

A CLASSIFICAÇÃO CONTÍNUA (FUZZY) NO MAPEAMENTO TEMÁTICO DE ÁREAS DE CERRADO

JORGE ALBERTO BUSTAMANTE BECERRA¹
MARISA DANTAS BITENCOURT¹

¹Laboratório de Ecologia da Paisagem e Conservação – LEPaC, Departamento de Ecologia - IB,
Universidade de São Paulo – USP

Matão - Travessa 14 - nº 321 CEP: 05508-900, Cidade Universitária, São Paulo - SP

jabb@ib.usp.br

tencourt@ib.usp.br

Abstract. This research focused at the *Cerrado* (Brazilian savanna) in São Paulo, Brazil. The aim is, first, the identification of land cover types using Fuzzy supervised classification and, second, the comparison of that classification with a conventional (hard) supervised classification. The algorithms used were the Maximum Likelihood classification and the Wang' fuzzy supervised classification. The Maximum Likelihood classification is based on the probability density function associated with a particular training site signature. The classification algorithm based on fuzzy representation consist of two major steps: estimate of fuzzy parameters from fuzzy training data and a fuzzy partition of spectral space. Encouraging results have been achieved in identifying types and proportions of land covers component in mixed pixels. Improvement in overall classification accuracy has been achieved using the fuzzy mean and fuzzy covariance matrix. For comparison, the hard maximum likelihood classification was also performed at the same data set.

Keywords: thematic mapping, *Cerrado* (brazilian savanna), fuzzy supervised classification, maximum likelihood classification.

1 Introdução

O mapeamento de fisionomias de cerrado, usando imagens multiespectrais, tem sido realizado utilizando, principalmente, a técnica de classificação supervisionada de imagens digitais. O resultado é um mapa com classes de vegetação discretas. No entanto, dependendo das características das classes no terreno e da resolução espacial da imagem, pode ser comum encontrar *pixels* misturados (Foody, 1996). Neste contexto, o objetivo deste trabalho visa identificar as variações fisionômicas do cerrado, ao longo de um gradiente contínuo, usando a técnica de *classificação supervisionada fuzzy* e comparar o desempenho desta classificação (*max_fuz*) com o resultado da classificação supervisionada por máxima verossimilhança (*maxver*).

A classificação convencional baseia-se no teorema de Bayes, que é uma função de densidade de probabilidade associada à assinatura espectral de uma área de treinamento conhecida (Eastman, 1999). Assim, uma das principais suposições desta abordagem Bayesiana é que as áreas de treinamento representam amostras puras das classes analisadas.

Para atingir o objetivo proposto, foi utilizado o procedimento implementado no Idrisi32 para o desenvolvimento de assinaturas *fuzzy*, baseado em Wang (1990) e fundamentada no conceito de conjuntos *fuzzy* (Zadeh, 1965, 1998). Onde cada pixel tem potencialmente algum grau de associação a mais que um conjunto (classe) de acordo com o grau de mistura destas classes dentro do pixel. Utilizando esse modelo são calculadas as médias, variâncias e covariâncias ponderadas, pelo grau de associação ao conjunto *fuzzy* das classes envolvidas.

Finalmente, para avaliar o grau com que cada *pixel* pertence a cada uma das classes de cobertura do solo é utilizado o classificador relativo (*soft classifier*) *Bayclass*. Assim, ao invés de tomar uma decisão rígida (*Booleana*) sobre a associação de cada *pixel* a uma determinada classe, o classificador relativo gera como saída uma imagem em formato de números reais para cada classe, a qual expressa a associação ao conjunto, numa escala de 0 a 1 (Eastman, 1999).

2 Material

Os materiais utilizados no mapeamento temático foram:

- Imagens de satélite Landsat5 TM Bandas 3, 4 e 5 de 20 de maio de 1999 (Figura 1).
- Registros das Características das classes identificadas no campo com ajuda de GPS.
- Material cartográfico das áreas de estudo.
- Material bibliográfico sobre características agropecuárias das áreas estudadas.
- Programas especializados (Idrisi32, ERmapper, Arc/View e Arc/Info, entre outros).

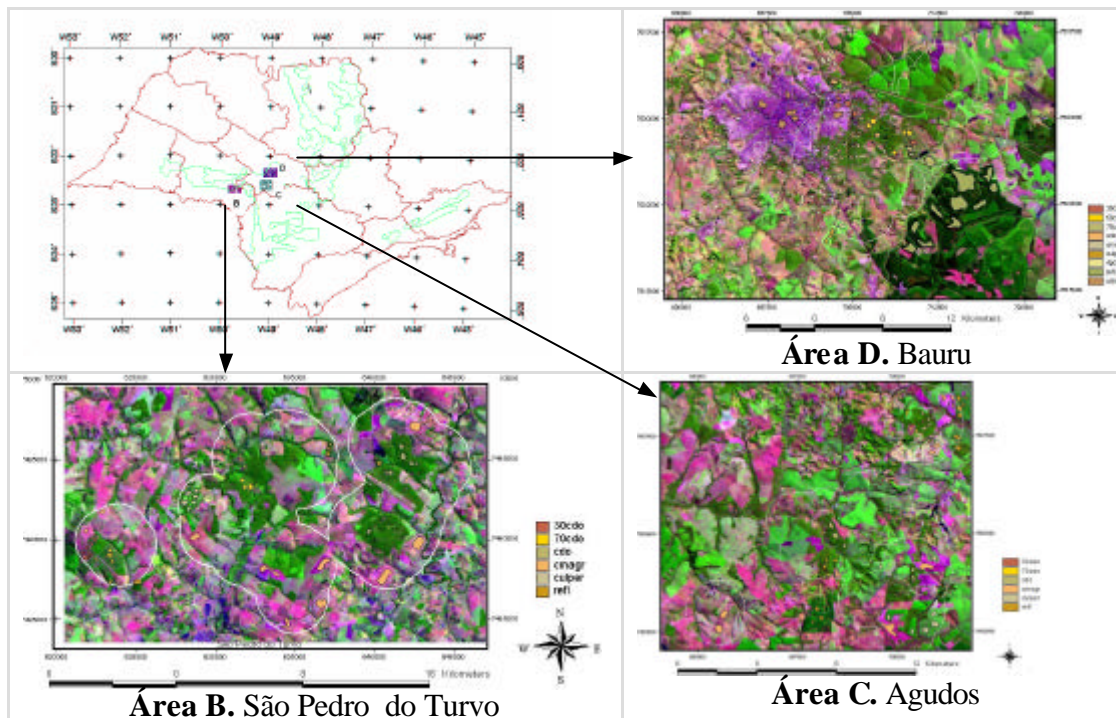


Figura 1. Localização da área de estudo no Estado de São Paulo, mostrando nas áreas definidas por linhas cor verde o Domínio do Cerrado (figura superior à esquerda). E imagens de satélite Landsat5 TM, bandas 3, 4 e 5, data 20 de maio de 1999, das três áreas de estudo. As linhas brancas dos polígonos indicam os limites das áreas de estudo. As classes indicadas nas legendas, correspondem às áreas de treinamento.

O estudo foi conduzido em três áreas, localizadas nos municípios de São Pedro do Turvo (área B), Agudos (área C) e Bauru (área D). No contexto do Estado de São Paulo, estas áreas situam-se na região centro ocidental (Figura 1) e correspondem a Áreas Prioritárias para a Conservação (SEMA, 1997).

3 Métodos

O objeto de estudo é o fragmento de cerrado mais o seu entorno, que corresponde a algum tipo de atividade agropecuária, no caso das Áreas B e C. Já no caso da Área D, o entorno pode corresponder a algum tipo de ocupação rural ou urbana.

A metodologia utilizada para realizar a classificação supervisionada e classificação *fuzzy*, das imagens correspondentes as três áreas de estudo, foi esquematizada no fluxograma da Figura 2.

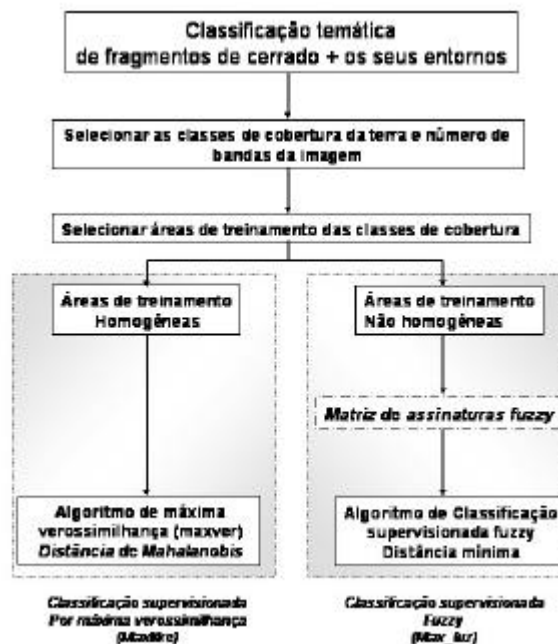


Figura 2. Fluxograma de comparação das classificações *supervisionada por máxima verossimilhança* e *supervisionada fuzzy*.

De acordo com o fluxograma acima, a matriz de assinaturas *fuzzy* (Wang, 1990) é a principal diferença entre a classificação *fuzzy* e a classificação por máxima verossimilhança.

A matriz de assinaturas *fuzzy* também é conhecida como matriz de desenvolvimento de assinaturas *fuzzy*, a qual está detalhada no item 3.1.

3.1 Matriz de desenvolvimento de assinaturas fuzzy

Foram elaboradas duas matrizes, uma para a área D (Tabela 1) e outra para as áreas B e C (Tabela 2). A matriz da área D, possui mais classes de cobertura da terra, isto devido às características da região, em virtude da proximidade da área urbana com os fragmentos de cerrado. Isto permite identificar as classes: cerrado degradado (dgcdo) e áreas de cerrado com 50% de cerrado e 50% de campo/agricultura (50cdo), além da classe áreas urbanas (urban).

Ainda na elaboração da matriz de assinaturas *fuzzy* da área D, foram utilizadas, além do material indicado no item materiais, aerofotografias da região que permitiram identificar as 9 classes de cobertura da terra.

Tabela 1. Matriz de desenvolvimento de assinaturas *fuzzy*, elaborado para a área D, contendo 9 classes de cobertura da terra

IDR_ID	cdo	dgcdo	cmagr	refl	urban	culper	70cdo	50cdo	30cdo
cdo	1	0	0	0	0	0	0	0	0
dgcdo	0	1	0	0	0	0	0	0	0
cmagr	0	0	1	0	0	0	0	0	0
refl	0	0	0	1	0	0	0	0	0
urban	0	0	0	0	1	0	0	0	0
culper	0	0	0	0	0	1	0	0	0
70cdo	0	0	0,3	0	0	0	0,7	0	0
50cdo	0	0	0,5	0	0	0	0	0,5	0
30cdo	0	0	0,7	0	0	0	0	0	0,3

Tabela 2. Matriz de desenvolvimento de assinaturas fuzzy, elaborado para as áreas B e C, contendo 6 classes de cobertura da terra.

IDR_ID	30cdo	70cdo	cdo	cmagr	Culper	refl
30cdo	0,3	0	0	0,7	0	0
70cdo	0	0,7	0	0,3	0	0
Cdo	0	0	1	0	0	0
Cmagr	0	0	0	1	0	0
Culper	0	0	0	0	1	0
Refl	0	0	0	0	0	1

3.2 Característica das classes identificadas no campo

As classes de cobertura da terra identificadas e utilizadas na elaboração das matrizes de partição ou assinaturas *fuzzy* das Tabelas 1 e 2, são:

cdo: áreas definidas como cerrado *ss* e cerrado

dgcd: áreas de cerrado *ss* degradado, principalmente por causa da presença de gado

cmagr: áreas que corresponde a fisionomia de campo e áreas utilizadas na agricultura

refl: áreas ocupadas por reflorestamento, principalmente de eucalipto e pinus.

urban: áreas que correspondente a região urbana e áreas de solo nu

culper: áreas com culturas permanentes

70cdo: áreas de transição com 70% de cdo e 30% de cmagr

50cdo: áreas de transição com 50% de cdo e 50% de cmagr

30cdo: São áreas de transição com 30% de cdo e 70% de cmagr

3.3 Incerteza da classificação

Para extrair a incerteza média associada a cada uma das classes de cobertura da terra indicadas na classificação supervisionada *fuzzy* (maxfuz), é utilizado o classificador relativo Bayclass. Este classificador relativo avalia o grau com que cada pixel pertence a cada uma das classes de cobertura da terra.

4 Resultados e discussões

A partir das classificações supervisionada (maxlike), *fuzzy* (max_fuz) e a análise da incerteza, os resultados e discussões obtidos para as três áreas de estudo são:

4.1 Classificação supervisionada versus classificação *fuzzy*

Com o intuito de comparar o desempenho dos dois classificadores, em cada área de estudo, é apresentado um histograma e o índice de concordância Kappa.

Este histograma é gerado a partir da matriz de tabulação cruzada de Maxlike com Max_fuz e mostra a variação ou similaridade, em termos percentuais para cada classe de cobertura analisada. Já a similaridade entre as duas imagens é evidenciada pelo índice de concordância de Kappa.

4.1.1 Área B

Nesta área, que corresponde à região de São Pedro do Turvo, os resultados dos dois classificadores, usando seis classes de cobertura da terra, pode ser observados nas Figuras 3 e 4. De acordo com o índice de concordância de Kappa, a similaridade entre as duas

classificações é da ordem de 0,8826, isto indica similaridade significativa, com algumas diferenças importantes que podemos observar no histograma da Figura 3.

Índice de Kappa: 0,8826

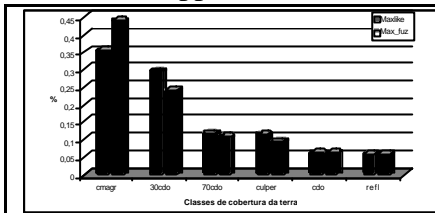


Figura 3. Histograma das variações percentuais das classes de cobertura da terra usando os métodos de classificação supervisionada (Maxlike) e classificação *fuzzy* (Max_fuz) na área de estudo B.

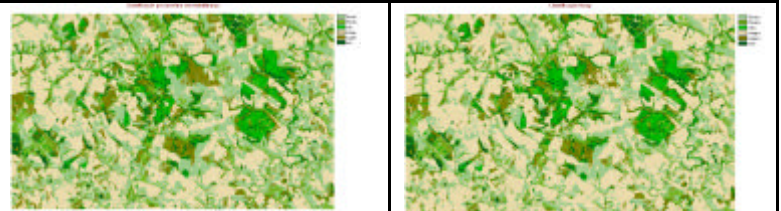


Figura 4. Resultado das classificações, supervisionada por máxima verossimilhança (parte superior da Figura) e classificação *fuzzy* (parte inferior), da área de estudo B.

Segundo Figura 3, existe um claro padrão, que indica duas classes dominantes, cmagr e 30cdo, tanto na classificação convencional, quanto na classificação *fuzzy*. Comparando o resultado das duas classificações, podemos observar, primeiro, duas classes de cobertura (cdo e refl) com resultados idênticos para as duas classificações, segundo, uma classe (cmagr) com mais pixels alocados pela classificação *fuzzy*, e terceiro, três classes (30cdo, 70cdo e culper) caracterizadas por apresentar mais pixels alocados pela classificação convencional do que com a classificação *fuzzy*.

Estes resultados podem ser compreendidos utilizando a Tabela 2, matriz de desenvolvimento de assinaturas *fuzzy*, onde observa-se que, nas colunas a classe cmagr recebe pixels das classes 30cdo e 70cdo, além da própria classe. Desta maneira, é explicada a maior proporção de pixels alocados pela classe cmagr, quando analisado o desempenho da classificação *fuzzy*.

Ainda na Tabela 2, quando analisamos as linhas, observamos que, a classe 30cdo e classe 70cdo, compartilham pixels com a classe cmagr. Explicando desta maneira a menor proporção de pixels alocados pelas duas classes, quando comparado o desempenho da classificação *fuzzy* com a classificação convencional.

4.1.2 Área C

Nesta área, que corresponde à região de Agudos, os resultados dos dois classificadores, usando seis classes de cobertura da terra, podem ser observados nas Figuras 5 e 6. A similaridade entre as duas classificações é da ordem de 0,8785, de acordo com o índice de concordância de Kappa, isto indica similaridade significativa, com algumas diferenças importantes que podemos observar no histograma da Figura 5.

Segundo Figura 5, existe um padrão, que indica duas classes dominantes, cmagr e 30cdo, para a classificação convencional. Já, na classificação *fuzzy* existe claro domínio da classe cmagr, sobre as outras classes de cobertura da terra.

Comparando o resultado das duas classificações, podemos observar, primeiro, duas classes de cobertura (cdo e refl) com resultados similares para as duas classificações, segundo, três classes (cmagr, 70cdo e culper) com mais pixels alocados pela classificação *fuzzy*, e terceiro, uma classe (30cdo) caracterizada por apresentar mais pixels alocados pela classificação convencional do que a classificação *fuzzy*.

Overall Kappa: 0,8785

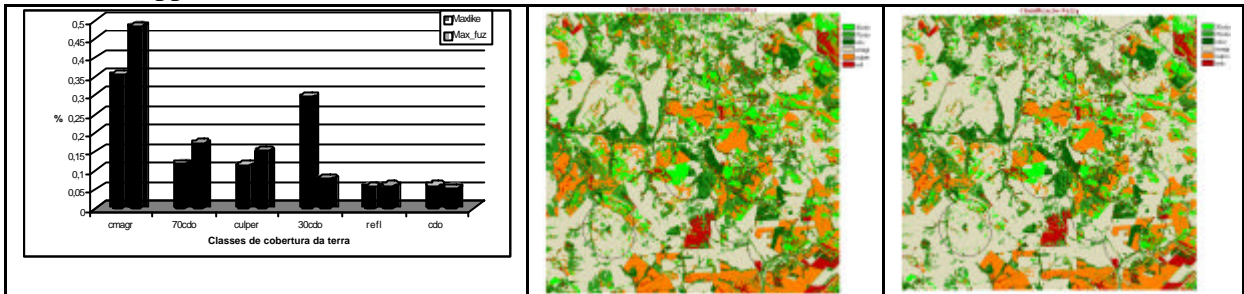


Figura 5. Histograma das variações percentuais das classes de cobertura da terra usando as métodos de classificação supervisionada (Maxlike) e classificação *fuzzy* (Max_fuz) na área de estudo C.

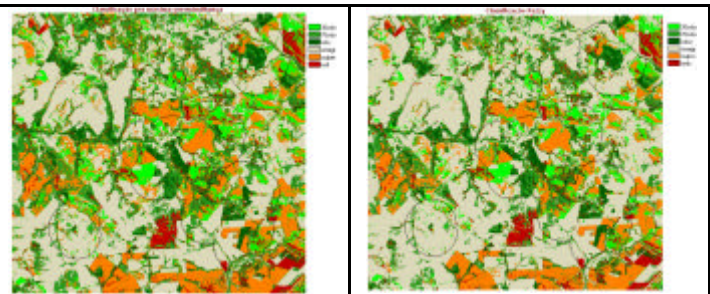


Figura 6. Resultado das classificações, supervisionada por máxima verossimilhança (parte superior da Figura) e classificação *fuzzy* (parte inferior), da área de estudo C.

A interpretação destes resultados para a área de estudo C, é similar à interpretação dos resultados obtidos na área B, utilizando a Tabela 2. Assim, é explicada, a seguinte relação: maior número de pixels alocados na classe cmagr contra menor número de pixels na classe 30cdo, quando comparado o desempenho da classificação *fuzzy* com a classificação convencional.

4.1.3 Área D

Nesta área, que corresponde à região de Bauru, os resultados dos dois classificadores, usando nove classes de cobertura da terra, pode ser observado nas Figuras 7 e 8. De acordo com o índice de concordância de Kappa, a similaridade entre as duas classificações é da ordem de 0,9307, isto indica similaridade significativa, a maior das três áreas estudadas, com algumas diferenças que podemos observar no histograma da Figura 7.

Índice de Kappa: 0,9307

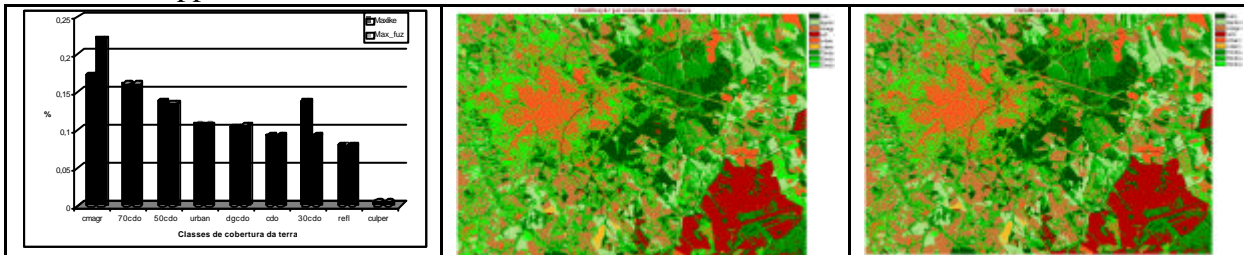


Figura 7. Histograma das porcentagens das classes de cobertura da terra da: classificação supervisionada (Maxlike) e classificação *fuzzy* (Max_fuz), na Área D.

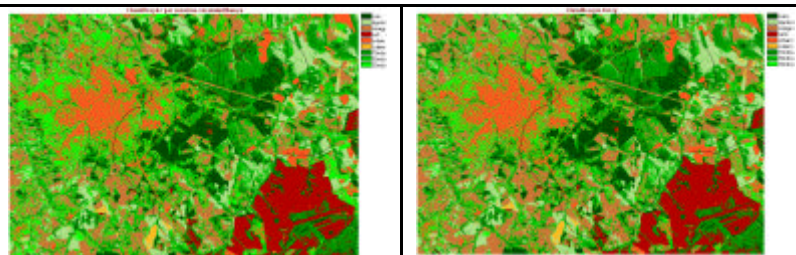


Figura 8. Resultado das classificações, supervisionada por máxima verossimilhança (parte superior da Figura) e classificação *fuzzy* (parte inferior), da área de estudo D.

Segundo a Figura 7, existe um padrão similar ao obtido nas áreas de estudo B e C, embora o número de classes de cobertura da terra seja maior nesta área de estudo. Isto é, no resultado da classificação *fuzzy*, domínio relativo da classe cmagr, sobre as outras classes. Comparando o resultado das duas classificações, podemos observar, primeiro, uma classe (cmagr) com mais pixels alocados pela classificação *fuzzy*, e segundo, outra classe (30cdo) caracterizada por apresentar mais pixels alocados pela classificação convencional do que a

classificação *fuzzy*. As outras restantes classes (70cdo, 50cdo, urban, dgdo, cdo, refl e culper), na comparação, apresentam pouca variação entre os resultados das duas classificações.

Estes resultados podem ser entendidos utilizando a Tabela 1, matriz de desenvolvimento de assinaturas *fuzzy*, para nove classes de cobertura da terra.

A interpretação desta tabela é semelhante à interpretação dos resultados obtidos na área C, onde também foi evidenciada a seguinte relação: maior número de pixels alocados na classe *cmagr* contra menor número de pixels na classe 30cdo, quando comparado o desempenho da classificação *fuzzy* com a classificação convencional.

4.2 Incerteza da classificação

Na imagem da área B, de acordo com a Tabela 6, incerteza associada a cada uma das classes de cobertura do solo, e a Figura 9, podemos observar que a classe com menor incerteza média é *cmagr* (0,11) e a classe com maior incerteza média é *culper* (0,36).

Tabelas 6, 7 e 8. Incerteza média associada a cada uma das classes de cobertura do solo, nas áreas de estudo

Tabela 6. Área B

ID	Classe	Média
1	<i>cdo</i>	0,243
2	<i>cmagr</i>	0,345
3	<i>refl</i>	0,348
4	<i>culper</i>	0,112
5	<i>70cdo</i>	0,361
6	<i>30cdo</i>	0,341

Tabela 7. Área C

ID	Classe	Média
1	<i>cdo</i>	0,375
2	<i>cmagr</i>	0,215
3	<i>refl</i>	0,242
4	<i>culper</i>	0,053
5	<i>70cdo</i>	0,259
6	<i>30cdo</i>	0,451

Tabela 8. Área D

ID	Classe	Média
1	<i>cdo</i>	0,325
2	<i>dgdo</i>	0,224
3	<i>cmagr</i>	0,281
4	<i>refl</i>	0,048
5	<i>urban</i>	0,221
6	<i>culper</i>	0,204
7	<i>70cdo</i>	0,436
8	<i>50cdo</i>	0,520
9	<i>30cdo</i>	0,357

Na área C, a incerteza média associada a cada uma das classes de cobertura da terra poder ser observada na Tabela 7. Onde a classe *cma*, igual que na área B, apresenta menor incerteza média (0,05), por outro lado, a maior incerteza média está associada à classe *refl* (0,45), igual ao encontrado na área B.

Na área D, onde são analisadas nove classes de cobertura, os resultados são diferentes (Tabela 8), assim, a menor incerteza média esta associada a classe *refl* (0,04) e a maior incerteza média esta associada a classe *50cdo* (0,52). Já, a classe *cmagr*, que nas duas áreas B e C possuiu a menor incerteza média, aqui está na ordem de 0,28.

Os resultados obtidos nas três áreas de estudo, quanto a incerteza da classificação, seguem a seguinte seqüência: classes caracterizadas por possuir menor heterogeneidade espectral, possuem menor incerteza média e indicam áreas de treinamento homogêneas. Esta característica é observada nas classes *cmagr*, nas áreas B e C, e *refl* na área D. Já, as classes com maior heterogeneidade espectral, indicam áreas de treinamento heterogêneas e possuem maior incerteza média. Comportamento observado nas classes *culper* (área B), *refl* (área C) e *50cdo* (área D).

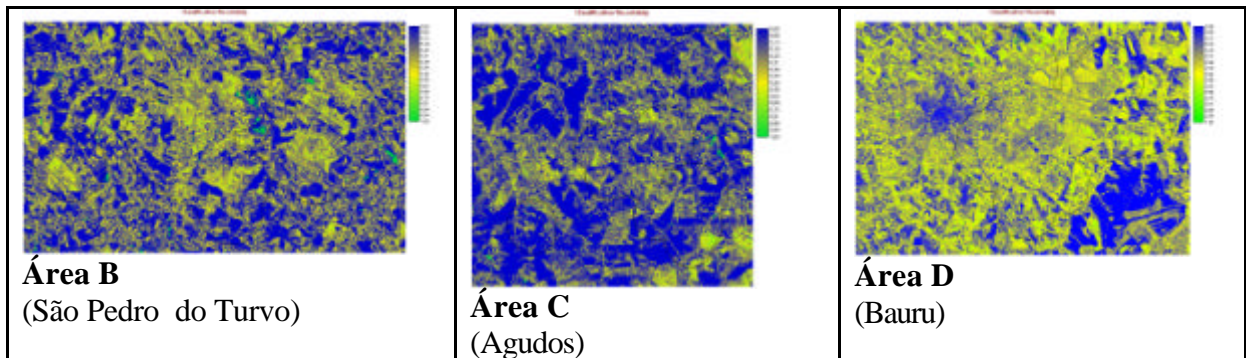


Figura 9. Resultado da Incerteza da classificação, das três áreas de estudo, calculada a partir do módulo Bayclass (Idrisi32), que indica a incerteza numa escala contínua que vai de 0 a 1. Sendo 1 (verde) a máxima incerteza e 0 (Azul) a mínima incerteza.

Quando observadas as três imagens da Figura 9, da incerteza da classificação de cada uma das áreas de estudo, podemos notar que, as áreas com tons azuis correspondem à menor incerteza na classificação. Já as áreas com tons verdes, correspondem às áreas com maior incerteza da classificação. Assim, podemos notar, na Figura 9, que a área de estudo D possui a maior incerteza da classificação. Este fato é explicado pelo maior número de classes de cobertura (9), os quais não foram discriminados adequadamente por causa de dois fatores. Primeiro, devido à limitação espectral das imagens utilizadas (Schwengerdt, 1997), somente três bandas; e segundo, devido à imprecisão na definição das áreas de treinamento.

Considerações finais

A classificação fuzzy permitiu identificar e classificar classes contendo pixels misturados, ou seja, gradientes vegetacionais. O sucesso da classificação fuzzy dependeu da identificação, no terreno, das áreas de transição e suas respectivas proporções.

Agradecimentos

Ao projeto Biota-FAPESP: Viabilidade de conservação de remanescentes de cerrado *lato sensu* no Estado de São Paulo (processo # 98/05251-0) e à FAPESP pela bolsa de doutorado.

Referências

- Eastman, J. R. 1999. *Idrisi32. Guide to GIS and image processing Vol 1 and Vol 2*. Clark University, Worcester, MA.
- Foody, G. M. Fuzzy modelling of vegetation from remote sensed imagery. *Ecological Modelling*, 85, 3-12, 1996.
- Secretaria do Meio Ambiente do Estado de São Paulo. *Cerrado: Bases para conservação e uso sustentável das áreas de cerrado do Estado de São Paulo*. PROBIO/SP 1997. 184p.
- Schwengerdt, R. A. *Remote Sensing. Models and Methods for Image Processing*. Second Edition. A. P. 1997.
- Wang, F. Fuzzy supervised classification of remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 28, 2, 194-201, 1990.
- Zadeh, L.A. Fuzzy set. *Information and Control*, 8pp. 338-353, 1965.
- _____. Fuzzy logic. *IEEE Computer*, pp. 83-92, 1998.