

## Avaliação da exatidão do mapeamento da cultura do café no Estado de Minas Gerais

Marcos Adami <sup>1</sup>  
Mauricio Alves Moreira <sup>1</sup>  
Marco Aurélio Barros <sup>2</sup>  
Vagner Azarias Martins <sup>1</sup>  
Bernardo Friedrich Theodor Rudorff <sup>1</sup>

<sup>1</sup> Instituto Nacional de Pesquisas Espaciais - INPE  
Caixa Postal 515 - 12245-970 - São José dos Campos - SP, Brasil  
{adami, mauricio, vagner, bernardo}@dsr.inpe.br

<sup>2</sup> GEOAMBIENTE Sensoriamento Remoto Ltda  
Av. Shishima Hifumi, nº 2.911, 2º andar - 12244-000 - São José dos Campos - SP, Brasil  
{marcoarelio.barros}@geoambiente.com.br

**Abstract.** Accurate and updated agricultural statistics on coffee crop area estimation, based on an objective method, is not available. Remote sensing images are an important source of information to estimate crop area, especially when crop identification procedure is carried out in an effective way. The objective of the present work is to evaluate a coffee crop map obtained from remote sensing images in the state of Minas Gerais. In the adopted method the samples were randomly selected within strata obtained in a two stage stratification. In the first stage five strata were obtained based on both percent of coffee area and regional crop management characteristics. In the second stage the strata were divided into areas with and without coffee crop. For each stratum 52 samples were randomly selected. Prior to the field work the samples were visually identified on both high spatial resolution images available in Google Earth and recent Landsat images used to perform the map. Several samples could be clearly identified as coffee or non-coffee reducing drastically the field work. Confusion matrix was used to provide the global accuracy index, the consumer accuracy and the producer accuracy for each stratum. Overall mapping accuracy for strata 1, 2, 3, 4 and 5 were 99%, 90%, 95%, 86% and 83%, respectively. Global mapping accuracy for Minas Gerais was 95%. It was observed that the mapping accuracy is related to: a) regional vegetation that may cause confusion with coffee; b) cultural practices; c) crop management; and d) relief.

**Palavras-chave:** remote sensing, sampling, error matrix, sensoriamento remoto, amostragem, matriz de erros.

### 1. Introdução

Dentre os inúmeros produtos agrícolas cultivados no território brasileiro, o café destaca-se desde o período colonial como gerador de emprego e renda. Para a safra de café de 2008, a Companhia Nacional de Abastecimento (CONAB) estima que o Brasil produza em torno de 35 milhões de sacas de café arábica e, que o Estado de Minas Gerais, será responsável por 66% do total desta produção (CONAB, 2008). Neste sentido, o conhecimento da área cultivada com café, em Minas Gerais, é fundamental para a formulação de uma estimativa precisa da safra cafeeira no Brasil.

As informações oficiais sobre área plantada com café são divulgadas pelo Instituto Brasileiro de Geografia e Estatística (IBGE) e pela CONAB. Além do setor público, o setor privado também procura obter informações a respeito da área plantada com café. Atualmente a forma que se obtém essa informação é inteiramente subjetiva, pois os dados são coletados através de entrevistas com agentes financiadores, comerciantes, produtores, etc. Além da subjetividade, essa informação não fornece a distribuição espacial das lavouras.

Para tornar essas informações objetivas, tem-se empenhado em desenvolver metodologias que utilizam imagens de satélites de sensoriamento remoto para realizar o mapeamento de áreas de café (Tardin et al., 1992; Moreira et al., 2004; Moreira et al., 2007), possibilitando dessa forma estimar a área e fornecer a localização das lavouras de café.

O uso de imagens de satélites para mapear lavouras de café tem vantagem e desvantagem. A vantagem neste caso é tornar a informação objetiva e, ao mesmo tempo, conhecer a localização das lavouras. No entanto, Moreira et al. (2007) relatam que o comportamento espectral do café tem alta variabilidade. Essa variação está relacionada com fatores da planta (sanidade, estado nutritivo e variedade), dos tratos culturais e do manejo da cultura. Em razão disso podem ocorrer erros que estão relacionados com a precisão e a exatidão do mapeamento. A precisão refere-se à comparação entre a geometria da forma mapeada e a forma do objeto no terreno e, a exatidão, refere-se à comparação entre a classe mapeada e a classe do objeto no terreno (Congalton e Green, 1999). Para aferir o mapeamento da cultura do café no Estado de Minas Gerais são necessárias informações objetivas a respeito da área plantada. As informações disponíveis são subjetivas, portanto não são passíveis de comparação estatísticas (Pino, 1999). Neste sentido, este trabalho tem por objetivo desenvolver uma metodologia para avaliar a exatidão do mapa de café elaborado por Moreira et al. (2007).

## 2. Material e Métodos

A área de estudo compreende o estado de Minas Gerais, localizado na região sudeste do Brasil, entre os paralelos de 14° e 23° Sul e os meridianos de 40° e 51° Oeste. A Figura 1 apresenta um mapa do estado de Minas Gerais dividido em Mesorregiões (PRODEMGE, 1996), com a distribuição espacial do café, que foi mapeado por Moreira et al. (2007), através da interpretação de imagens do Landsat/TM obtidas no período seco do ano de 2007.

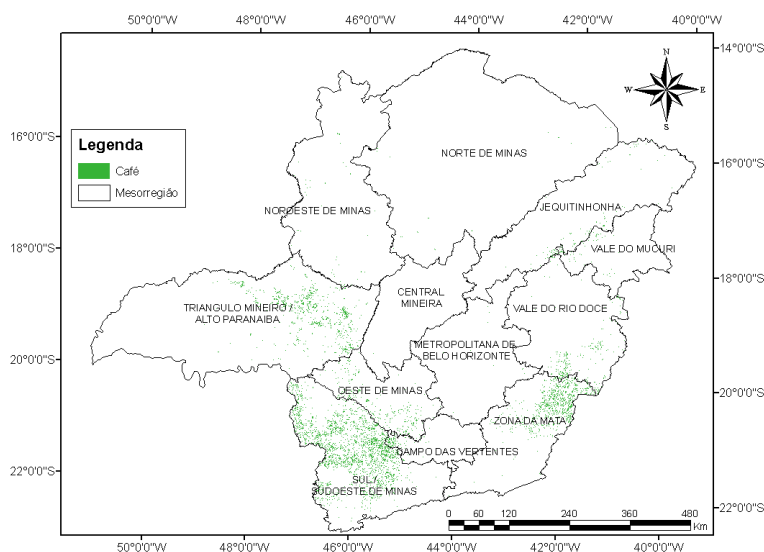


Figura 1. Mapa com a distribuição espacial da cultura do café no estado de Minas Gerais por mesorregiões.

Fontes: Moreira et al. (2007) e PRODEMGE (1996).

Nota-se que o café não se distribui uniformemente em todo o estado. É notória a maior concentração de áreas com café no Sul/Sudeste, Zona da Mata, Oeste de Minas e Triângulo Mineiro e Alto do Paranaíba (Figura 1) e segundo Moreira et al. (2007) estes dados estão em consonância com as estimativas oficiais. Por esta razão, optou-se por utilizar amostragem aleatória estratificada, conforme recomenda Congalton (1988), em dois estágios.

O primeiro estágio de estratificação foi realizado em função da área de café e das características regionais de cultivo. Para isto, utilizou-se o mapa e os resultados de área de café mapeados por município de Moreira et al. (2007). Foram definidos cinco estratos (H) no

primeiro estágio: i) Estrato 1 - onde está inserida a Mesorregião Sul/Sudoeste de Minas; ii) Estrato 2 – onde se insere a Mesorregião da Zona da Mata; iii) Estrato 3 – contendo parte das Mesorregiões Triângulo Mineiro e Alto do Paranaíba; iv) Estrato 4 – que contem as Mesorregiões do Vale do Rio Doce, Mucuri e Jequitinhonha e; v) Estrato 5 – para as demais áreas onde a cultura do café tem pouca representatividade.

No segundo estágio, foi realizada a intersecção entre o mapa de café, em cada um dos estratos, e uma grade regular, de 500 m de lado, obtendo-se dois conjuntos de segmentos: com e sem café (Figura 3). Para estes dois conjuntos, por estrato, foi calculado o tamanho amostral, através da função binomial (Congalton, 1988; Congalton e Green, 1999; Thompson, 2002), uma vez que o mapa possui somente duas classes, café e não-café. A função utilizada para estimativa do número de pontos amostrais, por estrato é apresentada na Equação 1:

$$n = \frac{(z_{\alpha/2})^2 pq}{E^2} \quad (1)$$

em que, n = número total de elementos a serem amostrados; z = distribuição normal padrão;  $\alpha$  = nível de significância; p = exatidão desejada; q = (1-p); E = um erro de amostragem.

Após a obter o tamanho amostral foi realizado o sorteio aleatório dos locais a serem visitados para coletar os dados sobre a ocupação do solo e a partir dessa informação determinar a exatidão do mapeamento. Esses locais, nas áreas da classe não-café, foram representados pelo centróide da grade regular. Na classe café estes locais foram representados pelo centróide resultante da intersecção entre o mapa de café e a grade regular. Esta abordagem foi realizada com o objetivo de minimizar o efeito de borda resultante de deslocamento de pixels, quer seja por questões ligadas ao georreferenciamento ou pela mistura espectral que pode ocorrer nos limites das lavouras.

Os pontos sorteados foram incorporados às imagens disponíveis no sítio do Google Earth ([www.googleearth.com](http://www.googleearth.com)) a fim de minimizar o trabalho de campo. Os pontos sorteados e sobrepostos às imagens de alta resolução espacial, quando disponíveis no Google Earth, foram interpretados na tela do computador para identificar o uso de cobertura do solo, o que pôde ser feito com relativa facilidade nestas imagens. Esse mesmo conjunto de pontos foi sobreposto às imagens Landsat/TM utilizadas no mapeamento do café para verificar se a ocupação do solo observada nas imagens Google Earth se mantinha. Quando não havia alteração do uso nas imagens Landsat/TM em relação às imagens de alta resolução disponíveis no Google Earth, os pontos eram classificados de acordo com a classe (café ou não café) identificada nas imagens de alta resolução. No caso de haver alteração na classe de uso, dúvida ou os pontos amostrados estarem fora da área em que se dispõe de imagens de alta resolução espacial do Google Earth, os mesmos foram identificados em visita a campo por um técnico da CONAB.

Para se obter a exatidão geral do mapa de café de Minas Gerais através da matriz de confusão, foram calculadas as estimativas de exatidão global ( $P_0$ ), exatidão do usuário ( $P_U$ ) e exatidão do produtor ( $P_P$ ) (Stehman, 1997; Congalton e Green, 1999; Smits et al., 1999), sendo que os valores de cada estrato foram ponderadas por um peso (w) dado pelo percentual da área de café por região, equações 2, 3 e 4, respectivamente.

$$P_0 = \frac{\sum_{h=1}^H \sum_{i=1}^m n_{ii} \cdot w_h}{\sum_{h=1}^H n \cdot w_h} \quad (2)$$

$$P_U = \frac{\sum_{h=1}^H n_{ii} \cdot w_h}{\sum_{h=1}^H \sum_{i=1}^m n_{i+} \cdot w_h} \quad (3)$$

$$P_p = \frac{\sum_{h=1}^H n_{ii} \cdot W_h}{\sum_{h=1}^H n_{+i} \cdot W_h} \quad (4)$$

em que:  $m$  é o número de categorias presentes na matriz de erro;  $n_{ii}$  representa os elementos corretamente classificados, presentes na diagonal principal;  $n_{i+}$  representa o somatório dos elementos da linha e  $n_{+i}$  representa o somatório dos elementos da coluna. Para a determinação da exatidão do mapeamento, para cada estrato, o peso torna-se igual a um (1).

Os materiais utilizados nesta pesquisa encontram-se listados na Tabela 1.

Tabela 1. Materiais utilizados na pesquisa.

Item	Descrição	Escala	Fonte
1	Limites municipais	1:1.500.000	PRODEMGE (1996)
2	Mapa da área de café	1:100.000	Moreira et al. (2007)
3	GPS de navegação		
4	<i>Google Earth</i>		

### 3. Resultados e Discussão

Com base na área de café mapeada por município e nas características regionais de cultivo foram obtidos cinco estratos, apresentados na Figura 2 e as áreas mapeadas com café e respectivos percentuais são apresentados na Tabela 2.

Tabela 2. Estratos, área mapeada e percentual com café.

Estrato	Área com café (ha)	Percentual (%)
Estrato 1	504.085	54,7
Estrato 2	189.404	20,6
Estrato 3	158.070	17,2
Estrato 4	30.967	3,4
Estrato 5	38.678	4,2
Total	921.204	100

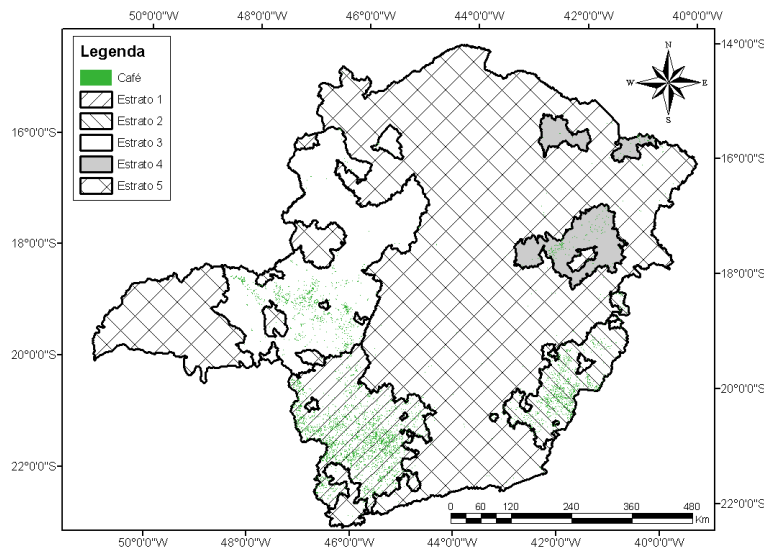


Figura 2. Estratos para avaliação do mapeamento.

Após obter os estratos realizou-se o segundo estágio da estratificação em que foram obtidos os segmentos com e sem lavouras de café, conforme mostrado na Figura 3.

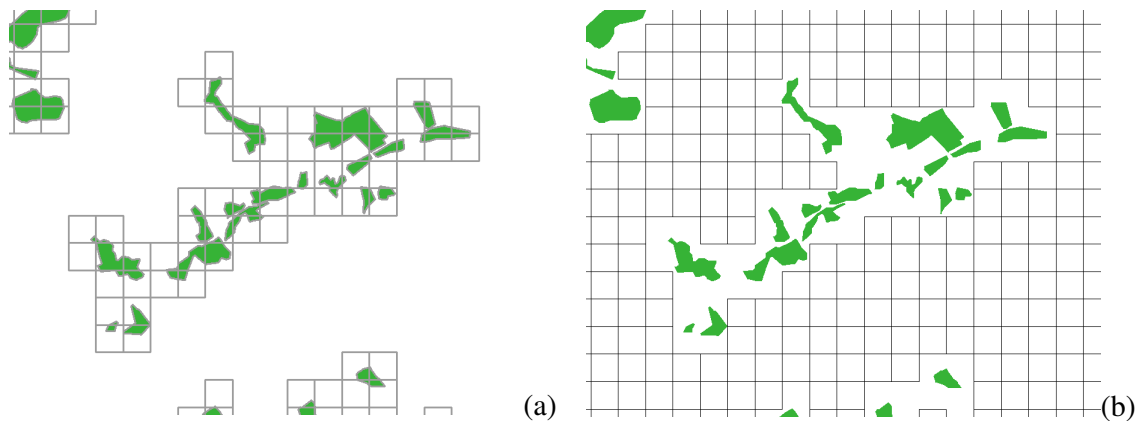


Figura 3. Resultado da intersecção entre o mapa de café e os segmentos. a) segmentos com área mapeada com café; b) segmentos sem área mapeada com café.

Com base na Equação 1, para cada um dos estratos foi calculado um tamanho amostral de 104 pontos, por estrato, utilizando os parâmetros de  $\alpha=,1$ ,  $p=0,95$  e  $E=0,05$ , sendo 52 pontos em lavouras de café e 52 pontos em áreas que não foram mapeadas com café, em cada estrado, distribuídos aleatoriamente, o que aumenta a confiança nas estimativas geradas pela amostragem. O total de pontos sorteados para todos os estratos foi 520. A espacialização dos pontos por estrato e para todo estado é apresentada na Figura 4.

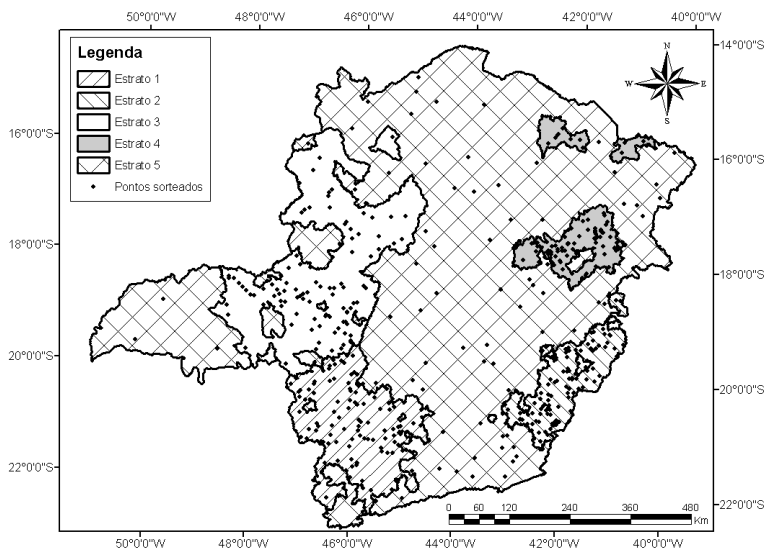


Figura 4. Localização dos pontos sorteados.

O Estrato 1 (Tabela 3), que corresponde a 8,8% da área territorial do estado de Minas Gerais, corresponde a 54,7% da área total de café no Estado (Tabela 2). Neste estrado a exatidão de mapeamento foi de 99,0%. Uma das causas dessa exatidão pode ser atribuída à característica de cultivo (manejo e trato cultural) uma vez que a cultura do café é a principal fonte de renda nesta região. Esse fato contribuiu para que o comportamento espectral do café diferencia-se dos demais alvos de ocupação do solo e facilitando a sua identificação nas imagens TM/Landsat.

Tabela 3. Matriz de confusão com as estimativas de exatidão total, do usuário e do produtor, para o Estrato 1.

Estrato 1				
	Dados de referência			Exatidão Usuário (%)
	Café	Não café	Total classificação	
Café	<b>52</b>	0	52	<b>100</b>
Não-café	1	<b>51</b>	52	<b>98,1</b>
Total de referência	53	51	<b>104</b>	
Exatidão do produtor (%)	<b>98,1</b>	<b>100</b>		
Exatidão geral (%)	<b>99,0</b>			

Nos Estratos 2, 3 e 4 a exatidão total de mapeamento foi de 90%, 95% e 86%, respectivamente (Tabelas 4, 5 e 6). O resultado obtido para o Estrato 2 (Tabela 4) correspondente à Mesorregião da Zona da Mata, teve uma exatidão geral de 90,4 %. Neste caso, o relevo acidentado é um fator crítico para o mapeamento do café. Pelos resultados obtidos nesses três Estratos percebe-se que o mapeamento apresentou um grau de dificuldade maior para o Estrato 4 (Tabela 6) onde está inserida a Mesorregião do Mucuri em que predominam áreas reflorestadas. O reflorestamento novo apresenta comportamento espectral muito semelhante ao café o que contribuiu para aumentar o erro de inclusão (Moreira et al., 2004). Além disso, nesta região o clima é bastante seco afetando o comportamento espectral do café.

Tabela 4. Matriz de confusão com as estimativas de exatidão total, do usuário e do produtor, para o Estrato 2.

Estrato 2				
	Dados de referência			Exatidão Usuário (%)
	Café	Não café	Total classificação	
Café	<b>47</b>	5	52	<b>90,4</b>
Não-café	5	<b>47</b>	52	<b>90,4</b>
Total de referência	52	52	<b>104</b>	
Exatidão do produtor (%)	<b>90,4</b>	<b>90,4</b>		
Exatidão Geral (%)	<b>90,4</b>			

Tabela 5. Matriz de confusão com as estimativas de exatidão total, do usuário e do produtor, para o Estrato 3.

Estrato 3				
	Dados de referência			Exatidão Usuário (%)
	Café	Não café	Total classificação	
Café	<b>48</b>	4	52	<b>92,3</b>
Não-café	1	<b>51</b>	52	<b>98,1</b>
Total de referência	49	55	<b>104</b>	
Exatidão do produtor (%)	<b>97,9</b>	<b>92,7</b>		
Exatidão Geral (%)	<b>95,2</b>			

Tabela 6. Matriz de confusão com as estimativas de exatidão total, do usuário e do produtor, para o Estrato 4.

Estrato 4				
Café	Dados de referência			Exatidão Usuário (%)
	Café	Não café	Total classificação	
Café	<b>39</b>	13	52	<b>75,0</b>
Não-café	1	<b>51</b>	52	<b>98,1</b>
Total de referência	40	64	<b>104</b>	
Exatidão do produtor (%)	<b>97,5</b>	<b>79,7</b>		
Exatidão Geral (%)	<b>86,5</b>			

A Tabela 7 apresenta o resultado de exatidão do mapeamento para o estrato 5 onde o café é pouco expressivo.

Tabela 7. Matriz de confusão com as estimativas de exatidão total, do usuário e do produtor, para o Estrato 5.

Estrato 5				
Café	Dados de referência			Exatidão Usuário (%)
	Café	Não café	Total classificação	
Café	<b>33</b>	19	52	<b>63,5</b>
Não-café	0	<b>52</b>	52	<b>100</b>
Total de referência	33	71	<b>104</b>	
Exatidão do produtor (%)	<b>100</b>	<b>73,2</b>		
Exatidão geral (%)	<b>81,7</b>			

Esse Estrato representa 69,0% da área territorial do estado de Minas Gerais. A área de café que foi mapeada por meio de imagens do TM/Landsat foi de 38.678 ha, o que representa 4,2% de todo café plantado em Minas Gerais. Para esse Estrato a exatidão geral do mapeamento foi de 81,7%. Ao observar os resultados contidos na Tabela 7 nota-se que dos 50 pontos identificados como café somente 33 foram realmente identificados a campo. Nesta situação o erro observado foi de inclusão (pontos de não-café que foram mapeados como café). A causa desse erro está associada à pouca expressividade do café neste Estrato e também à ocorrência de alguns tipos de cerrado e de pastagem com comportamento espectral semelhante ao café, de acordo com Moreira et al. (2004).

A Tabela 8 apresenta a matriz de confusão com as estimativas de exatidão total, do usuário e do produtor para toda a área mapeada. Pode-se verificar também que a exatidão total do mapeamento do café foi de 95,5%. Foi observado que a maioria das lavouras de café presente nas imagens de alta resolução disponível no sítio do Google Earth também estava presente nas TM/Landsat. Isto ocorre devido ao café ser uma cultura perene e as imagens de alta resolução disponíveis no sítio do Google Earth serem dos anos de 2003, 2004 e 2005. Estes fatores facilitam a interpretação dos pontos e reduziram a necessidade de trabalho de campo.

Tabela 8 – Matriz de confusão com as estimativas de exatidão total, do usuário e do produtor, para toda a área de estudo.

Café	Dados de referência			Exatidão Usuário (%)
	Café	Não café	Total classificação	
Café	<b>49</b>	3	52	<b>94,3</b>
Não-café	2	<b>50</b>	52	<b>96,7</b>
Total de referência	51	53	<b>104</b>	
Exatidão do produtor (%)	<b>96,5</b>	<b>94,5</b>		
Exatidão Geral (%)	<b>95,5</b>			

#### 4. Conclusões

De acordo com os resultados obtidos nesta pesquisa foi possível concluir que a exatidão de mapeamento de áreas de café está relacionada com: a) tipo de vegetação da região e que pode causar confusão com a cultura do café; b) condições de manejo e tratos culturais; e c) relevo. Por essa razão a exatidão de mapeamento deve ser feita por estratos. A exatidão geral do mapeamento da cultura do café no Estado de Minas Gerais foi de 95,5%. As imagens de alta resolução disponíveis no sítio do Google Earth reduziram a necessidade de trabalho de campo.

#### 4. Referências Bibliográficas

Companhia de Tecnologia da Informação do Estado de Minas Gerais (PRODEMGE). **Limite dos Municípios de Minas Gerais**. 1996. Escala 1:100.000. Disponível em:< <http://www.geominas.mg.gov.br/>>. Acesso em: 15 mar. 2005

Companhia Nacional de Abastecimento (CONAB) **Acompanhamento da safra brasileira – Café** (segunda estimativa – maio/2008). Disponível em:<<http://www.conab.gov.br>>. Acesso em: 16 de julho de 2008.

Congalton, R. G.; Green, K. **Assessing the accuracy of remotely sensed data: principles and practices**. New York: Lewis Publishers, 1999. 136 p.

Congalton, R.G. A comparison of sampling schemes used in generating error matrices for assessing the accuracy of maps generated from remotely sensed data. **Photogrammetric Engineering and Remote Sensing**, v. 54, n. 5, p. 593-600, 1988.

Moreira, M. A.; Barros, M. A.; Rosa, V. G. C.; Adami, M. Tecnologia de informação: imagens de satélite para o mapeamento de áreas de café de Minas Gerais. **Informe Agropecuário**, v. 28, n. 241, p. 27-37, 2007.

Moreira, M. A.; Adami, M.; Rudorff, B. F. T. Análise espectral e temporal da cultura do café em imagens LANDSAT. **Pesquisa Agropecuária Brasileira**, v. 39, n.3, 223-231 p. Mar 2004.

Pino, F. A. Estatísticas agrícolas para o século XXI. **Agricultura em São Paulo**. v.46, n.2, p.71-105. 1999.

Smits, P. C.; Dellepiane, S. G.; Schowengerdt, R. A. Quality assessment of image classification algorithms for land-cover mapping: a review and a proposal for a cost based approach. **International Journal of Remote Sensing**, v.20, n.8, p.1461-86, 1999.

Stehman, S. V. Selecting and interpreting measures of thematic classification accuracy. **Remote Sensing of Environment**, v.62, n.1, p.77– 89, 1997.

Tardin, A. T.; Assunção, G. V.; Soares, J. V. Análise preliminar de imagens TM visando a discriminação de café, citrus e cana-de-açúcar na região de Furnas, MG. **Pesquisa Agropecuária Brasileira**, v. 27, n. 9, p. 1355-61, 1992.

Thompson, S.K. **Sampling**, 2 ed. New York: John Wiley & Sons, Inc., 2002. 343 p.