



Ciência Rural

ISSN: 0103-8478

cienciarural@mail.ufsm.br

Universidade Federal de Santa Maria

Brasil

Chimelo Ruiz, Luis Fernando; ten Caten, Alexandre; Diniz Dalmolin, Ricardo Simão
Árvore de decisão e a densidade mínima de amostras no mapeamento da cobertura da terra

Ciência Rural, vol. 44, núm. 6, junio, 2014, pp. 1001-1007

Universidade Federal de Santa Maria

Santa Maria, Brasil

Disponible en: <http://www.redalyc.org/articulo.oa?id=33131128008>

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en redalyc.org

redalyc.org

Sistema de Información Científica

Red de Revistas Científicas de América Latina, el Caribe, España y Portugal
Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso abierto

Árvore de decisão e a densidade mínima de amostras no mapeamento da cobertura da terra

Decision tree and minimum sample density in land cover mapping

Luis Fernando Chimelo Ruiz^I Alexandre ten Caten^{II} Ricardo Simão Diniz Dalmolin^{III}

RESUMO

A cobertura da terra é uma informação espacial de extrema relevância para uma série de modelos, sendo utilizada para estimar a produção de sedimentos e para mensurar a potencialidade da paisagem em sequestrar carbono. A classificação da cobertura da terra pelo método de classificação supervisionado necessita de áreas de treino, já que essas áreas devem ser representativas para cada classe de cobertura da terra. Para o algoritmo de classificação por árvore de decisão (AD), a complexidade da AD resulta em diferentes valores de acurácia para os mapas temáticos. Desse modo, o objetivo deste estudo foi determinar a densidade mínima de amostras em um modelo por AD, a fim de discriminar as classes de cobertura da terra e avaliar o tamanho da AD gerada quanto ao seu número de folhas. Além disso, preocupou-se em identificar as classes da cobertura da terra de mais difícil classificação. Nesse contexto, foram utilizadas bandas da imagem do satélite RESOURCESAT-1 e índices espectrais. A densidade mínima de amostras variou entre 0,15 e 0,30% da área total para cada classe. Esse intervalo de amostragem possibilitou resultados melhores que 80% para o índice kappa. O menor agrupamento entre observações em uma mesma folha terminal foi de 45, e as classes mais difíceis de classificar foram floresta e lavoura de arroz, devido à semelhança espectral que as florestas sombreadas possuem com as lavouras de arroz irrigadas.

Palavras-chave: classificação supervisionada de imagem, mineração de dados, sensoriamento remoto, uso da terra.

ABSTRACT

Land cover is a spatial information of great relevance for a variety of models for estimating sediment yield and to measure the potential of the landscape carbon sequestration. The classification of land cover by the supervised method requires training areas, these areas must be representative of each class of land cover. For the classification decision tree (DT) algorithm, the complexity of DT, results in different values of accuracies for

thematic maps. Thus, the objective of this study was to estimate the minimum sample density in a DT model which would allow to discriminate land cover classes, evaluate the size of the generated DT model, as well as, identify the more difficult land cover class to be mapped. Satellite images from RESOURCESAT-1 as well as spectral indices were used in the study. The minimum sample density varied between 0.15 and 0.30% of the total area for each class, this sampling interval allowed better results than 80% for kappa index. The smallest grouping of observations in the same terminal leaf was 45 observations. In this study the most difficult land use classes to be mapped were forest and rice crops due to spectral similarity of shaded forests with irrigated rice crops.

Key words: supervised image classification, data mining, remote sensing, land use.

INTRODUÇÃO

Nos últimos 40 anos, houve uma evolução no desenvolvimento de instrumentos que possibilitam a coleta de um grande volume de dados, a exemplo do sensoriamento remoto com plataformas e sensores com crescente ganho em resolução temporal, espacial, radiométrica e espectral. Simultaneamente ao aumento do volume de dados sobre a superfície terrestre, a capacidade de processamento dos computadores e a evolução dos algoritmos de reconhecimento de padrões para a classificação da cobertura da terra tornaram-se também mais sofisticados.

Entre os algoritmos mais utilizados, podemos citar redes neurais artificiais (RNA) (FOODY & ARORA, 1997), máquina vetor de suporte (SVM) (CRISTIANINI & SHAWETAYLOR, 2001) e árvore

^IPrograma de Pós-graduação em Sensoriamento Remoto da Universidade Federal do Rio Grande do Sul (UFRGS), Av. Bento Gonçalves, 9500, Campus do Vale, Bairro Agronomia, 91501-970, Porto Alegre, RS, Brasil. E-mail: ruiz.ch@gmail.com. Autor para correspondência.

^{II}Departamento de Solos, Universidade Federal de Santa Maria (UFSM), Santa Maria, RS, Brasil.

^{III}Departamento de Ciências Biológicas e Veterinárias, Universidade Federal de Santa Catarina (UFSC), Curitibanos, SC, Brasil.

de decisão (AD) (FRIEDL & BRODLEY, 1997). Os métodos RNA e SVM exigem um maior tempo de processamento e são considerados “caixas pretas”, devido à dificuldade encontrada na interpretação dos seus resultados.

Os algoritmos de AD apresentam vantagens para a classificação da cobertura da terra, entre elas, podem ser citadas: a facilidade de interpretação dos resultados, o baixo tempo de processamento, a capacidade do modelo trabalhar com dados medidos em distintas escalas, o fato de não demandar distribuição normal do dados e a capacidade em processar as relações não lineares entre os preditores e as classes de cobertura da terra (FRIEDL & BRODLEY, 1997).

De acordo com as definições de QUINLAN (1996), o algoritmo AD pode ser entendido como uma hierarquia de nós que são conectados por ramos, uma vez que sua estrutura traduz uma árvore invertida, a qual se desenvolve da raiz para as folhas. O nó de decisão possui um teste para algum atributo, e cada ramo descendente corresponde a um possível valor para esse atributo, sendo a folha relacionada a uma classe. O caminho completo da árvore, que vai da raiz para a folha, determina uma regra de classificação.

Estudos demonstraram a eficácia desse classificador em dados de sensoriamento remoto. FRIEDL & BRODLEY (1997) compararam AD com os algoritmos máxima verossimilhança e com a função discriminante linear. Nesse caso, as maiores acurárias foram encontradas para o algoritmo AD. Em trabalho para classificar a cobertura da terra e espécies de vegetação, SASAKI et al. (2012) utilizaram AD por meio da integração de dados do escanner a laser e de imagens de alta resolução, obtendo valores do índice *kappa* para a classificação superior a 0,97.

Na classificação supervisionada, as características dos dados utilizados têm influência na classificação resultante. Fatores como o número de classes e o volume de dados para treinamento dos algoritmos são elementos críticos para a acurácia da classificação da cobertura da terra. Para o método AD, outra relação importante está na complexidade (tamanho) da árvore gerada com seu poder preditivo. Em geral, quanto mais complexa a AD, maior é seu potencial preditivo (PAL & MATHER, 2003). Por outro lado, AD grandes são de difícil implementação em programas de Sistema de Informação Geográfica.

Diante do exposto, este trabalho teve como objetivo: (i) determinar qual a densidade mínima de amostras necessárias para discriminar as

classes de cobertura da terra, (ii) avaliar o tamanho da AD gerada quanto ao seu número de folhas; e (iii) identificar as classes de cobertura de difícil classificação.

MATERIAL E MÉTODOS

Área de estudo

A bacia hidrográfica do rio Vacacaí Mirim abrange 1153,47km² e está localizada na região central do estado do Rio Grande do Sul, Brasil, entre as coordenadas geográficas 53° 46'30" e 53° 49'29" de longitude oeste e 29° 36'55" e 29° 39'50" de latitude sul. A geomorfologia do local é caracterizada pelo Planalto Médio, Rebordo do Planalto Sul-Riograndense e pela Depressão Central. Tanto no Planalto Médio como na Depressão Central, predominam áreas de agricultura, com destaque para soja na primeira e arroz irrigado na segunda. Em ambos os locais, foram observadas também áreas de campo e solo exposto. As áreas mais conservadas com presença de Mata Atlântica estão localizadas no Rebordo do Planalto. Ao percorrer a área de estudo em uma fase de reambulação, houve a formulação de um modelo mental da cobertura da terra que guiou as fases seguintes de vetorização dos conjuntos de dados de treinamento e de avaliação. As classes de cobertura da terra estabelecidas foram: campo, lavoura de soja, lavoura de arroz, floresta de Mata Atlântica, solo exposto e água.

Dados de sensoriamento remoto

Foram utilizados os dados das bandas do sensor LISS III, que estão a bordo da plataforma RESOURCESAT-1. A imagem utilizada foi do dia 12 de fevereiro de 2010, a cena abrange toda a área de estudo, com poucas nuvens e resolução espacial de 23,5 metros. A imagem foi corrigida geometricamente a partir de dez pontos de controle, utilizando a função polinomial de segunda ordem para interpolação espacial, obtendo-se o erro médio quadrático menor que um pixel. Para reamostragem dos valores dos pixels, foi utilizada a interpolação bilinear. A correção atmosférica da imagem foi efetuada por meio do método de subtração do objeto escuro (CHAVEZ, 1996).

Após a aplicação das correções na imagem, foram gerados índices espectrais. Para isso, empregaram-se as bandas espectrais 0,52-0,59μm (verde - VERDE); 0,62-0,68μm (vermelho - VERM); 0,77-0,86μm (infravermelho próximo - IVP) e 1,55-1,70μm (infravermelho médio - IVM). Os índices espectrais elaborados foram: Índice de Umidade por Diferença Normalizada para estresse hídrico

(NDWI1) (GAO, 1996), Índice de Umidade por Diferença Normalizada para lâmina d'água (NDWI2) (MCFEETERS, 1996), Razão Simples (RS), Índice de Vegetação por Diferença Normalizada (NDVI) (TUCKER, 1979) e Índice de Vegetação ajustado ao Solo (SAVI) (HUETE, 1988). O SAVI possui a constante L, que é um fator de ajuste de substrato do dossel. Neste trabalho, utilizou-se o valor de L=0,5, pois engloba maior variação de condições de vegetação (HUETE, 1988).

Árvore de decisão (AD)

Dentre os algoritmos de AD disponíveis no programa de mineração de dados Weka v3.6 (HALL et al., 2009), utilizou-se, neste trabalho, o J48. Esse algoritmo realiza a classificação de forma recursiva, dividindo a totalidade do conjunto de dados de treinamento em subconjuntos cada vez mais homogêneos entre si. O algoritmo interrompe o processo de segmentação dos dados até que um nível de semelhança entre as classes finais, previamente determinado pelo usuário, seja atingido. Essas subdivisões dos dados são baseadas em procedimentos estatísticos, considerando-se os erros dos nós e seus descendentes. A identificação da raiz e dos seus descendentes é dada através dos cálculos de entropia e razão do ganho da informação. A entropia avalia a capacidade de previsão de um atributo (ELMASRI & NAVATHE, 2005).

Divisão da área de treinamento e geração das AD

Com o intuito de gerar os modelos de AD e possibilitar a avaliação da qualidade dos mapas de cobertura da terra gerados, foram coletados dois grupos de dados (amostras) em uma composição de imagem B3G4R5. Ambos os grupos de dados foram gerados por vetorização de seis diferentes classes de cobertura, no programa QGIS 1.7.0, por meio das ferramentas de edição. O primeiro conjunto de dados foi utilizado como treinamento para gerar os modelos de AD. Cada uma das seis classes de cobertura da terra foi amostrada em um total de 5.000 pixels. Como o pixel de trabalho foi de 23,5m, tem-se um total de 552,25m² por pixel. Isso compreende um total de amostras de 16,57km² para cada uma das seis classes de cobertura da terra. O segundo conjunto, com vetorização completamente independente do conjunto de dados de treinamento, foi amostrado para representar os dados de avaliação. Nesse conjunto de dados, foram reunidos 19.497 pixels, perfazendo um total de 10,76km² para as seis classes de cobertura.

Visando gerar as diferentes densidades de amostragem, o conjunto de dados de treinamento

foi dividido, aleatoriamente, em conjuntos de dados de 250, 500, 750, 1.000, sucessivamente até o conjunto total de 5.000 pixels. Esse agrupamento dos dados em conjuntos de 250 pixels foi definido por meio de avaliações prévias, considerando a quantidade mínima de pixels em que a AD teve seu poder preditivo fortemente deteriorado. Ao todo, geraram-se 20 subgrupos de amostras a serem avaliados. Para cada subgrupo, foi criada uma classificação de imagem, e esse grupo foi confrontado com o conjunto de avaliação.

No desenvolvimento dos modelos de AD no programa Weka, foi selecionado “verdadeiro” para a poda (*reducedErrorPruning=True*), e o número mínimo de amostras (NMA) nas folhas terminais igual a dois (NMA=2). A poda serve para que ramos menos confiáveis sejam removidos. Esses ramos em maior quantidade podem ser decorrentes de ruídos do conjunto de treinamento. O processo de poda produz AD com melhor nível de precisão e menos tempo de processamento (HAN & KAMBER, 2001). Sempre que é feita uma divisão nos dados produzindo uma nova folha na árvore, e essa folha representa menos do que um NMA do conjunto de dados, o nó pai e os seus filhos são compactados em um único nó. Esse processo continua ao longo da criação de toda a árvore. A AD no formato texto, gerada pelo programa Weka, foi convertida para um formato condicional *if else* a fim de ser implementada no programa SAGA-GIS.

Na segunda etapa do estudo, foram avaliados distintos valores de NMA que poderiam ser reunidos em uma folha terminal da AD. Para isso, criaram-se 10 AD em que o NMA informado no programa Weka variou de cinco a 50, com intervalo de cinco em cinco. Esses valores foram escolhidos por compreenderem um intervalo significativo de valores. As novas classificações procedentes dessa etapa foram comparadas com os dados de avaliação. Em cada uma das fases do estudo, a avaliação da qualidade dos mapas gerados foi realizada pelo coeficiente *Kappa* (CONGALTON, 1991).

RESULTADOS E DISCUSSÃO

Efeitos do tamanho da área de treinamento e do número mínimo de amostras

A complexidade da AD aumenta de acordo com o tamanho da amostragem (Figura 1a). O acréscimo no número de folhas representa a dificuldade de o modelo classificar as classes de cobertura da terra, pois, quanto maior o número de amostras inseridas no algoritmo por AD, mais folhas serão necessárias para encontrar padrões nessas amostras. Isso resultará em maior tempo de processamento e em árvores mais

complexas.

Por outro lado, o aumento da densidade das amostras proporcionou melhores valores de acurácia (Figura 1b). Em uma análise mais detalhada, percebemos que até em uma amostragem de 2.500 pixels ($1,38\text{km}^2$ e 0,12% do total da área) os valores para o índice *kappa* foram menores que 50%. Esse comportamento no gráfico pode ser explicado pela semelhança espectral que as classes floresta, campo e lavoura de arroz possuem. Nesse caso, para capturar a variabilidade espectral dessas classes, é necessário um número maior de amostras. Assim, com a classificação utilizando 3.000 ($1,66\text{km}^2$ e 0,14% da área total) pixels por classe, os valores da confiabilidade aumentaram, chegando a mais 90%. Para as demais densidades de amostragem, o índice *kappa* manteve-se igualmente elevado. Esses resultados indicam haver uma densidade de amostragem mínima que deveria ser seguida em futuros trabalhos, quando AD for aplicada na classificação da cobertura da terra.

Resultados semelhantes foram

encontrados por PAL & MATHER (2003) no leste da Inglaterra, que utilizaram dados do Landsat ETM⁺ para sete classes de cobertura da terra em uma área de $91,18\text{km}^2$. Esses autores obtiveram valores de acurácia que variaram de 78,3% a 84,1% para as amostragens que possuíam 100 ($0,09\text{km}^2$ e 0,01% da área total) e 300 ($0,27\text{km}^2$ e 0,30% da área total) pixels por classe. Para amostragens de 350 ($0,31\text{km}^2$ e 0,34% da área total) e 400 ($0,36\text{km}^2$ e 0,39% da área total) pixels, as diferenças de acurácia não foram significativas. Neste trabalho, procurou-se determinar a menor área de treinamento possível, sem obter perdas significativas na acurácia. Nesse caso, o tamanho da área de treinamento escolhido foi de 3.500 pixels para cada classe. O total desses pixels abrange uma área de $1,93\text{km}^2$ e representa 0,17% de toda a área. Essa amostragem obteve o valor de 94,2% de índice *kappa* e 57 para o número de folhas. A diferença desses valores comparados à amostragem que utilizou 5.000 ($2,76\text{km}^2$ e 0,24% da área total) pixels foi sutil, visto que o valor de acurácia para a amostragem de 5.000 pixels foi de

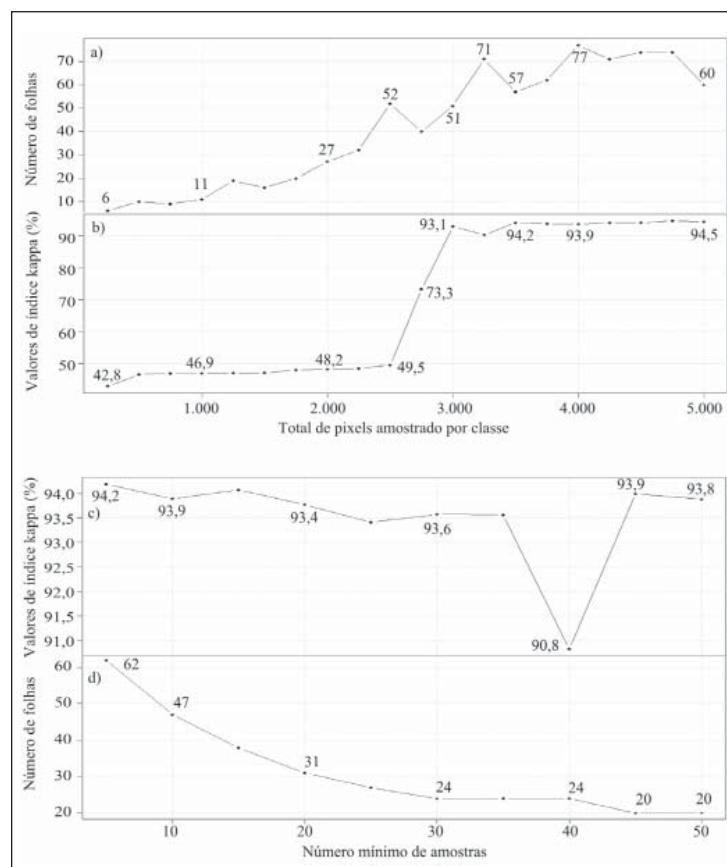


Figura 1 - a) número de folhas em relação ao total de pixels para cada classe, b) valores de índice *kappa* em relação ao total de pixels para cada classe, c) valores de índice *kappa* em relação ao NMA, d) número de folhas em relação ao NMA.

94,5%, e o número de folhas foi de 60. A importância de diminuir esses 1.500 pixels pode estar na coleta de amostras para classes que são menos representativas em uma imagem, por exemplo, a classe água. Outra questão está no tempo e custo de aquisição dessas amostras, que se torna mais dispendiosa quando esses dados são levantados a campo. Com isso, foi determinado que a densidade mínima de amostras na classificação da cobertura da terra por AD pode estar entre 0,15% a 0,30% da área total para cada classe.

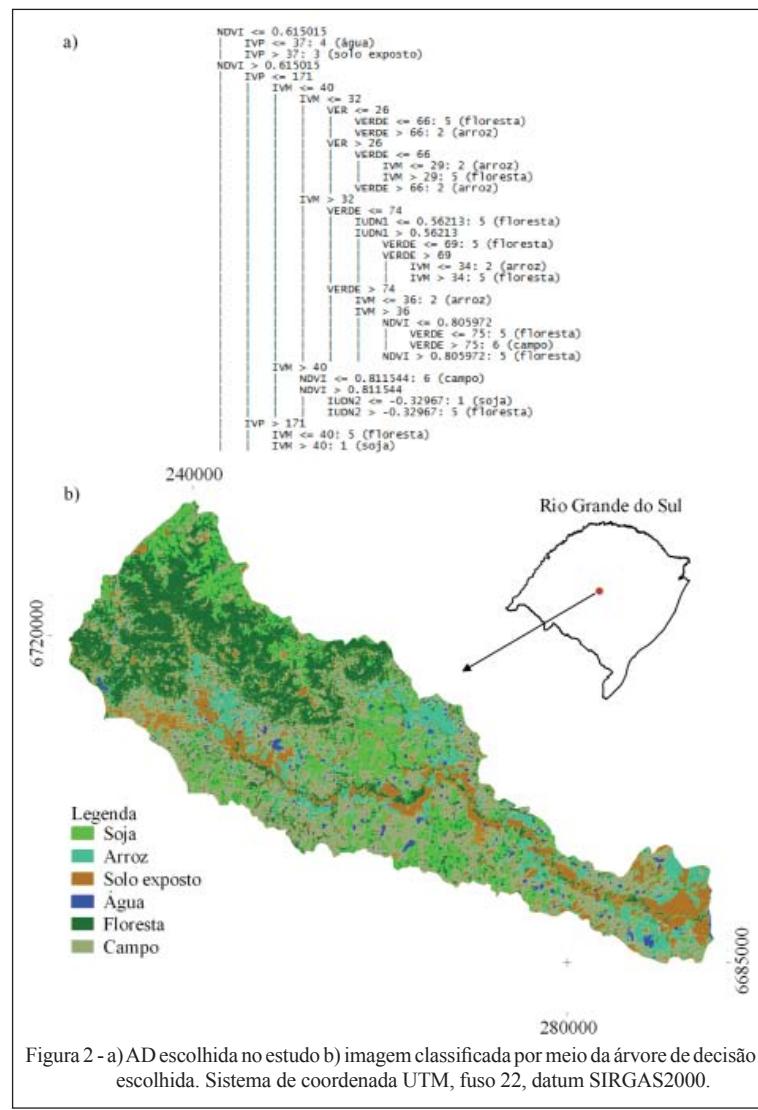
Com os melhores resultados encontrados para a amostragem de 3.500 pixels para cada classe (0,17% da área total), o próximo passo foi submetê-la a diferentes valores de NMA. O primeiro valor informado foi cinco e o último valor foi de 50, variando de cinco em cinco (Figura 1c). A complexidade da AD também está relacionada com o NMA, pois, quanto mais valores das amostras

forem agrupados nas folhas terminais, menores serão as AD. Na figura 1d, podem ser avaliados os números de folhas para as AD geradas anteriormente, apresentando a diminuição das folhas para maiores valores do NMA.

Para as AD com NMA de cinco e 50, verificou-se uma redução de 42 folhas. Quando analisados seus valores de índice *kappa*, a diferença foi menos representativa, isto é, de apenas 0,4%. Essa redução de folhas sem grandes perdas da confiança facilita a interpretação da AD resultante e diminui o tempo de processamento. Futuros trabalhos poderão avaliar a classificação por AD em relação ao tamanho da área treinamento com o número de classes da cobertura da terra.

Análise da classificação

Ao considerar o número de folhas em



relação ao índice *kappa*, a AD mais vantajosa neste estudo foi a gerada com 3.500 pixels e o NMA igual a 45. O atributo preditivo que forneceu a maior redução na entropia foi o NDVI, sendo escolhido como raiz da árvore (Figura 2a). As classes mais difíceis de classificar foram áreas de floresta e áreas de arroz irrigado, possivelmente, devido à semelhança espectral que áreas de floresta sombreadas possuem com as áreas úmidas de arroz irrigado, necessitando de mais folhas para classificá-las.

Na figura 2a, pode ser visualizada a AD utilizada para a classificação da cobertura da terra na área de estudo. A partir dela, observou-se que as classes floresta e lavouras de arroz resultaram no maior número de folhas. Já as coberturas da terra, água e solo exposto necessitaram de apenas uma folha para cada. O resultado da classificação originado dessa árvore está ilustrado na figura 2b. Alguns padrões em relação ao uso da terra podem ser observados no mapa classificado, em que as florestas (29,54%) estão situadas no rebordo, nas maiores declividades da área e próximo às drenagens. As áreas de lavoura de arroz (13,70%) que são áreas irrigadas, na sua maioria pelos rios da região, estão próximas às drenagens. O resultado mais controverso está nas áreas de solo exposto (12,64%), pois, sabendo que a imagem é datada no dia 12 de fevereiro, essas áreas de solo exposto deveriam estar com lavouras de arroz. O motivo disso não acontecer é devido a fortes precipitações ocorridas na região, acarretando enxurradas que provocam a remoção das plantas de arroz da lavoura e também o acúmulo de sedimentos, que resultou no aumento da área de solo exposto no mapa temático. Outra questão em relação à classe arroz, observada a campo, está na substituição desse cultivo por plantações de soja (12,02%). A classe de uso da terra mais representativa nesse mapa temático foi o uso campo (30,69%) e as menos representativas foram as áreas de água (1,41%).

CONCLUSÃO

A densidade mínima de amostras permaneceu no intervalo de 0,15% a 0,30% da área total para cada classe. Esse intervalo pode ser usado como guia no momento da determinação da área de treinamento.

O número mínimo de amostras igual a 45 foi o que proporcionou melhor relação entre tamanho da AD gerada e a exatidão da classificação. Além disso, a redução aplicada na AD propiciou identificar os usos da terra de difícil classificação, que, neste trabalho, foram as florestas e as lavouras de arroz. Nessas classes, foi necessário maior número de folhas para distingui-las

das demais.

AGRADECIMENTOS

Os autores agradecem ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelas bolsas de iniciação científica do primeiro autor e produtividade em pesquisa ao terceiro autor.

REFERÊNCIAS

- 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. Disponível em: <[http://dx.doi.org/10.1016/0034-4257\(91\)90048-B](http://dx.doi.org/10.1016/0034-4257(91)90048-B)>. Acesso em: 12 dez. 2012. doi: 10.1016/0034-4257(91)90048-B.
- CHAVEZ, JR., P.S. Image-based atmospheric corrections - revisited and improved. *Photogrammetric Engineering and Remote Sensing*, v.62, p.1025-1036, 1996. Disponível em: <http://www.asprs.org/a/publications/pers/96journal/september/1996_sep_1025-1036.pdf>. Acesso em: 12 nov. 2012.
- CRISTIANINI, N.; SHAWE-TAYLOR, J. *An introduction to support vector machines and other kernel based learning methods*. Cambridge: Cambridge University, 2001. 191p.
- ELMASRI, R.; NAVATHE, S.B. *Sistemas de banco de dados*. São Paulo: Addison, 2005. 744p.
- FOODY, G.M.; ARORA, M.K. An evaluation of some factors affecting the accuracy of classification by an artificial neural network. *International Journal of Remote Sensing*, v.18, p.799-810, 1997. Disponível em: <<http://dx.doi.org/10.1080/014311697218764>>. Acesso em: 10 dez. 2012. doi: 10.1080/014311697218764.
- FRIEDL, M.A.; BRODLEY, C.E. Decision tree classification of land cover from remotely sensed data. *Remote Sensing of Environment*, v.61, p.399-409, 1997. Disponível em: <[http://dx.doi.org/10.1016/S0034-4257\(97\)00049-7](http://dx.doi.org/10.1016/S0034-4257(97)00049-7)>. Acesso em: 11 dez. 2012. doi: 10.1016/S0034-4257(97)00049-7.
- GAO, B. NDWI – A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment*, v.58, p.257-266, 1996. Disponível em: <[http://dx.doi.org/10.1016/S0034-4257\(96\)00067-3](http://dx.doi.org/10.1016/S0034-4257(96)00067-3)>. Acesso em: 11 dez. 2012. doi: 10.1016/S0034-4257(96)00067-3.
- HAN, J.; KAMBER, M. *Data mining: concepts and techniques*. San Diego: Academic, 2001. 550p.
- HALL, M. et al. The WEKA data mining software: an update. *SIGKDD Explorations*, v.11, n.1, p.18 2009. Disponível em: <<http://www.cs.waikato.ac.nz/ml/weka/>>. Acesso em: 10 jan. 2013.
- HUETE, A.R. A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment*, v.25, p.295-309, 1988. Disponível em: <[http://dx.doi.org/10.1016/0034-4257\(88\)90106-X](http://dx.doi.org/10.1016/0034-4257(88)90106-X)>. Acesso em: 11 dez. 2012. doi: 10.1016/0034-4257(88)90106-X.
- MCFEEETERS, S.K. The use of the normalized difference water index (NDWI) in the delineation of open water features. *International Journal of Remote Sensing*, v.17, p.1425-1432, 1996. Disponível em: <<http://dx.doi.org/10.1080/01431169608948714>>. Acesso em: 10 dez. 2012. doi: 10.1080/01431169608948714.
- PAL, M.; MATHER, P.M. An assessment of the effectiveness of decision tree methods for land cover classification. *Remote*

- Sensing of Environment**, v.86, p.554-565, 2003. Disponível em: <[http://dx.doi.org/10.1016/S0034-4257\(03\)00132-9](http://dx.doi.org/10.1016/S0034-4257(03)00132-9)>. Acesso em: 11 dez. 2012. doi:S0034-4257(03)00132-9.
- QUINLAN, R. Improved use of continuous attributes in C4.5. **Journal of Artificial Intelligence Research**, v.4, p.77-90, 1996. Disponível em: <<http://dx.doi.org/10.1613/jair.279>>. Acesso em: 11 dez. 2012. doi: 10.1613/jair.279.
- SASAKI, T. et al. Object-based classification of land cover and tree species by integrating airborne LiDAR and high spatial resolution imagery data. **Landscape and Ecological Engineering**, v.8, p.157-171, 2012. Disponível em: <<http://dx.doi.org/10.1007/s11355-011-0158-z>>. Acesso em: 11 dez. 2012. doi: 10.1007/s11355-011-0158-z.
- TUCKER, C.J. Red and photographic infrared linear combinations for monitoring vegetation. **Remote Sensing of Environment**, v.8, p.127-150, 1979. Disponível em <[http://dx.doi.org/10.1016/0034-4257\(88\)90106-X](http://dx.doi.org/10.1016/0034-4257(88)90106-X)>. Acesso em: 11 dez. 2012. doi: 10.1016/0034-4257(88)90106-X.