

ESTUDO DE CORRELAÇÕES ENTRE DADOS RADIOMÉTRICOS COLETADOS EM CAMPO E EM NÍVEL ORBITAL

Márcio Pupin de Mello¹
Fábio Bueno de Lima¹
Gustavo Bayma Siqueira Silva¹
Denilson Ribeiro Viana¹

¹Instituto Nacional de Pesquisas Espaciais – INPE
Divisão de Sensoriamento Remoto – DSR
Caixa Postal 515 – 12.201-970 – São José dos Campos – SP, Brasil
{pupin, fbueno, bayma, denilson}@dsr.inpe.br

Resumo

Um dos objetivos da espectroradiometria aplicada à vegetação é o refinamento de modelos que possam estimar seus atributos biofísicos. Para atingir este objetivo, autores ressaltam a importância de se determinar relações quantitativas entre as características espectrais de alvos na superfície da Terra e o espectro registrado por um sensor orbital (a bordo de um satélite artificial). Neste contexto este trabalho teve como principal objetivo estudar as correlações entre dados de campo e orbitais assim como índices de vegetação (NDVI, DI e SAVI) gerados a partir desses dados, utilizando dois métodos de correlação: *Pearson* e *Spearman*. Foi usada uma imagem do sensor *Thematic Mapper (TM)*, a bordo do satélite *Landsat-5*, como dado orbital, assim como 40 amostras coletadas em campo concomitantemente à passagem do satélite. A área de estudo corresponde à uma região de pasto, aparentemente homogênea, localizada no município de Pirassununga, no estado de São Paulo – Brasil. Os dados de Fator de Reflectância Bidirecional (FRB) das amostras, na imagem, foram calculados por meio da média aritmética simples dos quatro pixels mais próximos ao pixel central (incluindo o próprio). Já os valores de FRB de campo foram agrupados, por meio da média aritmética, de acordo com os intervalos de comprimentos de onda das bandas do *TM*. As correlações entre os valores de FRB dos dois níveis de coleta de dados apresentaram-se baixos, com média de 12,67% e 15,62% para os índices de *Pearson* e *Spearman*, respectivamente. Para os índices de vegetação a única correlação significativa encontrada foi para o NDVI através do método *Spearman*.

Palavras-chave: Correlação entre dados de campo e orbitais; Fator de Reflectância Bidirecional; Índices de vegetação.

Abstract

One objective of spectroradiometrics applied to vegetation is the refinement of the models that can esteem its biophysics attributes. To reach this objective, authors stand out the importance of determining quantitative relations between the spectral characteristics of targets in the Earth's surface and the specter registered for a orbital sensor (on board an artificial satellite). In this context this study aimed to investigate the correlations between orbit and field data, as well three vegetation indexes (NDVI, DI and SAVI) generated with these data. Two methods of correlation were used: *Pearson* and *Spearman*. An image of the sensor *Thematic Mapper (TM)*, on board at *Landsat-5* satellite, was used as given orbital data, as well 40 samples collected in field concomitant satellite imagery. The study area corresponds to a grass region, apparently homogeneous, located in the city of Pirassununga, state of São Paulo – Brazil. The values of Bidirectional Reflectance Factor (FRB) of the samples, in the image, had been calculated by the simple arithmetic mean of four pixels next to central pixel (including the central one). On the other hand, the values of FRB field data had been grouped, by the arithmetic mean, in

accordance with the intervals of wave lengths in the TM bands. The correlations between the values of FRB in the two levels of data had been presented low, with average of 12.67% and 15.62% for the Pearson and Spearman indexes, respectively. For the vegetation indexes the only significant correlation found was for the NDVI through the Spearman method.

Key-words: Correlation between field and orbital data; Bidirectional Reflectance Factor; Vegetation index.

1 Introdução

A espectroscopia envolve o estudo das inter-relações entre os atributos biofísicos de determinados alvos e suas características espectrais. Quando aplicados à vegetação, esses estudos buscam o refinamento de modelos que possam estimar atributos biofísicos a partir da Radiação Eletromagnética (REM) refletida pelos dosséis (ou pela vegetação) que é medida por sensores remotamente situados.

Morelli et al (1993) descrevem a Estrutura Geométrica da Vegetação (EGV) como o maior determinante da distribuição angular da REM em coberturas vegetais. A EGV pode ser descrita quantitativamente por parâmetros biofísicos como a distribuição das plantas no terreno, o Índice de Área Foliar (IAF), o Índice de Cobertura (IC) e a Função de Distribuição Angular de Folhas (DAF).

Tais parâmetros variam espacialmente e temporalmente de acordo com o tipo de vegetação, o estágio de desenvolvimento e as condições da vegetação. No entanto, nem sempre a análise de dados espectrais mostra essa variação devido a: i) alterações na fonte de REM; ii) alterações na geometria de coleta; e iii) fatores de atenuação da atmosfera.

Um dos objetivos dos diversos índices de vegetação criados é minimizar essas influências. Entre os mais usados na literatura encontram-se o Índice de Vegetação da Diferença Normalizada (NDVI), Índice de Vegetação pela Diferença (ID) e o Índice de Vegetação Ajustado ao Solo (SAVI) (JAKSON et al, 1983; HUETE, 1988). Nesse contexto, Jakson et al. (1983) e Marsh e Lyon (1980) apontam duas preocupações: o primeiro trabalho destaca a necessidade de modelar corretamente os efeitos da atmosfera para evitar dados enviesados; já Marsh e Lyon (1980) destacam a importância de se determinar relações quantitativas entre as características espectrais de alvos na superfície e o espectro registrado por um sensor orbital.

Nesse contexto, este trabalho objetiva correlacionar dados de campo e orbitais, assim como verificar a existência de uma associação entre alguns índices de vegetação obtidos a partir desses dados.

2 Materiais e métodos

A área de estudo está localizada no município de Pirassununga, no Estado de São Paulo (Brasil), e encontra-se delimitada pelas coordenadas geográficas W 47°20'18.42" - S 22°03'49.20" e W 47°20'02.85" - S 22°04'03.62". O local corresponde a uma área de pasto com características aparentemente homogêneas de refletância (Figura 1).

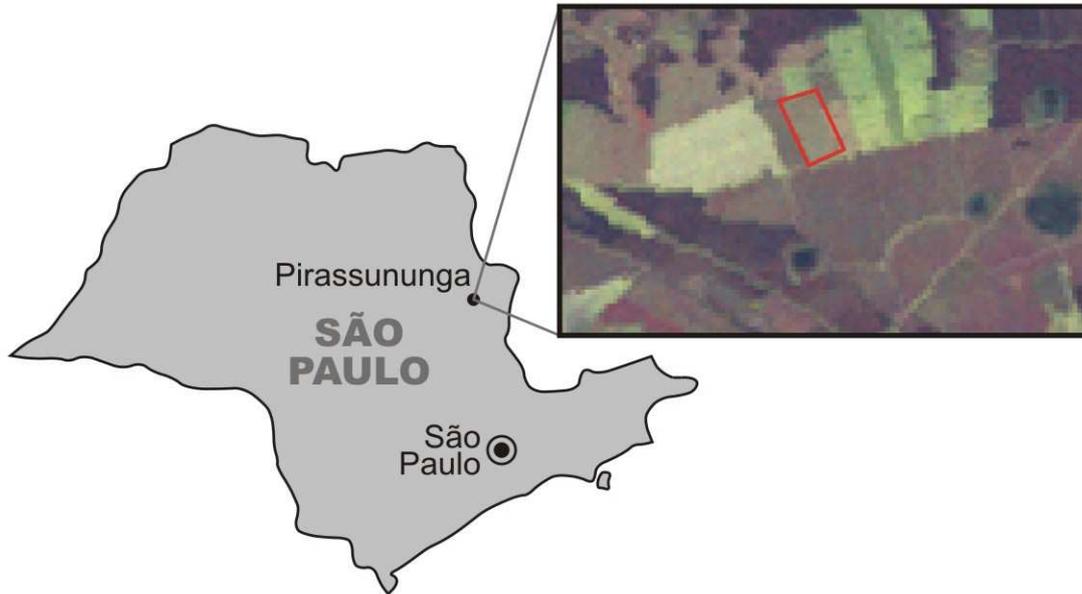


Figura 1: Localização em destaque da área de estudo (imagem *TM* - RGB-231).

Os dados de Fator de Refletância Bidirecional (FRB) de campo correspondem a um total de 40 amostras coletadas em 04/jun/2007, concomitantemente à passagem do satélite *Landsat-5*. Esses dados foram obtidos utilizando o espectrorradiômetro *FieldSpec Pro FR*. A faixa espectral abrangida pelo equipamento é de 350 a 2500 nm, com resolução espectral de 3 a 10 nm e o FOV (*field of view*) adotado foi de 25°. As amostras foram coletadas de 40 em 40 m, aproximadamente.

Para a comparação com os dados de campo, foi utilizada uma imagem do sensor *Thematic Mapper (TM)* da plataforma *Landsat-5*. O sensor *TM* possui resolução espacial de 30 m, resolução radiométrica de 8 bits e opera em sete bandas espectrais, das quais seis foram utilizados nesse trabalho (Tabela 1).

Tabela 1: Intervalos de comprimento de onda das bandas do sensor *TM*.

Banda <i>TM</i>	Região do Espectro	Limite Inicial (nm)	Limite Final (nm)
B1	Azul	452	518
B2	Verde	529	610
B3	Vermelho	624	693
B4	NIR	776	905
B5	SWIR	1.568	1.784
B7	SWIR	2.097	2.347

Os valores de Número Digital (ND) de cada uma das bandas da imagem foram convertidos para FRB aparente, a partir dos coeficientes apresentados por Chander e Markham (2003). Em seguida esses valores foram transformados em FRB de superfície por meio da aplicação do modelo de correção atmosférica 6S (VERMOTE et al., 1997).

As análises estatísticas foram realizadas utilizando o sistema livre *R* versão 2.6.1 (R DEVELOPMENT CORE TEAM, 2007). A manipulação e extração dos valores de FRB da imagem foram executadas utilizando o SPRING 4.3.2 (CÂMARA et al, 1996).

Os dados de FRB das amostras na imagem foram calculados por meio da média aritmética simples dos quatro *pixels* mais próximos ao *pixel* central (incluindo o próprio), como ilustra a *Figura 2*. A fim de comparar com dados da imagem, os valores de FRB de campo foram agrupados, por meio da média aritmética, de acordo com os intervalos de comprimentos de onda das bandas do *TM* (Tabela 1).

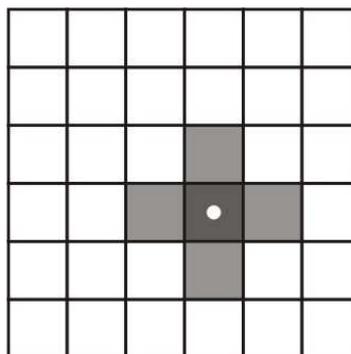


Figura 2: Ilustração dos *pixels* considerados no cálculo das médias de cada ponto na imagem orbital.

Além da comparação dos valores de reflectância dos pontos amostrados, foram calculados três índices de vegetação (*Normalized Difference Vegetation Index* [NDVI], *Difference Vegetation Index* [DI] e *Soil Adjusted Vegetation Index* [SAVI], cujas equações encontram-se na literatura; veja JENSEN, 2006) a partir dos dados de campo e

orbitais para que pudessem ser estudadas as relações entre esses índices obtidos a partir desses dois níveis de coleta de dados.

No desenvolvimento da metodologia os seguintes procedimentos foram realizados:

- 1 – Análise de correlação de *Pearson* e *Spearman* dos valores de FRB gerados a partir dos dados de campo e dados orbitais dos 40 pontos amostrais para cada banda;
- 2 – Estudo da correlação entre índices de vegetação (NDVI, DI e SAVI) gerados com dados dos dois níveis de coleta.

Os estudos de correlações nesse trabalho foram realizados segundo os coeficientes de *Pearson* e de *Spearman* (para mais detalhes veja BUNCHAFT e KELLNER, 2001). O índice de correlação de *Pearson* quantifica o quão linearmente associados estão dois conjuntos de dados. Já o índice de *Spearman* nada mais é que o índice de *Pearson* sobre o posto dos conjuntos de dados, sendo mais indicado quando os dados não seguem distribuição normal, quando existem valores discrepantes nos dados ou em casos onde a associação entre os dois conjuntos de dados não é linear (DYTHAM, 2003).

3 Resultados e discussões

Visando dar sustentabilidade às análises de correlação executadas, foi verificada a normalidade dos dados de campo e orbitais para cada uma das seis bandas, usando o teste de *Kolmogorov-Smirnov* do sistema *R* que aceitou a hipótese de que todos os dados têm distribuição normal ao nível de significância de 1%. Entretanto, para o mesmo nível de significância de 1%, no teste de *Shapiro-Wilk*, os dados de campo da banda B7, assim como os dados orbitais para as bandas B4 e B7, apresentaram distribuição significativamente diferente da normal.

Seguindo a seqüência das etapas enumeradas em “Materiais e Métodos” segue a descrição dos resultados e discussões acerca desses que foram encontradas nesse estudo.

3.1 Correlação entre reflectâncias dos pontos (campo e orbital) em cada banda

As análises de correlação entre os dados de campo e orbitais de FRB mostraram o quão delicado são as questões relacionadas à obtenção de informações remotas através de sensores em plataformas orbitais. Os valores médios de correlação foram 12,67% e 15,62% para os índices de *Pearson* e *Spearman*, respectivamente. É possível ver com clareza que a banda quatro (B4 – NIR, de 776 a 905 nm) foi a menos correlacionada, sendo até de correlação negativa para *Spearman*. Os índices de correlação para a FRB em cada uma das seis bandas podem ser vistos na *Figura 3* e seus valores estão relacionados na *Tabela 2*.

Tabela 2: Índices de correlação de *Pearson* e *Spearman* dos dados nos níveis de campo e orbital dos 40 pontos amostrais em cada banda.

	B1	B2	B3	B4	B5	B7
<i>Pearson</i>	0,18677	0,15198	0,12897	0,03632	0,13908	0,11702
<i>Spearman</i>	0,25119	0,12851	0,15257	-0,08654	0,18144	0,30986

Observou-se uma diferença mais acentuada entre os índices de *Pearson* e *Spearman* nos dados da banda B7. Essa ocorrência está diretamente relacionada com o fato dos dados dessa banda não apresentarem distribuição normal conforme verificado no teste de *Shapiro-Wilk* com 1% de significância.

Apesar dos diferentes valores de correlação todos eles não são significativamente diferentes de zero quando aplicado o teste estatístico *t* para índice de correlação com um nível de significância de 5%. Isso quer dizer que estatisticamente os

dados dos níveis campo e orbital não são correlacionados.

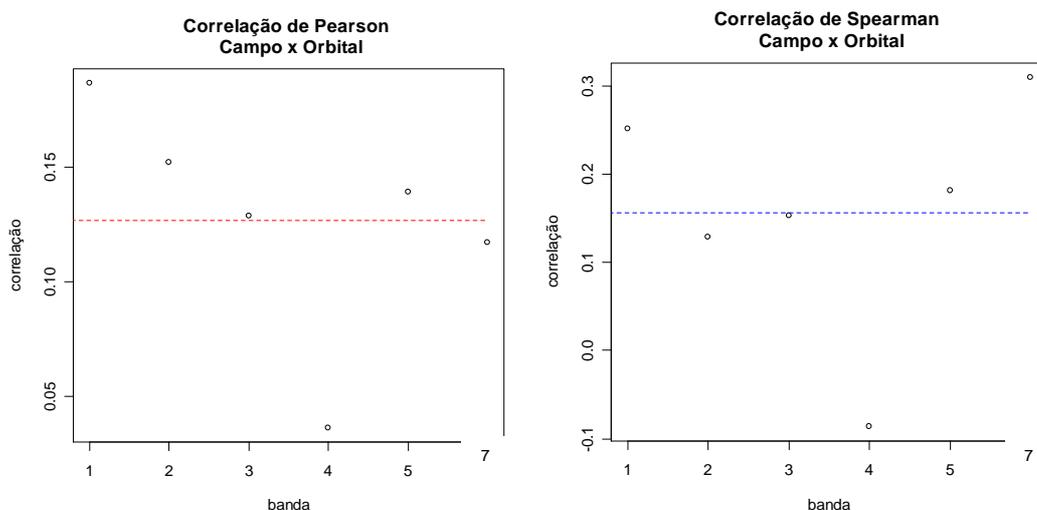


Figura 3: Gráficos com os valores das correlações dos valores de FRB para os dados de campo e orbitais nas 6 bandas (B1, B2, B3, B4, B5 e B7), usando os índices de *Pearson* e *Spearman*, respectivamente (a linha tracejada indica a média para ambos os gráficos).

Esses resultados diferem dos obtidos por Marsh e Lyon (1980) que encontraram correlações significativas comparando dados de campo e orbitais (com correção atmosférica). A principal razão dessa diferença está no fato de que Marsh e Lyon (1980) usaram dados provenientes de amostras significativamente heterogêneas enquanto que o presente trabalho fez uso de dados de amostras relativamente homogêneas. Isso acontece porque a própria diferença nas resoluções (principalmente a espacial) entre o sensor de coleta de dados em campo e o sensor orbital faz com que, em amostras homogêneas, a diferença entre a variância dos dados de campo e a variância dos dados orbitais (veja Figura 4) seja maior que em dados coletados de amostras heterogêneas.

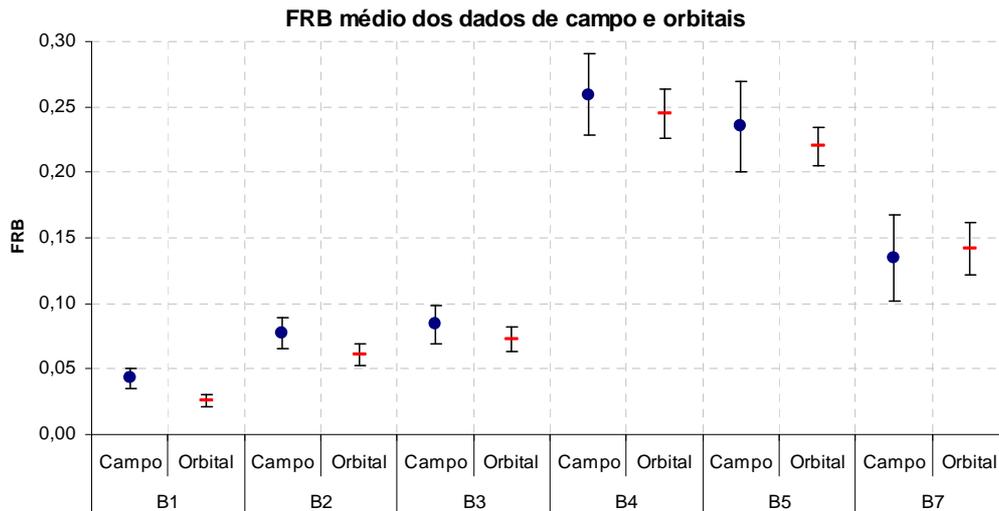


Figura 4: Gráfico dos FRB's médios dos dados de campo e orbitais em cada banda (as barras indicam o desvio padrão – que equivale à raiz quadrada da variância – nos 40 pontos amostrais).

3.2 Correlação entre índices de vegetação para os dois níveis (campo e orbital).

A Tabela 3 mostra os valores de correlação encontrados para os índices de vegetação estudados. Da mesma forma que para as correlações de FRB dos dados de campo e orbital dentro de cada banda, nenhuma das correlações dos índices foi significativa a 5% de significância no teste *t*, com exceção da correlação de *Spearman* para o NDVI. Isso provavelmente ocorreu pelo fato da banda B4 não apresentar normalidade na distribuição dos dados orbitais e, ou, pelo fato de que a associação entre os NDVI's, nos dois níveis, não ocorre de maneira linear.

Tabela 3: Índices de correlação de *Pearson* e *Spearman* dos índices de vegetação nos níveis de campo e orbital (valor significativamente diferente de zero segundo o teste *t* a 5% de significância é destacado em fundo cinza).

	NDVI	DI	SAVI
<i>Pearson</i>	0,17790	0,07994	0,12859
<i>Spearman</i>	0,52266	-0,02179	0,22451

Para efeito de visualização das diferenças entre os índices de correlação de *Pearson* e *Spearman* foi gerado um gráfico de dispersão dos valores de NDVI para os dados de campo (eixo das abscissas) e orbital (eixo das ordenadas) em cada um dos 40 pontos amostrados; e sobre os pontos plotados foram desenhadas duas retas que passam pelo centro geométrico da nuvem de pontos e são qualitativamente representativas dos dois índices de correlação supracitados (Figura 5).

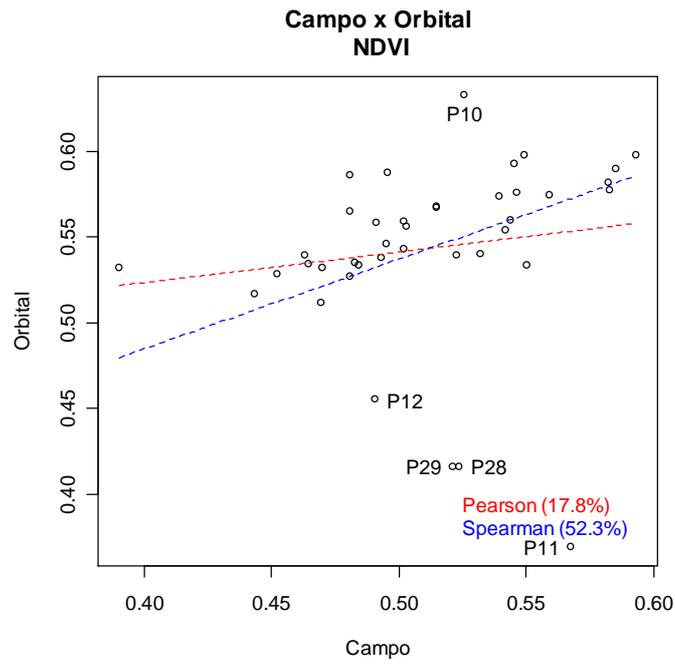


Figura 5: Gráficos de dispersão dos valores de NDVI calculados com dados de campo e orbitais e retas representativas dos índices de correlação de *Pearson* e *Spearman*.

4 Considerações finais

De maneira geral, os valores de FRB das 40 amostras, nos níveis de campo e orbital, apresentaram baixos valores de correlações de *Pearson* (média de 12,67%) e *Spearman* (média de 15,62%), sendo não significativos estatisticamente. Isso provavelmente é consequência da discrepância entre os valores das variâncias desses dados, uma vez que as diferenças entre os sistemas de coleta dos dois níveis (principalmente a resolução espacial) fazem com que os dados de campo apresentem maior variação que os orbitais.

Nos estudos de associações entre os índices de vegetação obtidos a partir de dados de campo e dados orbitais apenas o índice de correlação de *Spearman* aplicado ao NDVI foi estatisticamente significativo. Trata-se de um indício de associação não linear entre os NDVI's dos dados nos dois níveis e, ou violação de um pressuposto no cálculo dos coeficientes de correlação, que nesse caso, deve-se à falta de normalidade da banda B4 para os dados orbitais.

O estudo mostrou o quão delicadas são as relações entre dados de campo e orbitais, devido às diferenças existentes entre esses níveis, como as resoluções, influência atmosférica e geometria de coleta. Diante dessas peculiaridades é preciso ser cauteloso na manipulação de dados para que não se cometa equívocos na interpretação de seus significados e em suas aplicações.

5 Bibliografia

BUNCHAFT, G; KELLNER, S. R. O. **Estatística sem mistérios**. 2. ed. Petrópolis, RJ: Vozes, 2001. 230 p.

CÂMARA, G.; SOUZA, R. C. M.; FREITAS, U. M. GARRIDO, J. SPRING: integrating remote sensing and GIS by object-oriented data modeling. **Computers & Graphics**, v. 20, n. 3, p. 395-403, 1996.

CHANDER, G.; MARKHAM, B. Revised Landsat-5 TM radiometric calibration procedures and postcalibration dynamic ranges. **IEEE Transactions on Geoscience and Remote Sensing**, v. 41, n. 11, p. 2674-2677, 2003.

DYTHAM, C. **Choosing and using statistics: a biologist's guide**. 2. ed. Oxford, Inglaterra: Blackwell, 2003. 248p.

HUETE, A. R. A soil-adjusted vegetation index (SAVI). **Remote Sensing of Environment**, v. 25, n. 3, p. 295-309, 1988.

JACKSON, R. D.; SLATER, P. N.; PINTER JR., P. J. Discrimination of growth and water stress in wheat by various vegetation indices through clear and turbid atmospheres. **Remote Sensing of Environment**, v. 13, n. 3, p. 187-208, 1983.

JENSEN, J. R. **Remote sensing of the environment: an Earth resource perspective**. 2. ed. New Jersey, USA: Prentice Hall, 2006. 608 p.

MARSH, S. E., LYON, R. J. P. Quantitative relationships of near surface spectra to Landsat radiometric data. **Remote Sensing of Environment**, v. 10, n. 4, p. 241-261, 1980.

MORELLI, A. F.; FORMAGGIO, A. R.; EPIPHANIO, J. C. N. Influência da geometria de visada na reflectância espectral de duas gramíneas forrageiras. In: Simpósio Brasileiro de Sensoriamento Remoto, 7., 1993, Curitiba. **Anais...** São José dos Campos: INPE, 1993.

R Development Core Team. **R: A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Áustria, 2007. Disponível em: <<http://www.R-project.org>>. Acesso em: 27 nov. 2007.

VERMOTE, E. F.; TANRE, D.; DEUZE, J. L.; HERMAN, M.; MORCRETTE, J. J. Second simulation of the satellite signal in the solar spectrum, 6S: An overview. **IEEE Transactions on Geoscience and Remote Sensing**, v. 35, n. 3, p. 675-686, 1997.

Agradecimentos

Os autores agradecem aos colegas do INPE, em especial ao mestrando Luís Gonçalves Longo de Oliveira pelos dados gentilmente cedidos; ao pesquisador Flávio

Jorge Ponzoni pelas orientações; e aos pesquisadores Antônio Roberto Formaggio, Evlyn Márcia Leão de Moraes Novo e Lênio Soares Galvão, pelos valiosos conselhos.