

Mapeamento da flora nativa e das florestas de produção do Estado de Minas Gerais: procedimentos metodológicos e resultados parciais

Luis Marcelo Tavares de Carvalho ¹

¹ Universidade Federal de Lavras - UFLA/DCF/LEMAF
Caixa Postal 3037 - 37200-000 - Lavras - MG, Brasil
passarinho@ufla.br

Abstract. The objective of this work was to develop and apply a methodological procedure to map the native flora and the production forests of Minas Gerais, Brasil. Landsat TM and ETM+ images, field data and contour lines derived from topographic maps were used. The available data set was georeferenced to the Albers conic equal area projection. Clouds were removed and replaced by alternative data and the resulting near cloud free images were gathered in a state wide mosaic. The contour lines were used to build a digital terrain model for the whole state. All images were converted to at-satellite reflectance and used to calculate Tasseled Cap indices. Homogeneous mapping zones based on biophysical and reflectance characteristics were defined to be input to the classification stage. Decision tree algorithms were used in a supervised way in order to derive the rules to separate land cover classes. After classification, visual interpretation was employed to correct possible misclassification. The procedure was applied to two pilot areas and accuracy assessment was performed for one of them. Overall accuracy was 90.8% and the kappa coefficient was 0.8918.

Palavras-chave: decision trees, image processing, large area mapping, árvores de decisão, processamento de imagens, mapeamento de grandes áreas.

1. Introdução

O Governo Estadual de Minas Gerais através da Secretaria de Estado de Meio Ambiente e Desenvolvimento Sustentável (SEMAD) e do Instituto Estadual de Florestas (IEF) está promovendo o desenvolvimento da base de dados ambientais de Minas Gerais. A crescente necessidade de informações confiáveis sobre tamanha extensão territorial levou a escolha da utilização de técnicas de sensoriamento remoto para fins de mapeamento e monitoramento. No entanto, o estudo de grandes áreas para obter informações nas escalas de interesse para estas instituições usando as referidas técnicas ainda enfrenta desafios técnicos e práticos.

A integração de dados de múltiplas fontes, bem como o processamento de múltiplas cenas são exemplo de questões limitantes. A semelhança espectral entre as diversas classes de ocupação do solo quando retratadas pelos sensores TM e ETM+ constituem outra dificuldade importante. As alternativas sugeridas para melhorar a eficiência na utilização das imagens fornecidas por estes sensores se baseiam na inclusão de informações adicionais principalmente retratando aspectos topográficos, geológicos, temporais e de índices derivados das próprias imagens para realçar feições de interesse.

Motivado pelo exposto acima, o objetivo deste trabalho foi desenvolver e aplicar um procedimento metodológico para mapear a flora nativa e as florestas de produção de todo o estado de Minas Gerais.

2. Dados utilizados

As principais etapas deste trabalho foram baseadas em dados coletados pelos sensores TM e ETM+, que se encontram a bordo dos dois últimos satélites da série Landsat. Como demonstrado em estudos anteriores (Vogelmann et al., 2001; Carvalho et al., 2004; Oliveira, 2004), dados multitemporais são uma importante fonte de informação para o mapeamento da vegetação usando imagens de sensoriamento remoto. Esta informação proporciona uma maior capacidade de distinguir as diversas fitofisionomias que ocorrem no Estado. Assim, no intuito de captar variações fenológicas que ocorrem nas formações vegetais da área, foram utilizadas

imagens que retratam três épocas do ano (primavera, verão e inverno). As imagens foram adquiridas junto ao Instituto Nacional de Pesquisas Espaciais (INPE) e a seleção das mesmas foi feita levando em consideração a qualidade das imagens e uma quantidade de nuvens mínima ou nula quando possível. A **Figura 1** mostra o recobrimento do estado com as cenas imageadas pelos referidos sensores e a **Tabela 1** apresenta, para cada uma destas cenas, a relação das datas de tomada das imagens utilizadas.

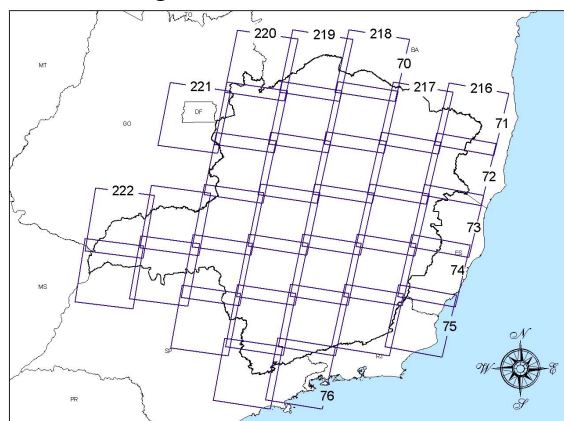


Figura 1. Identificação das cenas que cobrem o estado de Minas Gerais.

Tabela 1. Relação das imagens utilizadas neste trabalho.

Cena	Primavera	Verão	Inverno	Cena	Primavera	Verão	Inverno
216-71	30/09/1999	30/03/1999	29/06/2003	219-71	13/10/2002	06/03/2003	18/06/2003
216-72	10/10/2003	30/03/1999	29/06/2003	219-72	13/10/2002	06/03/2003	18/06/2003
216-73	05/10/2001	13/02/2003	29/06/2003	219-73	13/10/2002	16/02/2004	18/06/2003
216-74	16/10/1999	13/02/2003	29/06/2003	219-74	13/10/2002	06/03/2003	18/06/2003
216-75	19/10/2003	15/02/2001	29/06/2003	219-75	13/10/2002	15/02/2002	18/06/2003
217-71	24/09/2003	12/02/2002	09/06/2002	219-76	13/10/2002	02/02/2003	25/05/2003
217-72	28/10/2001	02/12/2002	20/06/2003	220-70	22/10/2000	24/01/2003	11/07/2003
217-73	04/10/2001	18/03/2001	06/07/2003	220-71	23/11/2000	24/01/2003	11/07/2003
217-74	28/10/2001	18/03/2001	20/06/2003	220-72	23/11/2000	24/01/2003	27/07/2003
217-75	04/10/2001	29/01/2001	06/07/2003	220-73	20/10/2002	25/02/2003	16/07/2002
218-70	06/10/2002	10/01/2003	27/06/2003	220-74	20/10/2002	25/02/2003	16/07/2002
218-71	06/10/2002	10/01/2003	27/06/2003	220-75	20/10/2002	25/02/2003	16/07/2002
218-72	06/10/2002	10/01/2003	27/06/2003	221-71	24/10/2001	02/04/2002	16/06/2003
218-73	06/10/2002	11/02/2003	27/06/2003	221-73	11/10/2002	17/01/2001	16/06/2003
218-74	06/10/2002	17/03/2001	27/06/2003	221-74	11/10/2002	17/01/2001	16/06/2003
218-75	14/10/1999	11/02/2003	27/06/2003	222-73	10/10/1999	27/01/2002	30/05/2003
218-76	27/10/2001	11/02/2003	27/06/2003	222-74	02/10/2002	27/01/2002	30/05/2003
219-70	05/10/2001	13/02/2003	29/06/2003				

Além das imagens, uma base de dados contendo as curvas de nível de Minas Gerais foi fornecida pela SEMAD para dar suporte ao mapeamento da vegetação. Esta base de dados foi digitalizada a partir dos fotolitos das cartas topográficas do IBGE, que se encontram nas escalas de 1:50.000 e 1:100.000 com distância vertical entre curvas de 20 e 50 m respectivamente.

As informações de campo foram obtidas através de incursões aéreas utilizando o helicóptero do IEF, que foram realizadas para retratar amostras das diversas formações vegetais da área e para verificar a acuracidade do trabalho. Equipamentos de videografia, fotografia digital e GPS foram montados na aeronave para coleta de dados.

Por fim, informações provenientes de diversos levantamentos realizados no passado por pesquisadores do Departamento de Ciências Florestais da UFLA, bem como informações coletadas e disponibilizadas pelos técnicos do IEF foram utilizadas para complementar as amostras usadas na classificação das imagens.

3. Pré-processamento

As imagens utilizadas neste trabalho foram adquiridas com correções geométricas e radiométricas padrão (nível de processamento 1G). Neste nível de correção, possíveis erros devido ao efeito do terreno são corrigidos utilizando um modelo digital de terreno e as imagens são registradas a um sistema de coordenadas. As bandas de 1 a 5 e 7 são reamostradas para um tamanho de pixel igual a 30 m utilizando o algoritmo de reamostragem por vizinho mais próximo. Os algoritmos usados para correção modelam características relacionadas ao sensor, ao plano focal e ao alinhamento dos detectores, usando dados gerados por computadores de bordo durante o imageamento. Correções radiométricas incluem remoção de ruídos de sensibilidade dos detectores, calibração para unidades de radiância e posterior conversão para números digitais. Maiores informações sobre as correções padrão podem ser encontradas em <http://landsat.usgs.gov/>.

3.1. Conversão dos números digitais para reflectância aparente

As correções descritas acima resultam em imagens cujos pixels representam as informações sobre a superfície da Terra em números digitais. O número digital (ou valor de brilho) é uma medida da energia que chega ao sensor (i.e., radiância), mas que foi escalonada para ser representada com 8 bits. Com o intuito de padronizar as imagens minimizando o impacto da geometria de iluminação (Markham & Barker, 1986; Irish, 2000; Huang et al., 2002), os números digitais foram convertidos novamente para radiância aparente (aquela que chega ao sensor após haver interagido com a atmosfera) e posteriormente para reflectância aparente usando parâmetros fornecidos pelo Serviço Geológico dos Estados Unidos (USGS). Para tanto foram utilizadas a **Equação 1** (Lillesand & Kiefer, 2000) e a **Equação 2** (Huang et al., 2002).

$$L_{\lambda} = ((LMAX_{\lambda} - LMIN_{\lambda}) / 255) \cdot ND + LMIN_{\lambda} \quad (1)$$

$$\rho_{\lambda} = \pi \cdot L_{\lambda} \cdot d^2 / ESUN_{\lambda} \cdot \text{sen} \theta_s \quad (2)$$

Onde,

λ = comprimento de onda

L_{λ} = radiância espectral aparente

ND = valor do pixel calibrado, dado em número digital

$LMIN_{\lambda}$ = radiância espectral escalonada para 0

$LMAX_{\lambda}$ = radiância espectral escalonada para 255

ρ_{λ} = reflectância aparente

d = distância entre a Terra e o Sol (tabela 2)

$ESUN_{\lambda}$ = irradiância solar exoatmosférica média (tabela 3)

θ_s = ângulo de elevação do Sol

Os resultados da conversão descrita acima são representados em 32 bits sobrecarregando muito a memória física dos computadores e o tempo de processamento. No intuito de reduzir o volume de dados, os valores de reflectância aparente foram truncados em 0,6375 e posteriormente multiplicados por 400 para produzir dados em 8 bits novamente.

3.2. Tratamento das imagens Landsat TM

Assim como as imagens provenientes do sensor ETM+, as imagens do sensor TM foram adquiridas com correções radiométricas e geométricas padrão gerando imagens TM em número digital. Tendo em vista que os sensores ETM+ e TM são compatíveis tanto

geometricamente quanto radiometricamente, todos os procedimentos de pré-processamento aplicados nas imagens ETM+ foram também aplicados nas imagens TM. No entanto, para tirar proveito da calibração radiométrica mais refinada dos números digitais das imagens ETM+, as imagens TM foram normalizadas usando a **Equação 3**.

$$ND_{(ETM)} = ND_{(TM)} \cdot a + b \quad (3)$$

Onde,

$ND_{(ETM)}$ = número digital das imagens TM normalizados para ETM+

$ND_{(TM)}$ = número digital original das imagens TM

a = inclinação da reta de regressão

b = intercepto da reta com o eixo das abscissas

Os valores de a e b para cada banda são apresentados na **Tabela 2**.

Tabela 2. Coeficientes da regressão para normalização das imagens TM.

Banda	a	b	Banda	a	b
1	0,9398	4,2934	4	1,4239	7,0320
2	1,7731	4,7289	5	0,9828	7,0185
3	1,5348	3,9796	7	1,3017	7,6568

Assim, usando os valores de ganhos e vieses apresentados na **Tabela 3** e a **Equação 4** para cálculo da radiância aparente, as imagens Landsat TM passam a ser tratadas como se fossem imagens Landsat ETM+ nos passos subseqüentes (Vogelmann et al., 2001).

Tabela 3. Ganho e viés usados para cálculo da radiância aparente das imagens TM.

Banda	ganho	viés	Banda	ganho	viés
1	0,7756863	-6,1999969	4	0,6372549	-5,1000061
2	0,7956862	-6,3999939	5	0,1257255	-0,9999981
3	0,6192157	-5,0000000	7	0,0437255	-0,3500004

$$L_{\lambda} = \text{ganho}_{\lambda} \cdot ND_{\lambda} + \text{viés}_{\lambda} \quad (4)$$

Onde,

λ = comprimento de onda

L_{λ} = radiância espectral aparente

ND_{λ} = número digital na banda de comprimento de onda λ

3.3. Georeferenciamento, reamostragem, remoção de nuvens e mosaico

Todos os dados utilizados neste trabalho, bem como os produtos finais, foram georeferenciados para a projeção cartográfica Albers de Igual Área (**Tabela 4**) e reamostrados para uma resolução espacial de 30 metros usando o algoritmo de interpolação bilinear.

Tabela 4. Parâmetros da projeção cartográfica utilizada.

Projeção:	Cônica Albers de Igual Área
Datum:	América do Sul 1969 (SAD-69)
Latitude do 1º Paralelo padrão:	5° 00' 00,0000" S
Latitude do 2º Paralelo padrão:	42° 00' 00,0000" S
Longitude do meridiano central:	60° 00' 00,0000" O
Latitude da origem da projeção:	32° 00' 00,0000" S
Falso leste no meridiano central:	0,0000000000 metros
Falso norte na origem da projeção:	0,0000000000 metros

O registro consiste na identificação de pontos correspondentes em duas imagens de determinado local, mas que apresentam geometrias diferentes. As imagens são transformadas para adquirir características de escala e projeção próprias de mapas ou de outra imagem usada como referência.

As imagens do inverno foram registradas usando apenas um ponto de controle por cena. No processo de registro foram utilizados parâmetros como efemérides, altitude do satélite, dados sobre o sistema de imageamento do sensor e do modelo de referência da Terra para o georeferenciamento.

As imagens do verão e da primavera foram registradas com as imagens do inverno usando de 15 a 20 pontos de controle coincidentes. Em todos os casos o RMSE (raiz do erro médio quadrático) foi menor que um pixel, indicando uma boa coincidência entre as imagens.

O registro foi feito utilizando sistemas de equações polinomiais de primeiro grau. As imagens Landsat TM e ETM+ sofreram rotações, escalonamento e translações, passando a ter o mesmo formato e dimensões das imagens georeferenciadas que foram utilizadas como referência.

Após todas as imagens estarem registradas e georeferenciadas ao sistema de coordenadas escolhido, procederam-se a identificação e substituição de pixels cobertos por nuvens. Foram utilizados pedaços de imagens do mesmo local, porém de outra data, que não apresentavam cobertura de nuvens para substituição.

Por fim, foram gerados mosaicos quase livres de nuvens com todas as cenas que recobrem o estado para as três datas consideradas neste trabalho (primavera, verão e inverno).

4. Desenvolvimento do Banco de Dados

4.1. Construção e Classificação do Modelo Digital de Terreno

A partir das curvas de nível foi gerado um modelo digital de elevação pelo método da triangulação Delaunay de um conjunto planar de pontos composto pelos vértices das curvas e pelos pontos cotados. A rede de triângulos irregulares geradas por este método foi então convertida para o formato raster por interpolação linear. O modelo resultante foi classificado em classes de altitude com amplitude de 50 metros.

4.2. Transformação *Tasseled Cap*

Os dados de reflectância normalizados e convertidos para 8 bits foram utilizados para o cálculo dos índices *brightness*, *greenness* e *wetness* da transformação *Tasseled Cap* usando os coeficientes derivados por Huang et al. (2002) (**Tabela 5**). Estes coeficientes são específicos para dados convertidos para reflectância aparente. *Tasseled Cap* é uma transformação linear que, através de rotações e translações dos dados, estabelece novos eixos no domínio espectral que captam a maior parte da informação contida nos dados em três novas variáveis diretamente relacionadas com características físicas da cena. A primeira nova variável, *brightness*, está relacionada com a reflectância dos solos. *Greenness*, a segunda variável, está fortemente relacionada com a quantidade de vegetação verde presente na cena. A terceira variável chamada *wetness*, está relacionada com a umidade do solo e do dossel. A transformação *Tasseled Cap* reduz significativamente o volume de dados a ser processado.

Tabela 5. Coeficientes usados na transformação *Tasseled Cap*.

	Banda 1	Banda 2	Banda 3	Banda 4	Banda 5	Banda 7
Brightness:	0,35612057	0,39722874	0,39040367	0,69658643	0,22862755	0,15959082
Greenness:	-0,33438846	-0,35444216	-0,45557981	0,69660177	-0,02421353	-0,26298637
Wetness:	0,26261884	0,21406704	0,09260517	0,06560172	-0,76286850	-0,53884970

A **Equação 5** foi utilizada para escalonar os valores das imagens *Tasseled Cap* (tc_{32bit}) para serem representados em 8 bits (tc_{8bit}).

$$tc_{8bit} = \text{arredondar}[(tc_{32bit} + \text{deslocamento}) * 255 / \text{intervalo}] \quad (5)$$

Os valores de deslocamento e intervalo utilizados na **Equação 5** estão listados na **Tabela 6**.

Tabela 6. Valores de deslocamento e intervalo utilizados para converter as imagens *Tasseled Cap* para 8 bits.

	Deslocamento	Intervalo
Brightness:	-20	380
Greenness:	100	255
Wetness:	170	320

4.3. Classificação da flora nativa e das florestas de produção

Utilizando um conjunto de atributos composto por 14 variáveis (bandas 3, 4, 5 e 7 do inverno, índices *Tasseled Cap* da primavera, inverno e verão, bem como classes de altitude do terreno), procedeu-se a classificação da ocupação atual do solo.

O mosaico gerado para o estado foi subdividido em zonas de mapeamento homogêneas com relação a características biofísicas e espectrais (Bauer et al., 1994; Lillesand, 1996). Esta estratificação foi realizada através de interpretação visual. Em cada zona de mapeamento, amostras foram coletadas de forma a representar as principais classes de ocupação do solo que ali ocorrem.

O método de classificação foi supervisionado e o algoritmo classificador utilizado foi o de árvores de decisão (Quinlan, 1986; Carvalho, 2001; Carvalho et al., 2004). As árvores de decisão baseiam-se em mecanismos de categorização através de divisões hierárquicas dos dados para chegar em subdivisões que representam as classes temáticas. O uso de árvores de decisão possibilita ainda que dados auxiliares, como as derivativas de um modelo de elevação, possam ser completamente integrados no processo de classificação. Além disso, o método é não-paramétrico e, portanto independente do formato da distribuição das assinaturas das classes; pode utilizar tanto dados contínuos como nominais; gera regras de classificação fáceis de serem interpretadas; é rápido e geralmente tão preciso quanto os demais classificadores normalmente utilizados.

Após a classificação das imagens, classes com pastagem, solo exposto, áreas urbanas etc, foram agrupadas em uma única classe chamada “outros”.

Por fim, cada zona de mapeamento foi interpretada visualmente para correção de possíveis erros de classificação.

4.4. Avaliação da acuracidade

Amostras diferentes daquelas utilizadas para treinar o classificador foram coletadas para verificar a precisão dos mapas. Estas amostras foram aleatorizadas proporcionalmente em estratos definidos pela área ocupada por cada classe. A ferramenta utilizada neste trabalho para determinar a acuracidade dos mapas gerados é a tradicional matriz de confusão e os índices dela derivados. Estes incluem a acuracidade geral e o índice *Kappa*.

5. Resultados

Os métodos descritos acima foram aplicados em duas áreas piloto (**Figura 2**): a bacia hidrográfica do Rio São Francisco em Minas Gerais e a área de abrangência do projeto

Promata/MG. Já a verificação da acuracidade (**Tabela 6**) foi feita apenas para a área de abrangência do projeto Promata/MG.

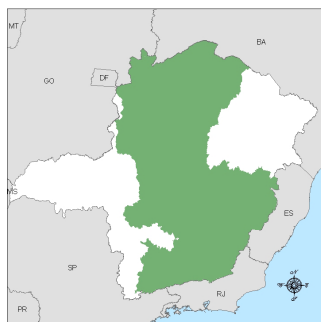


Figura 2. Áreas piloto onde foram aplicados os procedimentos descritos neste artigo.

A **Figura 3** apresenta os mapas gerados neste trabalho para as áreas piloto e a **Tabela 6** apresenta a matriz de confusão, a acuracidade geral e o índice Kappa para a área de abrangência do projeto Promata/MG.

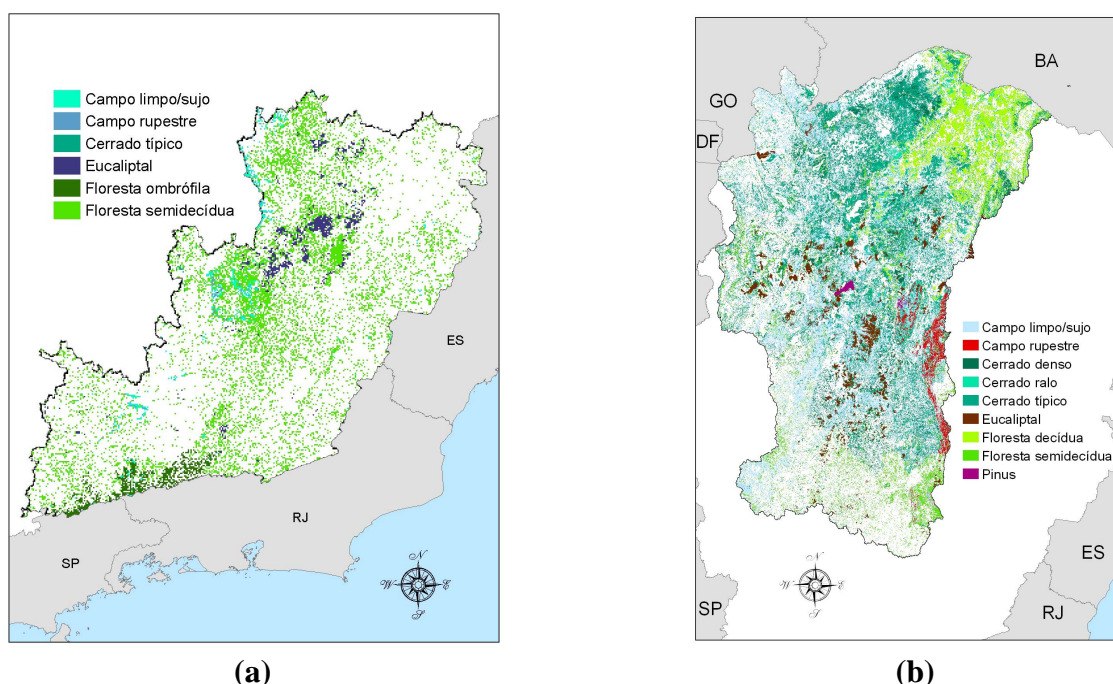


Figura 3. Mapas da flora nativa e florestas de produção. (a) Área de abrangência do projeto Promata/MG. (b) Bacia hidrográfica do Rio São Francisco em Minas Gerais

Tabela 6. Matriz de confusão da classificação da área de abrangência do Promata/MG.

Classe	Urbana	Outros	Floresta semidecídua	Campo	Cerrado típico	Água	Eucalipto	Floresta ombrófila	Campo rupestre	Total
Urbana	38	2								40
Outros	1	115		2					2	120
Flor. semidecídua		12	87				1			100
Campo				49	1					50
Cerrado típico		1		2	26				1	30
Água			1			29				30
Eucalipto		1	6				43			50
Flor. ombrófila		8						42		50
Campo rupestre	1	4							25	30
Total	40	143	94	53	27	29	44	42	28	500

Acuracidade geral: **90.8 %** Coeficiente Kappa: **0.8918**

6. Considerações Finais

Foi desenvolvido e está sendo aplicado um procedimento para processamento de imagens Landsat e mapeamento da flora nativa e das florestas de produção de todo o estado de Minas Gerais. A experiência adquirida até o momento mostra a necessidade de estudos acerca dos vários tipos e estágios de desenvolvimento das florestas de produção para melhorar a acuracidade de mapeamento destas formações vegetais. Similarmente, as diversas fitofisionomias do bioma cerrado merecem estudos mais aprofundados para definir o comportamento espectral das suas fases de regeneração.

O método apresentado neste trabalho tem se mostrado bastante eficiente e com acuracidade considerada excelente quando consideramos os resultados parciais expressos pelo índice *Kappa* para a área de abrangência do projeto Promata/MG.

Referências

- Bauer, M. E., Burk, T. E., Ek, A. R., Coppin, P. R., Lime, S. D., Walsh, T. A., Walters, D. K., Belfort, W., Heinzen, D. F. Satellite inventory of Minnesota forest resources. **Photogrammetric Engineering and Remote Sensing**, v. 60, n. 3, p. 287-298, 1994.
- Carvalho, L. M. T., Clevers, J. G. P. W., Skidmore, A. K., de Jong, S. M. Selection of imagery data and classifiers for mapping Brazilian semideciduous Atlantic forests. **International Journal of Applied Earth Observation and Geoinformation**, v. 5, p. 173-186, 2004.
- Carvalho, L. M. T. **Mapping and monitoring forest remnants: a multiscale analysis of spatio-temporal data**. 2001. 150 p. Tese (Doutorado em Sensoriamento Remoto) – Wageningen University, Wageningen, The Netherlands. 2001.
- Huang, C., Wylie, B., Homer, C., Yang, L., Zylstra, G. Derivation of a Tasseled Cap transformation based on Landsat 7 at-satellite reflectance. **International Journal of Remote Sensing**, v. 23, n. 8, p. 1741-1748, 2002.
- Irish, R. R. **Landsat 7 science data user's handbook**. 2000. (Report 430-15-01-003-0) – National Aeronautics and Space Administration. 2000.
- Lillesand, T. M., Kiefer, R. W. **Remote Sensing and Image Interpretation**. New York: John Wiley & Sons, 2000. 724 p.
- Lillesand, T. M. A protocol for satellite based land cover classification in the upper Midwest. In: Scott, M., Tear, T. H., Davis, F. W. (eds.) **Gap Analysis: A Landscape Approach to Biodiversity Planning**. ASPRS, 1996. 320 p.
- Markham, B. L., Barker, J. L. Landsat MSS and TM post-calibration dynamic ranges, exoatmospheric reflectances and at-satellite temperatures, **EOSAT Landsat Technical Notes**, v. 1, p. 3-8, 1986.
- Oliveira, L. T. **Fusão de imagens de sensoriamento remoto e mineração de dados geográficos para mapear as fitofisionomias do bioma cerrado**. 2004. 150 p. Dissertação (Mestrado em Manejo Ambiental) – Universidade Federal de Lavras, Lavras. 2004.
- Quinlan, J. R. Induction of decision trees, **Machine Learning**, v. 1, p. 81-106.1986.
- Vogelmann, J. E., Howard, S. M., Yang, L., Larson, C. R., Wylie, B. K., Van Driel, J. N. Completion of the 1990's national landcover data set for the conterminous United States. **Photogrammetric Engineering and Remote Sensing**, v. 67, n. 6, p. 650-662, 2001.