

Geotecnologias e redes neurais artificiais para o mapeamento da cultura de cana-de-açúcar por meio de serie temporal de EVI no estado do Paraná

Rennan Andres Paloschi¹
Clóvis Cechim Júnior¹
Jerry Adriani Johann¹

¹ Universidade Estadual do Oeste do Paraná - UNIOESTE
Caixa Postal 0701- 85819 -110 - Cascavel - PR, Brasil
rennan_paloschi@yahoo.com, {juniorcechim, jerry.johann}@hotmail.com

Abstract. The Brazilian agricultural sector has been marked by a new cycle of expansion of sugarcane cultivation. Currently, Brazil is the largest sugarcane producer and the state of Paraná is the fifth largest national producer. Therefore, a reliable estimation of sugarcane crop area is essential for the sugar-ethanol agribusiness. Geotechnologies has been employed on agricultural crop estimates to reduce the level of subjectivity of the traditional methods. A methodology that takes into account a multitemporal multispectral analysis was applied. This paper aims to perform the mapping the cultivation of sugarcane in the state of Paraná, southern Brazil, for the years 2012/2013 through classification with artificial neural network in a time series of images from the EVI AQUA and TERRA satellites by the MODIS (Moderate Resolution Imaging Spectroradiometer) sensor. The time series was smoothed with the filter Flat Smoother Series, was classified as artificial neural network and its classification had filtered with Sieve series and Clump series. Samples for classification based on Landsat-8 images with the aid of Thiessen polygons have been collected. In order to assess the quality of classification, validation accuracy for classification performance was accomplished with satellite images Landsat-8, obtaining a global accuracy of 89.5% and a kappa index of 0.79.

Palavras-chaves: Sensoriamento remoto, rede neural artificial, análise multitemporal, multispectral, MODIS, remote sensing, artificial neural network, multitemporal multispectral analysis.

1. Introdução

Os procedimentos para geração de estimativas de safra agrícola, assim como o conhecimento da sua distribuição espacial, constituem uma importante informação para o setor da agricultura. O setor agrícola brasileiro tem sido marcado por um novo ciclo de expansão do cultivo da cana-de-açúcar. O Brasil por ser o maior produtor de cana-de-açúcar do mundo, necessita de uma estimativa confiável de área plantada da cana-de-açúcar para fomentar o agronegócio sucroalcooleiro (Antunes, 2014).

O Paraná foi em 2013 o quinto maior produtor de cana-de-açúcar segundo a Connab (2013) o que mostra o a expansão da cultura no estado bem como sua importância.

A utilização de recursos computacionais para o mapeamento de culturas auxilia a tomada de decisão e possibilita avaliar grande quantidade de informações a respeito de área e expansão de culturas, permitindo a estruturação de estratégias mais adequadas para o aumento da produtividade ou até mesmo previsão de safras, sendo o custo de implantação um dos mais decisivos fatores na implantação de uma metodologia.

Neste sentido, as geotecnologias têm sido empregadas nas estimativas de safras agrícolas para diminuir o nível de subjetividade dos métodos tradicionais. As técnicas de sensoriamento tem se mostrado eficientes no monitoramento da cobertura vegetal, evidenciando as variações temporais, detecção de mudanças, dinâmica sazonal e fenologia. Nesse propósito, o sensor *Moderate Resolution Imaging Spectroradiometer* (Modis) é um espectrorradiômetro com 36 bandas, abrangendo as radiações do visível ao infravermelho, que permite a obtenção de dados relacionados à vegetação. Dentre os produtos disponibilizados gratuitamente para os satélites TERRA e AQUA, os índices de vegetação, presentes nos MOD13Q1.5 e MYD13Q1.5, respectivamente, permitem obter-se uma série temporal de imagens com resolução temporal de 8 dias e espacial de 250m (National Aeronautics and Space Administration, 2009).

Esse contexto da margem à exploração de recursos computacionais avançados, que extraíam informações de forma eficiente e mais precisa possível, como a utilização de Redes Neurais

Artificiais (RNA) como método de classificação supervisionada com o uso de séries espectro-temporais de imagens de índices de vegetação.

A RNA é uma abstração da rede neural real, observada no cérebro orgânico, e é vista como um modelo paramétrico não-linear. Segundo Santos et. al (2005) a ideia básica do paradigma RNA é construir um modelo composto por um grande número de unidades simples de processamento, chamadas de neurônios, com um grande número de conexões entre eles. A informação dos neurônios é então transmitida através de conexões denominadas sinapses ou pesos sinápticos.

A técnica de RNA pode ser resumida em duas partes: a arquitetura e o algoritmo de aprendizagem, treinamento. O treinamento da rede parte do princípio que o conhecimento sobre o problema está guardado dentro dos exemplos essencialmente disponíveis e apresentados a rede. O algoritmo de aprendizagem cria uma generalização dos dados e memoriza o conhecimento dentro dos parâmetros adaptáveis da rede, os pesos.

Santos et. al (2005) diz que a capacidade das RNA vai muito além de mapear entradas e saídas, é possível extrair informação das mesmas obtendo respostas para dados não apresentados anteriormente, porém essa informação não é fornecida de forma explícita, sendo necessária uma complexa análise dos pesos sinápticos.

Nesse contexto, o objetivo deste trabalho foi utilizar técnicas de processamento digital de imagens do índice de vegetação EVI do sensor MODIS, com redes neurais artificiais para a classificação e consequente mapeamento da cultura de cana-de-açúcar no Estado do Paraná, para o ano-safra 2012/2013.

2. Materiais e Métodos

A área de estudo compreende o estado do Paraná, na região Sul do Brasil, situado entre os paralelos 22° 29' S e 26° 43' S e os meridianos 48° 2' W e 54° 38' W (Figura 1).

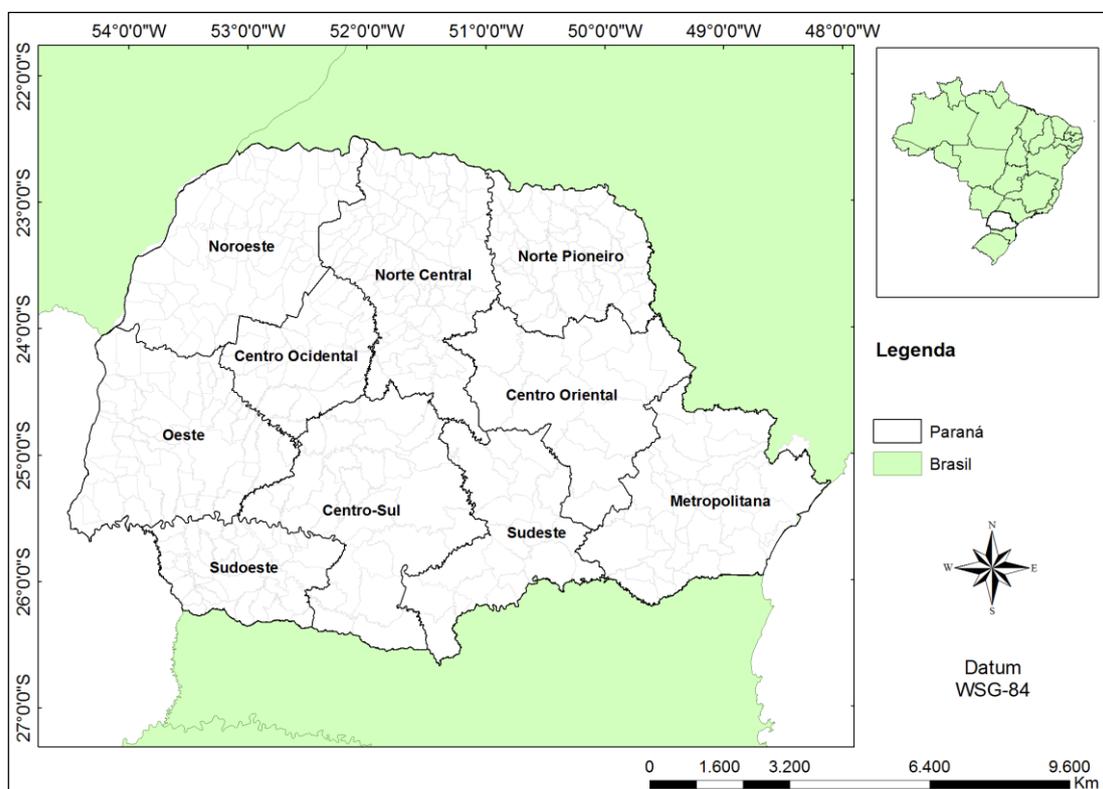


Figura 1: Localização da área de estudo, Paraná, Brasil.

Para este trabalho foram usados os dados do Banco de Produtos MODIS na base Estadual Brasileira (Embrapa Informática Agropecuária, 2014), responsável por armazenar e disponibilizar

na Internet imagens para uso, sem que seja necessário implementar processamentos complementares. Estas imagens são composições de EVI referentes ao período de 01/12/2012 à 27/12/2013, com resolução temporal de 8 dias (em alternância entre TERRA e AQUA, cada qual com resolução de 16 dias).

Para uma melhor análise do comportamento espectral, um cubo temporal de 92 imagens foi organizado (período de dois anos). Nesta série o filtro *Flat smoother series* foi adotado para suavizar ruídos provenientes de nuvens (picos em quedas bruscas no índice de EVI). Este filtro ajusta picos negativos no espectro, “fletando” o nível de queda no pico negativo pelo valor mínimo observado.

Para treinamento da rede neural foram mapeadas amostras de cana-de-açúcar e outros tipos de alvos que não eram de cana. O mapeamento realizou-se a partir de análise visual de imagens Landsat-8 em composição colorida de falsa cor, RGB564 (que evidencia a vegetação).

A fim de manter a resolução das amostras compatíveis com um pixel puro (sem mistura espectral de alvos) do MODIS (250 m), a partir do perímetro de 7 municípios (Douradina, Iporã, Mamboré, Mandaguaçu, Maringá, Ortigueira e Santo Antônio da Platina) distribuídos geograficamente pela região de cana do estado, fornecendo assim uma boa representação das amostras ao longo do estado, e de uma grade de pontos de cada pixel MODIS, foram construídos polígonos de Thiessen (Brassel e Reif, 1979). Neste trabalho polígonos de Thiessen, construídos no software R, foram utilizados com a finalidade de auxiliar a coleta de amostras, sendo eles sobrepostos ao mosaico Landsat-8, o que permitiu a coleta de pixels MODIS puros (amostras) em cada alvo (Figura 2).

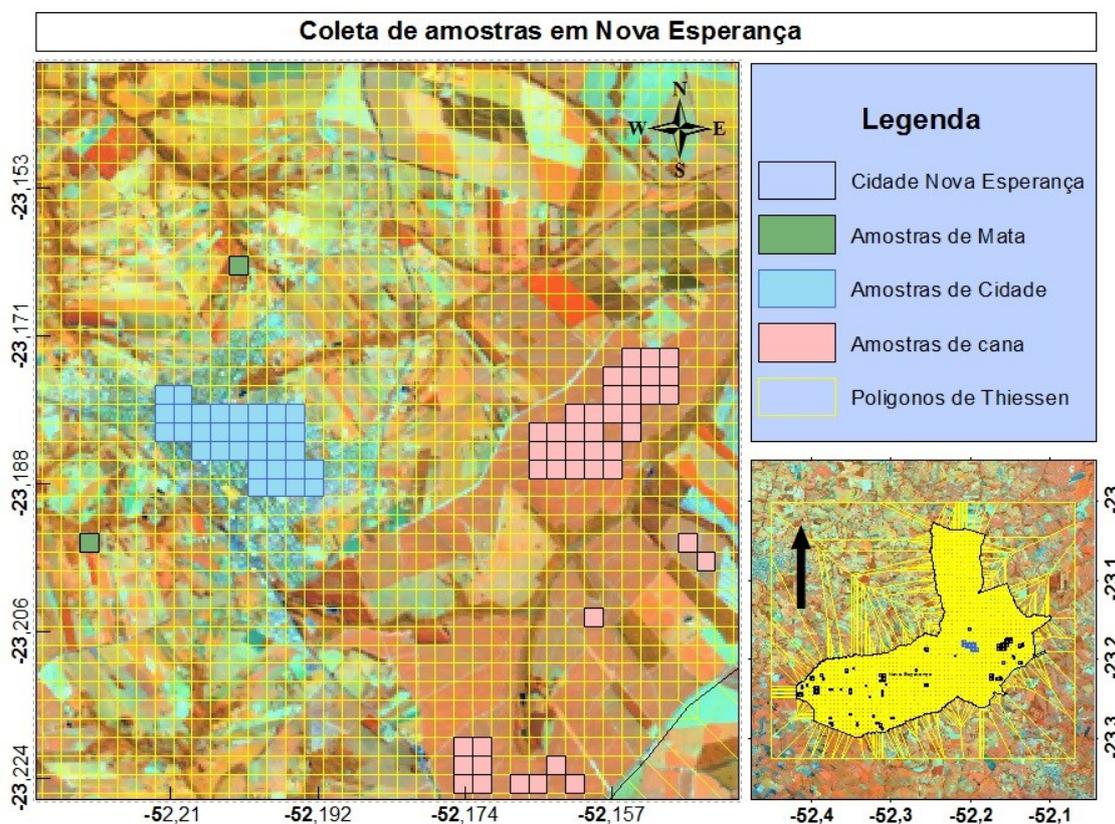


Figura 2: Coleta de amostras com polígonos de Thiessen no município de Nova Esperança-PR.

A rede neural artificial, construída no software ENVI, foi estruturada com 4 camadas de neurônios ocultas (*hidden layers*) (Figura 3), submetida ao treinamento com funções hiperbólicas

realizado pelo algoritmo *backpropagation*, algoritmo *default* do ENVI. Esse tipo de classificação mostrou bons resultados nos trabalhos efetuados por Braga et al. (2007) e Arruda et al. (2013).

A classificação da série temporal ocorreu pixel a pixel, automaticamente após o término do seu treinamento (Figura 3), considerando um buffer com raio de 50 km em volta das 32 usinas de cana-de-açúcar presentes no estado (Trentin et al., 2011), visualizado na Figura 1.

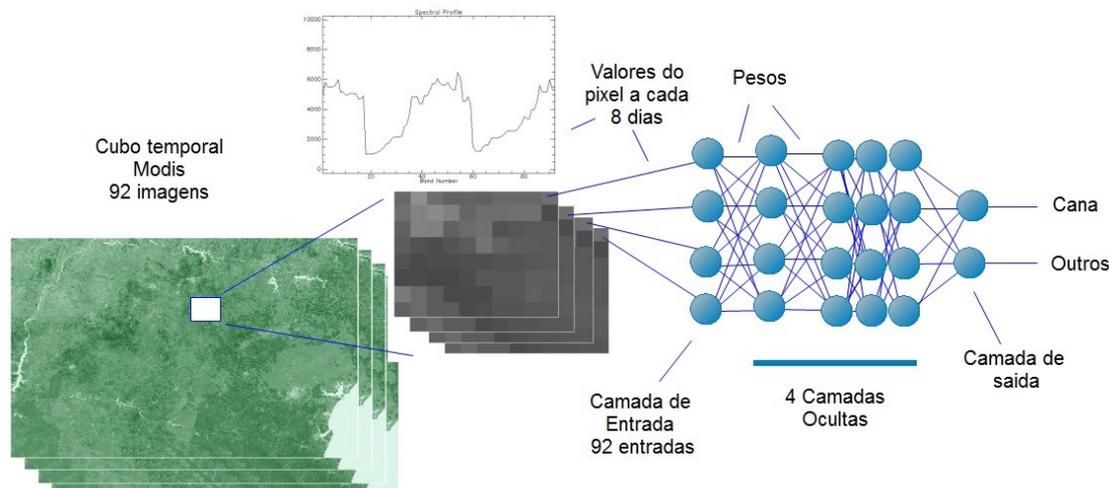


Figura 3: Estruturação da rede neural artificial para mapeamento de cana-de-açúcar no Paraná

Após a classificação foi utilizado o filtro *Sieve classes* removendo pixels isolados e o *Clump classes*, aumentando a definição dos talhões mapeados (Exelis Visual Information Solutions 2014).

Por meio de exclusão de áreas definidas como alvos permanentes (rios, cidades, etc.) foram subtraídas da classificação assim como em Trentin et al. (2011).

Para a validação do mapeamento gerado por meio de exatidão espacial (acurácia), foram usados o índice de Exatidão Global (EG), o índice de concordância Kappa (IK) (Congalton, 1991; Congalton e Green, 1999). Dessa forma para a obtenção do erro amostral foram sorteados de forma aleatória 400 pontos, sendo que destes 200 foram sorteados sobre a cana-de-açúcar, e os outros 200 foram sorteados sobre o alvo não cultura, se tratando de uma amostragem aleatória estratificada.

3. Resultados e Discussão

Para treinamento da rede neural foram mapeadas amostras de diferentes tipos de alvos, de forma que esta pudesse diferenciar os comportamentos e dessa forma diminuir erros de inclusão. Tal mapeamento de amostras foi feito com análise visual utilizando imagens de satélite Landsat-8 composição RGB564, de forma a evidenciar a vegetação facilitando a coleta de amostras.

Além de cana-de-açúcar, foram coletadas amostras de Cidade, Floresta/Mata, e outras culturas, que possuíam comportamentos diferentes da cana-de-açúcar, para aprendizado da rede neural artificial de forma que esta não confundisse tais alvos. À medida que conflitos entre os alvos foram visualmente identificados, novas amostras foram coletadas aumentando o conhecimento da rede e consequentemente sua capacidade de classificação. Como o comportamento espectral dos alvos pode apresentar padrões complexos e/ou diferenças sutis, estes não foram analisados, deixando a cargo do algoritmo de treinamento identificar tais diferenças e relações.

Diversas redes neurais artificiais (construídas no software ENVI), com diversos parâmetros, foram testadas. A rede de melhor acurácia foi estruturada com 4 camadas de neurônios ocultas (*hidden layers*), para que pudesse aprender o máximo possível do comportamento espectral dos alvos sem elevar muito a necessidade de recursos computacionais ou tempo de processamento. Ela foi submetida ao treinamento das amostras à um *training rate* de 0,0005 com funções hiperbólicas, que apresentou melhor resposta, de forma que houve o mínimo oscilações possíveis durante o

treinamento, realizado pelo algoritmo *backpropagation*, com condição de parada para o treinamento de um erro mínimo de 0,7, excluindo erros de inclusão sem perder a capacidade de generalização da rede.

Como a classificação ocorre pixel a pixel, sem considerar o comportamento do pixel vizinho, pixels isolados que não correspondem a área de cana-de-açúcar podem apresentar comportamentos parecidos, assim como, um determinado pixel no interior da cultura pode apresentar um comportamento atípico. Aplicando os filtros *Sieve classes* e *Clump classes*, foi observado diminuição nos pontos de inclusão, e maior definição nos talhões mapeados, já que pixels isolados foram removidos e talhões onde se tinha apenas um pixel não classificado em seu interior foi considerado como cultura de cana.

Foi então subtraída da classificação de áreas identificadas como alvos permanentes (cidades, rios, matas). Porém verificou-se que as áreas de conflito foram mínimas, ou mesmo inexistentes na maior parte da classificação. As maiores divergências foram observadas na classificação da cultura de cana-de-açúcar com áreas de pastagem devido ao comportamento espectral semelhante. O resultado final do mapeamento pode ser observado na Figura 4.

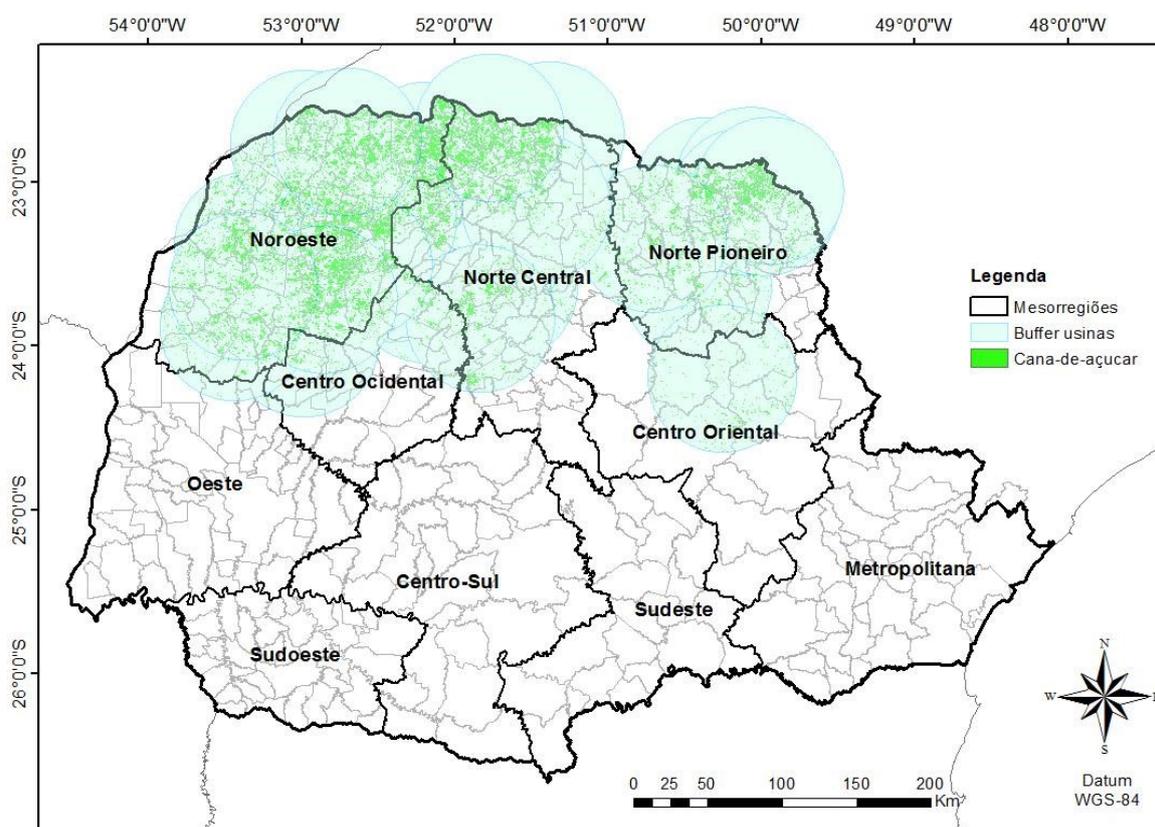


Figura 4: Mapeamento da cultura de cana-de-açúcar no Paraná: ano-safra 2012/2013.

A classificação foi avaliada pelo EG e pelo IK, obtidos pela contagem dos erros de inclusão e exclusão dos 400 pontos sorteados obtendo-se EG de 89,5% e IK de 0,79 que é classificado como muito bom, segundo Fonseca (2000).

4. Conclusão

Os polígonos de Thiessen na grade de pixels MODIS possibilitaram uma coleta de amostras mais precisa e rápida, acelerando o desenvolvimento do trabalho e melhorando sua qualidade.

O filtro *Flat Smoother Series* suavizou o perfil espectro-temporal do índice de vegetação sem perder informações significativas na identificação do perfil temporal, melhorando a acurácia da classificação.

A rede neural artificial possibilitou o mapeamento de cana-de-açúcar no estado do Paraná utilizando imagens MODIS de resolução 250m com uma EG de 89,5% e IK de 0,79.

Os filtros *Sieve classes* e *Clump classes* contribuíram para melhorar a qualidade da classificação removendo pixels isolados e definindo talhões.

5. Agradecimentos

À coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Capes e à Fundação Araucária, pelo apoio financeiro. Ao Laboratório de Estatística Aplicada (LEA) da UNIOESTE/Campus Cascavel, pela infra-estrutura disponibilizada para realização da pesquisa. Ao programa de pós-graduação em Engenharia Agrícola (PGEAGRI).

Referências Bibliográficas

ANTUNES J. F. G. **Geotecnologias e técnicas de inteligência artificial para monitoramento sistemático do cultivo de cana-de-açúcar**. 2014. 139 p. UNICAMP. Tese (doutorado Engenharia Agrícola) – Universidade Estadual de Campinas. Campinas. 2014.

ARRUDA, G. P. de; Dematte J. A. M.; Chagas C. S. Mapeamento digital de solos por redes neurais artificiais com base na relação solo paisagem. **Revista Brasileira de Ciência do Solo**, n. 37, p. 327-338, 2013.

BRAGA, A. L.; VIEIRA, C. A. O.; ASSIS, L. C. de. Identificação de extensas áreas de culturas agrícolas empregando uma abordagem espectro-temporal utilizando imagens MODIS. **Anais XIII Simpósio Brasileiro de Sensoriamento Remoto**, Florianópolis, Brasil, n. 21-26 abril, INPE, p. 5603-5610, 2007.

BRASSEL K. E.; REIF D. A Procedure to Generate Thiessen Polygons. **Geographical Analysis**, v. 11, n. 3, p. 289-303, 1979.

CONAB. Companhia nacional de abastecimento. **Acompanhamento da Safra Brasileira – Cana-de-açúcar** Disponível em : http://www.conab.gov.br/OlalaCMS/uploads/arquivos/14_04_10_09_00_57_boletim_cana_portugues_-_4o_lev_-_13.pdf . Acessado em 15.out.2014.

CONGALTON, R. G. A review of assessing the accuracy of classifications of remotely sensed data. **Remote Sensing of Environment**, v.37, p.35-46, 1991.

CONGALTON, R. G.; GREEN, K. **Assessing the accuracy of remotely sensed data: principles and practices**. Boca Raton: CRC Press, 1999. 160p.

EMBRAPA INFORMÁTICA AGROPECUÁRIA. **Banco de Produtos MODIS na base Estadual Brasileira**. Disponível em: <<http://www.modis.cnptia.embrapa.br/geonetwork/srv/pt/main.home>>. Acesso em 10.ago.2014.

EXELIS VISUAL INFORMATION SOLUTIONS. **Post Classification**. Disponível em: <http://www.exelisvis.com/docs/>. Acesso em 03.nov.2014.

FONSECA, L. M. G. **Processamento digital de imagens**. Instituto Nacional de Pesquisas Espaciais (INPE), 2000. 105 p.

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION. Technical specifications: Moderate Resolution Imaging Spectroradiometer (MODIS). 2009. Disponível em: <<http://modis.gsfc.nasa.gov/about/design.php>>. Acessado em 20.out.2014.

SANTOS, A. M. dos; SEIXAS, J. M. de; PEREIRA, B. B.; MEDRONHO, R. A.; Usando Redes Neurais Artificiais e Regressão Logística na Predição da Hepatite A. **Revista Brasileira de Epidemiologia**. Vol. 8, n.2, p. 127-126, 2005.

TRETIN, R.; LOHMANN, M.; DEPPE, F.; GRASSI, J. Mascaras de cultivo de cana-de-açúcar das safras 2008, 2009 e 2010 no estado do Paraná. **Anais XV Simpósio Brasileiro de Sensoriamento Remoto - SBSR**, Curitiba, PR, Brasil, abril, INPE, p. 607-614, 2011.