Aplicação dos Índices KAPPA & PABAK na validação da classificação automática de imagem de satélite em Feira de Santana-BA

Jocimara Souza Britto Lobão [1]
Washington de Jesus Sant'anna da França-Rocha [2]
Ardemírio de Barros Silva [3]

[1 2 3]Universidade Estadual de Feira de Santana – UEFS
Br 116, Km 3 – Campus Universitário
GP-SIG (Grupo de Pesquisas em Sistemas de Informações Geo-referenciadas) – UEFS
juci.lobao@ig.com.br [1]
wrocha@terra.com.br [2]
abarros@uefs.br [3]

Abstract. This work evaluates the Kappa Index use on accuracy assessment of the processings accomplished in a LANDSAT ETM+ sub-setting scene, regarding the municipal district of Feira de Santana, aiming at to delimit your urban area. In the analysis, is considered the inherent paradoxes to the Kappa Index and the application of PABAK is discussed, in the correction of the problems of marginal distribution of the data. The results demonstrate that although there is no distortion in the data for the case in study, the application of PABAK countersigns the calculated index owing, therefore, to be considered to avoid inherent paradoxes to the Kappa Index.

Palavras-chave: Kappa Index, Urban, accuracy, PABAK, Ìndice Kappa**,** acurácia, Urbano.

1- Introdução

Com o advento de novas tecnologias, a quantidade de dados e informações possíveis de serem adquiridas vem se tornando imensurável. Só com o Processamento Digital de Imagens de Satélite - PDIS, uma grande e variada quantidade informações são produzidas. No entanto, não é comum trabalhos que avaliem a acurácia desses dados. A validação dos dados gerados por meio de PDIS é de fundamental importância para que se possa integrá-lo num Sistema de Informações Geográficas SIG, evitando interpretações errôneas. Segundo Silva, (1999) a utilização de SIG, não garante soluções corretas se não houver controle da qualidade dos dados a serem incorporados. Nesta perspectiva, sem mensuração de erros, como avaliar a qualidade dos dados gerados e garantir que as informações abstraídas não se constituem de impróprios para o uso destinado?

Este trabalho contribui para evidenciar a importância da validação de dados adquiridos por meio de PDIS, visando a produção cartográfica, uma vez que a confiabilidade das análises geográficas decorrentes desses dados está diretamente ligada a sua qualidade. Aponta ainda algumas possibilidades para sua validação.

A acurácia dos dados geográficos deve ser observada levando-se me consideração três pontos fundamentais: a posição dos dados; a época de aquisição dos dados e a temática investigada. Nesta lógica, três tipos de erro diferentes devem ser avaliados: O erro posicional, o erro temporal e o erro temático (LINGNAU, *et. al*., 2003). Neste trabalho será discutido a metodologia para avaliação de erros temáticos, onde o índice KAPPA calculado com duas diferentes formas de construção do mapa de verdade de campo foi comparado aos resultados obtidos com o PABAK (*Prevalence and Bias Adjusted Kappa*).

Em trabalhos anteriores, delimitou áreas urbanas no município de Feira de Santana-BA, e realizou uma primeira estimativa de erro com base nas amostras de controle para as classificações, utilizando o Índice Kappa. (LOBÃO, 2003 <http://lagavulin.ltid.inpe.br:1905/col/ltid.inpe.br/sbsr/2003/03.27.11.57/doc/14_091.PDF>). Recentemente,

este trabalho foi ampliado apresentando, diferentes métodos de validação e distribuição de amostras em campo foram utilizadas para mensurar erros neste mapa (método randômico e dos itinerários). (LOBÃO, 2004 – no prelo).

Nesta perspectiva, o objetivo central deste trabalho foi analisar os resultados obtidos com base no Índice de Kappa, obtido por meio de desses dois procedimentos diferenciados de aquisição do mapa de verdade de campo e verificar se há paradoxos. Como referência utilizou-se o coeficiente PABAK.

A área teste localiza-se na zona limítrofe entre o semi-árido e a mata atlântica, a cerca de 100 km da capital entre as coordenadas 38º51’23’’WG e 39º18’32’’WG de longitude e 11º57’16’’S e 12º25’51’’S de latitude, na área correspondente ao município de Feira de Santana que é o maior do interior do estado da Bahia, com população estimada em 481,137 mil habitantes (IBGE, 2001). **(Figura 1**)

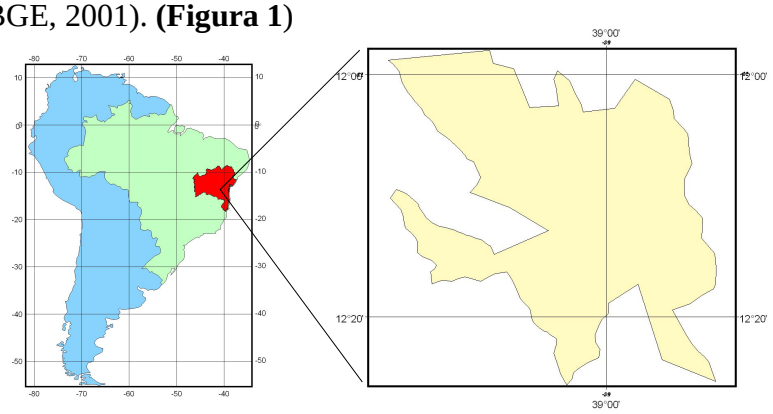


Figura 1 – Localização da área de estudo

2- Material e métodos

Para a realização deste trabalho utilizou-se como base os processamentos realizados por Lobão et. al. (2003) no recorte da cena 216/68 do sensor LANDSAT 7 ETM+, obtida em 05/10/2001 e resolução espacial de 30m. Foram utilizadas três bandas em fusão IHS (Hue, Saturation, Intensity) com uma banda pancromática com resolução de 15m. Utilizou-se também, como dados adicionais: 1 - o mapa digital dos limites do município (IBGE); 2 - mapas de limites, distritos e bairros do município de Feira de Santana, adquiridos junto a Prefeitura Municipal de Feira de Santana, PMFS/IBGE.

O trabalho foi realizado em três etapas distintas:

1. Revisão bibliográfica, com levantamento de dados sobre a região, sobre acurácia, validação e geoestatística.
2. Elaboração lógica/conceitual da metodologia a ser empregada (métodos de coleta e distribuição das amostras, seleção dos coeficientes utilizados, etc...)
3. Etapa de Laboratório, para tratamento, organização, correção, e processamento dos dados com base em modelos de qualidade.

O fluxograma abaixo (**Figura 2**) demonstra as etapas do trabalho.

Para implementar um modelo que avalie a qualidade dos dados Lobão *et al* (no prelo) testou dois procedimentos distintos de coletar os dados para a confecção do mapa de verdade de campo, (a distribuição aleatória de pontos de controle e o método dos itinerários) e partir daí a matriz de confusão. Com base nestes resultados, e da bibliografia consultada, observou-se que existia a possibilidade de paradoxos que são inerentes ao IK é que testou-se o PABAK

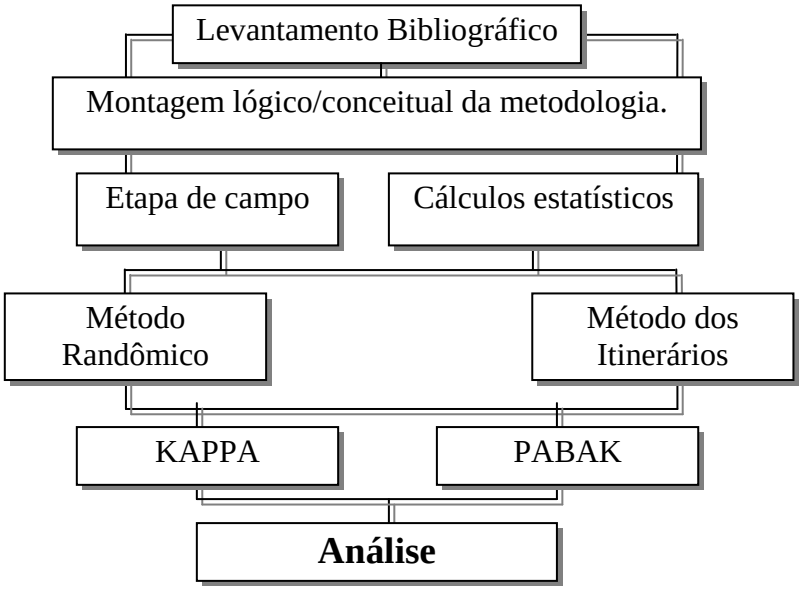


Figura 2- Fluxograma do processamento da imagem

3- Resultados e Discussões

O produto obtido por meio de processamento digital de imagens, pode variar em função da cena, da área de estudo, da época em que foi adquirida a imagem e dos métodos utilizados para extrair informações da imagem, mas estes são fundamentais para o alcance do objetivo a que o pesquisador se propõe. Diante dessa variação de possibilidades é praticamente impossível obter dados exatos. A incerteza e o erro são inerentes a qualquer aquisição de dados.

Em estatística, a precisão é uma grandeza que mede a dispersão, ou seja, a repetição dos dados próximos a um valor e o cálculo a acurácia está relacionada a veracidade de obtenção de dados em um intervalo de probabilidade especificado. Logo, medir a acuraria transcede o processo de mensurar a precisão dos dados.

Entendendo o erro, em sentido mais amplo, como a diferença entre um dado adquirido e o seu valor correto, mensurá-lo é imprescindível, a fim de se definir sua aplicabiliadade.

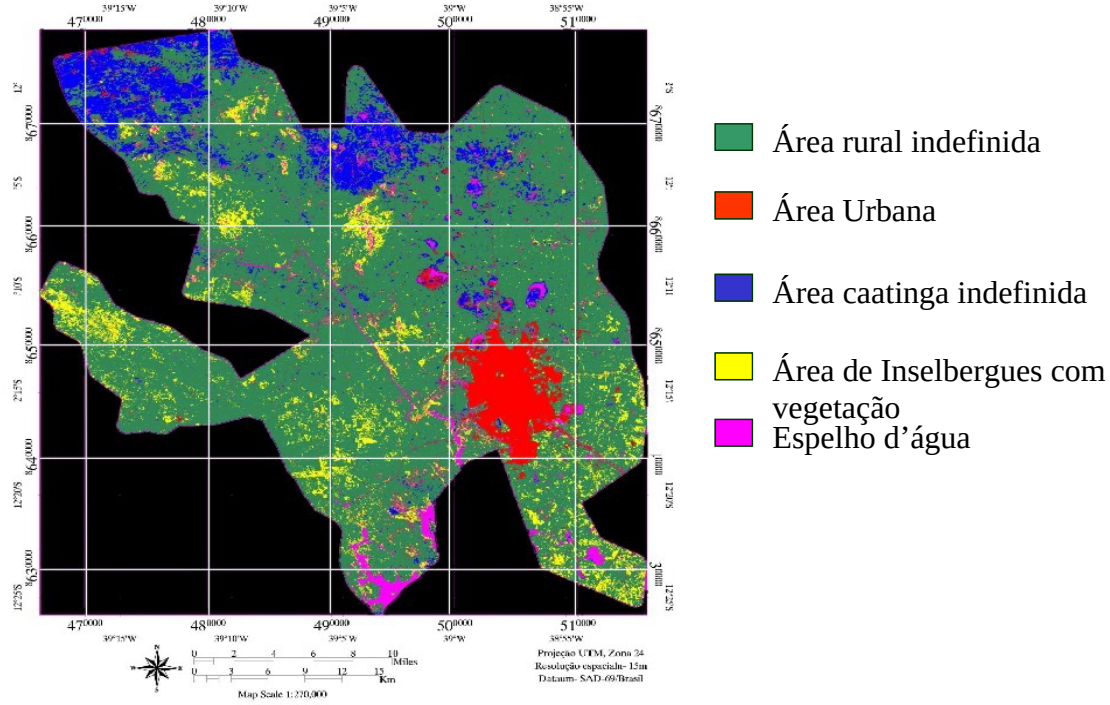
Neste estudo foi tomado como base a classificação supervisionada pelo método de Máxima verossimilhança (MAXVER), processado por Lobão, *et al* (2003) (**Figura 3**) que discute a dificuldade de separação automática de áreas urbanas, pois alvos urbanos são muito heterogêneos e similares a muitos outros não urbanos.

Embora o tema deste estudo seja o erro temático, não é possível dissociá-lo dos demais tipos de erros (posicional e temporal), pois se houver grande discordância posicional a validação estará comprometida. E como se trata de mapeamento de área urbana, onde há uma dinâmica acelerada em função do processo de urbanização, o erro temporal se torna muito relevante.

O erro posicional foi calculado por Lobão (2003), obtendo um RMS (Root Mean Square) de 0,99 pixel, ou seja, o intervalo de incerteza é inferior a 15m (tamanho do pixel processado).

Para o erro temporal tem-se uma diferença de um ano e meio entre a data de obtenção da imagem e os trabalhos de campo realizados. Embora se considere a dinamicidade com que o espaço urbano é modificado, não são perceptíveis as diferenças da imagem para o real. Nas campanhas de campo, áreas recentemente loteadas e/ou ocupadas foram identificadas a partir do depoimento dos moradores e desconsideradas como referências para a imagem de 2001.

O terceiro tipo de erro a ser considerado possui um peso mais significativo e é objeto de maior interesse neste trabalho – o erro temático.



Fonte: Lobão, 2003

Figura 3 – Imagem classificada pelo método MAXVER

Para validar o que foi constatado visualmente nas classificações, construiu-se a matriz de confusão, que representa a distribuição de percentagens de pixel classificados de forma correta ou erroneamente (CONGALTON, 1991).

Neste trabalho escolheu-se o Índice Kappa, para avaliar a acurácia temática por ser mais sensível as variações de erros de omissão e inclusão (COHEN, 1960). Sua grande vantagem é que para o seu cálculo não se incluem somente os elementos da diagonal principal e sim todos os elementos da matriz de erro (MOREIRA, 2001).

No cálculo do coeficiente Kappa é necessário a construção de um mapa de verdade de campo, para que se possa fazer uma tabulação cruzada indicando a proporção de casos presentes e/ou ausentes nos mapas: Mapa classificado (1) e o Mapa real (2). Considerando uma situação de com apenas duas classes o resultado é expresso em uma tabela onde:, a célula *a* indica a proporção dos casos em que o real (1) e o classificado (2) são corretos; *b*, quando (1) for correto e a (2) errado; *c*, quando a observação 1 for correto e a 2 errado; e a *d*, quando ambas forem errados (**tabela 1**).

Tabela 1 – Tabulação cruzada

		Mapa Classificado	
		Sim	Não
Mapa Real	Sim	a	b
	Não	c	d

Como mapa de verdade de campo, para construir a matriz de confusão, utilizou-se duas opções diferentes:

1- Pontos randômicos, dispersos por toda a área de estudo – PR;

2- Coleta de novos pontos de controle espacializados através do método dos itinerários – PI descrita por BELART, PINILLA & ARIZA, (2002).

Os métodos de construção desses mapas de verdade de campo são discutidos por Lobão *et al*, (no prelo).

O Índice de Kappa - IK, foi calculado inicialmente para todo o município incluindo todas classes mapeadas (**equação 1**) e posteriormente apenas para a classe urbano que é o foco deste trabalho (**equação 2**).

Equação1 – IK calculado para toda a área de estudo

$$Po = a+d$$
$$Pe = (a+b)(a+c) + (b+d)(c+d)$$
$$IK = (Po - Pe) / (1 - Pe)$$

Equação2 – IK calculado para a classe urbano

$$Pe_1 = a/T$$
$$Pe_2 = (a+c)/T$$
$$Pe_3 = (a+b)/T$$
$$IK = (Pe_1 - Pe_2 * Pe_3)/(Pe_2 - Pe_2 * Pe_3)$$

Os coefficientes calculados pelo IK são qualificados com base na tabela desenvolvida por Landis e Koch, 1977, in: Moreira, 2001 e revelam uma excelente qualidade para classificação (**tabela 2**).

Tabela 2- IK obtido pelos três métodos (ASTC, PR e PI)

Índice Kappa					
Amostras de treinamento - ASTC		Randômicos - PR		Itinerários - PI	
Total	urbano	Total	urbano	Total	urbano
0,74	-	0,92	0,91	0,88	0,92
Muito Boa	-	Excelente	Excelente	Excelente	Excelente

A qualidade do mapeamento já apontado desde o início com o IK obtido com as amostras de treinamento (calculadas sem trabalho de campo) (**tabela 2**) . Esta amostragem não é muito apropriada por tender a reproduzir os erros atribuídos na classificação (Lobão - no prelo). Por isso há necessidade da construção de mapas de verdade de campo com métodos mais adequados. Desta forma, são muito mais significativas as matrizes obtidas com os métodos: Randômico e dos Itinerários. Por outro lado, em função de sua forma de distribuição, a quantidade de pontos concentrados na área urbana é muito maior no método dos itinerários do que no método randômico. E por isso a indicação de maior acurácia para a classe urbano no método dos itinerários se torna mais relevante. Nesta perspectiva a variação do índice calculado para toda a área e para a classe urbano separadamente deve-se a especificidade dos métodos empregados.

Embora o IK, seja amplamente utilizado alguns paradoxos são descritos na utilização deste coeficiente em função da distribuição marginal dos dados na matriz de confusão. (HOEHLER, 2000; FEINSTEIN, 1990).

Observando os exemplos abaixo verifica-se a presença de alguns paradoxos associados com os efeitos existentes na proporções marginais da tabulação cruzada dos dados. (**tabela 3A, 3B e 3C**). Em todos os três casos apresentam uma idêntica concordância entre os dados

(Po = 300), entretanto, o IK apresenta-se muito baixo no exemplo 3A, médio no 3B e alto no 3C. Isso decorre da superioridade existente entre o número de "sim" e o número de "não". Logo, a proporção dos resultados positivos é substancialmente diferentes.

Tabela 3A, 3B e 3C – Exemplos de paradoxos do IK

3A	Sim	Não	
Sim	270	30	300
Não	0	0	0
	270	30	300

IK = 0,0

3B	Sim	Não	
Sim	240	30	270
Não	0	30	30
	240	60	300

IK = 0.62

3C	Sim	Não	
Sim	135	30	165
Não	0	135	235
	135	265	300

IK = 0,8

Os resultados do Índice Kappa acima demonstram-se insuficientes para assegurar a acurácia dos dados. Embora o IK seja universalmente utilizado para validar dados, nos exemplos anteriores sua aplicação, mostra-se insuficiente.

Paradoxos dessa natureza podem ser ajustados com a aplicação do PABAK (*Prevalence and Bias Ajusted Kappa*). HOEHLER, (2000) demonstra que quando o total do efeito de prevalência positiva resultante é substancialmente diferente de 50% podem ocorrer estes paradoxos, necessitando assim que a partir do IK, seja calculado o coeficiente PABAK.

O cálculo do PABAK é descrito na **equação 3**.

Equação 3 – PABAK

$$PABAK = 2Po - 1 = IK(1 - PI^2 + BI^2) + PI^2 - BI^2$$

Onde: $BI = b - c$

$PI = a - d$

Aplicando o coeficiente PABAK para os exemplos 3A, 3B e 3C tem-se:

Exemplo	Kappa	PABAK
3A	0.0	0.82
3B	0.62	0.82
3C	0.8	0.82

Verifica-se que embora o IK seja reconhecido como instrumento de validação é de fundamental importância que se considere os seus erros marginais inerentes e que sejam realizados outros testes para identificar esses paradoxos e ajustá-los.

Nesta perspectiva, calculou-se o coeficiente PABAK para o mapa classificado com o algoritmo MAXVER no município de Feira de Santana. (**Tabela 4**), com base nos dois métodos de distribuição das amostras.

Estes resultados comprovam que no caso específico deste mapa não há nenhum problema com a distribuição marginal dos dados, pois os resultados foram similares, diferindo apenas nas casas centesimais. Logo, a classificação automática em questão é um produto de excelente acurácia podendo ser incorporado a qualquer Sistema de Informações Geográficas sem que se corra o risco de imprecisões e especificamente para esta classificação não há paradoxos associados ao IK.

Tabela 4 – Comparação dos resultados ÍNDICE KAPPA & PABAK

	Randômicos - PR		Itinerários - PI	
	Total	urbano	Total	urbano
KAPPA	0,92	0,91	0,88	0,92
	Excelente	Excelente	Excelente	Excelente
PABAK	0,92	0,91	0,88	0,92
	Excelente	Excelente	Excelente	Excelente

5 – Considerações finais

É de fundamental importância agregar todos os elementos possíveis para se obter resultados confiáveis, que vão desde a bibliografia existente sobre os métodos e sobre a área de estudo, passando pela subjetividade do fotointerprete até a quantificação desses resultados.

Toda representação cartográfica é sempre uma aproximação da realidade e para que se possa melhor conhecer o espaço geográfico, as generalizações são necessárias. A quantificação é essencial para que se possam validar os resultados e assim perceber quão próximo da realidade se encontra esse produto.

Entretanto, os métodos utilizados para essa mensuração são determinantes, pois estes podem apresentar valores diferenciados para um mesmo produto. Logo, a escolha do método de validação a ser utilizado é de fundamental importância e deve atender ao objetivo específico. Não deve existir o "melhor" método, e sim o melhor método para determinado objetivo (Lobão – no prelo).

Qualquer metodologia a ser utilizada guarda as suas limitações e sua utilização deve preceder de adaptações a serem avaliadas em função da área de estudo, do objetivo proposto, do tempo e do custo na sua implementação.

Os resultados obtidos neste trabalho com o Índice Kappa adquiridos apontam para a necessidade do cálculo individualizado para cada classe, evitando-se assim generalizações maiores. A validação da classificação além de fundamental deve ser entendida como uma etapa do trabalho a ser avaliada em função de suas especificidades.

A determinação da acurácia dos dados validados deve ser uma constante em qualquer dado produzido. Entretanto, sua implementação ainda é custosa e dificilmente implementada e por isso mesmo pouco discutida. Novas alternativas de métodos, índices e formas de validação devem ser testadas para que estes procedimentos passem a ser rotineiros.

A aplicação do coeficiente PABAK, associado ao IK é importante para ajustar o efeito da distribuição marginal presente em algumas tabelas de cálculo do IK, como discutido anteriormente. Mesmo quando não ocorre desvio de prevalência em função da distribuição marginal (como no trabalho apresentado) sua aplicação é fundamental pois a simples visualização da distribuição dos dados não fornece subsídios suficientes para a identificação desses problemas. Logo, não é desejável que se conviva com incertezas desta natureza, quando soluções de fácil implementação já foram discutidas e trabalhadas em situações e áreas distintas do conhecimento.

A adaptação desta técnica a realidades diferenciadas é uma prática que se almeja para todo e qualquer dado gerado com a finalidade de compor um Sistema de Informações

Georreferenciadas de forma que as quantificações do erro, por meio do IK, ajustado pelo PABAK, possam fazer parte de seus metadados.

6. Referências

Belart, P.; Pinilla, C.; Ariza, F. J**., Control de Calidad de Clasificaciones de imágenes por el método de los itinerarios**. Grupo mde Investigación em Ingeniería Cartográfica. Escuela Politécnica Superior Universidad de Jaén C/ Virger de la Cabeza nº 223071 Jaén., 2002

Congalton, R. G.; Mead R. A. **A review of assessing the accuracy of classifications of remote sensed data. Remote Sensing of Environment**, v.37, n.1, p.35-46, 1991

Congalton , Russell g., **Assesssing the accuracy of remotely sensed data principles and practices**: Boca Raton : Lewis Publications , 1999

Feinstein A.R., Cicchetti D.V. : **High agreement but low kappa : I. The problems of Two Paradoxes**, J. Clin. Epidemiol., 1990, 43, 543-548

Hoehler, Fred K. **Bias and prevalence effects on kappa viewed in terms odf sensitivity and specificity** in: Journal of Clinical Epidemiology53 (2000) 499-503

Lobão, Jocimara. S. B. et al . **O Uso de Imagens de Satélite na Delimitação de Alvos Urbanos e Não-Urbanos em Feira de Santana-BA** In: Simpósio Brasileiro de Sensoriamento Remoto, 11., 2003, Belo Horizonte. Anais... São José dos campos: INPE, 2003. p. 1099-1106. Disponível na biblioteca digital URLib <http://lagavulin.ltid.inpe.br:1905/col/ltid.inpe.br/sbsr/2003/03.27.11.57/doc/14_091.PDF> Acesso em: 22 abr. 2004

Lobão, Jocimara S. B. *at al* **Acurácia da Classificação Automática em Imagens de Satélite pelos Métodos Randômicos e dos Itinerários: O Caso da Delimitação Urbana de Feira de Santana-Ba** In: Chaves, Joselisa M. e Franca Rocha, Washington do J. S. da. Geotecnologias: Trilahando Novos Caminhos, Salvador: SBG/FAPESB, 2004 (no prelo).

Lingnau, C. et al. **Avaliação da Acurácia de Mapas Temáticos Oriundos de Classificação Digital de Imagens**. In: http://cieg.ufpr.br/Producao/artigos/simgeo.htm, 2003

Moreira, Maurício Alves, **Fundamentos de Sensoriamento Remoto**. São José dos campos-SP: INPE, 2001

SILVA, Ardemirio Barros, **Sistemas de Informações Georeferrenciadas: conceitos e fundamentos**. São Paulo: Unicamp, 1999

SUDENE, Carta Topográfica de Santo Estevão 24-V-B-III, Escala 1:100.000, 1975