

CONCEPÇÃO DE UM SISTEMA DE CLASSIFICAÇÃO HIERÁRQUICO COM INSERÇÃO DE ABORDAGEM FUZZY NA CLASSIFICAÇÃO

FERNANDA DE ALMEIDA PRADO
MARIA DE LOURDES BUENO TRINDADE GALO

Universidade Estadual Paulista - Unesp
Faculdade de Ciências e Tecnologia - FCT
Departamento de Cartografia, Presidente Prudente - SP
prado_fernanda@yahoo.com.br, mlourdes@fct.unesp.br

RESUMO - O mapeamento da cobertura da terra constitui uma importante ferramenta para o planejamento regional. Contudo, os atuais mapeamentos atendem a propósitos muito específicos e desconsideram a estrutura hierárquica dos geossistemas. Dentro disso, este trabalho objetiva propor um sistema de classificação hierárquico para o mapeamento da cobertura da terra em escala regional, que contribua com sistemas clássicos existentes. Além disso, pretende-se testar esse sistema para um estudo de caso que contemple o uso de um método de classificação baseado em abordagem *fuzzy*, a qual tem se mostrado mais adequada que as abordagens convencionais. Para tanto, foi concebido um sistema de classificação hierárquico com três níveis de detalhamento e um estudo de caso foi definido com a especificação da área de teste e do projeto de classificação. Para isso foi feito, inicialmente, o georreferenciamento de uma imagem TM/Landsat-5 que abrange a área de teste. Posteriormente, foi aplicada uma abordagem de classificação *fuzzy* dessa imagem, a partir da qual foram geradas imagens de probabilidade para as classes de cobertura da terra e uma imagem de incerteza, assim como uma saída convencional, que representa o mapeamento temático da área de teste.

ABSTRACT - The land cover mapping is an important tool to the regional planning. However, the current mappings assist to very specific purposes and ignore the hierarchical structure of the geosystems. Thus, this paper aims at proposing a hierarchical classification system for the land cover mapping in regional scale that contributes with existent classic systems. Besides, it is intended to test that system for a case study that contemplates the use of a classification method based on fuzzy approach, which has shown more appropriate than the conventional approaches. Therefore, it was conceived a hierarchical classification system with three detail levels and a case study was defined with the specification of the test area and of the classification project. In order to do that, first, the georeferencing of a TM/Landsat-5 image that comprises the test area was carried out. Later, it was applied an approach of fuzzy classification of the image, starting from images of probability for the land cover classes and an uncertainty image were generated, as well as it was produced a conventional output that represents the thematic mapping of the test area.

1 INTRODUÇÃO

Nas últimas décadas, os sistemas de Sensoriamento Remoto têm sido amplamente utilizados no mapeamento da cobertura terrestre e no monitoramento dos recursos naturais. Isto se deve ao fato de que os dados obtidos a partir de satélites propiciam coberturas repetitivas da superfície terrestre em intervalos relativamente curtos. Além disso, esses dados podem ser processados rapidamente, através de técnicas de análise associadas à sistemas computacionais (VILELA et al., 2000).

Dessa forma, os dados de sensores remotos necessitam ser analisados para que sejam extraídas informações necessárias ao planejamento, manejo e monitoramento dos recursos. Essa análise torna-se

possível pelo fato de os fenômenos e materiais da superfície terrestre apresentarem comportamentos específicos ao longo do espectro eletromagnético, os quais podem, portanto, ser usados para identificá-los.

Convencionalmente, os materiais superficiais são representados a partir de mapeamentos temáticos, onde os dados de sensores remotos são caracterizados e associados à classes temáticas que definem o uso e a cobertura da terra. Esses mapeamentos, gerados com o uso de técnicas de classificação multiespectral, têm se tornado cada vez mais frequentes, devido à grande quantidade de imagens disponíveis geradas por satélites de média e alta resoluções.

Dentre as técnicas de classificação multiespectrais, as convencionais supervisionadas e não supervisionadas ainda são amplamente utilizadas, principalmente por

usuários não especialistas. Essas técnicas fazem uso da lógica *booleana* e classificam os dados de Sensoriamento Remoto associando cada *pixel* a uma única classe temática, desconsiderando o aspecto contínuo de muitos fenômenos geográficos. Contudo, nota-se uma busca por novas alternativas que garantam uma representação mais adequada da estrutura do ambiente terrestre.

Nesse sentido, Wang (1990) afirma que a teoria de conjunto *fuzzy* pode contribuir para uma melhor representação dos fenômenos geográficos. Uma classificação baseada em abordagem *fuzzy* permite caracterizar a transição entre tipos de uso e cobertura terrestre e lidar com a ocorrência de *pixels* mistos, muito frequentes, principalmente, em imagens de baixa e média resolução espacial.

Independente da técnica utilizada, classificar envolve a definição dos tipos de classes de interesse e de seus critérios, os quais devem ser claros e objetivos. Contudo, o que se vê, atualmente, é uma grande quantidade de mapeamentos temáticos com classes inapropriadas (ou com terminologia ambígua), voltados à propósitos específicos, ou com informação, muitas vezes, obsoleta (DI GREGÓRIO, 2004).

Normalmente, cada usuário define sua própria legenda de classificação contendo as classes de uso e cobertura da terra necessárias à sua aplicação. Assim, ainda que a nomenclatura das classes seja similar, em alguns casos, nem sempre as definições dos critérios para as classes são as mesmas.

Existem poucas contribuições na literatura para uma padronização de sistemas de classificação. Os pioneiros, nesse sentido, foram Anderson et al. (1976) que desenvolveram, na década de 70, o sistema utilizado pelo Serviço Geológico dos Estados Unidos (USGS) para a geração de mapas do território americano. Desde então, vários outros sistemas foram propostos, contudo, ainda assim, independentes e específicos às suas regiões.

O sistema de classificação do programa CORINE Land Cover, definido na década de 80 para o mapeamento de terras européias, também é uma referência no sentido de prover um sistema cuja hierarquia de classes possa ser aplicável em diferentes projetos de mapeamento, atendendo à demanda de diferentes usuários.

No Brasil, a iniciativa de definição de um sistema de classificação está restrita ao IBGE (Instituto Brasileiro de Geografia e Estatística), através do Projeto de Uso da Terra definido no final da década de 90, com vistas a sistematizar as informações sobre o meio ambiente.

Dessa forma, entre muitos sistemas de classificação existentes, cada qual proposto por uma entidade e voltado à um fim específico, o sistema utilizado pelo USGS continua sendo a grande referência teórica seguido, recentemente, pelo sistema europeu do programa CORINE. Ambos privilegiam um sistema voltado para a caracterização da cobertura da terra. Por outro lado, a proposição teórica brasileira, recém-melhorada e atualizada, tem como foco principal o uso da terra.

Nesse contexto, o objetivo do presente trabalho é conceber um sistema de classificação hierárquico para o mapeamento da cobertura da terra em escala regional, que contribua com os sistemas clássicos existentes, partindo de concepções já consagradas e realizando as adaptações necessárias. O referido sistema será testado para um estudo de caso que contempla o uso de um método de classificação baseado em abordagem *fuzzy*.

2 LÓGICA FUZZY E CLASSIFICAÇÃO DE IMAGENS MULTIESPECTRAIS

A matemática e a lógica tradicionais (ou *booleana*) já conseguiram mapear satisfatoriamente um incontável número de processos de cálculo e decisão. Porém, muitas experiências humanas necessitam de um enfoque mais abrangente do que a simples associação de falso ou verdadeiro, sim ou não, certo ou errado (LIMA JR., 2002).

Neste contexto, a teoria *fuzzy* surge como uma alternativa para representar modelos de raciocínio impreciso, necessários à tomada de decisões racionais em ambientes de incertezas e imprecisões, proposta por Lotfi Zadeh, em 1965, que a definiu como a parte da lógica matemática dedicada ao raciocínio incerto ou aproximado, utilizada para caracterizar classes que, por diversas razões, não possuem limites rígidos entre si (ZADEH, 1965).

Na lógica *booleana*, um conjunto A pertencente a um domínio X pode ser representado por uma função característica $A(x)$, cujos valores dessa função podem ser igual a 1, se $x \in X$, ou igual a 0, se $x \notin X$. Essa estrutura é restritiva, ideal para conjuntos com limites bem definidos pois, para qualquer objeto que pertença ao domínio X , pode-se determinar com exatidão se o mesmo pertence ou não ao conjunto A (ZADEH, 1965).

Por outro lado, um conjunto *fuzzy* é caracterizado por uma função de pertinência que mapeia os elementos de um espaço ou domínio X para um número real em $[0,1]$ ou, formalmente, $A: X \rightarrow [0,1]$. Dessa forma, um conjunto *fuzzy* apresenta-se como um conjunto de pares ordenados, em que o primeiro elemento é $x \in X$, e o segundo, $\mu_A(x)$, é o grau de pertinência ou a função de pertinência de x em A , que mapeia x no intervalo $[0,1]$, ou seja, $A = \{(x, \mu_A(x)) | x \in X\}$ (ZADEH, 1965).

Em classificação de imagens multiespectrais, normalmente, a informação de interesse é extraída por meio de métodos de classificação convencionais, isto é, baseados na lógica *booleana*, onde o *pixel* é atribuído a uma única classe de cobertura da terra, desconsiderando a existência de qualquer proporção de contribuição de outras classes para esse *pixel*.

Contudo, certos *pixels* podem constituir mais de um tipo de cobertura da terra, sendo denominados *pixels* mistos. A proporção de *pixels* mistos em uma imagem aumenta geralmente com a diminuição da resolução espacial do sensor. Assim, para dados de média e baixa resolução espacial usados em mapeamento nas escalas

regional e global, onde o Sensoriamento Remoto é a principal fonte de dados em mapeamentos temáticos, *pixels* mistos podem dominar as imagens (FOODY et al., 1997).

Para lidar com *pixels* mistos, as classes podem ser representadas como conjuntos *fuzzy*, atribuindo graus de pertinência aos *pixels* para indicar a extensão com que estes pertencem à cada classe, o que permite uma melhor representação para fenômenos geográficos.

O resultado de uma classificação *fuzzy* pode, então, ser expresso na forma de uma matriz de partição *fuzzy* (WANG, 1990):

$$\begin{bmatrix} f_{F_1(x_1)} & f_{F_1(x_2)} & \dots & f_{F_1(x_n)} \\ f_{F_2(x_1)} & f_{F_2(x_2)} & \dots & f_{F_2(x_n)} \\ \vdots & \vdots & \ddots & \vdots \\ f_{F_m(x_1)} & f_{F_m(x_2)} & \dots & f_{F_m(x_n)} \end{bmatrix} \quad (1)$$

onde: F_1, \dots, F_m representam as m classes espectrais; x_1, \dots, x_n são os *pixels*; n é o número de *pixels*; e f_{F_1}, \dots, f_{F_m} são as funções de pertinência para as classes.

Existem várias maneiras de derivar uma classificação *fuzzy*, dentre quais pode-se destacar a classificação relativa bayesiana, na qual são consideradas média e matriz de variância e covariância ponderadas, cujos pesos são graus de pertinência *fuzzy* estimados a partir de dados de treinamento para as classes (EASTMAN e LANEY, 2002). Assim, a média *fuzzy* (\vec{m}_c^*) é calculada como:

$$\vec{m}_c^* = \frac{\sum_{i=1}^n \mu_c(\vec{x}_i) \cdot \vec{x}_i}{\sum_{i=1}^n \mu_c(\vec{x}_i)}, \quad (2)$$

onde: n é o número total de *pixels* de treinamento; \vec{x}_i é o vetor de atributos do *pixel* de amostra; e μ_c é a função de pertinência da classe c (a proporção de classe de cobertura c na amostra de treinamento).

De modo similar, segundo Wang (1990), a matriz de covariância *fuzzy* (Σ_c^*) é calculada como:

$$\Sigma_c^* = \frac{\sum_{i=1}^n \mu_c(\vec{x}_i) \cdot (\vec{x}_i - \vec{m}_c^*) \cdot (\vec{x}_i - \vec{m}_c^*)^T}{\sum_{i=1}^n \mu_c(\vec{x}_i)}. \quad (3)$$

Uma vez que as estatísticas de treinamento são desenvolvidas, o estágio de classificação pode ser realizado. A pertinência de classe é definida com base na lógica da classificação de Máxima Verossimilhança, mas usando média e matriz de covariância *fuzzy* substituindo a média e a matriz de covariância convencionais, isto é:

$$\mu_c(\vec{x}) = \frac{p(\vec{x}|\omega_c)}{\sum_{i=1}^m p(\vec{x}|\omega_i)} \quad (4)$$

onde: $p(\vec{x}|\omega_i)$ é a função densidade de probabilidade normal para a classe i , apresentada na Equação 5:

$$p(\vec{x}|\omega_i) = \frac{1}{(2\pi)^{N/2} |\Sigma_i^*|^{1/2}} e^{-\frac{1}{2}(\vec{x}-\vec{m}_i^*)^T \Sigma_i^{-1}(\vec{x}-\vec{m}_i^*)}, \quad (5)$$

onde: N é o número de bandas espectrais utilizadas (WANG, 1990).

Nesse método de classificação a qualidade do mapeamento temático pode ser avaliada por medidas de incerteza, a qual é determinada por medidas de probabilidade associadas à posição de cada *pixel*.

3 SISTEMAS DE CLASSIFICAÇÃO DE USO E COBERTURA DA TERRA

O termo “cobertura da terra” refere-se à cobertura (bio)física observada sobre a superfície terrestre, isto é, aquilo que se pode extrair diretamente de imagens orbitais ou fotografias aéreas. Por outro lado, o “uso da terra” pode ser caracterizado pelas atividades desenvolvidas pelas pessoas que ocupam um tipo de cobertura da terra específico e interagem para mantê-lo ou modificá-lo (DI GREGÓRIO, 2004). Assim, por exemplo, as áreas de recreação existentes nas cidades são usos da terra que podem ser aplicados a diferentes tipos de cobertura da terra, desde áreas verdes até áreas construídas.

Sistemas de classificação podem ser apresentados em dois tipos básicos: hierárquico e não-hierárquico. A maioria dos sistemas é hierarquicamente estruturada, visto que, dessa forma, a classificação oferece maior consistência devido à habilidade de acomodar diferentes níveis de informação, iniciando com as classes mais amplas e, a partir dessas, realizando uma subdivisão sistemática em subclasses mais detalhadas (DI GREGÓRIO, 2004).

Em um sistema hierárquico, as classes nos níveis superiores e menos detalhados podem ser subdivididas em outras classes, nos níveis inferiores e de maior detalhamento. Assim, pode-se subdividir uma classe de cobertura da terra que representa as águas marinhas, que está num nível superior, em classes mais detalhadas, num nível inferior, denominadas baías, lagunas, mares e oceanos, por exemplo.

A estrutura das classes temáticas, portanto, é especificada pelo sistema de classificação. Esse sistema deve estar previsto dentro de um projeto de classificação, o qual contempla dois componentes críticos: (1) um conjunto de rótulos, isto, de nomes das classes estabelecidas para o sistema de classificação; e (2) um conjunto de regras ou definições que caracterizem os diferentes tipos de ocorrências a serem mapeadas. Sem um claro conjunto de regras, a atribuição dos rótulos aos tipos de cobertura torna-se arbitrária e sem consistência (CONGALTON e GREEN, 1999).

Dessa forma, um projeto de classificação é um meio de representar a informação espacial de uma maneira ordenada e lógica, sendo fundamental a qualquer projeto de mapeamento. O projeto de classificação

possibilita, para quem produz o mapa, caracterizar as feições da paisagem, e, para o usuário, reconhecê-las. O nível de detalhamento (isto é, número e complexidade das classes) estabelecido no projeto é direcionado pelo uso que será dado à informação representada no mapa e pelos atributos da terra que são discerníveis nos dados utilizados na sua elaboração (CONGALTON e GREEN, 1999).

A legenda da classificação consiste na aplicação de uma classificação em uma área específica usando uma escala de mapeamento definida e um conjunto de dados específico. Dessa forma, a legenda pode conter apenas um sub-conjunto de todas as classes possíveis estabelecidas por um sistema de classificação. Conforme Di Gregório (2004), a legenda é dependente da escala e da representação cartográfica, bem como dos dados e do método de mapeamento.

3.1 Programas de mapeamento em escala regional

Dentre os sistemas de classificação propostos pelos programas de mapeamento, destaque pode ser dado ao sistema criado por Anderson et al. (1976) e utilizado pelo Serviço Geológico dos Estados Unidos (USGS), e também ao sistema de classificação do programa CORINE *Land Cover*. No território brasileiro, merece destaque o Projeto de Uso da Terra elaborado pelo IBGE.

O sistema de Anderson et al. (1976) foi apresentado e publicado no Documento Técnico 964 do Serviço Geológico dos Estados Unidos, sendo utilizado pelo USGS para mapear o uso e a cobertura da terra dos Estados Unidos nas escalas 1:250.000 e 1:100.000. Esse sistema inclui dois níveis de detalhamento que fornecem informação em uma base de âmbito nacional, interestadual ou estadual. Corresponde a um sistema propositalmente aberto, permitindo flexibilidade para o desenvolvimento posterior de uma classificação mais detalhada, nos níveis III e IV, apropriada para atividades de planejamento e gerenciamento regionais, locais ou nacionais (ANDERSON et al., 1976).

O sistema de classificação do programa CORINE (Coordenação de Informação do Ambiente) foi implementado pela Agência Ambiental Européia (EEA), no período de 1985 a 1990, e surgiu da necessidade de se ter uma base de informações ambientais de todo o território europeu, de forma homogênea e suscetível a atualizações periódicas. A nomenclatura desse sistema é hierarquizada em três níveis, sendo que o primeiro nível compreende as maiores categorias de cobertura da terra para o planeta; o segundo nível é utilizado para escalas geográficas de 1:1.000.000 a 1:500.000; e o terceiro nível deverá ser utilizado em projetos na escala 1:100.000 (BÜTTNER et al., 2004).

O Projeto de Uso da Terra do IBGE surgiu da ausência de um referencial nacional para o estabelecimento de normas e critérios a serem utilizados na elaboração de mapeamento de uso e cobertura da terra. O sistema de classificação, que se encontra na sua segunda versão, foi definido com três níveis hierárquicos, como classes e subclasses, procurando formar legendas com o máximo de significado e o mínimo de níveis (IBGE, 2006).

4 MATERIAL E MÉTODO

4.1 Concepção do sistema de classificação hierárquico

O sistema de classificação hierárquico para o mapeamento em escala regional foi estabelecido, principalmente, com base nos sistemas de classificação propostos por Anderson et al. (1976) e pelo projeto CORINE *Land Cover*, assim como pelo Projeto de Uso da Terra elaborado pelo IBGE.

Esse sistema de classificação hierárquico foi definido em três níveis de detalhamento (Tabela 1), visando, principalmente, mapeamentos de cobertura da terra com imagens de Sensoriamento Remoto. O nível I está bem próximo do que foi proposto pelo sistema CORINE, abrangendo as maiores classes de cobertura da terra para o planeta e, assim, totalizando cinco categorias. O nível II foi proposto com base nos três principais sistemas considerados nesta pesquisa e foram feitas adaptações de forma a melhor caracterizar a cobertura da terra e os ecossistemas existentes no Brasil, sendo composto por quinze classes. O nível III foi definido também com embasamento no sistema CORINE, não apenas com relação à nomenclatura, mas também aos critérios de definição de classe, mas com contribuições dos sistemas propostos pelo USGS e pelo IBGE, com significativas adaptações ao território brasileiro, consistindo de quarenta e seis classes.

Observa-se, no sistema proposto, que o nível I está relacionado com uma escala global (exploratória) e os níveis II e III constituem uma escala regional, de reconhecimento e semi-detalle, respectivamente. Dessa forma, o sistema abrange níveis mais generalizados e de menor detalhamento, a partir dos quais, cada usuário pode detalhar os níveis IV e V de forma a atender às suas necessidades específicas de mapeamento.

Para extrair a informação de cobertura da terra do nível I do sistema hierárquico podem ser utilizadas imagens com resolução espacial superior a 80 m. Para o nível II podem ser utilizadas imagens com resolução entre 20 e 80 m e, para o nível III, entre 5 e 20 m, conforme o que recomendam Anderson et al. (1976).

Tabela 1 – Sistema de classificação hierárquico para o mapeamento da cobertura da terra em escala regional.

NÍVEL I	NÍVEL II	NÍVEL II
1 Superfície construída	1.1 Área urbanizada	1.1.1 Área edificada urbana 1.1.2 Área edificada rural 1.1.3 Área verde urbana 1.1.4 Rodovia, ferrovia e feições associadas
	1.2 Área degradada	1.2.1 Área de extração 1.2.2 Depósito tecnogênico
2 Área agrícola	2.1 Área de cultivo	2.1.1 Área de cultivo permanente 2.1.2 Área de cultivo temporário 2.1.3 Área de cultivo heterogêneo
	2.2 Pastagem	2.2.1 Pastagem herbácea 2.2.2 Pastagem mista 2.2.3 Pastagem degradada
	2.3 Silvicultura	2.3.1 Reflorestamento 2.3.2 Florestamento 2.3.3 Cultivo agroflorestal
3 Área úmida	3.1 Área úmida continental	3.1.1 Alagado interior
	3.2 Área úmida litorânea	3.2.1 Pântano salino 3.2.2 Restinga 3.2.3 Salina 3.2.4 Manguezal
4 Água	4.1 Água interior	4.1.1 Curso d'água 4.1.2 Corpo d'água
	4.2 Água marinha	4.2.1 Baía e estuário 4.2.2 Laguna costeira 4.2.3 Mar e oceano
5 Formação vegetal	5.1 Floresta	5.1.1 Floresta latifoliada 5.1.2 Floresta de coníferas 5.1.3 Formação de bambus e/ou palmeiras 5.1.4 Floresta mista
	5.2 Campo / Pradaria	5.2.1 Campo limpo 5.2.2 Campo sujo
	5.3 Cerrado / Savana	5.3.1 Cerrado típico 5.3.2 Campo cerrado 5.3.3 Cerradão
	5.4 Caatinga	5.4.1 Caatinga arbórea 5.4.2 Caatinga arbustiva 5.4.3 Caatinga herbácea
	5.5 Tundra	5.5.1 Tundra arbustiva 5.5.2 Tundra herbácea 5.5.3 Tundra de solo nu 5.5.4 Tundra úmida 5.5.5 Tundra mista
	5.6 Área com pouca e/ou nenhuma vegetação	5.6.1 Solo nu 5.6.2 Área queimada 5.6.3 Praia, duna e areal 5.6.4 Geleira e neve

4.2 Aplicação do sistema de classificação

4.2.1 Área de teste

A área de estudo adotada neste trabalho contempla o município de Presidente Prudente, o qual possui uma extensão territorial de 562 km² e uma população de, aproximadamente, 202.789 habitantes. Esse município está localizado a oeste do Estado de São Paulo, entre as

seguintes coordenadas geográficas: 21° 40' a 22° 15' de latitude sul e 51° 10' a 51° 32' de longitude oeste de Greenwich. Essa localização pode ser vista na Figura 1, onde é apresentada parte do Estado de São Paulo e a situação do município.

De acordo com o Mapa Geomorfológico do Estado de São Paulo, elaborado na escala 1:500.000, o município de Presidente Prudente encontra-se localizado na Bacia Sedimentar do Paraná (morfoestrutura) e no Planalto

Ocidental Paulista (morfoescultura), mais precisamente no Planalto Centro Ocidental. Especificamente para o município, predominam como formas de relevo as colinas médias e baixas, cujas altitudes variam entre 300 a 600 m e declividades médias entre 10% a 20%.

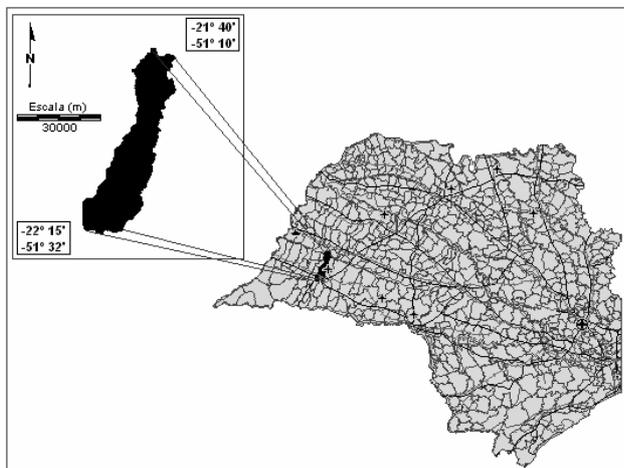


Figura 1 – Localização da área de estudo.

Presidente Prudente é a capital de sua região, constituindo um importante pólo administrativo e atrativo para os habitantes das cidades da região.

4.2.2 Especificação do projeto de classificação

A partir da definição do objetivo de uma aplicação específica pode-se identificar quais classes de informação serão mapeadas no estudo de caso. Para tanto, é preciso ter em mente qual o nível de detalhamento do mapeamento da cobertura da terra que se pretende obter. Nesse caso, foi estabelecido como prioridade extrair as informações descritas no nível II do sistema hierárquico.

Com isso foi possível identificar quais classes do nível II do sistema regional ocorrem na região de estudo. Essas classes irão compor a legenda da classificação. Assim, das quinze classes listadas no nível II do sistema, oito puderam ser identificadas no município de Presidente Prudente, sendo estas:

- área urbanizada: mancha urbana da cidade de Presidente Prudente e distritos;
- área degradada: lixão existente no município;
- área de cultivo: áreas de cultivo, permanentes e temporárias;
- pastagem: áreas de pastagem herbácea e mista;
- área úmida continental: áreas úmidas com pouca ou boa quantidade de vegetação e que são existentes no entorno de cursos d'água;
- água interior: corpos d'água e cursos d'água;
- floresta: áreas de floresta latifoliada;
- área com pouca e/ou nenhuma vegetação: áreas onde há a ocorrência de afloramentos rochosos, assim como as áreas de solo preparado para o plantio e/ou de cultivo recém-colhido.

Assim, definidas as classes de informação, deve-se especificar quais os tipos de dados necessários à extração dessa informação. Como descrito no item 4.1, imagens com resolução espacial entre 20 e 80 m devem ser satisfatórias para a finalidade em questão, de modo que optou-se pela imagem Landsat-5, de 30 m de resolução espacial, tomada pelo sensor TM (*Thematic Mapper*). Essa imagem foi adquirida a partir do site do INPE (Instituto Nacional de Pesquisas Espaciais) (<http://www.dgi.inpe.br/CDSR/>), sendo datada do dia 6 de setembro de 2007, cuja órbita é 222 e o ponto 75.

Dentre os métodos de classificação existentes, adotou-se uma abordagem *fuzzy* por se tratar de uma aplicação em escala regional, onde a forte ocorrência de *pixels* mistos é esperada. Com essa abordagem, pode-se analisar a acurácia temática da classificação a partir de medidas de incerteza geradas no processo de mapeamento.

4.3 Criação do Banco de Dados Geográficos e georreferenciamento da imagem

Após a definição do sistema hierárquico e do projeto de classificação, um Banco de Dados Geográficos foi criado utilizando o aplicativo SPRING (Sistema de Processamento de Informações Georreferenciadas). Nesse Banco de Dados foi definido um Projeto considerando o sistema de projeção UTM (Universal Transversa de Mercator) e o datum SAD69 (*South American Datum* 1969). As coordenadas do retângulo envolvente do Projeto, que abrange integralmente a área de estudo, são as seguintes: $X_1 = 447944,818$ m; $Y_1 = 7542217,323$ m; $X_2 = 480438,585$ m; $Y_2 = 7600376,028$ m.

O Projeto foi definido considerando o elipsóide associado ao datum SAD69, pois, para georreferenciar a imagem de média resolução espacial, foi adquirido um arquivo digital contendo a base do município de Presidente Prudente, georreferenciada em UTM/SAD69. Essa base está no formato vetorial e foi elaborada pela Prefeitura de Presidente Prudente, entre os anos de 1996 e 1997, a partir da restituição de fotografias aéreas datadas de 1995, na escala de 1:25.000.

Com isso, a imagem TM/Landsat-5 foi georreferenciada no aplicativo SPRING, utilizando transformação polinomial de 1º grau e reamostragem pelo método do vizinho mais próximo. Foram considerados 20 pontos de controle identificados na base vetorial e na imagem TM/Landsat-5 e o erro obtido foi de 0,55 *pixel* (16,5 metros).

4.4 Classificação *fuzzy*

Para realizar a classificação *fuzzy*, a imagem TM/Landsat-5 georreferenciada foi exportada do Banco de Dados criado no SPRING e convertida para o formato do aplicativo IDRISI, o qual permite trabalhar com métodos de classificação relativa.

A partir disso, a primeira etapa da classificação consistiu na coleta de amostras representativas das oito

classes de cobertura da terra identificadas para este trabalho. Durante esse processo, buscou-se amostras representativas que abrangessem a máxima variabilidade das classes, coletando elementos amostrais nas regiões *fuzzy*, isto é, de transição entre duas classes.

Em seguida, foi criada a matriz de partição *fuzzy*, a qual indica os graus de pertinência de cada local de treinamento para cada classe. Essa matriz é gerada na forma de uma tabela formada por *N* colunas (classes de cobertura da terra) e *M* linhas (classes de treinamento), sendo apresentada na Tabela 2:

Tabela 2 – Matriz de partição *fuzzy*.

M \ N	1	2	3	4	5	6	7	8
1	0.9	0	0	0	0.1	0	0	0
2	0	1	0	0	0	0	0	0
3	0	0	1	0	0	0	0	0
4	0	0	0	1	0	0	0	0
5	0	0.05	0	0	0.95	0	0	0
6	0	0	0	0	0	1	0	0
7	0	0	0	0	0.05	0	0.95	0
8	0	0	0.12	0.13	0	0	0	0.75

onde: 1 = área urbanizada; 2 = área de cultivo; 3 = água interior; 4 = floresta; 5 = área com pouca e/ou nenhuma vegetação; 6 = pastagem; 7 = área degradada; 8 = área úmida continental.

Dessa forma, para o conjunto das amostras coletadas para a classe 1 (área urbanizada), por exemplo, foi estimada uma proporção de 0.9 para a classe propriamente dita e uma proporção de ocorrência da classe 5 (área com pouca e/ou nenhuma vegetação) equivalente a 0.1.

Posteriormente, foi realizada a extração de assinaturas para as classes. Esse procedimento gera uma saída onde é atribuído a cada *pixel* um peso proporcional ao seu grau de pertinência na determinação da média, variância e covariância de cada banda para cada classe. Assim, um *pixel* que é predominantemente composto de floresta, por exemplo, terá um grande peso na determinação da assinatura referente a essa classe, mas um peso baixo na determinação da assinatura para outras classes constituintes.

Com isso pôde-se realizar a classificação relativa bayesiana. A saída do processamento constitui um conjunto de imagens (uma por classe) que expressa a probabilidade dos *pixels* pertencerem a cada uma das classes de cobertura da terra, bem como uma imagem adicional que representa o grau de incerteza da classificação.

Finalmente, foi gerada uma saída rígida (convencional) para a classificação relativa. Ou seja, a partir das imagens de probabilidade geradas para cada *pixel* em cada classe, gera-se uma única imagem classificada atribuindo o *pixel* à classe na qual este apresenta maior probabilidade de pertencer, ignorando as demais proporções de contribuição de outras classes.

5 RESULTADOS

5.1 Imagem georreferenciada e caracterização das feições

Com o georreferenciamento das bandas da imagem TM/Landsat-5, estas puderam ser importadas para o Banco de Dados criado no SPRING. Além disso, foi importado também o limite vetorial do município de Presidente Prudente disponibilizado por CPTI (1999). Na Figura 2 é mostrada a composição colorida B3G4R5 da imagem TM/Landsat-5 contrastada, assim como a sobreposição do limite vetorial do município.

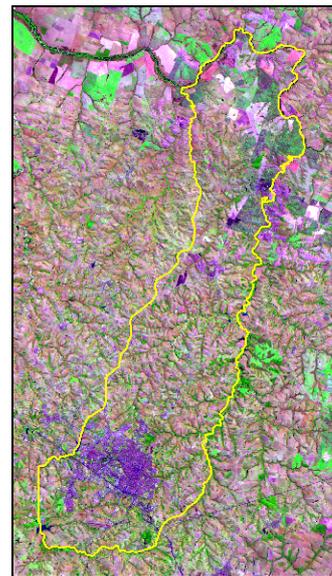
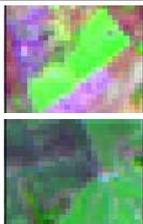
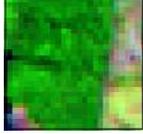


Figura 2 – Composição colorida B3G4R5 da imagem TM/Landsat-5 e sobreposição do limite vetorial do município.

Observando a composição colorida da imagem orbital pôde-se identificar os diferentes padrões de cada feição de interesse relacionada à cada uma das classes de cobertura da terra. A identificação das feições foi realizada com base nos elementos de interpretação de imagem (forma, textura e localização), incluindo a composição colorida considerada. A Tabela 3 apresenta a caracterização das feições regionais de interesse.

Tabela 3 – Caracterização das feições.

Classes	Amostras na imagem	Descrições
Área urbanizada		Está localizada na região sul do limite municipal; apresenta forma regular, textura grossa e cor variando do roxo ao lilás.
Área degradada		Está localizada à esquerda da Rodovia Raposo Tavares (SP-270), próxima à Cidade da Criança; apresenta forma regular, textura grossa e cor lilás.

Área de cultivo		No caso de cultura recém-plantada, apresenta forma regular, textura lisa e cor verde clara. Quando se trata de cultura próxima da colheita, mantém forma irregular, com textura mais rugosa e cor de verde mais escuro.
Pastagem		Apresenta diferentes tipos de formas, com textura geralmente lisa e cores variadas, desde áreas avermelhadas, arroxeadas e amareladas.
Área úmida continental		Apresenta diferentes tipos de formas, com textura geralmente lisa e cores tendendo ao marrom escuro.
Água interior		Apresenta forma irregular, com textura lisa e cor azul escuro, na ausência de sedimentos em suspensão, ou cor azul claro, para corpo d'água com a presença de sedimentos em suspensão.
Floresta		Apresenta forma irregular, textura rugosa e cor verde escuro.
Área com pouca e/ou nenhuma vegetação		Apresenta geralmente forma regular, textura lisa e cor branca, rosa, lilás ou roxa.

5.2 Geração de imagens classificadas e imagem de incerteza

Com a realização da classificação relativa bayesiana foi gerada uma imagem classificada para cada classe temática. Essa imagem indica a probabilidade “a posteriori” que cada *pixel* tem de pertencer às classes de interesse. Na Figuras 3 são apresentadas as imagens de probabilidade “a posteriori” para as classes “área urbanizada” (a) e “pastagem” (b).

Além das imagens de probabilidade foi gerada uma imagem de incerteza, indicando espacialmente as regiões com maior potencial de erro ao gerar uma única saída convencional. Na Figura 4 pode ser vista a incerteza da área de estudo classificada, com destaque para uma região onde verificou-se que a incerteza variou entre 20% e 50%.

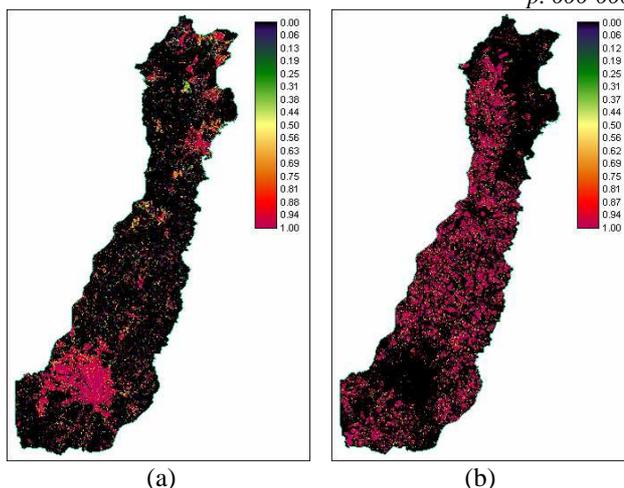


Figura 3 – Imagens de probabilidade “a posteriori” para as classes “área urbanizada” (a) e “pastagem” (b).

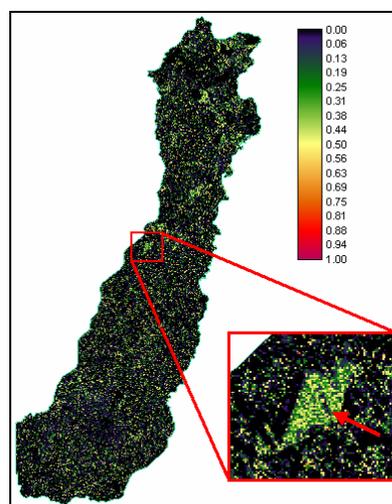


Figura 4 – Incerteza da área de estudo classificada, com destaque para uma região com maior potencial de erro na classificação.

O recorte da imagem de incerteza mostrado na Figura 4 ilustra uma região com maior probabilidade de erro de classificação, o que se comprovou ao gerar uma saída convencional a partir das imagens de probabilidade “a posteriori”. Na imagem classificada final pode-se verificar que essa região, a qual corresponde à classe “área com pouca e/ou nenhuma vegetação”, mostra locais erroneamente classificados como pertencente à classe “área urbanizada”.

Finalmente, após verificar e editar algumas inconsistências (erros de classificação) na saída convencional foi gerado o mapeamento temático da cobertura da terra do município de Presidente Prudente (Figura 5), a partir das classes definidas no nível II do sistema de classificação hierárquico.

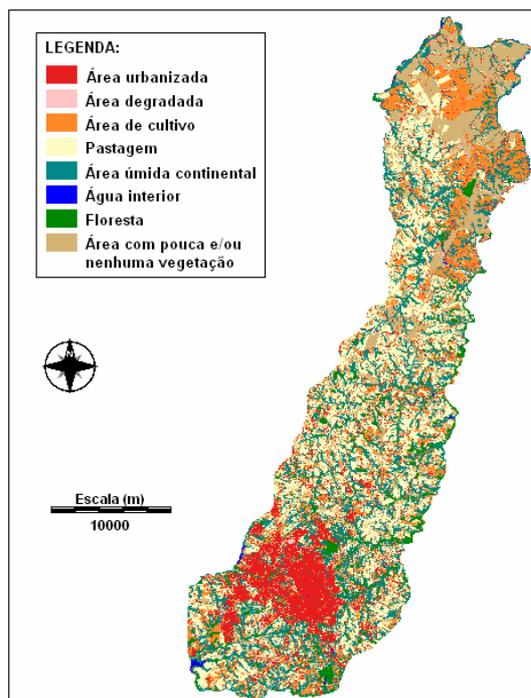


Figura 5 – Mapeamento temático da cobertura da terra do município de Presidente Prudente (SP).

6 CONSIDERAÇÕES FINAIS

O sistema de classificação hierárquico proposto para o mapeamento da cobertura da terra em escala regional é abrangente, contendo um grande conjunto de classes, que pode suprir a necessidade de uma variedade de aplicações e é adequado a diferentes escalas e níveis de detalhamento.

Como indicam os resultados obtidos, a utilização da imagem de média resolução espacial TM/Landsat-5 permitiu realizar o mapeamento temático do município de Presidente Prudente (SP) para o nível II de detalhamento definido no sistema hierárquico.

A classificação da imagem orbital utilizando abordagem *fuzzy* permitiu não apenas identificar a situação da cobertura da terra na área de estudo, mas também estimar as proporções de incerteza do mapeamento. Com isso, verificou-se que os locais de maior incerteza, em geral, proporcionaram maiores erros ao atribuir o *pixel* a uma dada classe de cobertura da terra.

Finalmente, ressalta-se que foi observado que a cobertura predominante no município é pastagem, com, aproximadamente, 26% do total da cobertura da área de estudo. Além disso, classes como “área úmida continental”, “área com pouca e/ou nenhuma vegetação” e “área de cultivo” também são significativas no município. A primeira está relacionada principalmente com as áreas úmidas no entorno de cursos d’água; a segunda, com solos preparados para o plantio; e a terceira, com cultivo temporário, principalmente, com destaque para plantação de cana-de-açúcar.

AGRADECIMENTOS

As autoras deste trabalho agradecem à FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo) pelo financiamento da pesquisa.

REFERÊNCIAS

- ANDERSON, J. R. et al. **A land use and land cover classification system for use with remote sensor data**. Geological Survey Professional Paper 964. Washington: USGS, 1976.
- BÜTTNER, G et al. The CORINE Land Cover 2000 Project. **EARSel eProceedings**, v. 3, n. 3, p. 331-346, 2004. Disponível em: <<http://www.e proceedings.org/>>. Acesso em: 12 jan. 2008.
- CONGALTON, R. G.; GREEN, K. **Assessing the accuracy of remotely sensed data: principles and practices**. New York: Lewis Publishers, 1999.
- COOPERATIVA DE SERVIÇOS, PESQUISAS TECNOLÓGICAS E INDUSTRIAIS (CPTI). **Diagnóstico da situação dos recursos hídricos da UGRHI – 22 Pontal do Paranapanema**: relatório zero. São Paulo: CPTI, 1999. CD-ROM.
- DI GREGORIO, A. **Land Cover Classification System (LCCS), version 2: classification concepts and user manual**. Rome: FAO, 2004.
- EASTMAN, J. R.; LANEY, R. M. Bayesian soft classification for sub-pixel analysis: a critical evaluation. **Photogrammetric Engineering and Remote Sensing**, v. 68, n. 11, p. 1149-1154, 2002.
- FOODY, G. M. et al. Non-linear mixture modelling without end-members using an artificial neural network. **International Journal of Remote Sensing**, v. 18, n. 4, p. 937-953, 1997.
- INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). **Manual técnico de uso da terra**. Manuais Técnicos em Geociências, n. 7. 2. ed. Rio de Janeiro: IBGE, 2006.
- LIMA JR., O. S. **Análise de pontos por função fuzzy**. 2002. 166 f. Dissertação (Mestrado em Informática Aplicada) – Fundação Educacional Edson Queiroz, Universidade de Fortaleza, Fortaleza.
- VILELA, M. F. et al. Avaliação de técnicas de realce e classificação digital na elaboração de um mapa de uso da terra mediante uma imagem TM/LANDSAT-5. **Revista Árvore**, v. 24, n. 2, p. 161-173, 2000.
- WANG, F. Fuzzy supervised classification of remote sensing images. **IEEE Transactions on Geoscience and Remote Sensing**, v. 28, n. 2, p. 194-201, 1990.
- ZADEH, L. A. Fuzzy sets. **Information and Control**, v. 8, p. 339-353, 1965.