estudos, eventualmente de equipes distintas, mas otimizando a utilização de espaço de armazenamento e da largura de banda de rede com a execução agendada em horários pré-determinados e um único *download* para cenas coincidentes em diferentes *datasets*.

Tabela 1 – Levantamento de imagens Landsat 7 (CC <10% e tipo de correção L1T).

Path/Row	Período							
	1999-2000	2001-2002	2003-2004	2005-2006	2007-2008	2009-2010	2011-2012	2013-2014
223/81	10 / 15	7 / 11	0 / 6	0 / 7	1 / 3	2 / 3	0 / 0	0 / 0
223/82	10 / 16	4 / 8	1 / 4	3 / 5	2 / 4	4 / 5	1 / 1	0 / 0
224/80	2 / 10	5 / 12	8 / 9	2 / 3	2 / 3	3 / 6	2 / 3	0 / 0
224/81	3 / 9	3 / 9	1 / 10	1 / 2	0 / 2	5 / 6	2 / 3	0 / 0
224/82	3 / 7	1 / 6	1 / 6	2 / 4	2 / 1	4 / 6	1 / 2	0 / 0
225/80	2 / 12	5 / 12	1 / 10	1 / 6	2 / 2	5 / 7	0 / 2	0 / 0
225/81	7 / 11	4 / 8	0 / 8	1 / 6	2 / 2	6 / 6	0 / 2	0 / 0

A Tabela 1 apresenta o número de imagens úteis identificadas manualmente e por meio da ferramenta, tendo como filtros a data, o *Path/Row* da cena, o percentual de cobertura de nuvem e o tipo de correção. Conforme pode ser observado nesta tabela, no período 1999-00 foram identificadas 37 imagens manualmente, enquanto que a ferramenta permitiu o uso de 80 imagens. Ou seja, neste período, o método manual usado originalmente descartava 43 imagens que poderiam ser usadas nas etapas posteriores no método de classificação. De forma geral, pode-se perceber que, com exceção de 2013-14 e de um possível erro humano identificado no período 2007-2008 para *Path/Row* 224/82, a quantidade de imagens que deixaram de ser descartadas foi significativa a partir do emprego da ferramenta desenvolvida.

Complementarmente, notou-se que o uso de XML foi um facilitador, prevenindo erros e simplificando a sua identificação por parte dos usuários. Isso se deve à existência do arquivo de validação, conforme o padrão XSD (*XML Schema Definition*), o qual descreve a estrutura do arquivo XML determinando quais elementos são obrigatórios, qual o número mínimo e máximo de ocorrências de determinados elementos, tipos de dados, dentre outras definições.

Por fim, destaca-se que o método usado por Trindade et al. (2016) é um caso de aplicação da ferramenta. Outras aplicações de SR que envolvam pesquisa, *download* e pré-processamento de imagens Landsat podem igualmente se beneficiar de suas características.

CONCLUSÕES

O emprego crescente da Ciência da Informação em soluções de Pecuária de Precisão se mostra cada vez mais promissor, contribuindo para o entendimento dos processos que interferem nas relações solo-planta-animal. Neste contexto o estudo da dinâmica dos Campos se mostra

fundamental, podendo este ser realizado a partir de métodos de Sensoriamento Remoto que envolvem a análise de grandes quantidades de imagens de satélite, passíveis de se tornarem mais eficientes com o uso de uma ferramenta computacional. A ferramenta OpenSLID, apresentada neste trabalho, demonstrou ser capaz de maximizar os resultados em termos do número de imagens úteis para o estudo, evitando erros humanos envolvidos no processo e otimizando o uso de espaço de armazenamento e da largura de banda de rede.

REFERÊNCIAS

BERCKMANS, D. Precision livestock farming (PLF). *Computers and Electronics in Agriculture*, v. 62, n. 1, p. 1, 2008.

CEMA. *Precision Farming: key technologies & concepts*. 2017. Acesso em: 11 mai. 2017. Disponível em: <<http://www.cema-agri.org/page/precision-farming-key-technologies-concepts>>.

DIAS, M. et al. *Manual de impactos ambientais: orientações básicas sobre aspectos ambientais de atividades produtivas*. Fortaleza, CE: Banco do Nordeste, 1999.

Landsat-Download. *Automated download of LANDSAT data from USGS website*. 2017. Acesso em: 11 mai. 2017. Disponível em: <<https://github.com/olivierhagolle/LANDSAT-Download>>.

Landsat Util. *A utility to search, download and process Landsat 8 satellite imagery*. 2017. Acesso em: 11 mai. 2017. Disponível em: <<https://github.com/developmentseed/landsat-util>>.

MELO, G. de A.; SILVA, S. F. da; GUIMARÃES, A. M. Perspectivas de utilização do sensoriamento remoto na agricultura de precisão. In: *IX Congresso Brasileiro De Agroinformática*. Cuiabá, MT: [s.n.], 2013. Acesso em: 11 mai. 2017. Disponível em: <http://200.129.241.80/sbiagro/sbianais/paginas/trabalhos/118801_1.pdf>.

OVERBECK, G. E. et al. Brazil's neglected biome: the South Brazilian Campos. *Perspectives in Plant Ecology, Evolution and Systematics*, Elsevier, v. 9, n. 2, p. 101–116, 2007.

TRINDADE, J. P. P. et al. Dinâmica dos Campos da Campanha, Fronteira Oeste e Missões do Rio Grande do Sul. In: *XXVI Congresso Brasileiro de Zootecnia (ZOOTEC 2016)*. Santa Maria, RS: [s.n.], 2016. v. 26.