

# ANÁLISE DE DADOS SAR PARA ESTIMATIVA VOLUMÉTRICA DA FLORESTA TROPICAL EM REGIME DE EXPLORAÇÃO MADEIREIRA

## SAR DATA ANALYSIS FOR TIMBER VOLUME ESTIMATES IN SELECTIVELY LOGGED TROPICAL FOREST

Fábio Guimarães Gonçalves<sup>1</sup>, João Roberto dos Santos<sup>1</sup>, Ana Lia de Barros Lopes<sup>1</sup>

<sup>1</sup> Instituto Nacional de Pesquisas Espaciais – INPE, Av. dos Astronautas, 1758, CEP 12.227-010  
São José dos Campos, SP, Brasil  
e-mail: {fablogg, jroberto, analia}@dsr.inpe.br

### RESUMO

Neste estudo, são apresentados resultados preliminares da avaliação do potencial de dados de radar de abertura sintética polarimétrica (PolSAR) para a estimativa volumétrica da floresta tropical sob exploração seletiva. Um modelo de regressão linear múltipla foi ajustado para a estimativa do volume comercial de madeira (VCM), a partir de um conjunto de dados coletados na Floresta Nacional do Tapajós. As variáveis explicativas selecionadas incluíram os valores mínimos de coeficiente de retroespalhamento (polarizações HH e HV) e amplitude (polarização HH), observados em parcelas amostrais de 20 x 100 m. Para validação do modelo, foram espacializadas estimativas do VCM numa área de exploração de 3.220 ha. Uma análise conjunta dos resultados a base de dados da exploração indicou não haver diferença estatística entre os valores de VCM observados e estimados. O diagnóstico do modelo de regressão selecionado indicou consistência na capacidade preditiva, quando avaliado com um amplo conjunto de dados independentes.

**Palavras-chave:** sensoriamento remoto, SAR, inventário florestal, exploração seletiva, Amazônia.

### ABSTRACT

In this study, preliminary results are shown on the retrieval of timber volume over a selectively logged tropical forest from polarimetric synthetic aperture radar (PolSAR) data. A multiple linear regression model was fitted using data from the Tapajos National Forest to predict timber volume. The selected predictor variables included minimum values of backscattering coefficients (HH and HV polarizations) and amplitude (HH polarization) reported in 20 x 100 m sample plots. The model validation process was done through spatialization of the timber volume estimates over a 3.220 ha exploitation area. The spatialization was analyzed by statistic tests using a field database. The results showed no statistically significant difference among the observed and estimated timber volume values. The regression model obtained show good predictive ability when checked against independent data.

**Key words:** remote sensing, SAR, forest inventory, selective logging, Amazon.

### INTRODUÇÃO

A exploração florestal seletiva na Amazônia brasileira vem despertando interesse particular da comunidade científica nos últimos anos, tornando-se uma temática significativa de investigação. Pesquisas vêm sendo desenvolvidas com a utilização de dados de sensoriamento remoto, especialmente no que se refere ao estabelecimento de metodologias mais acuradas para a detecção da ocorrência desta atividade (GRAÇA, 2004; ASNER et al., 2005). No entanto, as maiores incertezas no monitoramento destas áreas ainda residem na falta de conhecimento da quantidade de madeira que está sendo explorada.

A quantificação de determinados parâmetros biofísicos da cobertura florestal tem sido amplamente realizada com abordagens que empregam dados de sensoriamento remoto por microondas (LE TOAN et al., 1992; DOBSON et al., 1995; SANTOS et al., 2003). Nas regiões tropicais estes dados são especialmente úteis, em face da limitação dos sensores ópticos com as restrições impostas no imageamento pela constante presença

de nuvens. Contudo, o entendimento da dinâmica de interação do sinal radar com a estrutura complexa das florestas tropicais ainda é incipiente.

Em setembro de 2005, foi conduzida uma campanha com o radar de abertura sintética aerotransportado do Sistema de Proteção da Amazônia (SAR-R99B), adquirindo dados em banda X (polarização HH) e L (polarimétrica), sobre área de influência da Floresta Nacional do Tapajós. Neste trabalho, são apresentados alguns resultados preliminares da avaliação do potencial de determinadas variáveis SAR polarimétricas para a estimativa do volume comercial de madeira, em uma unidade sob o regime de exploração florestal seletiva.

## MATERIAL E MÉTODOS

### CARACTERIZAÇÃO DA ÁREA DE ESTUDO

A Floresta Nacional do Tapajós (FNT) está localizada no Estado do Pará, ao sul da cidade de Santarém. O clima da região, segundo a classificação de Köppen, é do tipo Amw, com temperatura média anual de 26° C. A concentração de chuvas ocorre entre fevereiro e maio, com uma precipitação média anual variando entre 1.750 e 2.000 mm. A topografia local, pouco acidentada, se apresenta com formas suavemente onduladas e onduladas, predominando na área o Latossolo Amarelo Distrófico. A vegetação da FNT é classificada como Floresta Ombrófila Densa, com a ocorrência de árvores emergentes ou com cobertura arbórea uniforme (RADAMBRASIL, 1976).

Este estudo foi conduzido em uma unidade de manejo florestal com aproximadamente 3.220 ha, instalada para a exploração sustentada de madeira em tora, conforme o plano de manejo florestal da FNT (IBAMA/PNUD, 1997). Nesta unidade foram estabelecidos 38 blocos de exploração, com uma área média de 84,7 ha, onde foram extraídos até 40 m<sup>3</sup> de madeira por hectare entre 1999 e 2003. O ciclo de corte previsto no plano de manejo é de 18 a 22 anos, esperando-se um incremento médio anual de 2,5 m<sup>3</sup>.ha<sup>-1</sup>.

### IMAGEAMENTO E PROCESSAMENTO SAR

Neste estudo foram empregados dados polarimétricos (PolSAR) adquiridos na banda L (1,28 GHz), a uma altitude média de vôo de 37.000 pés, integrando uma faixa de 20 km de largura. Na fase de processamento, os dados PolSAR L foram sintetizados com 5 looks para a redução do ruído speckle. As imagens amplitude (A) resultantes apresentaram resolução espacial de 5 m e passaram pelos processos de correção do padrão da antena e calibração radiométrica, para possibilitar a obtenção dos coeficientes de retroespalhamento ( $\sigma^0$ ). Para tanto, foram utilizadas as respostas radiométricas de 12 refletores de canto triédricos, alocados no campo durante a campanha de aquisição dos dados SAR.

As imagens A e  $\sigma^0$ , obtidas nas polarizações HH, HV e VV, foram georreferenciadas com a utilização das coordenadas dos refletores de canto, obtidas em campo com GPS geodésico. Foi utilizado o modelo polinomial de primeiro grau e a reamostragem pelo método do vizinho mais próximo, obtendo-se um erro quadrático médio inferior a um pixel da imagem.

### AQUISIÇÃO E TRATAMENTO DOS DADOS DE CAMPO

Um inventário florestal foi conduzido paralelamente ao imageamento SAR-R99B, na referida área de estudo, de forma a possibilitar o ajuste de um modelo de regressão para a estimativa do volume comercial de madeira a partir dos dados PolSAR L. Para representar uma porção significativa da variabilidade existente no interior da unidade de manejo florestal, foram estabelecidas 30 parcelas amostrais de 0,2 ha (20 x 100 m) em blocos com diferentes idades e graus de exploração. Estas parcelas foram delimitadas em campo por amostragem aleatória estratificada e posicionadas geograficamente com o auxílio do Sistema de Posicionamento Global (GPS) e uso de uma estação total. No interior das parcelas amostrais foi realizado o inventário de todos os indivíduos arbóreos com diâmetro à altura do peito (DAP) maior ou igual a 10 cm. Porém, no presente estudo foram consideradas apenas as árvores com DAP  $\geq 35$  cm, para correspondência

com aquelas inventariadas na unidade de manejo florestal. Além do diâmetro, foram realizadas estimativas visuais da altura comercial e delimitadas as posições dos indivíduos.

Com o objetivo de melhorar as estimativas realizadas para a altura comercial das árvores, foram também registradas 277 mensurações com o emprego de um clinômetro eletrônico. A partir deste conjunto de dados foi ajustado o modelo de regressão linear simples (1), em que a medida realizada por estimativa visual ( $H_{EV}$ ) é utilizada para explicar a medida obtida com o clinômetro eletrônico ( $H_{CE}$ ). O modelo resultante apresentou um coeficiente de determinação igual a 91% e um quadrado médio dos resíduos igual a 1,02 m, sendo empregado para uma estimativa mais acurada das alturas das demais árvores inventariadas.

$$H_{CE} = e^{0,184481} \cdot H_{EV}^{0,948037} \quad (1)$$

A composição volumétrica das parcelas amostrais foi determinada pelo método indireto com a utilização da equação alométrica (2), desenvolvida por QUEIROZ (1984) para a FNT. Com esta equação, o volume comercial de madeira de cada indivíduo contido na parcela ( $VCM_{ind}$ ) foi calculado, em metros cúbicos, em função dos seus valores de diâmetro à altura do peito ( $DAP$ ) e altura comercial ( $HC$ ), fornecidos em metros. Os volumes individuais resultantes foram então somados, de forma a se determinar o volume total de cada parcela amostral.

$$VCM_{ind} = 0,077476 + 0,517897 \cdot (DAP^2 \cdot HC) \quad (2)$$

## INTEGRAÇÃO E ANÁLISE DOS DADOS

Para se obter um modelo de estimativa do VCM a partir dos dados PolSAR L, foram aplicadas técnicas de análise de regressão linear múltipla. No ajuste do modelo, cada parcela inventariada constituiu uma amostra. O conjunto de variáveis explicativas incluiu o mínimo, o máximo, a média, o desvio padrão e o coeficiente de variação dos valores de A (amplitude) e  $\sigma^0$  (coeficiente de retroespalhamento) encontrados para cada observação, nas três polarizações (HH, HV e VV).

As análises de regressão foram realizadas com fundamentação na teoria apresentada por NETER et al. (1996). Na análise exploratória dos dados, diagramas de dispersão e matrizes de correlação foram utilizados para verificar a necessidade de aplicação de transformações nas variáveis e auxiliar a identificação das variáveis explicativas com maior potencial de contribuição para o modelo. Para se obter os melhores subconjuntos de variáveis explicativas, foram avaliados os efeitos de interações e aplicadas técnicas de seleção baseadas no coeficiente de determinação ajustado. Uma análise de *outliers* foi conduzida com base nos métodos DFFITS e *Cook's Distance*. Em adição, foi realizada uma análise de resíduos, com métodos gráficos e testes formais, para avaliar a aptidão do modelo selecionado.

O modelo de regressão, obtido com os procedimentos descritos, foi empregado para a espacialização das estimativas de VCM na área da unidade de manejo florestal. Para validar as estimativas realizadas, foi conduzida uma análise conjunta a base de dados dos 38 blocos de exploração, fornecida pelo IBAMA (Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis). Uma vez que o inventário e a exploração dos blocos foram realizados em épocas distintas, ambos antes da campanha de aquisição dos dados SAR em 2005, foi empregada uma equação de ajuste para atualização da base de referência:

$$VCM_{obs} = VCM_{inv} - VCM_{exp} + 2,5 \cdot AREA \cdot (2005 - ANO_{inv}) \quad (3)$$

Com a equação (3), o VCM observado ( $VCM_{obs}$ ) em um dado bloco, em 2005, foi calculado em função do VCM inventariado ( $VCM_{inv}$ ), do VCM explorado ( $VCM_{exp}$ ), da área do bloco em hectares ( $AREA$ ) e do ano de realização do inventário ( $ANO_{inv}$ ). O valor 2,5 refere-se ao incremento médio anual no volume, por hectare, esperado em função das práticas de manejo que foram adotadas na unidade (IBAMA/PNUD, 1997).

As estimativas SAR do VCM ( $VCM_{sar}$ ) foram comparadas, em média, aos valores atualizados da base de referência, com a utilização do teste t para amostras dependentes (teste t pareado). A estatística deste teste é apresentada em (4). Seu cálculo foi baseado na média ( $\bar{X}_{Y-\hat{Y}}$ ) e desvio padrão ( $S_{Y-\hat{Y}}$ ) das diferenças entre o VCM observado e estimado, nas  $n$  amostras analisadas. O nível de significância adotado para este teste foi de 5%.

$$t_{calc} = \frac{\bar{X}_{Y-\hat{Y}}}{\frac{S_{Y-\hat{Y}}}{\sqrt{n}}} \sim t_{n-1} \quad (4)$$

A análise da acurácia das estimativas, em nível de bloco de exploração, foi realizada com base em análise gráfica e no cálculo do erro quadrático médio (EQM) entre os valores observados e estimados de VCM:

$$EQM = \sqrt{\frac{1}{n} \cdot \sum_{i=1}^{38} (Y - \hat{Y})^2} \quad (5)$$

## RESULTADOS E DISCUSSÃO

### ANÁLISE DE REGRESSÃO

A matriz de correlação do conjunto total de dados é apresentada na Tabela 1. De maneira geral, os resultados apontaram baixos valores de coeficiente de correlação linear entre as variáveis polarimétricas e o VCM. Os valores mínimos apresentaram as maiores correlações para todas as variáveis analisadas. O mínimo, o máximo e a média dos valores de  $A$  e  $\sigma^\circ$  apresentaram, na polarização HH, suas melhores correlações. Resultados similares foram encontrados por GAMA et al. (2005), ao avaliar as correlações entre o VCM e o  $\sigma^\circ$  médio de dados PolSAR adquiridos em banda P, também na FNT. Do total de 30 variáveis analisadas, apenas 3 (valores mínimos de  $A_{HH}$ ,  $\sigma^\circ_{HH}$  e  $\sigma^\circ_{HV}$ ) apresentaram correlações significativas com o VCM. Os diagramas de dispersão das variáveis explicativas *versus* o VCM não indicaram a necessidade de aplicação de transformações para a linearização dos dados, de forma a melhorar os resultados. Nesta etapa, as variáveis que apresentaram correlações inferiores a 0,3 foram descartadas das análises subseqüentes, por proporcionarem uma fraca contribuição na explicação da variável dependente. Desta forma, 8 variáveis polarimétricas foram pré-selecionadas ao final da análise exploratória.

Tabela 1. Matriz de correlação entre as variáveis polarimétricas e o VCM.

Variável SAR	Mínimo	Máximo	Média	Desvio	CV%
$A_{HH}$	0,50	0,34	0,28	0,08	-0,22
$A_{HV}$	0,33	-0,05	0,22	-0,15	-0,32
$A_{VV}$	0,24	0,12	0,22	0,07	-0,15
$\sigma^\circ_{HH}$	0,49	0,28	0,24	-0,30	-0,22
$\sigma^\circ_{HV}$	0,36	-0,07	0,23	-0,31	0,15
$\sigma^\circ_{VV}$	0,24	0,05	0,20	-0,15	-0,23

\*Em vermelho, correlações significativas ao nível de 5% de significância.

Em função do número limitado de amostras, optou-se por conduzir a seleção de variáveis com a formação de modelos com no máximo 3 parâmetros. Com base no coeficiente de determinação ajustado, o melhor subconjunto de variáveis explicativas, escolhido entre todas as possíveis combinações das variáveis pré-selecionadas e suas interações, incluiu os valores mínimos de  $\sigma^\circ_{HH}$ ,  $\sigma^\circ_{HV}$  e  $\sigma^\circ_{HH} * A_{HH}$ .

A análise de *outliers* pelo método DFFITS indicou a presença de 7 casos influentes. Entretanto, não foram detectados casos influentes pelo método *Cook's Distance*. Tais resultados sugerem que os *outliers* identificados produzem influência nas próprias estimativas, mas não influenciam todos os  $n$  valores

ajustados. Ao se retirar os *outliers* identificados, não ocorreram alterações significativas nos parâmetros do modelo de regressão. Em adição, o valor do coeficiente de determinação sofreu um acréscimo de 0,36 para 0,49. Com base nestes resultados, optou-se por não conservar os *outliers* influentes. O modelo selecionado pode ser observado em (6).

$$VCM_{sar} = -139,853695 + 0,092800 \cdot (\sigma_{HH}^0 * \sigma_{HV}^0) - 0,002685 \cdot (\sigma_{HH}^0 * A_{HH}) \quad (6)$$

Na Figura 1 são apresentados diagramas de dispersão dos resíduos da regressão *versus* as variáveis explicativas empregadas no modelo. A análise destes diagramas indicou que o ajuste linear foi adequado aos dados. Este fato foi observado pela ausência de tendências sistemáticas na distribuição dos resíduos ao longo da reta horizontal centrada na origem. Ainda pela análise destes diagramas, foi observado que a variância dos resíduos não variou com o nível das variáveis explicativas. O teste de Levene modificado confirmou a constância da variância dos resíduos, ao nível de 5% de significância (P=0,308).

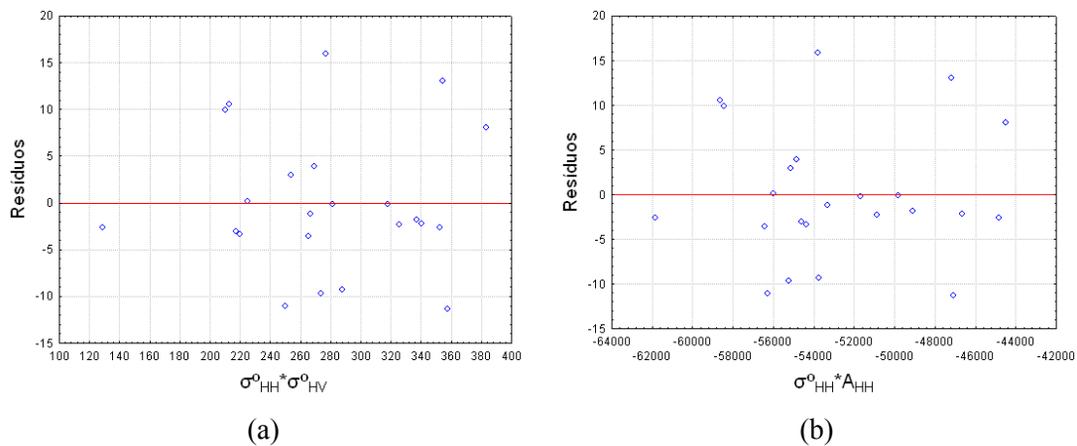


Figura 1. Diagramas de dispersão dos resíduos do modelo *versus* (a)  $\sigma_{HH}^0 * \sigma_{HV}^0$  e (b)  $\sigma_{HH}^0 * A_{HH}$ .

O *box plot* e o gráfico de probabilidade normal dos resíduos da regressão são apresentados na Figura 2a e 2b, respectivamente. A análise do *box plot* não sugere maiores problemas com relação à normalidade dos resíduos, indicando uma simetria razoável dos dados. Porém, a seqüência não linear do gráfico de probabilidade normal sugere que a distribuição dos resíduos não é gaussiana. Os testes formais de Kolmogorov-Smirnov (P>0,2) e Shapiro-Wilk (P=0,115) confirmaram a normalidade dos resíduos, ao nível de 5% de significância.

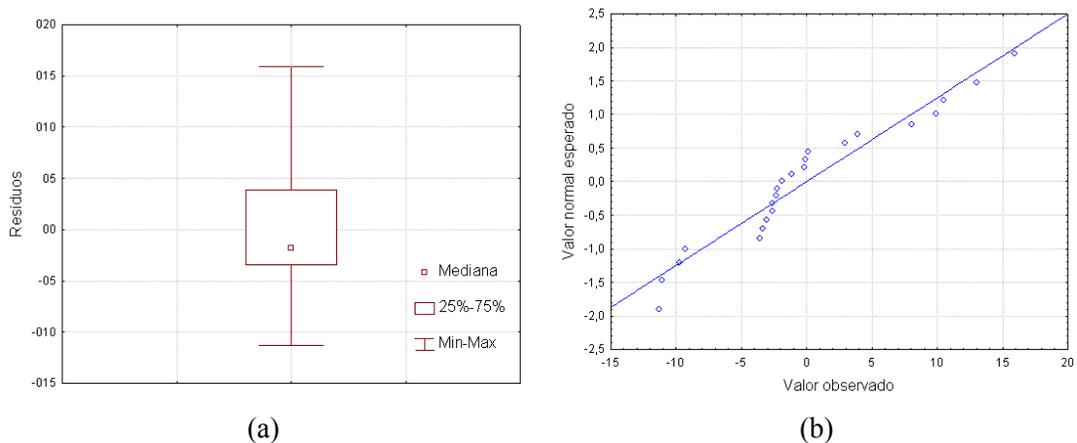


Figura 2. Resíduos: (a) *Box plot* e (b) gráfico de probabilidade normal.

## VALIDAÇÃO DO MODELO

O  $VCM_{sar}$  médio, computado para os 38 blocos de exploração, foi igual a  $135,40 \text{ m}^3.\text{ha}^{-1}$ . A base de referência, atualizada pela equação (3), resultou num  $VCM_{obs}$  médio igual a  $135,47 \text{ m}^3.\text{ha}^{-1}$ . Apesar do baixo coeficiente de determinação do modelo de regressão selecionado, foi obtido um erro, na média da estimativa do VCM, de apenas  $0,07 \text{ m}^3.\text{ha}^{-1}$ . Através do teste t para amostras dependentes, pôde-se confirmar a igualdade das médias do  $VCM_{sar}$  e do  $VCM_{obs}$ , adotando-se um nível de significância de 5% ( $P=0,989$ ).

Na Figura 3 é apresentado o ajuste linear dos valores de VCM observados e estimados, para os 38 blocos de exploração. O VCM de 79% dos blocos foi estimado com erros inferiores a 20%. O erro quadrático médio entre os valores observados e estimados foi de  $28,3 \text{ m}^3.\text{ha}^{-1}$ . Quando os *outliers* destacados na Figura 3 não foram considerados no cálculo, este valor foi reduzido para  $17,9 \text{ m}^3.\text{ha}^{-1}$ . As estimativas decorrentes destas amostras consideradas *outliers* apresentaram os maiores desvios encontrados em relação aos valores observados (38,6% e 41,1%). Entretanto, não foram identificadas razões aparentes que justificassem estes comportamentos. Os resultados da validação demonstraram que as estimativas do modelo de regressão selecionado foram consistentes, não só para a média do VCM, como também para a estimativa individual de cada bloco explorado.

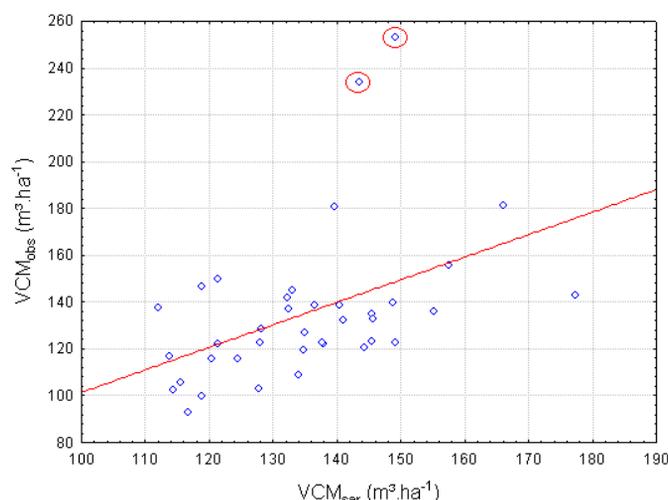


Figura 3. Diagrama de dispersão do  $VCM_{obs}$  versus o  $VCM_{sar}$ .

## CONSIDERAÇÕES FINAIS

Neste estudo, foi comprovado o potencial de dados PolSAR L na estimativa volumétrica da floresta tropical, em regime de exploração seletiva. O diagnóstico do modelo de regressão selecionado indicou consistência na capacidade preditiva, quando avaliado com um amplo conjunto de dados independentes.

É importante notar que os resultados apresentados neste trabalho são preliminares. Análises estão sendo conduzidas com a inclusão da banda X e outras variáveis polarimétricas importantes na modelagem. Como exemplo, espera-se que a utilização de índices biofísicos, somados a informação de fase, resulte em estimativas pontuais de VCM mais acuradas. Por outro lado, esforços estão sendo empenhados para a geração de modelos menos complexos, que também produzam informações estatisticamente confiáveis, mas que apresentem um significado físico que possa ser explicado.

## AGRADECIMENTOS

Este trabalho foi parcialmente desenvolvido com apoio de bolsas de estudo CAPES e de pesquisa CNPq. Os autores agradecem o Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis (IBAMA) pelo fornecimento da base de dados dos blocos de exploração e ao LBA/Santarém pelo suporte logístico.

Agradecimentos são estendidos ao Dr. José Cláudio Mura (INPE) pelo auxílio no processamento dos dados SAR-R99B.

## REFERÊNCIAS BIBLIOGRÁFICAS

ASNER, G.P. et al. Selective logging in the Brazilian Amazon. *Science*, v. 310, n. 5747, p. 480-482, 2005.

DOBSON, M.C. et al. Estimation of forest biophysical characteristics in Northern Michigan with SIR-C/X-SAR. *IEEE Transactions on Geoscience and Remote Sensing*, v. 33, n. 4, p. 877-895, 1995.

GAMA, F.F. et al. Aplicação de dados polarimétricos de radar aerotransportado (banda P) para estimativa de parâmetros estruturais de florestas tropicais. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 7, 2005. Goiânia. *Anais do XII Simpósio Brasileiro de Sensoriamento Remoto*, Goiânia, 2005. p. 4413-4420.

GRAÇA, P.M.L.A. *Monitoramento e caracterização de áreas submetidas à exploração florestal na Amazônia por técnicas de detecção de mudanças*. 2004. 262p. Tese – Instituto Nacional de Pesquisas Espaciais, São José dos Campos. 2004.

INSTITUTO BRASILEIRO DO MEIO AMBIENTE E DOS RECURSOS NATURAIS - IBAMA/PNUD. *Plano de manejo florestal da floresta nacional do Tapajós, Pará, Brasil*. Belém: 1997. 109 p.

LE TOAN, T.L. et al. Relating forest biomass to SAR data. *IEEE Transactions on Geoscience and Remote Sensing*, v. 30, n. 2, p. 403-411, 1992.

NETER, J. et al. *Applied linear statistical models*. 4.ed. Boston: McGraw-Hill, 1996. 1408p.

QUEIROZ, W.T. *Análise de fatores pelo método da máxima verossimilhança: aplicação ao estudo da estrutura de florestas tropicais*. 1984. Tese – Universidade de São Paulo, Escola Superior de Agricultura Luíz de Queiroz, Piracicaba. 1984.

RADAMBRASIL (1976). Folha SA. 21 Santarém; geologia, geomorfologia, solos, vegetação e uso potencial da terra. Rio de Janeiro, 522p. (Levantamento de recursos naturais, 10)

SANTOS, J.R. et al. Airborne P-band SAR applied to the aboveground biomass studies in the Brazilian tropical rainforest. *Remote Sensing of Environment*, v. 87, n. 4, p. 482-493, 2003.