

GUILHERME SILVERIO AQUINO DE SOUZA

*MÁQUINA DE VETOR DE SUPORTE APLICADA A DADOS DE
SENSÓRIAMENTO REMOTO PARA A ESTIMAÇÃO DE VOLUME E
BIOMASSA AÉREA DE POVOAMENTOS DE EUCALIPTO*

Dissertação apresentada à Universidade Federal de Viçosa como parte das exigências do Programa de Pós-Graduação em Ciência Florestal para obtenção do título de *Magister Scientiae*.

VIÇOSA
MINAS GERAIS – BRASIL
2016

**Ficha catalográfica preparada pela Biblioteca Central da Universidade
Federal de Viçosa - Câmpus Viçosa**

T

S729m
2016

Souza, Guilherme Silverio Aquino de, 1991-
Máquina de vetor de suporte aplicada a dados de
sensoriamento remoto para a estimação de volume e biomassa
aérea de povoamentos de eucalipto / Guilherme Silverio Aquino
de Souza. – Viçosa, MG, 2016.
vii, 81f : il. ; 29 cm.

Inclui apêndices.

Orientador: Vicente Paulo Soares.

Dissertação (mestrado) - Universidade Federal de Viçosa.

Referências bibliográficas: f.67-76.

1. Levamentos florestais. 2. Biomassa florestal. 3. Eucalipto
- Povoamento. 4. Florestas - Medição. 5. Algoritmos .

I. Universidade Federal de Viçosa. Departamento de Engenharia
Florestal. Programa de Pós-graduação em Ciência Florestal.

II. Título.

CDD 22. ed. 634.5246

GUILHERME SILVERIO AQUINO DE SOUZA

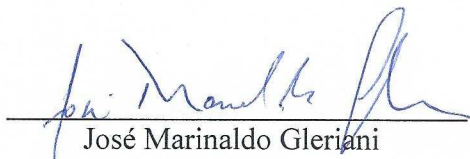
**MÁQUINA DE VETOR DE SUPORTE APLICADA A DADOS DE
SENSORIAMENTO REMOTO PARA A ESTIMAÇÃO DE VOLUME E
BIOMASSA AÉREA DE POVOAMENTOS DE EUCALIPTO**

Dissertação apresentada à Universidade Federal de Viçosa como parte das exigências do Programa de Pós-Graduação em Ciência Florestal para obtenção do título de *Magister Scientiae*.

APROVADA: 27 de abril de 2016.


Antônio Santana Ferraz


Cibele Hummel do Amaral


José Marinaldo Gleriani


Vicente Paulo Soares
(Orientador)

AGRADECIMENTOS

A Deus.

Ao apoio incondicional da minha família sul-mato-grossense.

Ao professor Vicente Paulo Soares pela oportunidade proporcionada em cursar o mestrado.

Ao professor Helio Garcia Leite pela amizade e conselhos durante minha jornada em Viçosa.

Ao departamento e aos professores da Engenharia Florestal da Universidade Federal de Viçosa pelos ensinamentos e oportunidade da pós-graduação.

À secretaria do Programa de Pós-graduação em Ciência Florestal da Universidade Federal de Viçosa. Meus agradecimentos aos colegas Alexandre Amorim e Dilson Garcia.

À Fundação de Amparo à Pesquisa do Estado de Minas Gerais – FAPEMIG – pelo apoio financeiro.

Aos ótimos amigos adquiridos ao longo do mestrado, em especial Gabriel Garnica, Rafaella Lima, Márcio Machado, Wagner Canal, Diogo Cosenza, Reginaldo Medeiros, Adriana Gómez, Crismeire Isbaex, José Batuíra, Leonardo Fardin, Vicente Júnior, Gustavo Fagundes, e a todos aqueles que de alguma forma contribuíram para essa jornada maravilhosa.

BIOGRAFIA

Guilherme Silverio Aquino de Souza é natural de Aquidauana, Mato Grosso do Sul. Filho de Luzia Lourenço da Cruz Silverio e Idílio Aquino de Souza, nasceu em 13 de setembro de 1991.

Concluiu o Ensino Médio na Escola Estadual Cândido Mariano em 2008. No mesmo ano foi selecionado no programa Jovens Embaixadores da Embaixada Norte-americana para um período de imersão nos Estados Unidos da América.

Em 2009, iniciou o curso recém-criado de Engenharia Florestal pela Universidade Estadual do Mato Grosso do Sul no campus de Aquidauana, fazendo parte da primeira turma, formada em 2013.

Em 2013, estagiou no Departamento de Engenharia Florestal da Universidade Federal de Viçosa na área de pesquisa em manejo florestal e estatística aplicada, sendo orientado pelo Professor Doutor Helio Garcia Leite.

Em 2014, iniciou o programa de mestrado em Ciência Florestal na Universidade Federal de Viçosa, sendo orientado pelo Professor Doutor Vicente Paulo Soares trabalhando com sensoriamento remoto aplicado ao manejo florestal.

SUMÁRIO

<i>RESUMO</i>	<i>vi</i>
<i>ABSTRACT</i>	<i>vii</i>
<i>1 INTRODUÇÃO</i>	<i>1</i>
<i>2 REVISÃO BIBLIOGRÁFICA</i>	<i>5</i>
2.1 <i>Fundamentos do Sensoriamento Remoto</i>	<i>5</i>
2.2 <i>Sensores Ópticos</i>	<i>6</i>
2.3 <i>Sensores RADAR</i>	<i>7</i>
2.4 <i>Estimativas de Parâmetros Biofísicos Florestais por SR</i>	<i>9</i>
2.4.1 <i>SR para estimativa de volume de povoamentos florestais</i>	<i>10</i>
2.4.2 <i>SR para estimativa de biomassa aérea de povoamentos florestais</i>	<i>13</i>
2.5 <i>Satélite ALOS</i>	<i>15</i>
2.6 <i>Máquina de Vetores de Suporte</i>	<i>16</i>
2.6.1 <i>Máquina de Vetor de Suporte para Classificação</i>	<i>17</i>
2.6.2 <i>Máquina de Vetor de Suporte para Regressão</i>	<i>24</i>
<i>3 MATERIAL E MÉTODOS</i>	<i>29</i>
3.1 <i>Áreas de Estudo</i>	<i>29</i>
3.2 <i>Procedimentos metodológicos empregados</i>	<i>30</i>
3.2.1 <i>Imagens e softwares</i>	<i>30</i>
3.2.2 <i>Inventário florestal</i>	<i>31</i>
3.2.3 <i>Seleção das áreas amostrais e extração dos valores numéricos</i>	<i>31</i>
3.3 <i>Estimativas a partir das imagens AVNIR-2 e PALSAR</i>	<i>32</i>
3.4 <i>Máquina de Vetor de Suporte (MVS) para regressão</i>	<i>33</i>
3.5 <i>Teste de Validação e Aplicação das MVS</i>	<i>35</i>
<i>4 RESULTADOS</i>	<i>38</i>
4.1 <i>Treinamento das MVS para estimativa de volume de madeira</i>	<i>38</i>
4.1.1 <i>Treinamento das MVS com 75% dos dados</i>	<i>38</i>
4.1.2 <i>Redução do número de amostras para o treinamento das MVS</i>	<i>41</i>
4.1.3 <i>Redução do número de variáveis de entrada</i>	<i>45</i>
4.2 <i>Treinamento das MVS para estimativa de biomassa aérea</i>	<i>50</i>
4.2.1 <i>Treinamento das MVS com 75% dos dados</i>	<i>50</i>
4.2.2 <i>Redução do número de amostras para o treinamento das MVS</i>	<i>53</i>
4.2.3 <i>Redução do número de variáveis de entrada</i>	<i>57</i>
<i>5 DISCUSSÃO</i>	<i>62</i>

<i>6 CONCLUSÕES</i>	66
<i>7 BIBLIOGRAFIA</i>	67
<i>APÊNDICES</i>	77

RESUMO

SOUZA, Guilherme Silverio Aquino de, M. Sc., Universidade Federal de Viçosa, abril de 2016. *Máquina de Vetor de Suporte Aplicada a Dados de Sensoriamento Remoto para Estimativa de Volume e Biomassa Aérea de Povoamentos de Eucalipto*. Orientador: Vicente Paulo Soares. Co-orientador: Helio Garcia Leite.

O objetivo deste estudo foi avaliar a eficiência do algoritmo de aprendizagem supervisionado Máquina de Vetores de Suporte (MVS) para regressão, em dados obtidos pelos sensores AVNIR-2 e o PALSAR a bordo do satélite ALOS, para estimativa de volume de madeira e biomassa aérea de povoamentos florestais de eucalipto. O trabalho também avaliou a redução do número de dados para treinamento, bem como a busca de estruturas de MVS mais simples. O estudo foi conduzido em áreas de reflorestamento na região leste do estado de Minas Gerais. Foram testados dois tipos de regressão ϵ -MVS e ν -MVS, bem como as seguintes funções *kernel*: linear, polinomial, RBF (*Gaussian Radial Basis Function*) e sigmoide. Foram utilizadas 190 parcelas, sendo 75% (142) para treinamento das MVS. A integração de todos os dados dos sensores AVNIR-2 e PALSAR, como variáveis de entrada das MVS, geraram estimativas para generalizações com os maiores valores de coeficiente de correlação ($r_{\hat{y}y}$) e menores valores da raiz quadrada do erro quadrático médio (RQEM), tanto para a variável volume quanto para a biomassa aérea. As ν -MVS com a função RBF resultaram em maior exatidão, e foram então utilizadas para as análises subsequentes. Em um treinamento com 95 parcelas obteve-se exatidão satisfatória nas generalizações para ambas as variáveis alvo. Para estimar volume dos povoamentos, generalizações relativamente satisfatórias dos dados foram encontradas combinando as bandas AV3 ou AV4 do sensor AVNIR-2 com uma polarização cruzada (L_{HV} ou L_{VH}) do sensor PALSAR. A combinação da banda AV4 e a polarização L_{VH} apresentou um $r_{\hat{y}y}$ de 0,991 e RQEM de 4,120 m³, equivalente a 1,617% da média das parcelas testadas. Já para biomassa aérea boas generalizações foram alcançadas combinando a banda AV4 do sensor AVNIR-2 com a polarização L_{HV} ou L_{VV} . No melhor desempenho de MVS para estimativas de biomassa aérea, utilizando a combinação AV4 e L_{VV} , obteve-se um $r_{\hat{y}y}$ de 0,951 e RQEM de 5,338 t.ha⁻¹, o equivalente a 2,095% da média das parcelas testadas. A análise gráfica dos resíduos se mostrou primordial para a seleção de MVS com menor viés quantitativo. A simplicidade de ajuste das MVS, atrelada ao poder de generalização dos dados mostrou o potencial operacional do algoritmo nas atividades de quantificação e mapeamento dos recursos florestais.

ABSTRACT

SOUZA, Guilherme Silverio Aquino de, M. Sc., Universidade Federal de Viçosa, april 2016. *Predicting *Eucalyptus* spp. Stand Volume and Above Ground Biomass using Support Vector Regression with Remote Sensing Data*. Adviser: Vicente Paulo Soares. Co-advisor: Helio Garcia Leite.

The key intention of this study was to evaluate the performance of Support Vector Regression (SVR) algorithm on *Eucalyptus* spp. timber stand volume and Above Ground Biomass (AGB) retrieval based on ALOS imagery. The sensors used were AVNIR-2 and PALSAR for optical and RADAR dataset respectively. This study also examined SVRs with a reduced number of inputs and training samples in order to reduce the amount of field measures. This study was carried out with 190 sample plots of eucalyptus stands from eastern region of Minas Gerais State, Brazil. The generalization ability of ϵ -SVR and ν -SVE were compared, as well as the performance of following kernel functions: linear, polynomial, Gaussian Radial Basis Function (RBF) and sigmoid. Initially, 142 samples were used for the training of SVRs and the remaining samples were used for the generalization test. The best generalization performances for volume and AGB estimates were obtained using ν -SVR with RBF as kernel function, and by combining all inputs from AVNIR-2 and PALSAR. This setting was used for subsequent generalized data analysis. Results with no loss of accuracy were obtained using just 95 training samples with a reduced number of inputs. The combination of AV3 or AV4 from AVNIR-2 with a cross-polarized input (L_{HV} ou L_{VH}) from PALSAR showed good performances for stand volume timber retrieval. The best result for volume output variable obtained a coefficient correlation ($r_{\hat{y}y}$) value of 0,991 and 4,120 m³ of root-mean-square error (RMSE). For AGB, the combination of AV4 optic data with L_{HV} or L_{VV} polarization input showed the best results. The best performance for AGB retrieval obtained a $r_{\hat{y}y}$ value of 0,951 and 5,338 t/ha of RMSE. The residual graphic analysis was essential for selecting less biased SVRs. Therefore, the good generalization ability places the SVR as potentially useful approach for remote sensing-based forest parameters retrieval.

1 INTRODUÇÃO

Um dos marcos do início da história do sensoriamento remoto orbital é o lançamento do satélite de observação de recursos terrestres ERTS-1 (*Earth Resources Technology Satellite*) em 1972, posteriormente renomeado de Landsat-1. A partir de então, e com o advento da internet, uma grande quantidade de dados de imageamento orbital começou a ser gradativamente disponibilizado aos milhares de pesquisadores e usuários ao redor do mundo. Vários outros programas e aprimoramentos surgiram, fornecendo imagens com diferentes resoluções espectrais, radiométricas, espaciais e temporais, permitindo novas abordagens no mapeamento e monitoramento dos recursos terrestres, bem como dos recursos naturais (CAMPBELL; WYNNE, 2011).

Sobre a crescente preocupação com as mudanças climáticas, essas novas abordagens possibilitaram o aprimoramento de estudos florestais para os seus mapeamentos, quantificações e estimativas, subsidiando projetos que visam reduzir emissões de carbono e aprimorar as práticas de manejo florestal. O ganho de eficiência no monitoramento de biomassa florestal, por exemplo, diante da sua extrema relevância para o entendimento do ciclo de carbono global, instituiu avanços para viabilizar negociações de crédito de carbono. Da mesma forma, é o ganho de eficiência no manejo de plantios florestais (AVTAR; SUZUKI; SAWADA, 2014).

Os plantios comerciais florestais constituem sumidouros de carbono atmosférico e diminuem a pressão da demanda de produtos florestais madeireiros e não-madeireiros de florestas naturais, ajudando a reduzir de certa forma o desmatamento (PUTZ et al., 2008). As boas práticas de manejo florestal, bem como sua eficiência, representam um dos métodos mais eficientes para redução de emissão de gases efeito estufa. Assim, tornam-se passíveis de serem implementados nas políticas de projetos REDD+ (*Reducing Emission from Deforestation and Forest Degradation Plus*), cujos objetivos visam auxiliar a redução de CO₂ evitando o avanço do desmatamento e protegendo a biodiversidade, resguardando os serviços ambientais (DUDLEY, 2009).

Somado a este contexto, o aumento do poderio computacional e o desenvolvimento de técnicas de processamento de dados digitais, vêm viabilizando estudos dos recursos naturais cada vez mais precisos e acurados (WANG et al., 2010; LI; XU; GUO, 2014).

Dados fornecidos por sensores orbitais que atuam na faixa óptica do espectro eletromagnético (intervalos do visível e infravermelho), vêm sendo utilizados no

mapeamento e monitoramento de diferentes formações florestais. Dentre os disponíveis destacam-se aqueles obtidos por sensores dos satélites Landsat, Ikonos, CBERS, Quickbird, SPOT, GeoEye, etc. Contudo, a cobertura de nuvens limita a aquisição destes dados principalmente em regiões tropicais, onde a nebulosidade é maior em certas épocas do ano. Além disto, a interação da radiação da região óptica com vegetação registrada nas imagens limita a análise às camadas superiores do dossel florestal, dificultando a estimativa de variáveis biofísicas relacionadas à estrutura vertical da floresta (HEDLEY et al., 2012).

Como uma forma de contornar as limitações de sensores que operam na faixa óptica, os estudos de parâmetros biofísicos florestais passaram a empregar dados de sensoriamento remoto por micro-ondas (DOBSON et al., 1992; LE TOAN et al., 1992). Atualmente, estão em operação os seguintes satélites de radar de abertura sintética: RADARSAT 2, PALSAR, ENVISAT, Cosmos SkyMed eo Terra SAR-X. No Brasil, um satélite com sensor radarmétrico que adquiriu um grande volume de dados foi o *Advanced Land Observing Satellite* (ALOS) com o sensor *Phased Array L-band Synthetic Aperture Radar* (PALSAR) que opera na faixa da banda L, com comprimento de onda de 23 cm (IGARASHI, 2001).

Dados radarmétricos podem fornecer informações multifacetadas no tocante aos seus parâmetros de aquisição: frequência, ângulo de incidência, linha de base de interferometria e polarimetria. Dados de múltiplas polarimetrias podem informar a geometria de espalhamento dos alvos, morfologia de árvores, bem como a umidade do solo sob o dossel florestal (ALI et al., 2015). Além disso, o sensoriamento remoto por radar apresenta independência de iluminação solar e proporciona a obtenção de imagens sob as mais variadas condições meteorológicas, mostrando-se uma eficiente ferramenta de suporte no processo de caracterização, inventário e monitoramento florestal.

Estimativa de variáveis por sensoriamento remoto (SR), seja para modelos empíricos ou teóricos, constituem uma problemática de certa complexidade e não-linearidade (MOSES et al., 2015; ZHENG; MOSKAL, 2009). Uma categoria de poderosos métodos para regressão, que tem sido eficientemente introduzida no campo das estimativas de variáveis geo e biofísicas florestais, nas últimas duas décadas, são técnicas não lineares de aprendizagem de máquina (*Machine Learning*). Devido às avançadas estratégias de aprendizagem, tais técnicas exploram informações contidas no conjunto de amostras e são capazes de predizer valores até mesmo de complexos casos não-lineares. Uma outra vantagem é o fato de não existir pressuposições da distribuição dos dados para

a sua aplicação. Dessa forma, o processo de modelagem pode integrar dados vindos de diferentes fontes, com distribuições até mesmo não conhecidas, e retornar previsões consistentes das variáveis alvos de estudo (ALI et al., 2015).

As Máquinas de Vetores de Suporte (MVS), Redes Neurais Artificiais (RNA), Árvores de decisão, *Naive Bayes*, *Stochastic Gradient Boosting* (SGB), *Random Forest* (RF), dentre outros, são exemplos de técnicas matemáticas de aprendizagem de máquina com aplicação em diversos campos para solucionar casos de classificação - variável de saída categórica- ou regressão - variável de saída contínua (SAMMUT; WEBB, 2010).

Na ciência florestal, as RNA vem mostrando eficiência para casos de classificação (GONG et al., 2013; OLIVEIRA et al., 2010; PAX-LENNEY et al., 2001) e regressão (BINOTI, 2012; FERRAZ et al., 2014). Já as MVS destacam-se principalmente para estudos envolvendo classificação (CHEHADE et al., 2009; COSENZA et al., 2015; HUANG et al., 2008). Entretanto, foram reformuladas para estimarem valores contínuos (VAPNIK, 1995) sendo, neste caso, denominadas também de MVS para regressão, MVS com aproximação por regressão ou Regressão de Vetor Suporte, em inglês, *Support Vector Regression*.

Após a reformulação das MVS em 1995, estudos vêm mostrando sua eficácia para a modelagem de atributos da vegetação (DURBHA; KING; YOUNAN, 2007). Os resultados ressaltam vantagens como a capacidade de generalização e a eficácia em lidar com ruídos dos dados mesmo com um número limitado de unidades de amostra. Para povoamentos florestais de eucalipto, Magalhães (2014) obteve êxito na aplicação da MVS para regressão na prognose de produção. Porém, desconhecem-se estudos que buscam estimar parâmetros florestais de eucalipto a partir de SR com aplicação das MVS.

Assim, o presente trabalho pretendeu avaliar a eficiência da Máquina de Vetor Suporte para Regressão em dados ópticos e de RADAR, respectivamente fornecidos pelos sensores AVNIR-2 e PALSAR do satélite ALOS, para estimar volume de madeira e biomassa aérea de povoamentos de eucalipto. Os objetivos específicos do estudo foram:

- Avaliar estimativas geradas pelas MVS buscando a melhor configuração e os melhores dados de entrada para estimar volume de madeira e biomassa aérea de povoamentos de eucalipto;
- Com base nos dados inventariados, avaliar a capacidade de generalização dos dados pelas MVS, com um número mínimo de parcelas para o treinamento, sem perda de exatidão.

- Avaliar estimativas geradas pelas MVS buscando estruturas mais simples para estimar volume de madeira e biomassa aérea florestal de povoamentos de eucalipto.

2 REVISÃO BIBLIOGRÁFICA

2.1 Fundamentos do Sensoriamento Remoto

Para Lillesand, Kiefer e Chipman (2015), o Sensoriamento Remoto (SR) constitui a arte e a ciência de estudar a matéria sem ter contato físico com a mesma, baseando-se somente na interação da radiação eletromagnética (REM) com a matéria.

Jensen (2000) descreve a REM como uma forma dinâmica de energia com comportamento ondulatório e campo eletromagnético ortogonal a propagação da onda (Figura 1). Os parâmetros que caracterizam o movimento ondulatório são: comprimento de onda (λ), frequência (f) e velocidade (c). A uma velocidade constante de aproximadamente $3 \cdot 10^8$ m/s, λ é inversamente proporcional a f . Segundo o mesmo autor, o comprimento de onda é formalmente definido como a distância média entre dois máximos (ou mínimos) consecutivos de padrão periódico. Com bases nestes parâmetros de onda, é possível ranqueá-las no espectro eletromagnético (Figura 2).

As técnicas de sensoriamento remoto permitem a aquisição de imagens da superfície terrestre nas várias regiões do espectro eletromagnético. Uma das principais características de imagens *remotamente* sensoriadas é a região de comprimento de onda representada do espectro. Este intervalo de comprimento de onda (ou frequência) do espectro eletromagnético é comumente chamado de banda, canal ou região (CAMPBELL; WYNNE, 2011; JENSEN, 2000).

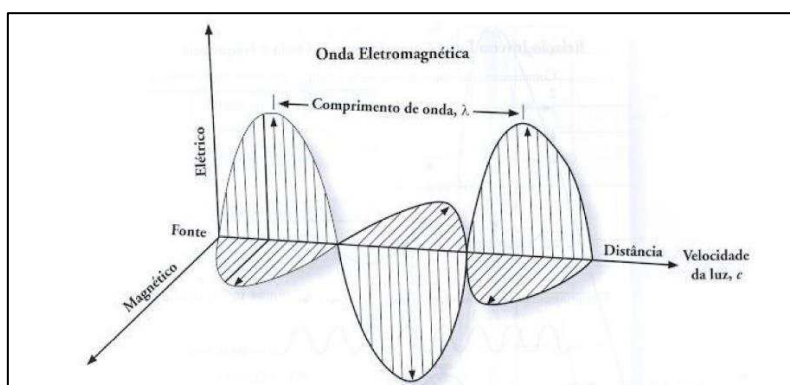


Figura 1 – Representação de uma onda Eletromagnética composta por vetores elétrico e magnético, que são ortogonais entre si. As ondas deslocam-se da fonte à velocidade da luz ($3 \cdot 10^8$ m.s⁻¹). Fonte: Jensen (2000).

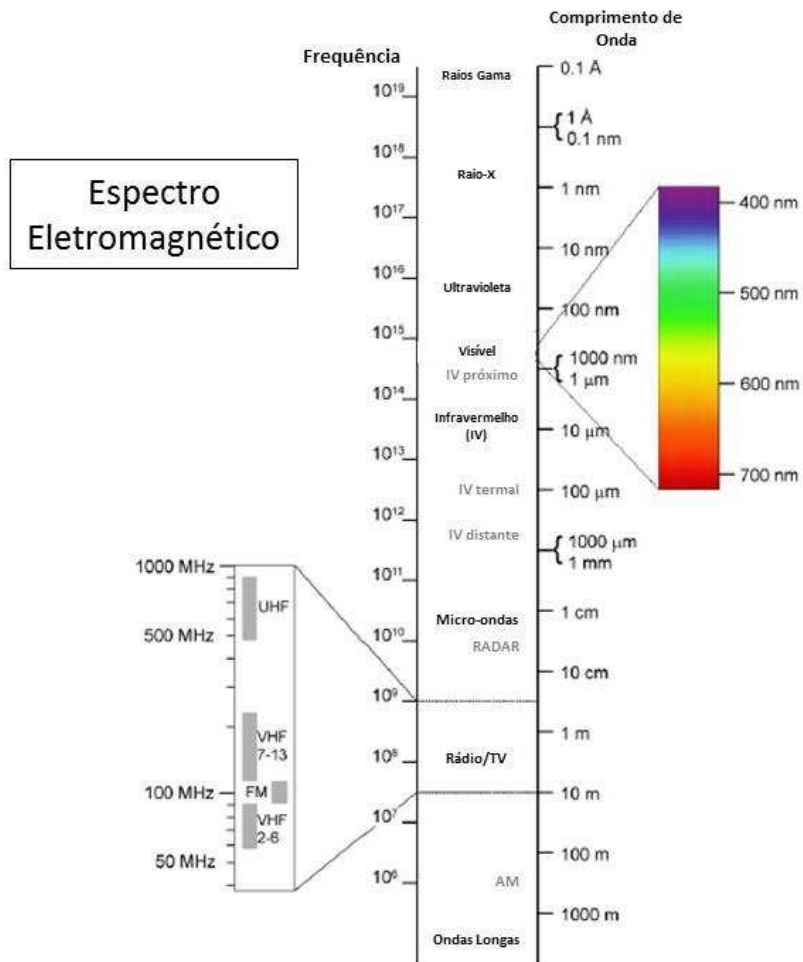


Figura 2 – Espectro Eletromagnético. Os intervalos de comprimento de onda do espectro eletromagnético são comumente chamados de banda, canal ou região. (Fonte: <https://www.mcgoodwin.net/pages/rsess421.pdf>)

As três principais regiões em estudos de imagens orbitais para recursos terrestres são: a região do visível, a do infravermelho e a das micro-ondas. As duas primeiras podem ser trabalhadas com sensores que não emitem a REM, denominados sensores passivos, que registram valores relativos a reflectância e emissividade das ondas EM. Para micro-ondas existem os sensores radarmétricos ativos que são capazes de emitir REM, e mensurar o retroespalhamento de onda da superfície terrestre (GAL-EL-HAK, 2008).

2.2 Sensores Ópticos

A radiação eletromagnética denominada “óptica” se limita à região do espectro eletromagnético que considera o intervalo de comprimento de onda do visível ao infravermelho próximo, médio e termal (0,3 a 15 μm). Segundo Campbell; Winnie (2011)

a região ótica define os comprimentos de onda que podem ser refletidos e refratados em equipamentos com lentes e espelhos.

Os objetos na superfície terrestre como vegetação, água e o solo refletem, absorvem e transmitem radiação eletromagnética em porções que variam com o comprimento da onda, de acordo com suas características bio-físico-químicas (Figura 3). Na região de luz visível a vegetação (verde e sadia) reflete mais energia na faixa correspondente ao verde, por isso a enxergamos assim. Porém é na região do infravermelho próximo que a vegetação reflete mais energia e se diferencia dos demais objetos. Em termos gerais, o solo possui baixa variação de energia refletida de acordo com os comprimentos de onda. A água limpa reflete pouca energia na região do visível e praticamente nenhuma energia na região do infravermelho, já a água turbida (com partículas em suspensão) reflete mais energia no visível e pouco no infravermelho próximo (MOREIRA, 2011).

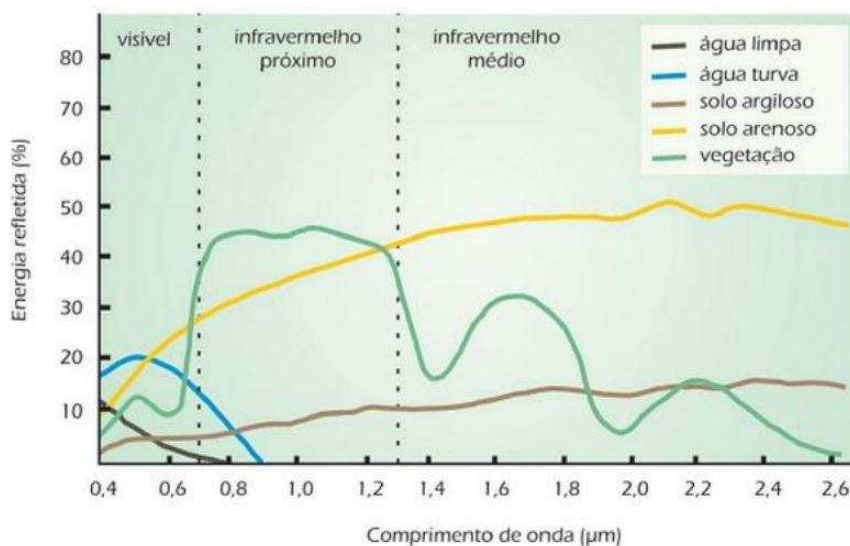


Figura 3 – Curva Espectral da vegetação, da água e do solo. Fonte: Florenzano (2002)

Para o sensoriamento remoto, a energia eletromagnética com esses comprimentos de onda, relativamente menores, estão associada a uma maior quantidade de energia e podem ser mais facilmente medida, porém vale destacar que sofrem grande interferência da atmosfera (CAMPBELL; WYNNE, 2011).

2.3 Sensores RADAR

O Sensoriamento Remoto por RADAR (*Radio Detection and Ranging* - Detecção e Telemetria por Rádio) opera em comprimentos de ondas maiores em relação aos das

bandas ópticas do espectro eletromagnético, fornecendo informações complementares de maciços florestais (FERNANDEZ-ORDÓNEZ; SORIA-RUIZ; LEBLON, 2009). Os comprimentos de onda de operação de sensores radarmétricos variam de 1mm a 1m, além disso apresentam independência das condições de iluminação solar e são muito pouco afetados por condições meteorológicas atmosféricas como nuvens, poeira ou aerossóis (FREEMAN, 1996).

A utilização de sensores RADAR por abertura sintética (*Synthetic Aperture RADAR – SAR*) para estimar determinados parâmetros biofísicos e, por conseguinte, obter um maior conhecimento fisionômico-estrutural da vegetação, tem sido intensificada nos últimos anos, devido, de um lado, à maior interação do sinal de micro-ondas com componentes do interior da copa das árvores, e de outro devido às melhorias tecnológicas oferecidas, como polarimetria, interferometria e uma melhor qualidade da radiometria nas imagens SAR (GAMA et al., 2006).

Os primeiros radares imageadores do tipo RADAR de abertura real apresentavam limitações devido as dimensões efetivas de antena da qual eram determinados pulsos de micro-ondas e resolução azimutal. Porém, com os sensores RADAR de abertura sintética (*SAR – Synthetic Aperture RADAR*) essas limitações foram contornadas através de tecnologias que simulam uma antena com cem vezes o seu tamanho real (CCRS, 2007).

Sensores SAR foram primeiramente desenvolvidos para uso militar (FERNANDEZ-ORDÓNEZ; SORIA-RUIZ; LEBLON, 2009). Atuais sensores SAR orbitais para uso civil operam na banda X (TerraSAR-X), banda C (ERS1/2, ENVISAT, RADARSAT-1/2), banda L (JERS e ALOS-PALSAR) e banda P (BIOMASS). Segundo os mesmos autores, as micro-ondas que incidem sobre superfícies naturais podem ser: (i) difusamente espalhadas e registradas no sensor como níveis de retroespalhamento, (ii) difusamente transmitidas da superfície para camadas subsuperficiais, (iii) ou serem refletidas especularmente em casos de superfícies lisas.

O espalhamento das micro-ondas sofre influências de componentes superficiais, quando ocorrem na superfície dos alvos, e volumétricos, quando ocorrem no interior dos objetos-alvos e incluem o espalhamento entre os elementos do interior. No caso do dossel florestal os elementos que interferem no espalhamento volumétrico podem ser galhos e folhas (LECKIE; RANSON, 1998). Segundo o mesmo autor, a despolarização da onda incidente e a geração de eco em polarização distinta constitui um dos resultados do espalhamento volumétrico.

O retroespalhamento das micro-ondas pelo dossel florestal compreende um fenômeno um tanto complexo. A importância de cada componente no interior do dossel no espalhamento global das micro-ondas incidentes dependem sobretudo do comprimento de onda (FERNANDEZ-ORDÓNEZ; SORIA-RUIZ; LEBLON, 2009). O tamanho e a densidade dos elementos vegetais possuem relação direta com o tipo de interação com as microndas incidentes. A Figura 4, apresenta o retroespalhamento dos componentes florestais para diferentes bandas de RADAR, na qual observa-se que a banda L configura-se como uma banda potencial para estudos volumétricos, visto a sua maior interação com o fuste das árvores, além de possuir maior penetração no dossel (DOBSON, 2000).

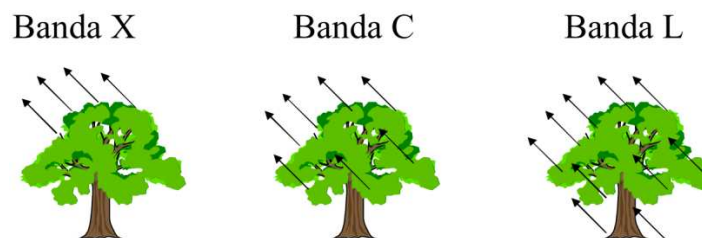


Figura 4 – Retroespalhamento pela vegetação para diferentes bandas de RADAR. Fonte: Adaptado de Leckie; Ranson (1998).

Um outro avanço nos sensores SAR foi a capacidade de mensurar a polarização das ondas eletromagnéticas. A polarização do sinal de RADAR é caracterizada basicamente pela orientação do campo elétrico da energia emitida e recebido pela antena (RICHARDS, 2009). Os sensores radarmétricos podem mensurar o continuamente a descrição completa do retorno das ondas emitidas, sendo que este sinal retroespalhado pode ser registrado em quatro combinações de polarização: onda transmitida e recebida horizontalmente (HH), onda transmitida e recebida verticalmente (VV), onda transmitida horizontalmente e recebida verticalmente (HV) e onda transmitida verticalmente e recebida horizontalmente (VH) (COOPS, 2002; RICHARDS, 2009).

2.4 Estimativas de Parâmetros Biofísicos Florestais por SR

Paralelo ao desenvolvimento tecnológico computacional e sistemas de processamento de dados, a aplicação e abordagens de dados de SR orbital em estudos na

estimação de parâmetros biofísicos florestais apresentou ao longo das últimas décadas avanços substanciais (TIMOTHY et al., 2016). Num primeiro momento com lançamento do satélite Landsat 1 em 1972, torna-se possível aplicação do registro orbital óptico em trabalhos com parâmetros biofísicos florestais (DENÈGRE; 2013; ROY et al., 2014). Anos mais tarde, o aprimoramento e disponibilização para o uso civil dos SAR orbitais, permitem estudos com aplicação de RADAR (RICHARDS, 2009). A partir dos anos 2000, começa-se a delinear os primeiros traços do atual cenário de pesquisas com avanços no imageamento orbital hiperespectral e advento da tecnologia LiDAR – Light Detection And Ranging (CHUVIECO; HUETE, 2009; LIM et al., 2003; TREITZ; HOWARTH, 1999). De meados da década de 2000-2010 ao presente momento observa-se uma quantidade crescente de trabalhos com a integração ou fusão dados de diferentes sensores e diferentes abordagens de modelagem de dados, como a inteligência artificial (ALI et al. 2015; DEL FRATE; GLEASON;IM, 2012; HYDE et al. 2006; LUCAS et al., 2006; SOLIMINI, 2004).

A seguir, é apresentada em ordem cronológica, uma compilação de trabalhos que utilizaram o SR para estimar volume de madeira e biomassa aérea com foco em povoamentos homogêneos, e explicitando os três momentos citados de utilização dos sensores ópticos e radarmétricos nas últimas décadas.

2.4.1 SR para estimativa de volume de povoamentos florestais

Trotter, Dymond e Goulding (1997) testaram a capacidade de dados ópticos do Landsat-TM em estimar o volume de madeira estocado em povoamentos florestais de espécies coníferas utilizando análise de regressão. Numa aproximação por pixels, a relação entre os dados Landsat e os valores mensurados de volume de madeira foi significativa, porém fraca, com valores de r^2 em torno de 0,3 e RQEM de $100 \text{ m}^3\text{ha}^{-1}$. Considerando a média dos pixels para os talhões, as estimativas com acurácia aceitável foram obtidas para povoamentos com no mínimo 40ha (RQEM de $46 \text{ m}^3\text{ha}^{-1}$).

Bauermann (2008), também utilizando dados ópticos e análise de regressão múltipla, estimou o volume de povoamentos de algumas espécies de eucaliptos no estado do Rio Grande de Sul. As imagens multispectriais foram adquiridas do sensor CCD a bordo do satélite CBERS-2, a partir das quais foram gerados índices de vegetação, variáveis texturais, morfológicas e cromáticas, totalizando 72 atributos. Com as variáveis independentes testadas, precisou-se de um mínimo de 40 talhões (unidade amostrais) para

resultados estáveis de estimativas de volume. Das 72 variáveis de entrada, o uso de 28 não correlacionadas foi suficiente para a modelagem a nível de talhão. As estimativas de volume apresentaram r^2 acima de 0,900. Com a análise de regressão e os dados de SR, o estudo apresentou ganhos de 45% no coeficiente de correlação e diminuição de até 77% na RQEM, em comparação ao volume estimado por regressão simples utilizando somente a idade como parâmetro.

Gebreslasie et al. (2011) utilizou dados ópticos multispectrais e pancromáticos do satélite IKONOS para gerar índices texturais e então estimar atributos estruturais de povoamentos florestais de eucalipto na África do Sul a nível de parcela. Os parâmetros estruturais de floresta consistiam em densidade de árvores, diâmetro a altura do peito (DAP), altura média total, área basal e o volume de madeira estocada. A modelagem dos dados foi realizada por meio de regressão linear múltipla e RNA. O coeficiente de determinação para as estimativas dos atributos estudados foram maiores quando utilizados índices texturais a partir de dados pancromáticos. Mais especificamente, para as estimativas de volume, utilizando a regressão linear múltipla alcançou-se um r^2 ajustado de 0,880 e utilizando RNA, obteve-se um r^2 de 0,920.

Dube et al. (2014) aplicou dois algoritmos de aprendizagem de máquina - *Stochastic Gradient Boosting* (SGB) e regressão *Random Forest* (RF) – em dados ópticos de alta resolução espacial do RapidEye para estimar volume de madeira de povoamentos de eucalipto e pinus na África do Sul. As variáveis de entrada compreenderam as bandas do azul, verde, vermelho, vermelho de transição e 13 índices de vegetação. Para cada algoritmo foi analisado a importância de cada variável de entrada das estimativas geradas. Com o algoritmo SGB utilizando as variáveis de entrada mais importantes, os autores chegaram a um r^2 de 0,80, 0,88 e 0,79 para as espécies *Eucalyptus grandis*, *E. dunii*, *Pinnus taeda*, respectivamente. Da mesma forma, com RF utilizando as variáveis consideradas mais importantes, alcançou-se um r^2 de 0,76, 0,79 e 0,80 para *E. grandis*, *E. dunii*, *P. taeda*, respectivamente.

Dube et al. (2015), utilizando dados das bandas ópticas do satélite SPOT-5, estimaram o volume de povoamentos florestais de eucalipto em Zululand na África do Sul. Além dos dados multispectrais, foram utilizadas como variáveis independentes 14 índices espectrais de vegetação, a idade do povoamento e dados de pluviosidade. O processo de modelagem comparou a regressão linear múltipla e dois algoritmos de aprendizagem de máquina, o SGB e RF. Utilizando todas as variáveis de entrada, a regressão linear múltipla, e os algoritmos SGB e RF geraram respectivamente um r^2 de

0,65, 0,78 e 0,76. Com os modelos utilizando as variáveis de maior importância, o algoritmo SGB obteve o melhor resultado com um r^2 de 0,83 e RQEM de 29,58 m^3ha^{-1} , correspondente a 17,63 % da média das parcelas. De modo geral, a modelagem por regressão linear múltipla gerou estimativas com menor acurácia em comparação aos dois algoritmos testados. Os autores do trabalho ainda destacam a importância da integração de dados remotamente sensoriados e dados auxiliares como a idade do povoamento e pluviosidade, para o aumento da acurácia dos modelos preditivos de parâmetros biofísicos dos povoamentos florestais.

Gama et al. (2006), utilizando regressão multivariada e dados radarmétricos provenientes de um SAR aerotransportado (OrbiSAR-1) operando nas bandas P e X, estimaram volume de povoamentos florestais de eucalipto na região sudeste do Brasil. Os modelos para estimativas de volume apresentaram forte correlação com dados de interferometria empregando as duas bandas, com erro aproximadamente de 10 %. Os autores também verificaram que os modelos desenvolvidos com estes dados obtiveram coeficientes de determinação de 84 a 88% maior em relação as estimativas do inventário florestal das áreas.

Utilizando modelos *Water Cloud* (WC) (ATTEMA; ULABY, 1978) e dados do sensor orbital PALSAR do satélite ALOS, Amrutkar et al. (2012) geraram índices de vegetação por radar para estimar parâmetros biofísicos de povoamentos florestais mistos, que incluíam comerciais (*Eucalyptus* spp.) e nativos da Floresta Gujarat, na Índia. Os autores utilizaram métodos não destrutivos para quantificar o volume das parcelas, e obtiveram um coeficiente de determinação de 0,573.

Santoro, Eriksson e Fransson (2015) também utilizaram ALOS-PALSAR para estimar volume para povoamentos de florestas naturais boreal e hemi-boreal na Suécia. A pesquisa detectou efeito das condições ambientais, polarização e ângulo de incidência na sensibilidade do sensor ao retroespalhamento do sinal. O volume de madeira foi mais bem estimado usando dados polarizados HV com ambientes sem a presença de gelo e ângulo de incidência de 34,3°, apresentando um RQEM relativo de 44%.

Integrando dados orbitais ópticos e radarmétricos, Ismail et al. (2015), através de regressão linear múltipla, estimaram volumes de madeira de povoamentos florestais comercial de eucalipto. Os autores utilizaram dados do SPOT-4 e do sensor radarmétrico ALOS-PALSAR para estimar a biomassa aérea e volume de povoamentos de eucalipto implantados e sob regime de talhadia. A integração dos dados retornou uma estimativa com r^2 de 0,68 para plantios implantados e 0,55 para talhadia. Entretanto, com a adição

da idade do povoamento como variável independente houve um ganho de acurácia de 15% para plantios implantados e sob 27% por regime de talhadia. Para os respectivos plantios, os modelos finais retornaram a um r^2 de 0,83 e 0,82, e um RQEM de $31,71\text{m}^3\text{ha}^{-1}$ e $27,70\text{m}^3\text{ha}^{-1}$.

2.4.2 SR para estimativa de biomassa aérea de povoamentos florestais

Watzlawick et al. (2006) aplicando regressão múltipla a dados ópticos do satélite IKONOS II, estimaram a biomassa aérea e carbono orgânico de plantações florestais de *Pinus taeda* L. no estado do Paraná. A biomassa aérea foi quantificada por métodos destrutivos que por análise estatística de matriz de correlação foi a variável dependente e avaliada em função de variáveis provenientes das imagens orbitais tais como: valores médios de reflectância das bandas MS-1 (azul), MS-2(verde), MS-3(vermelho) e MS-4 (infravermelho próximo), Razão de bandas (MS-4/MS-3), Índice de Vegetação por Diferença Normalizada (NDVI) e Índice de Vegetação Ajustado ao Solo (SAVI). As variáveis que apresentaram melhores correlações a biomassa aérea foram o MS-4, RAZÃO, NDVI e SAVI, com valores de correlação entre 0,70 a 0,80. Segundo os autores os resultados da aplicação de dados multiespectrais bem como o uso dos índices gerados, mostraram-se satisfatórios e recomendam a abordagem para as estimativas quantitativas de biomassa aérea e carbono de povoamentos de *Pinus taeda*.

Utilizando dados radarmétricos das bandas P, L e C, Beaudoin et al. (1994) verificaram fortes correlações do retroespalhamento (polarizações HH, HV, VV) com a biomassa do tronco de povoamentos de *Pinus spp*. Segundo os autores a banda L se mostrou muito eficiente para estimar a biomassa do tronco, porém a banda P apresentou maior correlação. Da mesma forma a polarização HV apresentou correlações mais fortes com parâmetros florestais avaliados (está errado – esse artigo é um *survey* e dá um belo parágrafo explicativo, ver se é apropriado para o andar do texto).

Gama, Santos e Mura (2010), avaliaram a relação da biomassa de povoamentos de eucalipto e dados radarmétricos utilizando regressão múltipla e sensor aerotransportado OrbiSAR-1 operando nas bandas X e P na região Sul do Brasil. As predições de biomassa apresentaram erro inferior a 10 % e as melhores variáveis para compor o modelo foram: a altura interferométrica, que consiste na diferença do modelo digital de altura interferométrica gerado pelas bandas de trabalho (X e P); e índice de

espalhamento do dossel (*Canopy Scattering Index* – CSI) que é a relação do L_{VV} e a soma de L_{VV} e L_{HH} .

Para povoamentos de Palma na Malásia Oriental, Morel et al. (2011) utilizaram dados polarimétricos do sensor ALOS-PALSAR para estimar biomassa aérea. Os autores verificaram que a estimativa direta de biomassa aérea através de modelos de regressão simples foi sensível às condições ambientais e estrutura da floresta. A aplicação da análise de regressão para estimativa de biomassa aérea por dados polarizados (L_{HV}) retornou um r^2 de 0,630.

Amrutkar et al. (2012) geraram índices de vegetação por radar através do ALOS-PALSAR com o qual estimou parâmetros biofísicos de povoamentos florestais, que incluíam comerciais (*Eucalyptus spp.*) e nativos da Floresta Gujarat, na Índia. Os autores utilizaram métodos não destrutivos para quantificar a biomassa aérea florestal, e com os modelos WC obtiveram um r^2 de 0,617.

Baghdadi et al. (2015) analisaram a performance do algoritmo para regressão RF com dados da banda L do sensor ALOS-PALSAR para predições de biomassa aérea de diferentes povoamentos de eucalipto no Brasil. A idade e o sinal polarizado HV foram as variáveis que melhor explicaram a quantidade de biomassa aérea dos povoamentos. Somente o uso de dados radarmétricos não foram eficientes para prever corretamente a biomassa aérea de povoamentos de eucalipto, apresentando um coeficiente de determinação inferior a 0,5 ($r^2 < 0,500$) e RQEM maior do que 46,7 t/ha. Com a adição da variável idade este índice aumentou para 0,880 e o RQEM caiu de 22,7 t/ha para 18,9t/ha.

Behera et al. (2016) utilizaram modelos WC e dados polarimétricos do sensor ALOS-PALSAR para estimativa de biomassa aérea de povoamentos florestais de *Tectona grandis* e *Shorea robusta*. Foram obtidas correlações significativas entre todos os coeficientes de retroespalhamento e os valores reais de biomassa aérea para as duas espécies, com um coeficiente de determinação máximo para polarização cruzada (HV) para *S. robusta*. Porém para sinais co-polarizados HH foi observado fraca correlação para florestas de *T. grandis*.

A fim de estimar biomassa aérea de plantios homogêneos de caju (*Anacardium occidentale*) e seringueira (*Hevea brasiliensis*) no Cambódia, Avtar, Suzuki e Sawada (2014) utilizaram regressão linear múltipla e integraram dados ópticos e radarmétricos, respectivamente, dos sensores PALSAR e AVNIR-2, a bordo do satélite ALOS. Os retroespalhamentos de todas as polarizações do PALSAR apresentaram respostas

diferentes para os dois tipos de plantações pelas diferenças dos parâmetros biofísicos. Ainda assim, os autores conseguiram mostrar que o retroespalhamento dos dados radarmétricos apresentou forte correlação com dados mensurados em campo.

2.5 Satélite ALOS

O satélite ALOS (*Advanced Land Observing Satellite*), Figura 5, foi um dos principais satélites já feito pelo Japão, ficando em órbita durante 5 anos. Foi lançado em 2006, pela Agência Espacial Japonesa (Japan Aerospace Exploration Agency - JAXA), como parte do programa japonês de observação terrestre, visando o monitoramento de desastres ambientais, e o levantamento de recursos naturais (ROSENQVIST et al., 2007).

O satélite possui três sensores: (i) o *Panchromatic Remote-Sensing Instrument for Stereo Mapping* (PRISM), que é o sensor com uma banda pancromática (ii) o *Advanced Visible and Near Infrared Radiometer type 2* (AVNIR-2), como sensor multiespectral óptico e (iii) *Phased Array type L-band Synthetic Aperture Radar* (PALSAR) como sensor ativo radarmétrico SAR (ROSENQVIST et al., 2007).

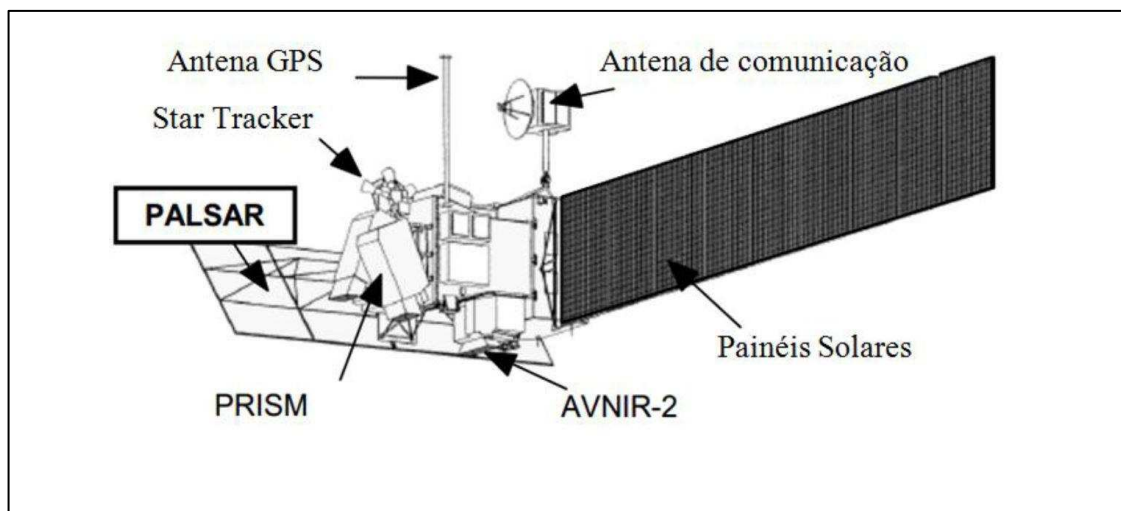


Figura 5 – Satélite ALOS. (JAXA, 2016)

O sensor AVNIR-2 fornece imagens multispectrais, em 4 bandas abrangendo a região espectral do visível e infra-vermelho próximo (Quadro 1). É o sucessor do AVNIR que estava a bordo do Advanced Earth Observing Satellite (ADEOS), lançado em agosto de 1996 (ROSENQVIST et al., 2007).

Quadro 1 – Características do sensor AVNIR-2. (JAXA, 2016)

Número de bandas	4
Comprimento de onda (μm)	Banda 1: 0,42 ~ 0,50
	Banda 2: 0,52 ~ 0,60
	Banda 3: 0,61 ~ 0,69
	Banda 4: 0,76 ~ 0,89
Resolução espacial	10 m (nadir)
Largura da faixa	70 Km (nadir)
Ângulo de apontamento	$\pm 44^\circ$
Resolução radiométrica	8 bits

O PALSAR é um sensor radarmétrico ativo, para observações diurnas e noturnas da Terra sem a interferência da nebulosidade, que opera na banda L (Quadro 2). Possui um modo polarimétrico que é capaz de imagear em duas e quatro polarizações, respectivamente, HH e HV, e HH, HV, VV e VH (ROSENQVIST et al., 2007).

Quadro 2 – Características do sensor PALSAR (JAXA, 2016).

Modo	Fino		ScanSAR	Polarimétrico
Frequência de centro	1270 MHz (Banda - L)			
Largura de banda	28 MHz	14 MHz	14,28 MHz	14 MHz
Polarização	HH ou VV	HH+HV ou VV+VH	HH ou VV	HH+HV+VH+V V
Ângulo de Incidência	8 ~ 60°	8 ~ 60°	18 ~ 43°	8 ~ 30°
Realce em alcance (range)	7 ~ 44 m	14 ~ 88 m	100 m (multi look)	24 ~ 89 m
Largura da faixa	40 ~ 70 Km	40 ~ 70 Km	250 ~ 350 Km	20 ~ 65 Km
Taxa de Transmissão de	240 Mbps	240 Mbps	120 Mbps, 240 Mbps	240 Mbps

2.6 Máquina de Vetores de Suporte

As Máquinas de Vetores de Suporte (*Support Vector Machines*) compreendem um algoritmo de aprendizagem de máquina supervisionado, isto é, um algoritmo com capacidade de explorar um conjunto de dados de treinamento a fim de descobrir relações entre os atributos (variáveis independentes) e uma variável alvo (dependente). O

aprendizado supervisionado (*supervised learning*) corresponde a formalização da premissa básica de aprendizado por exemplos (HUANG; LEARNED-MILLER, 2014).

Alicerçado sobre os princípios da teoria de aprendizagem estatística desenvolvida por Vapnik e Chervonenkis (1974), após um treinamento de um subconjunto de dados, o algoritmo possui propriedades que permite uma boa generalização de dados não antes explorados (SMOLA, A. J.; SCHÖLKOPF, 2004).

Para compreender o funcionamento do algoritmo, faz-se necessário detalhar alguns conceitos como:

- hiperplano de separação;
- vetores de suporte;
- margem máxima;
- hiperplano de separação ótima;
- margem suave e
- Funções *kernel*.

Será utilizado para conceituação e exemplificação um caso de classificação binária, e posterior descrição das MVS para regressão.

2.6.1 Máquina de Vetor de Suporte para Classificação

O hiperplano de separação (*decision boundary*) constitui uma função n -dimensional capaz de separar os dados de treinamento nas classes desejáveis (Figura 6). A função de classificação pode corresponder a um ponto, no caso de uma classificação unidimensional (Figura 6a), ou mesmo uma reta (Figura 6b) ou um plano (Figura 6c) no caso das dimensões R^2 e R^3 respectivamente (NOBLE, 2006). É chamado hiperplano, pois pode-se extrapolar a classificação para um espaço dimensional muito maior e encontrar um limite de classificação entre os casos.

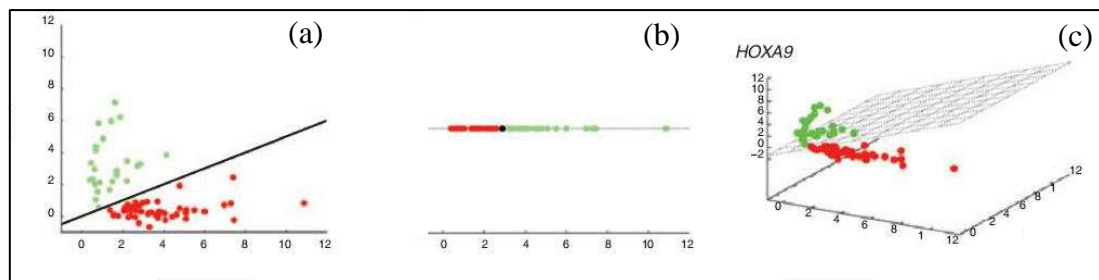


Figura 6 – Hiperplano de Separação para as Dimensões R^2 , R^1 e R^3 . Fonte: Noble (2006).

Para a classificação do conjunto de dados bidimensionais em suas classes, suponha L pontos de treinamento, onde cada entrada x_i possua D atributos (isto é, com dimensão D) e possa ser classificado como $y_i = -1$ ou $y_i = +1$, assim o conjunto de dados de treinamento pode ser descrito:

$$\{x_i, y_i\} \text{ onde } i = 1, \dots, L, y_i \in \{-1, 1\}, x \in \mathcal{R}^D \quad (1)$$

Assumindo que os dados sejam linearmente separáveis, isso significa dizer que uma linha pode ser traçada em um plano com eixos x_1 e x_2 separando duas classes quando $D=2$, representando o hiperplano de separação.

O hiperplano pode ser descrito por uma função $f(x) = w_0 + w_1x_1 + \dots + w_nx_n = 0$. Fletcher (2009) demonstra equação por $w \cdot x + b = 0$, onde w_0 é uma constante descrita como b , por convenção (VAPNIK, 1995). A equação também pode ser descrita matricialmente por $w^T x + b = 0$.

Dado um hiperplano de separação de classificação binária (Figura 7) e considerando os conceitos de álgebra linear:

- w é um vetor de parâmetros normal ao hiperplano
- $\frac{b}{\|w\|}$ é a distância perpendicular do hiperplano a origem

Os *vetores de suporte* (*support vectors*) são os pontos de treinamentos mais próximos ao hiperplano e o objetivo das MVS é orientar o hiperplano de modo que ele fique o mais distante possível dos vetores de suporte de ambas as classes (FLETCHER, 2009).

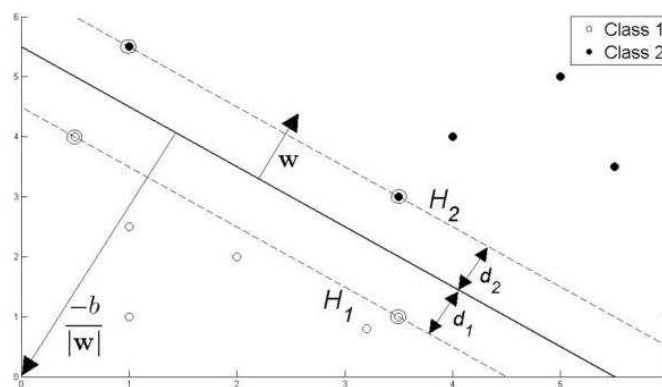


Figura 7 – Hiperplano de separação linear para um caso de classificação binária. Fonte: Fletcher (2009).

Considerando as inequações para o caso desta classificação, onde +1 representa a classe 1 e -1 a classe 2, tem-se:

$$x_i \cdot w + b \geq +1, \text{ para } y_i = +1 \quad (2)$$

$$x_i \cdot w + b \leq -1, \text{ para } y_i = -1 \quad (3)$$

Quando combinadas, retornam a seguinte inequação:

$$y_i(x_i \cdot w + b) - 1 \leq 0 \quad \forall_i \quad (4)$$

Se considerarmos somente os vetores de suporte na Figura 7, então os dois planos que estão sobre os pontos citados, H_1 e H_2 podem ser descritos por:

$$x_i \cdot w + b \geq 1, \text{ para } H_1 \quad (5)$$

$$x_i \cdot w + b \leq -1, \text{ para } H_2 \quad (6)$$

Ainda na Figura 7 pode-se definir d_1 e d_2 como as distâncias de H_1 e H_2 , respectivamente, ao hiperplano de separação. Estas equidistâncias somadas formam o comprimento da margem da MVS.

O algoritmo trabalha de forma a maximizar a distância entre o hiperplano e os pontos mais próximos a ele, através da maximização do cálculo da norma dos vetores, minimizando assim o erro dos dados de treinamento e ainda dados desconhecidos.

Em um outro exemplo de classificação binária, mostrado na Figura 8, muitas funções ou hipóteses de hiperplanos podem ser usadas para separar e classificar os dados. Porém, os vetores de suporte auxiliam no cálculo em ambos os lados do classificador, maximizando a distância do hiperplano de todos os vetores de suporte. Essa distância entre o hiperplano e os pontos, chamada de *margem máxima* (*maximum margin*), define a região de segurança da classificação de dados desconhecidos. O hiperplano que possui essa margem máxima, calculada com a maximização das norma dos vetores, é chamado de *hiperplano ótimo* (*optimal decision boundary*) (NOBLE, 2006).

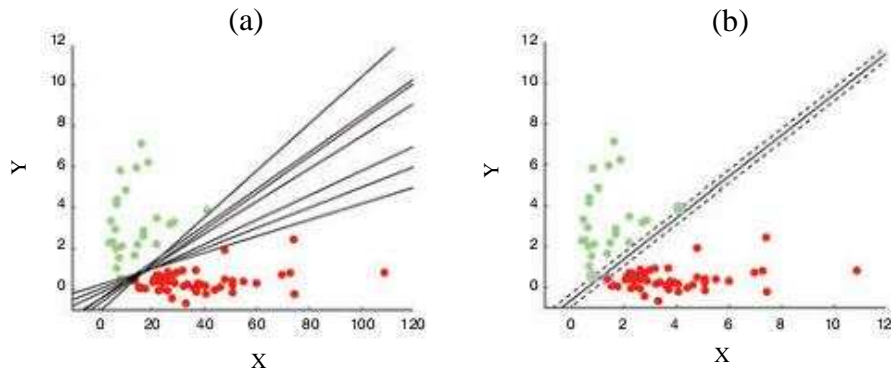


Figura 8 – A Figura 8a representa os possíveis hiperplanos de separação em um caso de classificação binária. A Figura 8b mostra o hiperplano de separação com margem máxima, calculada a partir dos vetores de suporte destacados. Fonte: Noble (2006).

De volta ao exemplo da Figura 7, para orientar o hiperplano a ser o mais distante possível dos vetores suporte é necessário maximizar a margem (d_1+d_2). A geometria de vetores simples mostra que a margem corresponde a $\frac{1}{\|w\|}$, e maximizando o termo sujeito a restrição (3), tem-se:

$$\text{Minimize } \|w\|, \quad \text{tal que } y_i(x_i \cdot w + b) - 1 \leq 0 \quad \forall_i \quad (7)$$

Minimizar $\|w\|$ é equivalente a minimizar o termo $\frac{1}{2} \|w\|^2$, o que possibilita tornar a otimização solucionável por Programação Quadrática (*Quadratic Programming – QP*). Logo, tem-se:

$$\text{Minimize } \frac{1}{2} \|w\|^2, \quad \text{tal que } y_i(x_i \cdot w + b) - 1 \leq 0 \quad \forall_i \quad (8)$$

ou então sob a forma matricial

$$\text{Minimize } \frac{1}{2} w^T w, \quad \text{tal que } y(w^T x + b) - 1 \leq 0 \quad \forall_i \quad (9)$$

Em muitos casos a classificação não pode ser facilmente separada por uma linha reta como mostrado nos exemplos até então, pela natureza do fenômeno estudado ou mesmo por apresentarem ruídos, *outliers*. Assim, em conjunto de dados pode-se encontrar pontos fora da região de classificação de uma determinada classe. Para lidar com esses pontos, o algoritmo MSV é modificado, adicionando-se uma *margem suave* (*soft margin*).

Esta modificação constitui a inclusão de um parâmetro que suaviza as margens de separação do classificador permitindo que alguns dados de aprendizagem permaneçam entre elas e também alguns erros de classificação (LORENA; CARVALHO, 2007). Em outras palavras, este parâmetro irá controlar a quantidade de exemplos que poderão violar a função classificadora bem como a distância destes até o hiperplano de separação (NOBLE, 2006). Essencialmente, isto permite com que *outliers* sejam ignorados e possam estar em regiões de classificação diferentes sem alterar o resultado final.

Assumindo que os mesmos dados do exemplo da Figura 7 não possam ser separáveis por um hiperplano linear (Figura 9), uma variável folga (*slack variables*) pode ser adicionada para as restrições (8) (9), permitindo alguns erros classificatórios. Sendo esta, descrita por ξ_i , para $i = 1, \dots, L$.

$$x_i \cdot w + b \geq +1 - \xi_i, \text{ para } y_i = +1 \quad (10)$$

$$x_i \cdot w + b \leq -1 + \xi_i, \text{ para } y_i = -1 \quad (11)$$

$$\xi_i \geq 0 \forall_i \quad (12)$$

Quando combinadas, retornam a seguinte desigualdade:

$$y_i(x_i \cdot w + b) - 1 + \xi_i \leq 0, \text{ onde } \xi_i \geq 0 \forall_i \quad (13)$$

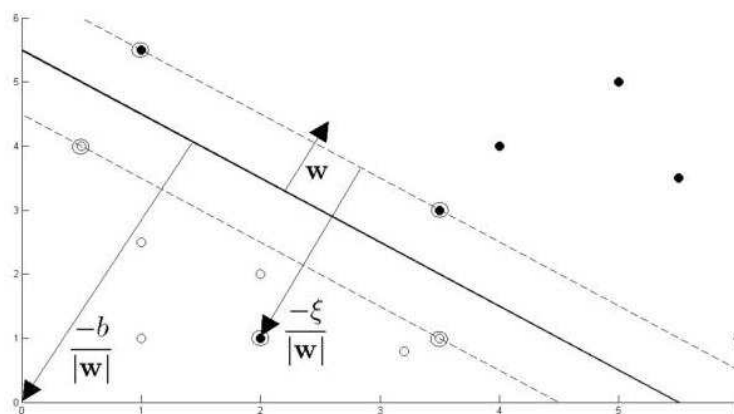


Figura 9 – Classificação binária para um caso em que os dados não podem ser separados totalmente por um hiperplano linear. Fonte: Fletcher (2009).

Nesta margem suave (*soft margin*), pontos incorretamente classificados, isto é, que estão do lado incorreto do hiperplano, são penalizados. Estas penalidades são maiores conforme o aumento de sua distância com o hiperplano. Com o objetivo de diminuir classificações erradas, uma maneira sensata de adaptar a função objetivo parte da idéia de que alguns erros são esperados, porém a margem de aceite desses erros pode ser minimizada. Sendo esta margem de aceite definida por ξ_i , tem-se:

$$\text{Minimize } \frac{1}{2} \|w\|^2 + C \sum_{i=1}^L \xi_i \quad (14)$$

$$\text{Sujeita a } y_i(x_i \cdot w + b) - 1 + \xi_i \leq 0 \quad \forall_i \quad (15)$$

Ou então sob a forma matricial:

$$\text{Minimize } \frac{1}{2} w^T w + C \sum_{i=1}^L \xi_i \quad (16)$$

$$\text{Sujeita a } y_i(w^T x + b) - 1 + \xi_i \leq 0 \quad \forall_i \quad (17)$$

em que C é o parâmetro de regulação do algoritmo de aprendizagem de máquina. A constante $C > 0$ funciona como termo compensatório, contrabalanceando a complexidade da estrutura das margens e os valores de desvios ξ , evitando o *overfitting*. O *overfitting* é o termo designado para o ajuste da MVS demasiadamente em acordo com dados de treinamento, impossibilitando a MVS treinada de obter bons desempenhos para novos dados, isto é, impossibilitando a sua habilidade de generalização (FLETCHER, 2009).

Em casos de classificações não-lineares, a utilização de um hiperplano no corrente espaço dos dados de entrada (*Input Space*) pode se tornar complexa, exigindo que o separador também seja não-linear. Para solucionar este problema, o algoritmo MVS realiza um mapeamento dos dados em um novo espaço de maior dimensão, denominado espaço de atributos (*Feature Space*), tornando os linearmente separáveis. As funções capazes de mapear os dados para o espaço de características são denominadas de funções *Kernels* (SMOLA, A. J.; SCHÖLKOPF, 2004).

As funções *kernels* constituem “funções critérios” para ditar o valor de possibilidade de cada dado de entrada pertencer a determinada classificação, levando em conta sua localização no plano e distância dos demais dados de mesmo valor categórico. Ou seja, baseiam-se no produto interno dos dados de entrada e os reformula em um espaço dimensional maior através de uma função não-linear capaz de mapear $x \rightarrow \Phi(x)$ (Figura 10).

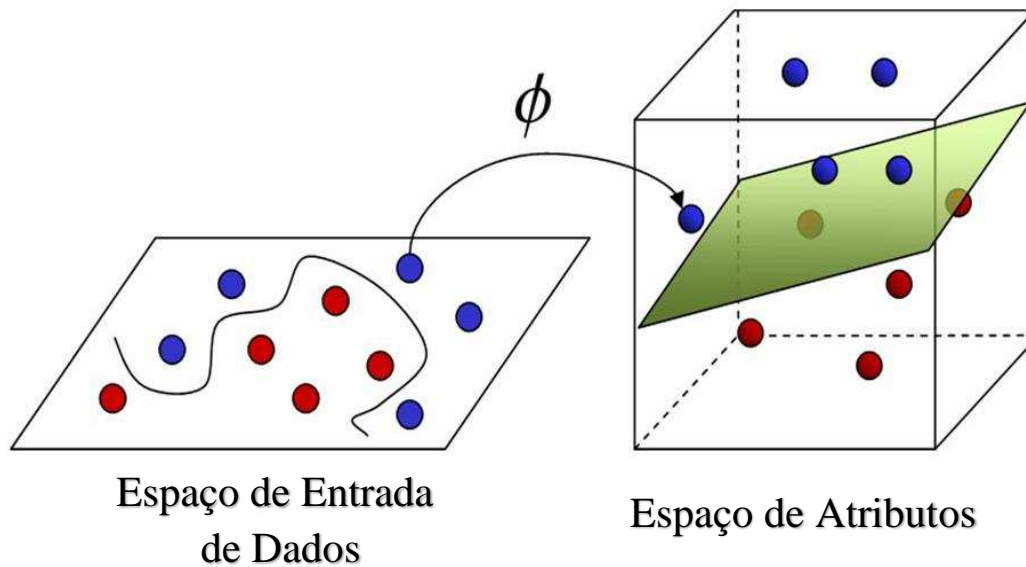


Figura 10 – Representação do funcionamento de uma funções *kernel* das Máquinas de Vetores de Suporte para classificação. Transformação de um problema não separável linearmente para o espaço de atributos (*feature space*) onde pode ser separado por um hiperplano linear. (Fonte: <http://www.imtech.res.in/raghava/rbpred/svm.jpg>).

Descreve-se a função *kernel* por $K(x_i, x_j) = \Phi(x_i) \cdot \Phi(x_j)$, onde k , a função *kernel*, representa o produto interno dos dados de entrada mapeados em um espaço de características maior, transformados por Φ (FLETCHER, 2009; SCHÖLKOPF; SMOLA, 2002; SMOLA; SCHÖLKOPF, 2004). O quadro 3 apresenta as principais funções *kernels* usadas pelo algoritmo MVS.

Quadro 3 – Principais *Kernels* utilizados nas MVS. Fonte: Adaptado de Lorena (2006).

<i>Tipo de Kernel</i>	<i>Função $K(x_i, x_j)$ correspondente</i>	<i>Comentários</i>
Linear	$x_i^T x_j$	-
Polinomial	$(x_i^T x_j + k)^p$	p e k deve ser especificada pelo usuário
Gaussiano (RBF)	$exp\left(-\frac{\ x_i - x_j\ ^2}{2\sigma^2}\right)$	A amplitude σ^2 é especificada pelo usuário
Sigmoide	$tanh(\gamma \cdot x_i^T x_j + k)$	Utilizado somente para alguns valores de γ e k

Dessa forma, para MVS considerando funções *kernels*, tem-se

$$\text{Minimize } \frac{1}{2} \|w\|^2 + C \sum_{i=1}^L \xi_i \quad (18)$$

$$\text{Sujeita a } y_i(\Phi(x_i) \cdot w + b) - 1 + \xi_i \leq 0 \quad \forall_i \quad (19)$$

ou então sob a forma matricial

$$\text{Minimize } \frac{1}{2} \mathbf{w}^T \mathbf{w} + C \sum_{i=1}^L \xi_i \quad (20)$$

$$\text{Sujeita a } y_i(\mathbf{w}^T \Phi(\mathbf{x}_i) + b) - 1 + \xi_i \leq 0 \quad \forall_i \quad (21)$$

Onde $K(x_i, x_j) = \Phi(x_i) \cdot \Phi(x_j)$. Isto é, a função *kernel* K , representa o produto interno dos dados de entrada mapeados em um espaço de características maior, transformados por Φ .

A função RBF (*Radial Basis Function*) é a opção mais amplamente utilizada para as MVS, devido as respostas finitas e localizadas por toda a variação do eixo x real. A Figura 12, mostra um exemplo de funcionamento da função *kernel* RBF.

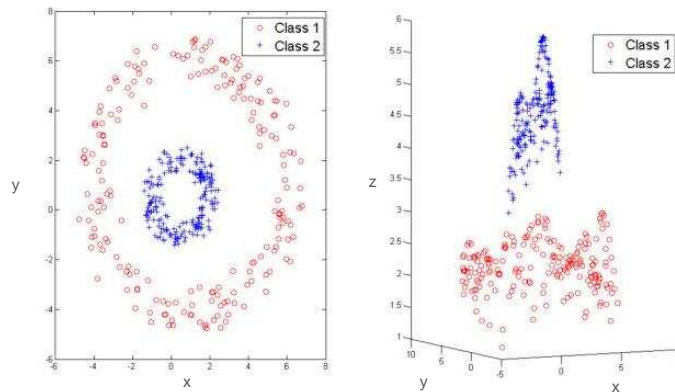


Figura 12 – Dados de um problema de classificação binária não-linear re-mapeados em um espaço de características tridimensional usando RBF. Fonte: Fletcher (2009).

Segundo Li et al. (2012), a MVS comparada a muitos métodos multivariados possui excelentes propriedades de generalização de problemas podendo ser estendida à aplicação de casos não-lineares considerando a utilização da função *kernel* apropriada.

2.6.2 Máquina de Vetor de Suporte para Regressão

2.6.2.1 Tipo I (ϵ -intensive)

As SVMs foram adaptadas para regressão (SCHÖLKOPF et al., 2000) retornando valores contínuos de saída. A primeira reformulação partiu de Vapnik (1995), com Regressão Vetor de Suporte ϵ -intensiva, em inglês *ϵ -intensive Support Vector Regression* ou somente ϵ -SVR.

Para entender a ϵ -SVR, suponha um conjunto de dados de treinamento $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\} \subset X \times \mathbb{R}$, onde X denota o espaço de padrões dos dados de entrada. Analogamente a margem máxima nas MVS classificatórias, o objetivo da ϵ -SVR é encontrar uma função $f(x)$ com uma margem de erro mínimo aceitável ϵ e esta margem ao longo de toda função deve possuir o mínimo de oscilações possíveis. O parâmetro ϵ é calculado e traçado com base nos pontos mais próximos ao hiperplano que agora é a própria função linear no espaço de características de dimensão N (FLETCHER, 2009).

$$f(x) = \langle w, x_i \rangle + b, \text{ com } w \in X, b \in \mathbb{R} \quad (22)$$

Onde $\langle w, x_i \rangle$ denota o produto interno em X , e x_i é parametrizado por w .

A “planura” (*flatness*) máxima da margem é alcançada ajustando w para que seja o menor possível. Uma maneira de alcançar altos valores para w , é minimizando sua norma, isto é, $\|w\|^2$. Assim pode-se otimizar um problema de solução convexa:

$$\text{Minimize } \frac{1}{2} \|w\|^2 \quad (23)$$

$$\text{Sujeito à } \begin{cases} y_i - \langle w, x_i \rangle - b \leq \epsilon \\ \langle w, x_i \rangle + b - y_i \leq \epsilon \end{cases} \quad (24)$$

A restrição (24) diz que a função f aproxima os pares (x_i, y_i) com uma precisão constante ϵ e ao mesmo tempo torna o problema de solução convexa possível. Porém, muitas vezes a restrição pode inviabilizar a solução do problema se não considerar parte dos erros.

Ainda que um erro mínimo constante seja estabelecido, muitas vezes os dados de treinamentos extrapolam a margem de precisão. Estes dados podem ser incorporados nos cálculos, analogamente ao conceito da função custo de “**margem suave**”. Assim, inclui-se variáveis folgas (*slack variables*) ξ_i, ξ_i^* na otimização tornando o problema solucionável (FLETCHER, 2009; SCHÖLKOPF; SMOLA, 2002).

$$\text{Minimize } \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) \quad (25)$$

$$\text{Sujeito à } \begin{cases} y_i - \langle w, x_i \rangle - b \leq \varepsilon + \xi_i \\ \langle w, x_i \rangle + b - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0, \quad i = 1, \dots, n \end{cases} \quad (26)$$

Também podendo ser escrito da seguinte forma:

$$\text{Minimize } \frac{1}{2} w^T w + C \sum_{i=1}^n (\xi_i + \xi_i^*) \quad (27)$$

$$\text{Sujeito à } \begin{cases} y_i - w^T x_i - b \leq \varepsilon + \xi_i \\ w^T x_i + b - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0, \quad i = 1, \dots, n \end{cases} \quad (28)$$

A constante $C > 0$ funciona como termo compensatório, contrabalanceando o nivelamento das margens e valores de desvios acima do ε que o tolerado, evitando o *overfitting*. Isso corresponde em lidar com a função custo ε -intensiva descrita como:

$$|\xi|_\varepsilon := \begin{cases} 0, & \text{Se } |\xi| \leq \varepsilon \\ |\xi| - \varepsilon & \text{do contrário} \end{cases} \quad (29)$$

A Figura 13 descreve o ajuste da margem suave no caso de um MVS linear. Apenas dados fora da região sombreada contribuem para o custo citado, sendo os desvios penalizados linearmente.

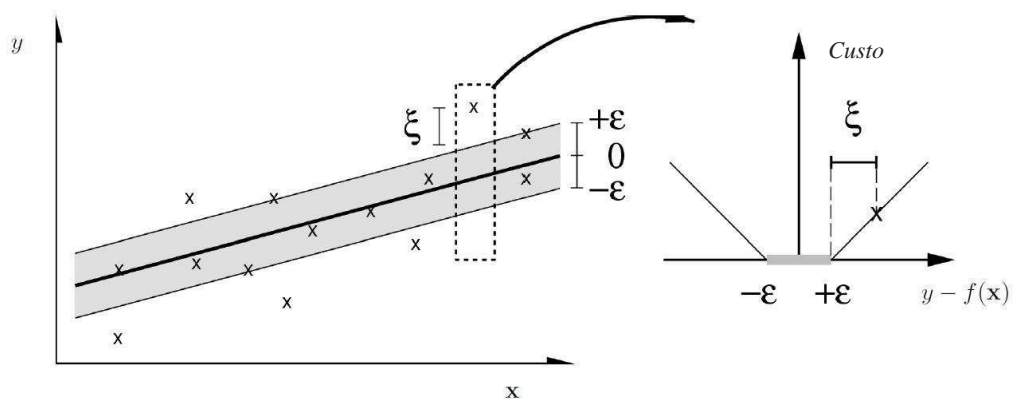


Figura 13 - Margem suave (*soft margin*) ajustada para uma MVS linear, representada pela área exterior a margem máxima (*maximum margin*). Fonte: Schölkopf; Smola (2002).

Para funções não lineares, ϵ -SVRs fazem o uso de funções *kernels* para mapear os dados em um espaço dimensional para linearização ou simplificação da complexidade dos dados. Dessa forma, dados de entradas x_i são processados pelo mapeamento $\Phi X \rightarrow F$ para um espaço de atributos F , como descritos por Aizerman; Braverman; Rozner (1964) e Nilsson (1965). Assim, tem-se ϵ -SVR para casos não lineares:

$$\text{Minimize } \frac{1}{2} \mathbf{w}^T \mathbf{w} + C \sum_{i=1}^n (\xi_i + \xi_i^*) \quad (30)$$

$$\text{Sujeito à } \begin{cases} y_i - \mathbf{w}^T \Phi(x_i) - b \leq \epsilon + \xi_i \\ \mathbf{w}^T \Phi(x_i) + b - y_i \leq \epsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0, i = 1, \dots, n \end{cases} \quad (31)$$

em que $\Phi(x_i)$ representa a i -ésima entrada com o valor transformado pela função *kernel*. $K(x_i, x_j) = \Phi(x_i) \cdot \Phi(x_j)$, isto é, a função *kernel*, representa um produto interno dos dados de entrada mapeados no maior espaço dimensional de atributos pela transformação do parâmetro Φ .

2.6.2.2 Tipo II (ν -SVM)

A ϵ -SVR contém uma dificuldade em que o próprio usuário deve selecionar ou determinar uma precisão ϵ . Para contornar este problema, Schölkopf et al. (1999) introduziram um novo parâmetro ν com o qual o usuário passa a controlar o número de vetores e o erro de treino. Assim, a precisão ϵ passa a ser ajustada automaticamente pelo sistema. Reformulada, passou a ser denominada também por Máquina de Vetor de Suporte- ν (ν -MVS), em inglês *ν -support vector machine* (ν -SVM) ou mesmo *ν -support vector regression* (ν -SVR).

As ν -SVMs representam uma nova classe de SVM e podem lidar com problemas de regressão e classificação. ν -SVMs para regressão é definida por: dado um conjunto de dados $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$, em que $x_i \in R^n$ correspondem aos dados de entrada (*inputs*) e $y_i \in R^l$ variável alvo de saída (*output*), tem-se

$$\text{Minimize } \frac{1}{2} \mathbf{w}^T \mathbf{w} + C \left(v\epsilon + \frac{1}{l} \sum_{i=1}^l (\xi_i + \xi_i^*) \right) \quad (32)$$

$$\text{Sujeito à } \begin{cases} y_i - \mathbf{w}^T \Phi(x_i) - b \leq \epsilon + \xi_i \\ \mathbf{w}^T \Phi(x_i) + b - y_i \leq \epsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0, i = 1, \dots, n, \epsilon \geq 0 \end{cases} \quad (33)$$

em que, $0 \leq v \leq 1$, C é o parâmetro de regulação e os vetores de treinamento x_i são mapeados em um espaço dimensional maior pelas funções *kernels* Φ .

3 MATERIAL E MÉTODOS

O presente estudo foi uma continuidade dos trabalhos realizados por Oliveira (2011) e Domingues (2015). Assim, o pré-processamento dos dados é descrito por Oliveira (2011) e neste trabalho são apresentadas as informações pertinentes para compreensão da aplicação das MVS.

3.1 Áreas de Estudo

Este estudo foi conduzido em áreas de reflorestamento de eucaliptos localizadas na região do Vale do Rio Doce, nos municípios de Belo Oriente, Santana do Paraíso, Ipaba e Caratinga, a leste do Estado de Minas Gerais (MG). Estas áreas estão localizadas entre os paralelos $19^{\circ}14'29,74''\text{S}$ e $19^{\circ}34'45,24''\text{S}$ e os meridianos $42^{\circ}32'48,06''\text{W}$ e $42^{\circ}20'52,81''\text{W}$ (Figura 14).

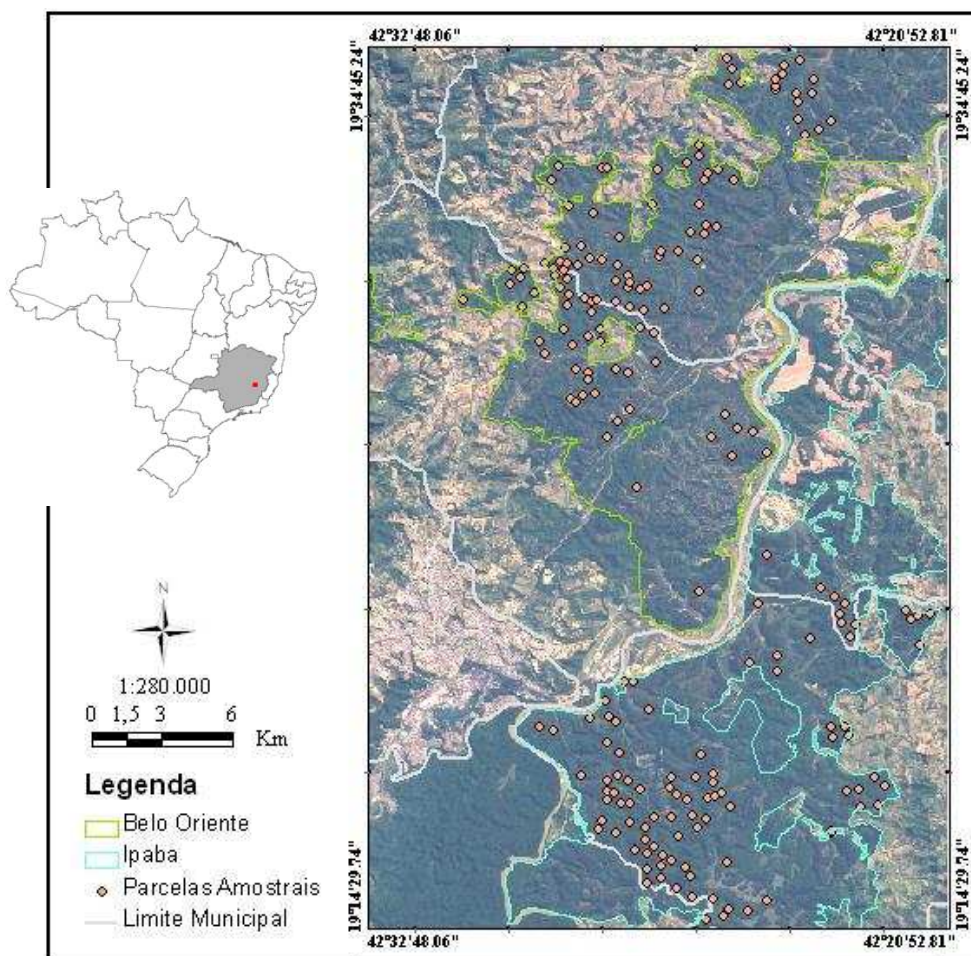


Figura 14 – Localização da área de estudo nos municípios de Belo Oriente, Santana do Paraíso, Ipaba e Caratinga-MG.

A área de estudo é composta por plantações de *Eucalyptus grandis*, com espaçamento de 3 x 2 m, com idades de 4 a 8 anos e altura média de 25,9 m, cujo plantio recobre uma área total de 837,12 km². As áreas estão localizadas em um relevo variando de plano a levemente ondulado.

O clima da região, segundo a classificação de Köppen, é do tipo Aw, sendo considerado tropical chuvoso-mesotérmico, com precipitação média anual de 1.153 mm. A temperatura média anual varia de 20 a 23°C. A área de estudo está inserida na bacia do Rio Doce. O solo da região é classificado como um Neossolo Flúvio distrófico típico, situado em área de baixada e relevo levemente ondulado, formada por sedimentos de afluentes do rio Doce (IBGE, 2001).

3.2 Procedimentos metodológicos empregados

3.2.1 Imagens e softwares

Para o desenvolvimento da pesquisa foram utilizadas imagens provenientes dos sensores AVNIR-2 e PALSAR a bordo do satélite ALOS. As imagens multiespectrais provenientes do sensor AVNIR-2, nas quatro bandas que abrangem o visível e o infravermelho próximo, foram obtidas no dia 27 de maio de 2009. Com um grau de visada de cerca de 44° e as imagens geradas apresentam resolução espacial de 10 m. As imagens obtidas do sensor PALSAR foram obtidas no dia 02 de maio de 2009, sendo utilizadas quatro tipos de polarizações sob um ângulo de incidência variando de 8 a 30°.

As imagens adquiridas apresentavam correção geométrica prévia de acordo com o nível de processamento solicitado (Nível 1B2-G), sendo referenciadas a partir de um arquivo vetorial corrigido, contendo as estradas e os corredores entres os talhões dos projetos localizados na área de estudo. Um dos objetivos foi fazer com que as linhas contendo as estradas sobrepusessem às visualizadas nas imagens. Foi empregada também a técnica de correção atmosférica pela subtração de DN's dos "pixels" escuros conforme Chavez (1975).

Quanto ao RADAR, as imagens originais, com resolução radiométrica de 16 bits, foram convertidas para 8 bits, gerando imagens polarimétricas (HH, HV, VV e VH) na mesma resolução da AVNIR-2. As imagens originais também foram convertidas para o coeficiente de retroespalhamento (HENDERSON; LEWIS, 1998). A conversão entre

amplitude (DN) e seção transversal de radar normalizada (dB), foi efetuada empregando a equação de Shimada et al. (2006), apresentada na sequência:

$$\sigma^o = 10 * \log_{10} * [DN^2] + CF \quad (34)$$

Onde CF (fator de conversão) = -83,0

Foram empregados os softwares ERDAS Imagine 2011 (Earth Resource Data Analysis System Inc., Atlanta, GA) para conversão dos dados em formato CEOS para o formato IMAGE e para o refinamento da correção geométrica; ENVI (Exelis Visual Information Solutions, Boulder, Colorado) para o tratamento dos dados de radar; ArcGIS (Environmental Systems Research Institute, Redlands, California) para a localização das parcelas amostrais e extração dos valores de DN e STATISTICA 12 (StatSoft Dell Software Inc., Tulsa, Oklahoma) para as análises estatísticas.

3.2.2 Inventário florestal

O inventário florestal referente à área de estudo foi realizado no período de fevereiro a setembro de 2009, dentro do cronograma de inventário contínuo. Foram realizadas medidas dos parâmetros dendrométricos da altura e diâmetro a 1,30m de altura (*dap*), para modelagem do volume de madeira, incluindo também a obtenção de medidas de biomassa por método destrutivo, derivada de árvores pré-selecionadas de representatividade do sítio.

3.2.3 Seleção das áreas amostrais e extração dos valores numéricos

Foram utilizadas 190 parcelas, de um total de 1924 parcelas, selecionadas seguindo os critérios: a) situadas em áreas imageadas pelos 2 sensores AVNIR-2 e PALSAR; b) parcelas sem coberturas de nuvens e sombras na imagem AVNIR; c) para a imagem do sensor PALSAR as parcelas deveriam estar fora das áreas de sombras de radar; d) Parcelas com idade superior a cinco anos.

Como as parcelas inventariadas são quadradas, foi gerado uma máscara quadrada de área igual ao das parcelas. A sobreposição da máscara nas imagens mostra que cada parcela abrangia diferentes número de pixels e valores numéricos, sendo alguns deles inteiramente e outros parcialmente inseridos dentro da máscara, dependendo da imagem

usada e da área da parcela. A extração dos valores numéricos correspondente a cada parcela foi realizada através da interseção da máscara com as imagens. Em seguida foi utilizada uma média ponderada entre os valores numéricos e as áreas dos pixels selecionados. A Figura 15 ilustra o modo de extração do valor numérico da imagem pela máscara sobreposta à imagem do sensor AVNIR-2.

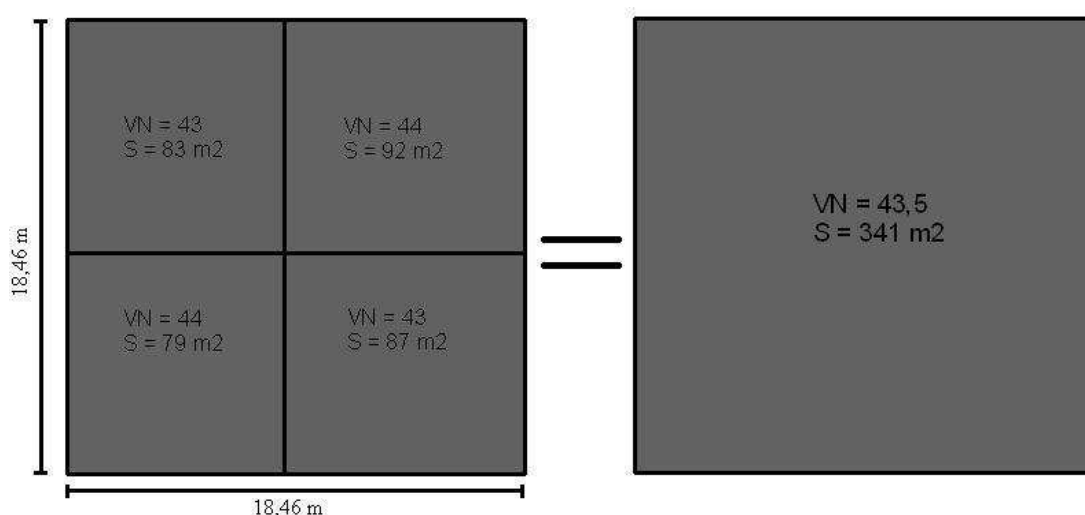


Figura 15 - Máscara para extração do valor numérico da amostra.

3.3 Estimativas a partir das imagens AVNIR-2 e PALSAR

Neste estudo o algoritmo empregado para estimar volume de madeira e biomassa aérea a partir das imagens digitais foi Máquina de Vetores de Suporte (MVS) para regressão. Para tal, foi considerado como variável dependente os dados de volume e biomassa obtidos por meio do inventário florestal e, como independentes, as variáveis digitais extraídas das 4 bandas das imagens AVNIR-2 (AV1, AV2, AV3 e AV4) e das 4 polarizações da banda L do sensor PALSAR (L_{HH} , L_{HV} , L_{VV} e L_{VH}) como mostrado nas Tabelas 1 e 2.

Tabela 1 – Variáveis AVNIR-2 utilizadas nas análises estatísticas

Variável	Descrição
AV1	Valor numérico na banda 1 : 0,42 ~ 0,50 μm (Azul)
AV2	Valor numérico na banda 2: 0,52 ~ 0,60 μm (Verde)
AV3	Valor numérico na banda 3: 0,61 ~ 0,69 μm (Vermelho)
AV4	Valor numérico na banda 4: 0,76 ~0,89 μm (Infravermelho Próximo)

Tabela 2 – Variáveis PALSAR utilizadas na análises estatísticas.

Variável	Descrição
L _{HH}	Valor numérico na polarização HH
L _{HV}	Valor numérico na polarização HV
L _{VV}	Valor numérico na polarização VV
L _{VH}	Valor numérico na polarização VH

3.4 Máquina de Vetor de Suporte (MVS) para regressão

O treinamento das MVS foi do tipo supervisionado, ou seja, o valor de saída desejado foi fornecido à MVS que comparou o erro (diferença entre saída desejada e calculada) com o valor de saída calculado. O erro do cálculo é minimizado pela MVS que modifica seus parâmetros internos a fim de se chegar em uma situação de erro mínimo e então, na solução ótima.

O algoritmo foi aplicado com o *software* STATISTICA 12 (StatSoft Dell Software Inc., Tulsa, Oklahoma), explorando todo o potencial do pacote do software, onde foram avaliados dois tipos de MVS para regressão:

- Tipo I (ϵ -SVM para regressão ou ϵ -SVR)

$$\begin{aligned} & \text{Minimize} \quad \frac{1}{2} \mathbf{w}^T \mathbf{w} + C \sum_{i=1}^n (\xi_i + \xi_i^*) \\ & \text{Sujeito à} \quad \begin{cases} y_i - \mathbf{w}^T \Phi(x_i) - b \leq \varepsilon + \xi_i \\ \mathbf{w}^T \Phi(x_i) + b - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0, \quad i = 1, \dots, n \end{cases} \end{aligned}$$

em que C é termo de regulação, w é o vetor dos parâmetros, b é a constante e ξ a variável folga do erro fora da margem de aceite ε , calculada pelo vetores de suporte. O índice i rotula os n casos de treinamentos. Já o termo $\Phi(x_i)$ representa a transformação dos dados de entrada (independentes) pelas funções *kernels* $K(x_i, x_j)$ no espaço de características. Sendo $K(X_i, X_j) = \Phi(x_i) \cdot \Phi(x_j)$.

- Tipo II (v-SVM para regressão ou v-SVM)

$$\begin{aligned} \text{Minimize} \quad & \frac{1}{2} \mathbf{w}^T \mathbf{w} + C \left(v\epsilon + \frac{1}{l} \sum_{i=1}^l (\xi_i + \xi_i^*) \right) \\ \text{Sujeito à} \quad & \begin{cases} y_i - \mathbf{w}^T \Phi(x_i) - b \leq \epsilon + \xi_i \\ \mathbf{w}^T \Phi(x_i) + b - y_i \leq \epsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0, \quad i = 1, \dots, n, \epsilon \geq 0 \end{cases} \end{aligned}$$

em que C é termo de regulação, w é o vetor dos parâmetros, b é a constante e ξ a variável folga do erro fora da margem de aceite ϵ , calculada com base nos vetores de suporte e otimizada pelo parâmetro v . O índice i rotula os n casos de treinamentos. Já o termo $\Phi(x_i)$ representa a transformação dos dados de entrada (independentes) pelas funções *kernels* $K(x_i, x_j)$ no espaço de características. Sendo $K(X_i, X_j) = \Phi(x_i) \cdot \Phi(x_j)$.

$$K(X_i, X_j) \left\{ \begin{array}{l} X_i \cdot X_j \\ (-\gamma X_i \cdot X_j + k)^d \\ \exp(-\gamma |X_i - X_j|^2) \\ \tanh(\gamma X_i \cdot X_j + k) \end{array} \right. \left. \begin{array}{l} \text{Linear} \\ \text{Polinomial} \\ \text{Gaussiana (RBF)} \\ \text{Sigmoide} \end{array} \right\}$$

em que $K(X_i, X_j)$ é válido para as MVS tipo I e II. Os valores de γ (gama), d (grau) e k (coeficiente) representam os parâmetros ajustáveis para a configuração das MVS pelo usuário.

Neste trabalho, o pacote MVS usado pelo software STATISTICA 12.0 ajusta automaticamente por validação-cruzada, os parâmetros: C (termo de regulação), ϵ (margem máxima de aceite) e v (parâmetro otimizador de ϵ), ficando a cargo do usuário a configuração dos parâmetros das funções *kernels*. Dessa forma, as diferentes configurações de MVS foram avaliadas e os melhores desempenhos foram escolhidos de acordo com valores das estatísticas: coeficiente de correlação entre os dados reais e estimados ($r_{y\hat{y}}$) e a raiz quadrada do erro quadrático médio (RQEM). Foram usadas também as estatísticas para avaliação da qualidade dos ajustes, as quais são descritas na sequência:

$$r_{y\hat{y}} = \frac{n^{-1} \sum_{i=1}^n (\hat{Y}_i - \hat{Y}_m)(Y_i - \bar{Y})}{\sqrt{[n^{-1} \sum_{i=1}^n (\hat{Y}_i - \hat{Y}_m)^2][n^{-1} \sum_{i=1}^n (Y_i - \bar{Y})^2]}}$$

$$\text{RQEM} = \sqrt{n^{-1} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2}$$

$$\text{Bias} = n^{-1} \sum_{i=1}^n (\hat{Y}_i - Y_i)$$

$$\text{MDA} = n^{-1} \sum_{i=1}^n |\hat{Y}_i - Y_i|$$

em que n é o número de observações; Y_i o i -ésimo valor observado; \hat{Y}_i o i -ésimo valor estimado; \bar{Y} = média dos valores observados e \hat{Y}_m a média das estimativas de \hat{Y}_i .

3.5 Teste de Validação e Aplicação das MVS

Mediante a configuração adequada para as MVS, estimou-se o volume de madeira e a biomassa aérea estimados a partir dos sensores AVNIR-2 e/ou PALSAR do satélite ALOS. Considerando os resultados encontrados por Oliveira (2011) e Domingues (2015), o presente estudo utilizou como critérios de aplicação das MVS, a validação das generalizações. Assim, foi usado um teste de validação modificado em relação ao proposto por Guimarães (1994).

Em Guimarães (1994), os testes que envolvessem comparações entre variáveis incluídas na formulação do modelo eram considerados de validação. Já os que envolviam variáveis para modelos de prognose, eram considerados testes de aplicação. Analogamente, no presente estudo, foi considerado teste de validação, se fossem envolvidas variáveis das MVS treinadas, e teste de aplicação, se fossem envolvidas variáveis das MVS testadas ou generalizadas. Como o objetivo deste trabalho foi testar a capacidade de generalização das MVS treinadas, foram avaliados os testes de aplicação para as MVS testadas.

Um dos critérios para avaliar a precisão dos valores estimados e aplicabilidade das MVS treinadas, foi o teste F (CAMPOS; LEITE, 2013; GUIMARÃES, 1994). O teste F avalia valores observados e estimados sob a hipótese $H_0: [\beta_0 \beta_1] = [0 \ 1]$ segundo o modelo linear

$$Y_j = \beta_0 + \beta_1 Y_1 + \varepsilon$$

em que :

Y_j = valor estimado

Y_i = valor observado;

β_0 e β_1 = parâmetros; e

ε = erro aleatório, $\varepsilon \sim \text{NID}(0, \sigma^2)$

A estatística de teste é dada por:

$$F(H_0) = \frac{(C' \hat{\beta})(y_1' y_1)^{-1} (C' \hat{\beta} - \theta)}{2. QM Res}$$

em que:

C = matriz identidade de ordem 2;

$$\hat{\beta} = [\hat{\beta}_0 \hat{\beta}_1];$$

$$\theta' = [0 \ 1]$$

A não rejeição de H_0 ($F(H_0) < F_\alpha(2, n-2)$) implica que os valores estimados e observados são estatisticamente coincidentes, perfazendo uma linha reta que passa pela origem ($\beta_0 = 0$), com declividade igual a 45° ($\beta_1 = 1$).

Campos; Leite (2013) ressaltam que esta estatística $F(H_0)$ isoladamente pode não gerar resultados conclusivos. Esta pode se tornar inconsistente quando os valores estimados estiverem muito próximos dos observados, pelo fato de ter como denominador o quadrado médio do resíduo (*QM Res*), que nesse caso será muito reduzido.

Assim, outros critérios adotados neste estudo para avaliar a precisão dos valores estimados, basearam-se na análise de correlação e no valor absoluto do erro médio percentual das estimativas das generalizações. Neste estudo, optou-se por aceitar a aplicabilidade das MVS, sempre que a correlação entre os valores estimados e observados fossem maiores que 0,900 e, simultaneamente, os valores absolutos dos erros médios percentuais das estimativas fossem menores que 10%. Assim, a interpretação dos resultados baseou-se nas seguintes combinações possíveis:

Caso	Teste F (H_0)	$r_{y\hat{y}}$	Erro (%)	Situação
1	n.s.	$r_{y\hat{y}} \geq 0,900$	$\leq 10\%$	Ideal
2	n.s.	$r_{y\hat{y}} \geq 0,900$	$> 10\%$	Aceitável
3	n.s.	$r_{y\hat{y}} < 0,900$	$\leq 10\%$	Inaceitável
4	n.s.	$r_{y\hat{y}} < 0,900$	$> 10\%$	Inaceitável
5	*	$r_{y\hat{y}} \geq 0,900$	$\leq 10\%$	Aceitável
6	*	$r_{y\hat{y}} \geq 0,900$	$> 10\%$	Inaceitável
7	*	$r_{y\hat{y}} < 0,900$	$\leq 10\%$	Inaceitável
8	*	$r_{y\hat{y}} < 0,900$	$> 10\%$	Inaceitável

Adaptado de Guimarães (1994). Em que, n.s. = não significativo ao nível de 5% de probabilidade; * = significativo ao nível de 5% de probabilidade.

Onde o |Erro (%)| é representado pela seguinte equação:

$$|Erro(\%)| = 100 \cdot n^{-1} \sum_{i=1}^n \left(\frac{\hat{Y}_i - Y_i}{Y_i} \right)$$

em que n é o número de observações; Y_i o i -ésimo valor observado; \hat{Y}_i o i -ésimo valor estimado; \bar{Y} = média dos valores observados e \hat{Y}_m a média das estimativas de \hat{Y}_i .

As estimativas foram consideradas adequadas, quando as comparações planejadas fossem coincidentes ao caso 1, 2 e 5. O caso 2, é considerado aceitável, pois presume-se que pode ocorrer em função de uma intensidade de amostragem insuficiente ou mesmo de erros de não amostragem. O caso 5, a forte correlação entre os valores reais e estimados implica na redução da variância residual, tornando o teste F extremamente sensível para rejeitar a hipótese de nulidade. Portanto mesmo com a significância do teste F, esta alternativa 5 pode ser aceita (CAMPOS; LEITE, 2013).

Os treinamentos das MVS foram realizados com: 75%, 60%, 50%, 40%, 25% e 10% do total de parcelas. Respectivamente, as generalizações foram realizadas utilizando-se: 25%, 40%, 50%, 60%, 75% e 90% do total de parcelas.

4 RESULTADOS

4.1 Treinamento das MVS para estimativa de volume de madeira

4.1.1 Treinamento das MVS com 75% dos dados

Foram treinadas cerca de 500 MVS utilizando 75 % dos dados (142 parcelas) para estimar o volume de madeira de povoamentos de eucalipto. Os 25% restantes, para cada ajuste, foram utilizados para o teste de generalização. Avaliou-se o desempenho de diferentes configurações e variáveis de entrada, a fim de se encontrar os desempenhos com melhores estimativas. Para uma mesma configuração, ajustou-se os parâmetros das MVS até retornarem os maiores índices de correlação e menores erros de generalização.

Na Tabela 3, são apresentadas as configurações das MVS com as melhores performances segundo o tipo de regressão e função *kernel* para predição de volume a partir de variáveis de entrada do sensor AVNIR-2. Nesta tabela, pode-se verificar que as MVS que obtiveram os melhores desempenhos foram as que utilizaram a regressão tipo II e função *kernel* RBF, com maiores valores de $r_{\hat{y}y}$ e menores valores de RQEM. As outras funções *kernel* apresentaram os piores desempenhos, com destaque para a linear e sigmoide, que obtiveram fraca correlação para as estimativas de treinamento. A Regressão tipo II apresentou em geral melhor desempenho em relação a regressão tipo I.

Tabela 3 – Configurações das MSV selecionadas para cada tipo de regressão e função *kernel* para a predição de volume (m³) utilizando dados do sensor AVNIR-2.

Variáveis de Entrada	Reg.	Vetores		Função <i>Kernel</i>				$r_{\hat{y}y}$			RQEM (m ³)	
		M	S	Tipo	Grau	Gama	Coefficiente	(total)	(treino)	(teste)	(treino)	(teste)
AVNIR-2	I	30	11	Lin	-	-	-	0,503	0,536	0,391	44,530	46,200
	II	18	12		-	-	-	0,563	0,563	0,404	26,198	24,055
AVNIR-2	I	120	118	Pol	3,000	1,000	1,000	0,741	0,785	0,572	15,704	25,323
	II	31	13		3,000	1,100	1,000	0,740	0,790	0,549	14,807	24,754
AVNIR-2	I	54	22	RBF	-	0,800	-	0,919	0,931	0,813	4,632	7,035
	II	53	18		-	0,900	-	0,919	0,953	0,833	4,332	6,310
AVNIR-2	I	62	60	Sig	-	0,050	4,000	0,502	0,547	0,321	21,369	28,630
	II	78	76		-	0,060	4,000	0,510	0,544	0,394	21,394	28,670

em que : Reg I = ϵ -SVM; Reg II = ν -SVM; Lin = Linear, Pol = Polinomial, RBF = Gaussian Radial Basis Function, Sig = Sigmoide. AVNIR-2 corresponde as bandas multiespectrais ópticas: AV1, AV2, AV3 e AV4; RQEM = Raiz Quadrada do Erro Quadrático Médio; Vetores M: margem máxima; Vetores S: margem suave.

Já para as MVS utilizando como variáveis de entrada as polarizações do sensor PALSAR, os desempenhos apresentaram média correlação das estimativas de volume (Tabela 4). Os valores de $r_{\hat{y}y}$ para as generalizações foram maiores do que os de

treinamentos, mostrando uma distribuição distinta dos dados do sensor PALSAR. Os índices de correlação para as generalizações variaram entre 0,729-0,748, mostrando pouca influência das configurações nos resultados dos ajustes. A melhor configuração neste caso foi MVS utilizando regressão tipo II e função *kernel* RBF.

Tabela 4 – Configurações das MSV selecionadas para cada tipo de regressão e função *kernel* para a predição de volume (m³) utilizando dados do sensor PALSAR.

Variáveis de Entrada	Reg.	Vetores		Função <i>Kernel</i>			r_{yy}			RQEM (m ³)		
		M	S	Tipo	Grau	Gama	Coefficiente	(total)	(treino)	(teste)	(treino)	(teste)
PALSAR	I	123	118	Lin	-	-	-	0,701	0,701	0,729	22,611	19,870
	II	81	74		-	-	-	0,702	0,700	0,730	21,687	19,176
PALSAR	I	106	98	Pol	2,000	1,300	2,000	0,709	0,709	0,731	22,327	19,833
	II	114	101		2,000	1,300	2,000	0,708	0,707	0,731	22,543	19,601
PALSAR	I	94	55	RBF	-	0,100	-	0,711	0,709	0,748	21,079	19,424
	II	78	39		-	0,100	-	0,719	0,718	0,748	21,030	18,584
PALSAR	I	81	66	Sig	-	0,080	0,000	0,690	0,684	0,743	25,198	20,599
	II	88	45		-	0,070	0,000	0,695	0,690	0,743	26,276	22,133

em que: Reg I = ϵ -SVM; Reg II = ν -SVM; Lin = Linear, Pol = Polinomial, RBF = Gaussian Radial Basis Function, Sig = Sigmoide; PALSAR remete as polarizações L_{HH}, L_{VV}, L_{HV}, L_{VH}; RQEM = Raíz Quadrada do Erro Quadrático Médio; Vetores M: margem máxima; Vetores S: margem suave.

As MVS que utilizaram simultaneamente dados de ambos sensores como variáveis de entrada apresentaram, com a função RBF, as melhores estimativas em comparação aos dados não integrados (Tabela 5). Como pode ser observado na tabela 5, as funções *kernel* Polinomial e RBF apresentaram as maiores correlações e menores valores de erros.

Tabela 5 – Configurações das MSV selecionadas para cada tipo de regressão e função *kernel* na predição de volume (m³) com dados dos sensores AV-2/PALSAR.

Variáveis de Entrada	Reg.	Vetores		Função <i>Kernel</i>			r_{yy}			RQEM (m ³)		
		M	S	Tipo	Grau	Gama	Coefficiente	(total)	(treino)	(teste)	(treino)	(teste)
AV-2 + PALSAR	I	7	2	Lin	-	-	-	0,721	0,725	0,705	18,789	20,347
	II	37	26		-	-	-	0,793	0,788	0,820	15,463	13,088
AV-2 + PALSAR	I	60	52	Pol	2,000	0,050	1,000	0,807	0,803	0,831	9,831	9,823
	II	95	64		3,000	0,200	2,000	0,885	0,895	0,854	8,122	8,036
AV-2 + PALSAR	I	102	15	RBF	-	2,200	-	0,958	0,834	0,848	2,272	6,599
	II	115	18		-	2,250	-	0,979	0,998	0,923	1,509	4,546
AV-2 + PALSAR	I	90	44	Sig	-	0,100	0,000	0,784	0,777	0,820	12,924	12,167
	II	89	42		-	0,020	0,000	0,800	0,795	0,829	8,287	9,868

em que: Reg I = ϵ -SVM; Reg II = ν -SVM; Lin = Linear, Pol = Polinomial, RBF = Gaussian Radial Basis Function, Sig = Sigmoide. AV-2 = bandas multiespectrais ópticas do sensor AVNIR-2, são elas: AV1, AV2, AV3 e AV4; PALSAR remete as polarizações L_{HH}, L_{VV}, L_{HV}, L_{VH}; RQEM = Raíz Quadrada do Erro Quadrático Médio; Vetores M: margem máxima; Vetores S: margem suave.

As funções linear e sigmoide apresentaram resultados inferiores para treinamento dos dados e consequente generalização. De modo geral, o melhor desempenho com as variáveis de entrada integradas foi utilizando na configuração a regressão tipo II e função *kernel* RBF.

4.1.1.1 Teste de validação e aplicação

Procedeu-se a análise da qualidade de ajuste das MVS com os melhores desempenhos (RBF), e subsequente teste de validação dos treinamentos e de aplicação das generalizações. As MVS foram consideradas aplicáveis quando suas generalizações atenderam os critérios do teste de Guimarães (1994) modificado proposto na metodologia do trabalho.

Pela Tabela 6, nota-se que a integração dos dados permitiu uma correlação mais forte das estimativas de generalização, além de menores valores residuais e Bias. A utilização somente das quatro polarizações do sensor PALSAR, gerou estimativas de generalização com maiores Bias, além de maiores valores residuais (RQEM e MDA). As bandas multiespectrais do AVNIR-2 apresentaram generalização com desempenho intermediário, e menores valores de Bias.

Tabela 6 – Análise da qualidade dos ajustes dos modelos MVS treinados para estimativa de volume de madeira a partir dos sensores AVNIR-2 e PALSAR.

Variável (Output)	Sensores (Input)*	Kernels	r_{yy} (teste)	RQEM (teste)		Bias (teste)		MDA (teste)	
				(m ³)	(%)	(m ³)	(%)	(m ³)	(%)
Volume	AV-2	RBF	0,833	6,310	2,494	0,172	0,068	7,028	2,778
	PAL	RBF	0,748	18,584	7,345	3,943	1,558	14,851	5,869
	AV-2+PAL	RBF	0,923	4,546	1,797	-1,508	-0,596	5,410	2,138

*AV-2: AVNIR-2 (AV1, AV2, AV3, AV4); PAL: PALSAR (L_{HH}, L_{VV}, L_{HV}, L_{VH}); RBF = função *kernel* Gaussiana Radial Basis Function

Segundo o teste de validação para os treinamentos, tabela 7, apenas o treinamento com as variáveis de entrada PALSAR foi considerado inaceitável. Já utilização do sensor AVNIR-2 para os treinamentos, mesmo sendo considerada válida, não foi aplicável para generalização dos dados. Somente a integração das variáveis de entrada apresentou treino válido e aceitável generalização para os dados de volume. A generalização aceitável foi a que apresentou RQEM de 4,546 m³ ou 1,797% da média de volume das parcelas

testadas, e valores relativos de Bias e MDA, respectivamente, -0,596% e 2,138% (Tabela 6).

Tabela 7 – Teste de validação e aplicação das MVS selecionadas com as diferentes variáveis de entrada na estimativa de volume de madeira (m³) de povoamentos de eucalipto.

	Estatística	AVNIR-2	PALSAR	AVNIR-2+PALSAR
Treinamento				
	F (H ₀)	17,6741*	97,6313*	0,8348 ^{ns}
	r _{yŷ}	0,953	0,718	0,998
	Erro(%)	0,003	2,100	-0,001
	Situação	Aceitável	Inaceitável	Aceitável
Generalização				
	F (H ₀)	18,043*	18,2876*	17,139*
	r _{yŷ}	0,892	0,748	0,923
	Erro(%)	0,225	2,100	-0,280
	Situação	Inaceitável	Inaceitável	Aceitável

* Teste F(H₀) significativo 5% de probabilidade; ^{ns} Teste F(H₀) não significativo a 5% de probabilidade.

Na Figura 16, situam-se os gráficos residuais das generalizações das MVS treinadas sem e com a integração das variáveis de entradas. Utilizando o sensor AVNIR-2 (Figuras 16a, 16d e 16g), a generalização apresentou estimativas enviesadas positivamente para as parcelas com os menores volumes de madeira. Já a utilização do sensor PALSAR (Figuras 16b, 16e e 16h) como já observado com o maior valor de MDA, apresentou maior amplitude dos erros com relação aos outros dois tipos de entradas. As Figuras 16f, 16c e 16i, mostram que a integração dos sensores gerou estimativas com valores residuais mais uniformes para todas as parcelas testadas, cujo resultado corrobora com o teste de aplicação da tabela 7.

Portanto, pela análise de qualidade de ajuste, análise gráfica dos resíduos e teste de validação e aplicação, as estimativas de volume de madeira foram consistentes com uma MVS com 142 parcelas amostrais, configurada com regressão tipo II, função *kernel* RBF, utilizando-se como variáveis de entrada a integração dos todos os dados fornecidos pelos sensores AVNIR-2 e PALSAR. Os gráficos residuais dos treinamentos das MVS podem ser encontrados no APÊNDICE A.

4.1.2 Redução do número de amostras para o treinamento das MVS

O treinamento com a variação do número de parcelas amostrais buscar verificar se a estimação de volume de madeira de eucalipto pode ser consistente com um número menor de amostras. Utilizando-se a SMV tipo II e a função Gaussiana RBF como função

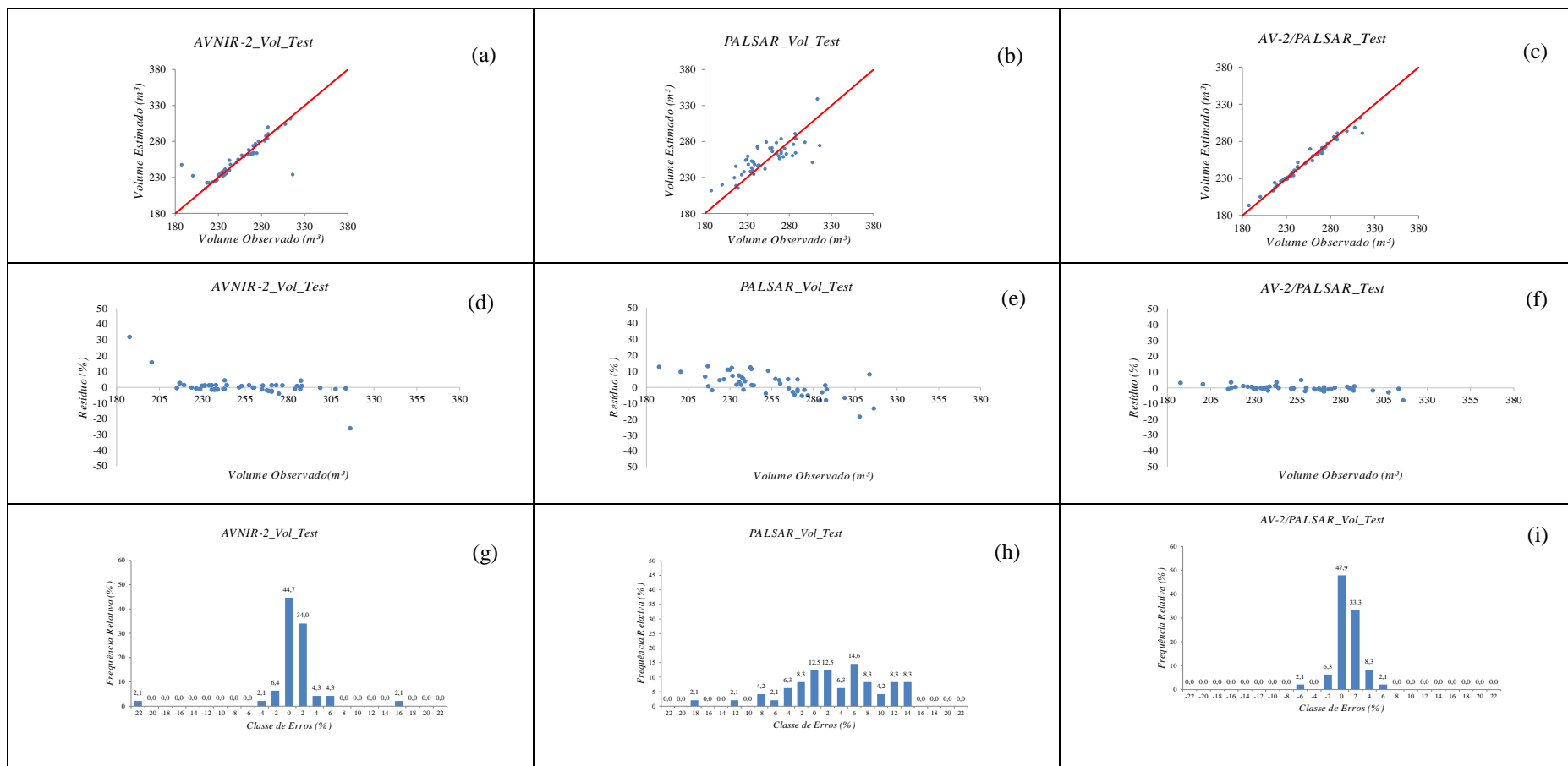


Figura 16 – Gráficos de resíduos das generalizações das MVS para as previsões de volume utilizando dados dos sensores AVNIR-2 e PALSAR. As Figuras 16a, 16b e 16c representam os gráficos da relação dos valores observados e estimados pelas MVS, utilizando como dados de entrada, respectivamente, os dados dos sensores: AVNIR-2, PALSAR, AVNIR-2+PALSAR. As Figuras 16d, 16e e 16f representam os gráficos de dispersão dos resíduos percentuais das estimativas das MVS utilizando como dados de entrada, respectivamente, os dados dos sensores: AVNIR-2, PALSAR, AVNIR-2+PALSAR. As Figuras 16g, 16h e 16i representam os gráficos de frequência de classes de erros das estimativas das MVS utilizando como dados de entrada, respectivamente, os dados dos sensores: AVNIR-2, PALSAR, AVNIR-2+PALSAR.

kernel, treinou-se uma MVS ajustando-se o parâmetro γ para os diferentes dados de entrada e número de parcelas para treinamento, como mostra a tabela 8. A redução do número de amostras para treinamento, com base no total de parcelas (190), consistiu em: 60% (114 parcelas), 50% (95 parcelas), 40% (76 parcelas), 25% (47 parcelas), 10 % (19 parcelas).

A Tabela 8 mostra que ao se reduzir o número de parcelas para o treinamento das MVS, os melhores desempenhos continuaram sendo com a utilização simultânea das bandas multiespectrais do AVNIR-2 e as polarizações do sensor PALSAR, assim como observado com o treinamento usando 142 parcelas.

Tabela 8 – Melhores MVS selecionadas para avaliar a estimativa de volume de madeira (m³) de povoamentos de eucalipto, com a redução do número de parcelas amostrais para treinamento.

Número de Amostras	Variáveis de Entrada	MSV		Vet.	Treinamento		Generalização	
		<i>Kernel</i>	(γ)		r_{yy}	RQEM(m ³)	r_{yy}	RQEM(m ³)
142	AVNIR-2	II - RBF	0,900	60 (32)	0,931	4,332	0,833	6,310
142	PALSAR	II - RBF	1,300	65 (59)	0,718	21,030	0,748	18,584
142	AV+P	II - RBF	2,250	115(18)	0,998	1,509	0,923	4,546
114	AVNIR-2	II - RBF	0,800	72 (64)	0,922	5,421	0,777	19,779
114	PALSAR	II - RBF	1,350	87 (76)	0,644	22,487	0,695	20,142
114	AV+P	II - RBF	2,200	52 (25)	0,997	1,665	0,932	9,631
95	AVNIR-2	II - RBF	0,250	60 (51)	0,931	4,447	0,794	20,433
95	PALSAR	II - RBF	1,450	50 (42)	0,555	22,155	0,601	22,956
95	AV+P	II - RBF	2,050	32(16)	0,998	1,748	0,984	9,366
76	AVNIR-2	II - RBF	0,250	46 (39)	0,884	6,677	0,645	20,446
76	PALSAR	II - RBF	1,550	40 (33)	0,556	30,884	0,244	45,274
76	AV+P	II - RBF	2,250	25 (13)	0,984	3,252	0,889	13,088
47	AVNIR-2	II - RBF	0,250	23 (13)	0,899	7,845	0,555	30,435
47	PALSAR	II - RBF	1,850	10 (7)	0,524	30,955	0,278	39,441
47	AV+P	II - RBF	1,850	30(16)	0,987	2,005	0,791	15,956
19	AVNIR-2	II - RBF	0,250	9 (2)	0,870	6,652	0,322	39,955
19	PALSAR	II - RBF	1,300	14 (8)	0,505	30,241	0,241	40,199
19	AV+P	II - RBF	2,200	10(4)	0,991	2,544	0,468	26,227

em que: AV+P: combinação dos dados dos sensores AVNIR-2 e PALSAR; II-RBF: combinação da configuração regressão tipo II e função *kernel* Gaussiana Radial Basis function. Vet.: Vetores de margem máxima e margem suave (parênteses).

4.1.2.1 Teste de validação e aplicação

Procedeu-se a análise das MVS com a combinação das variáveis de entrada, cujos desempenhos foram os melhores com a redução de amostras. A Tabela 9, mostra as estatísticas de qualidade de ajuste das generalizações das MVS treinadas com a redução

do número de parcelas amostrais para treinamento. Observou-se que conforme o número de amostras diminui, reduz-se também o número de vetores responsáveis por determinar, internamente ao algoritmo, as margens de aceite de erro e relevância de *outliers* (margem máxima e margem suave). Mesmo com treinos apresentando fortes correlações, com um menor número de vetores, as generalizações apresentaram desempenhos inferiores, com correlações cada vez mais fracas e maiores valores de erro (RQEM, Bias e MDA).

Tabela 9 – Melhores MVS selecionadas para estimar volume de madeira de povoamentos de eucalipto a partir de dados dos sensores AVNIR-2 e PALSAR, com a redução do número de parcelas amostrais para treinamento.

N	MVS	$r_{\hat{y}\hat{y}}$	$r_{\hat{y}\hat{y}}$	Vet.	RQEM (teste)		Bias (teste)		MDA (teste)	
		(treino)	(teste)		(m ³)	(%)	(m ³)	(%)	(m ³)	(%)
142	II-RBF	0,998	0,923	102 (50)	4,546	1,797	-1,508	-0,596	5,410	2,138
114	II-RBF	0,997	0,932	52 (25)	9,631	3,821	0,226	0,090	6,648	2,638
95	II-RBF	0,998	0,949	32 (16)	9,366	3,676	-0,574	-0,225	6,482	2,545
76	II-RBF	0,984	0,889	25 (13)	13,088	5,147	0,524	0,206	12,037	4,734
47	II-RBF	0,987	0,7910	30 (16)	15,956	6,262	0,099	0,039	10,022	3,933
19	II-RBF	0,991	0,468	10 (4)	26,227	10,305	4,947	1,944	23,049	9,057

em que: N equivale ao número de parcelas amostrais para treinamento das MVS. II-RBF é a combinação da configuração regressão tipo II e função *kernel* Gaussiana Radial Basis function. Vet.: Vetores de margem máxima e margem suave (parênteses).

A Tabela 10 mostra que as MVS são muito eficientes para o treinamento dos dados, com índices de correlação alto e treinamento aceitável até mesmo com 19 parcelas, 10% do total da base de dados para volume de madeira. Porém, ainda que validados, nem todos os treinos foram considerados aplicáveis. Somente os treinamentos com número mínimo de 95 parcelas foram aceitáveis para generalização dos dados, isto é, aceitáveis para aplicação. Estas generalizações aceitas apresentaram RQEM menor que 9,366 m³ (3,676% da média) e MDA abaixo de 2,545% da média das parcelas testadas (Tabela 9).

Tabela 10 - Teste de validação e aplicação das melhores MVS selecionadas, com a redução do número de parcelas amostrais para treinamento, na estimativa de volume de madeira (m³) de povoamentos de eucalipto.

Est.	142	114	95	76	47	19
Treino						
F(H ₀)	0,8348 ^{ns}	4,767*	4,045*	5,132*	0,2578 ^{ns}	19,338*
$r_{\hat{y}\hat{y}}$	0,998	0,997	0,998	0,984	0,987	0,991
Erro(%)	0,001	0,001	0,001	0,001	0,001	0,011
Situação	Aceitável	Aceitável	Aceitável	Aceitável	Aceitável	Aceitável
Generaliz.						
F(H ₀)	17,139*	13,822*	17,784*	18,685*	43,276*	439,103*
$r_{\hat{y}\hat{y}}$	0,949	0,932	0,949	0,889	0,7910	0,468
Erro(%)	-0,005	0,463	-0,018	0,489	0,465	3,031
Situação	Aceitável	Aceitável	Aceitável	Inaceitável	Inaceitável	Inaceitável

* Teste F(H₀) significativo 5% de probabilidade; ^{ns} Teste F(H₀) não significativo a 5% de probabilidade.

A Figura 17 mostra os gráficos residuais das generalizações das MVS, reduzindo-se o número de amostras para treinamento. Pela figura, verifica-se que as generalizações com menos de 95 parcelas apresentaram maior amplitude dos resíduos, resultado que mais uma vez corrobora com o teste de aplicação das MVS. Os gráficos residuais dos correspondentes treinamentos podem ser encontrados no APÊNDICE B

Dessa forma, pode-se afirmar que o mínimo de 95 parcelas foi suficiente para estimativas consistentes de volume de madeira de povoamentos de eucalipto, com MVS configuradas com a regressão tipo II, função *kernel* RBF e os dados dos sensores AVNIR-2 e PALSAR com a integração das variáveis de entrada.

4.1.3 Redução do número de variáveis de entrada

Ao contrário da regressão *stepwise* e das RNAs trabalhadas, respectivamente, por Oliveira (2011) e Domingues (2015), o pacote do STATISTICA 13 para as MVS não explicita quais variáveis de entrada tiveram maior importância para os resultados das estimativas. Com base nos resultados já encontrados até aqui para predição de volume, e com objetivo de explorar a capacidade de estimação das MVS com uma estrutura mais simples, testou-se as MVS com o mínimo de variáveis de entrada.

O treino das MVS consistiu na configuração regressão tipo II e função RBF como *kernel*, 95 parcelas amostrais tanto para treino quanto para teste de generalização dos ajustes e a combinação de variáveis de entrada de ambos sensores trabalhados. Os treinamentos e generalizações com o menor número possível de variáveis de entrada que retornaram ajuste ótimo são apresentados na Tabela 11, bem como suas estatísticas de qualidade de ajuste. De uma maneira geral as estimativas das generalizações apresentaram forte correlação, RQEM abaixo de 5%, com Bias e MDA não ultrapassando 2% do valor da média geral das parcelas testadas.

4.1.3.1 Teste de validação e aplicação

A Tabela 12 mostra que segundo o teste de validação e aplicação de Guimarães (1994) modificado, todas as MVS treinadas com pelo menos uma banda de cada sensor resultou em MVS válidas e aplicáveis. Verifica-se também que tanto os treinos quanto as

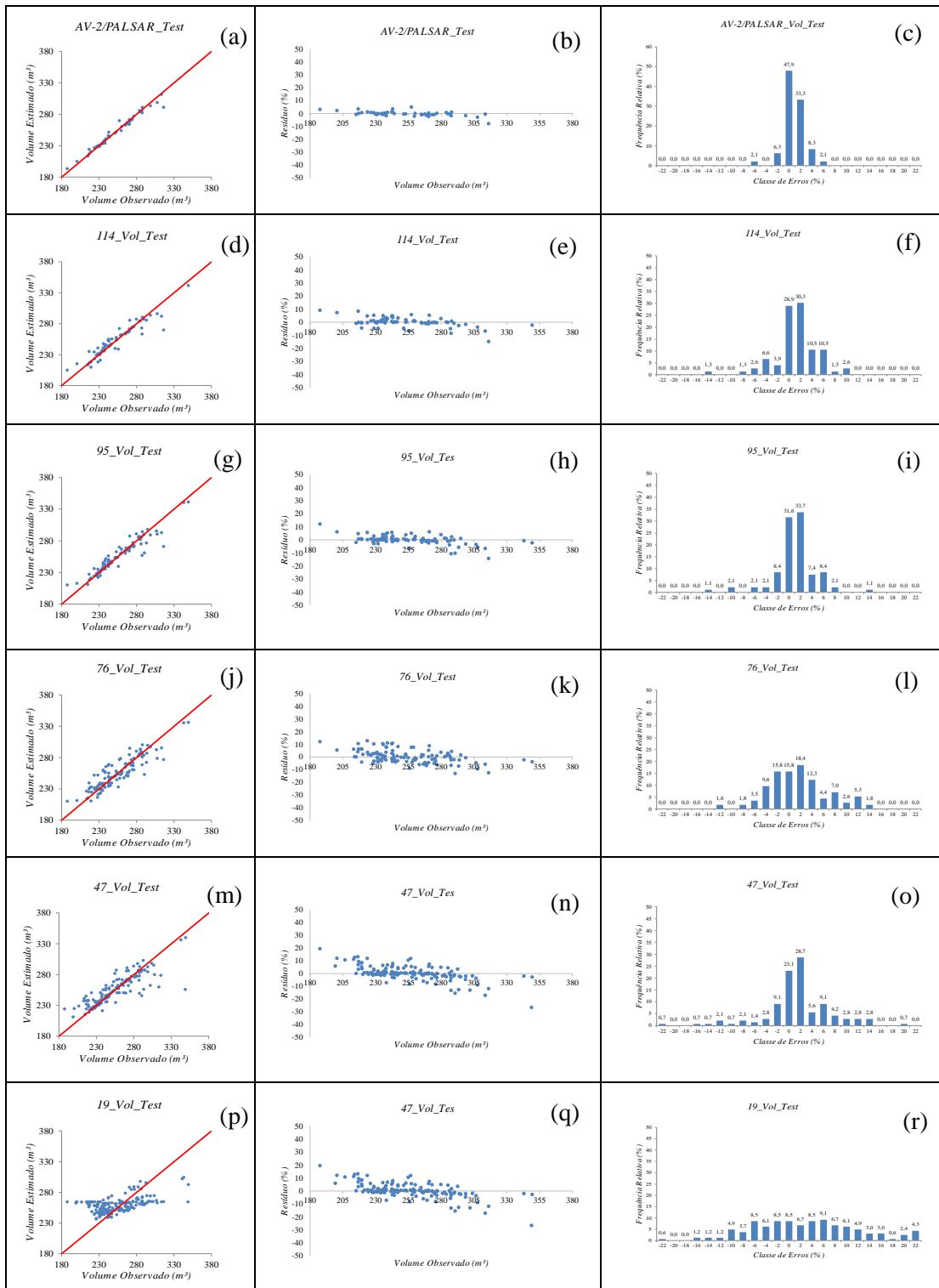


Figura 17 – Gráficos de resíduos das generalizações das MVS para as predições de volume utilizando dados dos sensores AVNIR-2 e PALSAR com a redução do número de parcelas amostrais para treinamento. As Figuras 17a, 17d, 17g, 17j, 17m e 17p representam os gráficos da relação dos valores observados e estimados pelas MVS treinadas, respectivamente, com: 142, 114, 95, 76, 47 e 19 parcelas. As Figuras 17b, 17e, 17h, 17k, 17n e 17q, representam os gráficos de dispersão dos resíduos percentuais das estimativas das MVS treinadas, respectivamente, com: 142, 114, 95, 76, 47 e 19 parcelas. As Figuras 17c, 17f, 17i, 17l, 17o e 17r representam os gráficos de frequência de classes de erros das estimativas das MVS treinadas, respectivamente, com: 142, 114, 95, 76, 47 e 19 parcelas.

generalizações das MVS com um menor número de variáveis de entrada resultam em estimativas de maior correlação com dados reais, do que os encontrados com o uso das 8 variáveis combinadas simultaneamente dos sensores AVNIR-2 e PALSAR.

Tabela 11 – MVS com duas variáveis combinadas dos sensores AVNIR-2 e PALSAR, utilizando 95 parcelas amostrais para treinamento, para estimativas de volume de povoamentos de eucalipto e estatísticas e qualidade de ajuste.

Variáveis de Entrada	MVS		$r_{y\hat{y}}$	$r_{\hat{y}y}$	RQEM (teste)		Bias (teste)	MDA (teste)
	Kernel	(γ)	(treino)	(teste)	(m ³)	(%)	(%)	(%)
AV1,L _{HH}	II-RBF	8,750	0,982	0,927	10,936	4,293	-0,200	1,950
AV1,L _{HV}	II-RBF	10,250	0,987	0,955	8,920	3,501	0,310	1,770
AV1,L _{VV}	II-RBF	15,750	0,997	0,947	9,887	3,881	0,788	1,855
AV1,L _{VH}	II-RBF	18,000	0,988	0,967	7,857	3,084	-0,956	1,352
AV2,L _{HH}	II-RBF	18,000	0,974	0,949	9,233	3,624	-1,760	1,420
AV2,L _{HV}	II-RBF	20,000	0,991	0,982	5,584	2,192	-0,400	1,170
AV2,L _{VV}	II-RBF	15,250	0,987	0,938	10,730	4,212	1,350	1,921
AV2,L _{VH}	II-RBF	20,000	0,990	0,981	5,857	2,299	-0,400	1,323
AV3,L _{HH}	II-RBF	10,150	0,992	0,941	10,070	3,953	-2,95	1,472
AV3,L _{HV}	II-RBF	25,000	0,998	0,966	7,759	3,046	-0,596	1,349
AV3,L _{VV}	II-RBF	25,500	0,988	0,945	9,948	3,905	0,978	1,86
AV3,L _{VH}	II-RBF	10,000	0,988	0,957	8,541	3,353	-0,355	1,401
AV4,L _{HH}	II-RBF	15,000	0,999	0,978	6,257	2,456	0,854	0,942
AV4,L _{HV}	II-RBF	17,250	0,999	0,990	4,369	1,715	-0,314	0,630
AV4,L _{VV}	II-RBF	17,500	0,991	0,963	8,371	3,286	0,765	1,309
AV4,L _{VH}	II-RBF	12,250	0,999	0,991	4,120	1,617	-0,139	0,542

em que : γ é o parâmetro ajustável da função *kernel* RBF; II-RBF: combinação da configuração regressão tipo II e função *kernel* Gaussiana Radial Basis function

Apesar das estimativas de generalizações serem aceitas pelo teste de aplicação proposto, algumas MVS utilizando 2 variáveis de entrada apresentaram Bias considerável para parcelas de menores volumes de madeira como mostra a Figura 18. Entretanto, os treinamentos com as bandas AV-3 e AV-4 quando combinadas com um variável de polarização cruzada (L_{HV} ou L_{VH}) foram capazes de gerar estimativas com menores Bias para as parcelas com menores volumes de madeira.

A análise gráfica nesta etapa da avaliação se mostrou essencial para a escolha de MVS que gerassem estimativas com menores Bias para parcelas com menores volumes de madeira, ainda que o teste de validação e aplicação acusassem a viabilidade de uso das MVS rejeitadas.

No presente trabalho, as MVS mais simples capazes de gerar estimativas consistentes do volume de povoamentos de eucalipto, foram aquelas que utilizaram a MVS regressão tipo II, função *kernel* RBF, 95 parcelas para treinamento, e duas variáveis de entrada, combinando as bandas AV3 ou AV4 do sensor AVNIR-2 com polarizações cruzadas (L_{HV} ou L_{VH}) do PALSAR. O melhor resultado, com a combinação AV4,L_{VH}, apresentou um $r_{\hat{y}y}$ de 0,991 e RQEM de 4,120 m³, equivalente a 1,617% da média das parcelas testadas.

Tabela 12 – Teste de validação e aplicação para as melhores e mais simples MVS ajustadas, com duas variáveis de entrada para a estimativa de volume de madeira de povoamentos de eucalipto.

Variáveis de Entrada	MSV (γ)	Treinamento				Generalização			
		$r_{\hat{y}y}$	F (H ₀)	Erro(%)	Situação	$r_{\hat{y}y}$	F (H ₀)	Erro(%)	Situação
AV1,L _{HH}	8,750	0,982	4,171*	0,001	Aceitável	0,927	20,014*	0,005	Aceitável
AV1,L _{HV}	10,250	0,987	3,227 ^{ns}	0,001	Aceitável	0,955	10,411*	0,001	Aceitável
AV1,L _{VV}	15,750	0,997	2,984 ^{ns}	0,001	Aceitável	0,947	15,045*	0,004	Aceitável
AV1,L _{VH}	18,000	0,988	3,124 ^{ns}	0,001	Aceitável	0,967	8,556 ^{ns}	0,001	Aceitável
AV2,L _{HH}	18,000	0,974	6,811*	0,001	Aceitável	0,949	11,277*	0,001	Aceitável
AV2,L _{HV}	20,000	0,991	3,001 ^{ns}	0,001	Aceitável	0,982	5,524 ^{ns}	0,001	Aceitável
AV2,L _{VV}	15,250	0,987	3,068 ^{ns}	0,001	Aceitável	0,938	18,541*	0,004	Aceitável
AV2,L _{VH}	20,000	0,990	2,886 ^{ns}	0,001	Aceitável	0,981	5,778 ^{ns}	0,001	Aceitável
AV3,L _{HH}	10,150	0,992	2,072 ^{ns}	0,001	Aceitável	0,941	16,193*	0,005	Aceitável
AV3,L _{HV}	25,000	0,998	1,278 ^{ns}	0,001	Aceitável	0,966	9,686 ^{ns}	0,004	Aceitável
AV3,L _{VV}	25,500	0,988	3,025 ^{ns}	0,001	Aceitável	0,945	14,527*	0,003	Aceitável
AV3,L _{VH}	10,000	0,988	3,337 ^{ns}	0,001	Aceitável	0,957	10,499*	0,001	Aceitável
AV4,L _{HH}	15,000	0,999	1,145 ^{ns}	0,001	Aceitável	0,978	7,548 ^{ns}	0,002	Aceitável
AV4,L _{HV}	17,250	0,999	1,002 ^{ns}	0,001	Aceitável	0,990	4,151 ^{ns}	0,001	Aceitável
AV4,L _{VV}	17,500	0,991	3,021 ^{ns}	0,001	Aceitável	0,963	7,047*	0,002	Aceitável
AV4,L _{VH}	12,250	0,999	1,003 ^{ns}	0,001	Aceitável	0,991	4,588 ^{ns}	0,001	Aceitável

em que: γ é o parâmetro ajustável da função *kernel* RBF; * Teste F(H₀) significativo 5% de probabilidade; ^{ns} Teste F(H₀) não significativo a 5% de probabilidade.

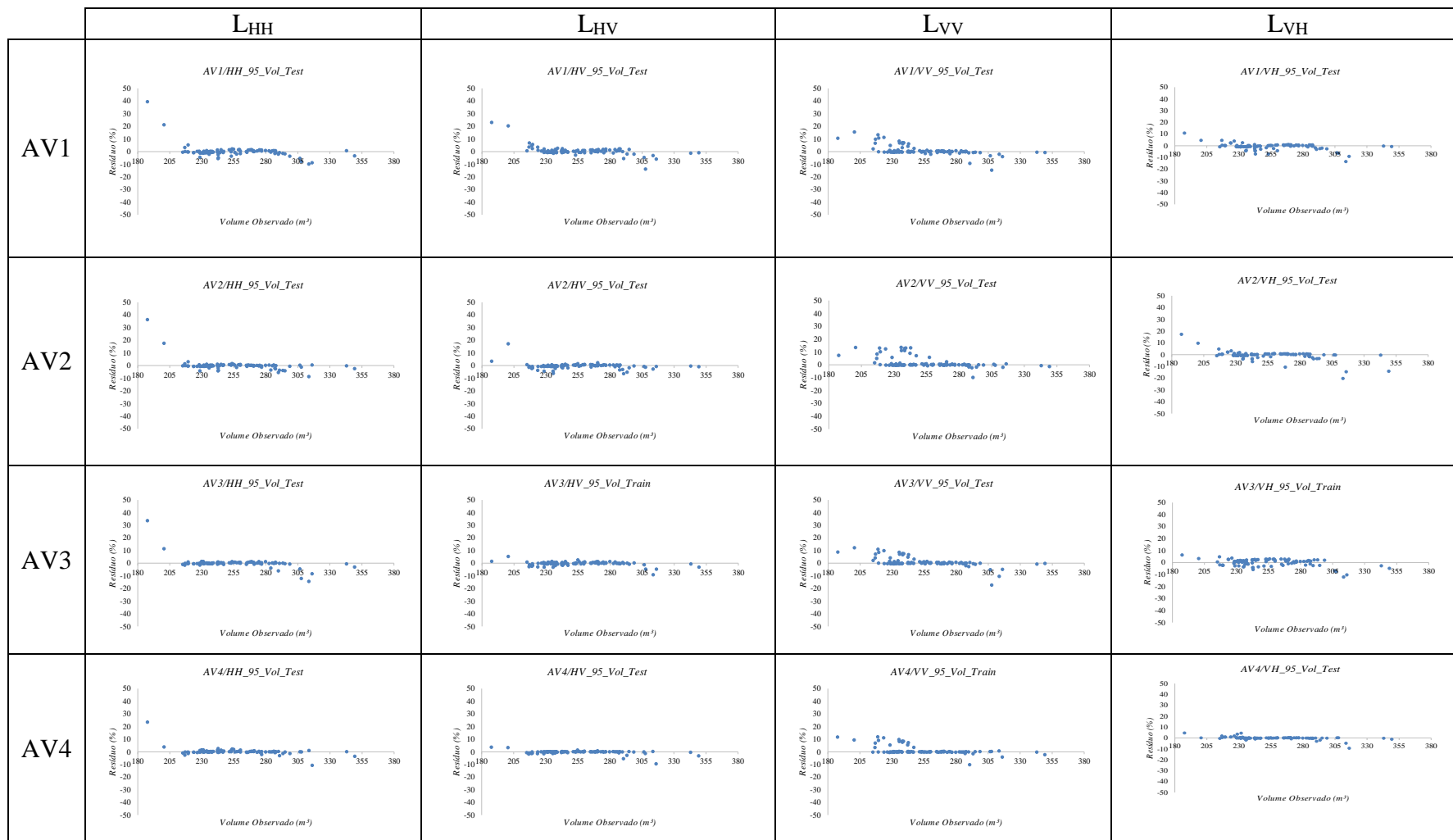


Figura 18 - Gráficos de resíduos (%) das generalizações das MVS para as predições de volume de madeira (m³) utilizando a combinação de dados dos sensores AVNIR-2 e PALSAR, com 95 parcelas amostrais para treinamento.

4.2 Treinamento das MVS para estimativa de biomassa aérea

4.2.1 Treinamento das MVS com 75% dos dados

Todos os procedimentos de análise para as predições de volume de madeira foram igualmente aplicados para as predições de biomassa aérea. Nesta etapa inicial, foram treinadas cerca de 500 MVS utilizando 75 % das unidades amostrais (142 parcelas) para estimar a biomassa aérea de povoamentos de eucalipto. Os 25% restantes foram utilizados para o teste de generalização de cada MVS. Avaliou-se a performance de diferentes configurações e variáveis de entrada, a fim de se encontrar os melhores desempenhos. Para uma mesma configuração, ajustou-se os parâmetros das MVS até retornarem os maiores índices de correlação e menor erro de generalização.

Na Tabela 13, são apresentadas as características das MVS com melhores performances segundo o tipo de regressão e função *kernel* para predição de biomassa aérea a partir do sensor AVNIR-2. Verificou-se que as MVS que obtiveram melhores desempenhos foram as que utilizaram a regressão tipo II e função *kernel* RBF e Polinomial, com maior destaque para a função RBF que apresentou maior coeficiente de correlação.

Tabela 13 – Configurações das MSV selecionadas para cada tipo de regressão e função

Variáveis de Entrada	Reg.	Vetores		Função <i>Kernel</i>				r_{yy}			RQEM (t/ha)	
		M	S	Tipo	Grau	Gama	Coeficiente	(total)	(treino)	(teste)	(treino)	(teste)
AVNIR-2	I	17	13	Lin	-	-	-	0,505	0,540	0,387	15,119	16,078
	II	18	13		-	-	-	0,505	0,539	0,387	15,135	16,089
AVNIR-2	I	105	96	Pol	2,000	0,300	2,000	0,677	0,718	0,522	12,053	13,870
	II	83	69		3,000	0,300	1,000	0,727	0,769	0,563	11,065	13,202
AVNIR-2	I	73	61	RBF	-	0,900	-	0,821	0,868	0,745	8,705	9,135
	II	60	32		-	1,100	-	0,907	0,949	0,783	5,737	8,151
AVNIR-2	I	114	113	Sig	-	0,050	1,000	0,503	0,533	0,404	14,708	14,470
	II	79	75		-	0,050	1,000	0,490	0,519	0,390	14,809	14,599

kernel para a predição de biomassa aérea (t/ha) de povoamentos de eucalipto utilizando dados do sensor AVNIR-2.

Onde: Reg I = ϵ -SVM; Reg II = ν -SVM; Lin = Linear, Pol = Polinomial, RBF = Gaussian Radial Basis Function, Sig = Sigmoide. AVNIR-2: bandas multispectrais AV1, AV2, AV3 e AV4; RQEM = Raíz Quadrada do Erro Quadrático Médio; Vetores M: margem máxima; Vetores S: margem suave.

Utilizando as polarizações do sensor PALSAR, os desempenhos nas predições de biomassa apresentaram fraca correlação, assim como os encontrados para volume (Tabela 14). Os valores de r_{yy} para as generalizações também foram maiores do que para os

treinamentos, mostrando uma distribuição não uniforme dos dados do sensor PALSAR. Os índices de correlação entre as funções *kernels* e tipo de regressão pouco variaram, mostrando mais uma vez pouca influência da configuração no resultado dos ajustes usando este sensor. A configuração com melhor performance foi utilizando regressão tipo II e função *kernel* RBF.

Tabela 14 – Configurações das MSV selecionadas para cada tipo de regressão e função *kernel* para a predição de biomassa aérea (t/ha) de povoamentos de eucalipto

Variáveis de Entrada	Reg.	Vetores		Função <i>Kernel</i>				r_{yy}			RQEM (t/ha)	
		M	S	Tipo	Grau	Gama	Coefficiente	(total)	(treino)	(teste)	(treino)	(teste)
PALSAR	I	108	100	Lin	-	-	-	0,706	0,723	0,665	12,032	11,848
	II	81	73		-	-	-	0,703	0,721	0,659	12,033	11,991
PALSAR	I	74	68	Pol	3	0,05	1	0,693	0,702	0,686	12,355	11,634
	II	65	59		3	0,03	1	0,692	0,701	0,686	12,326	11,803
PALSAR	I	93	29	RBF	-	0,10	-	0,707	0,691	0,700	12,037	11,280
	II	81	43		-	0,1	-	0,707	0,728	0,701	12,028	8,884
PALSAR	I	102	29	Sig	-	0,05	1	0,696	0,706	0,686	12,219	12,057
	II	105	103		-	0,07	1	0,695	0,704	0,691	12,257	12,001

utilizando dados do sensor PALSAR.

em que: Reg I = ϵ -SVM; Reg II = ν -SVM; Lin = Linear, Pol = Polinomial, RBF = Gaussian Radial Basis Function, Sig = Sigmoide; PALSAR remete as polarizações L_{HH} , L_{VV} , L_{HV} , L_{VH} ; RQEM = Raíz Quadrada do Erro Quadrático Médio; Vetores M: margem máxima; Vetores S: margem suave.

Utilizando simultaneamente os dados AVNIR-2 e PALSAR como variáveis de entrada, as MVS com a função RBF apresentaram as melhores estimativas com relação aos dados separados (Tabela 15), com coeficientes de correlação acima de 0,900. A função *kernel* linear e a sigmoide apresentaram os piores resultados de ajuste. A Regressão tipo II apresentou em geral melhor performance em relação a regressão tipo I.

Tabela 15 – Configurações das MSV selecionadas para cada tipo de regressão e função *kernel* para a predição de biomassa aérea (t/ha) de povoamentos de eucalipto utilizando dados dos sensores AVNIR-2/PALSAR.

Variáveis de Entrada	Reg.	Vetores		Função <i>Kernel</i>				r_{yy}			RQEM (t/ha)	
		M	S	Tipo	Grau	Gama	Coefficiente	(total)	(treino)	(teste)	(treino)	(teste)
AV-2 + PALSAR	I	53	37	Lin	-	-	-	0,793	0,792	0,806	10,584	9,607
	II	20	10		-	-	-	0,772	0,771	0,782	10,991	10,023
AV-2 + PALSAR	I	100	83	Pol	2,000	0,100	1,000	0,813	0,817	0,807	9,980	9,591
	II	68	57		2,000	0,040	1,000	0,800	0,800	0,807	10,378	9,743
AV-2 + PALSAR	I	111	59	RBF	-	2,050	-	0,948	0,966	0,901	5,399	6,763
	II	102	50		-	2,250	-	0,961	0,981	0,922	3,101	5,670
AV-2 + PALSAR	I	51	44	Sig	-	0,070	0,000	0,788	0,786	0,805	10,828	9,738
	II	80	73		-	0,100	0,000	0,782	0,780	0,801	10,829	9,799

em que: Reg I = ϵ -SVM; Reg II = ν -SVM; Lin = Linear, Pol = Polinomial, RBF = Gaussian Radial Basis Function, Sig = Sigmoide. AV-2 = bandas multiespectrais ópticas do sensor AVNIR-2, são elas: AV1, AV2, AV3 e AV4; PALSAR remete as polarizações L_{HH} , L_{VV} , L_{HV} , L_{VH} ; RQEM = Raíz Quadrada do Erro Quadrático Médio; Vetores M: margem máxima; Vetores S: margem suave.

Assim, a melhor configuração de MVS para estimar biomassa aérea de povoamentos de eucalipto, mais uma vez, consistiu no uso da regressão tipo II e função *kernel* RBF.

4.2.1.1 Teste de validação e aplicação

Procedeu-se a análise da qualidade de ajuste das MVS com os melhores desempenhos, com posterior teste de validação e aplicação. Pela Tabela 16, observa-se que a integração dos dados permitiu uma correlação mais forte das estimativas de generalização, além de menores valores residuais e Bias. A utilização somente das quatro polarizações do sensor PALSAR, gerou um Bias de 1,229 t/ha o que corresponde a 0,966% da média das parcelas testadas, além de maior RQEM e MDA. Com resultados semelhantes a predição de volume, as bandas multiespectrais do AVNIR-2 apresentaram generalização com desempenho também intermediário, ainda que com o menor Bias.

Tabela 16 – Análise da qualidade dos ajustes dos modelos MVS treinados para estimativa de biomassa aérea de povoamentos de eucalipto a partir dos sensores AVNIR-2 e PALSAR.

Variável (Output)	Sensores (Input)*	MVS	r_{yy} (teste)	RQEM		Bias		MDA	
				(t/ha)	(%)	(t/ha)	(%)	(t/ha)	(%)
Biomassa	AV-2	II - RBF	0,783	8,151	6,406	-0,774	-0,608	6,531	5,133
	PAL	II - RBF	0,701	8,884	6,982	1,229	0,966	9,018	7,088
Aérea	AV-2/PAL	II - RBF	0,922	5,670	4,456	-0,979	-0,769	4,346	3,416

*AV-2: AVNIR-2 (AV1, AV2, AV3, AV4); PAL: PALSAR (L_{HH} , L_{VV} , L_{HV} , L_{VH}); RBF = Função *Kernel* Gaussiana Radial Basis Function.

A Tabela 17 mostra os teste de validação e aplicação, onde verifica-se que somente o treinamento não aceitável foi o que utilizou o sensor PALSAR. O treinamento com dados do sensor AVNIR-2 foi considerado válido, entretanto sua generalização não foi aceita. Somente a integração das variáveis de entrada apresentou treino válido, e aceitável generalização das estimativas para biomassa aérea. A generalização aceitável pelo teste proposto apresentou RQEM de 5,670 t/ha o que representa 4,456% da média de biomassa aérea das parcelas testadas. Da mesma forma, valores relativos de MDA para essa generalização foi de 3,416% (Tabela 16).

Tabela 17 - Teste de validação e aplicação das MVS selecionadas com as diferentes variáveis de entrada na estimativa de biomassa aérea (t/ha) de povoamentos de eucalipto.

	Estatística	AVNIR-2	PALSAR	AVNIR-2+PALSAR
Treinamento				
	F (H ₀)	20,189*	242,421*	7,9897*
	r _{ŷŷ}	0,949	0,728	0,981
	Erro(%)	0,008	0,003	0,001
	Situação	Aceitável	Inaceitável	Aceitável
Generalização				
	F (H ₀)	29,462*	87,772*	24,981*
	r _{ŷŷ}	0,783	0,701	0,922
	Erro(%)	0,207	1,964	0,002
	Situação	Inaceitável	Inaceitável	Aceitável

* Teste F(H₀) significativo 5% de probabilidade; ^{ns} Teste F(H₀) não significativo a 5% de probabilidade.

A Figura 19 apresenta os gráficos residuais das generalizações das MVS treinadas sem e com a integração das variáveis de entradas para estimar biomassa aérea. Utilizando o sensor AVNIR-2 (Figuras 19a, 19d e 19g), a generalização apresentou claras tendências para as parcelas com os menores valores de biomassa aérea, com como alguns pontos com os maiores valores. A utilização do sensor PALSAR (Figuras 19b, 19e e 19h) apresentou fortes tendências positivas para as parcelas com os menores valores e tendências negativas para os maiores valores. As Figuras 19f, 19c e 19i, mostram que a integração dos sensores gerou estimativas com valores residuais mais uniformes para todas as parcelas testadas, corroborando com o teste de aplicação (Tabela 17).

Dessa forma, pela análise de qualidade de ajuste, análise gráfica dos resíduos e teste de validação e aplicação, as estimativas de biomassa aérea foram consistentes com uma MVS com 142 parcelas amostrais, configurada com regressão tipo II, função *kernel* RBF, utilizando-se como variáveis de entrada a integração dos 8 dados fornecidos pelos sensores AVNIR-2 e PALSAR.

4.2.2 Redução do número de amostras para o treinamento das MVS

Com a redução do número de parcelas amostrais para treinamento, buscou-se verificar se a estimação de biomassa aérea de povoamento eucalipto poderia ser consistente com um número menor de parcelas inventariadas. Utilizando-se a SMV tipo II e a função RBF como função *kernel*, treinou-se MVS ajustando-se o parâmetro γ para os diferentes dados de entrada e número de parcelas para treinamento, como mostra a Tabela 18. A redução do número de amostras para treinamento, com base no total de parcelas (190), consistiu em: 60% (114 parcelas), 50% (95 parcelas), 40% (76 parcelas),

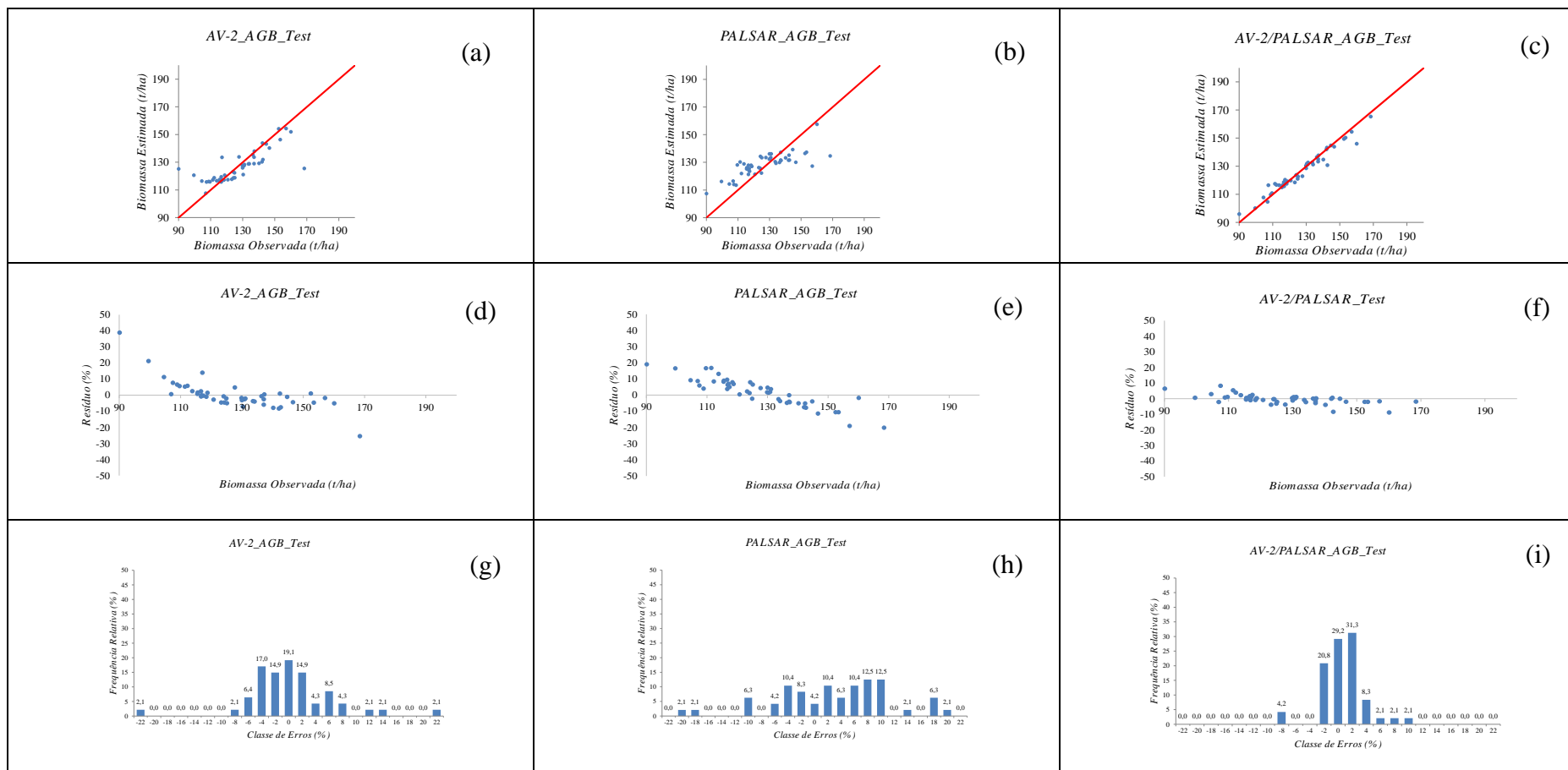


Figura 19 – Gráficos de resíduos das generalizações das MVS para as predições de biomassa aérea (t/ha) utilizando dados dos sensores AVNIR-2 e PALSAR. As Figuras 19a, 19b e 19c representam os gráficos da relação dos valores observados e estimados pelas MVS, utilizando como dados de entrada, respectivamente, os dados dos sensores: AVNIR-2, PALSAR, AVNIR-2+PALSAR. As Figuras 19d, 19e e 19f representam os gráficos de dispersão dos resíduos percentuais das estimativas das MVS utilizando como dados de entrada, respectivamente, os dados dos sensores: AVNIR-2, PALSAR, AVNIR-2+PALSAR. As Figuras 19g, 19h e 19i representam os gráficos de frequência de classes de erros das estimativas das MVS utilizando como dados de entrada, respectivamente, os dados dos sensores: AVNIR-2, PALSAR, AVNIR-2+PALSAR

25% (47 parcelas), 10% (19 parcelas).

Pela Tabela 18 verifica-se que com a redução do número de parcelas amostrais para o treinamento das MVS, os melhores desempenhos continuaram sendo com a utilização simultânea dos dados dos sensores AVNIR-2 e PALSAR como variáveis de entrada, assim como observado nos treinamentos anteriores para volume e também biomassa com 142 parcelas.

Tabela 18 – Melhores MVS selecionadas para avaliar a estimativa de biomassa aérea (t/ha) de povoamentos de eucalipto, com a redução do número de parcelas amostrais para treinamento.

Número de Amostras	Variáveis de Entrada	MSV		Vet.	Treinamento		Generalização	
		<i>Kernel</i>	γ		r_{yy}	RQEM(t/ha)	r_{yy}	RQEM(t/ha)
142	AVNIR-2	II - RBF	1,100	60 (32)	0,949	5,737	0,783	8,151
142	PALSAR	II - RBF	0,070	65 (59)	0,728	12,028	0,701	8,884
142	AV+P	II - RBF	2,250	115(18)	0,981	3,101	0,922	5,670
114	AVNIR-2	II - RBF	0,800	72 (64)	0,900	7,122	0,788	8,541
114	PALSAR	II - RBF	1,350	87 (76)	0,778	8,973	0,705	9,056
114	AV+P	II - RBF	2,20	52 (25)	0,988	2,055	0,904	5,502
95	AVNIR-2	II - RBF	0,250	60 (51)	0,902	7,025	0,741	9,014
95	PALSAR	II - RBF	1,450	50 (42)	0,764	9,110	0,644	10,322
95	AV+P	II - RBF	2,050	32(16)	0,992	1,203	0,901	5,851
76	AVNIR-2	II - RBF	0,250	46 (39)	0,886	7,044	0,678	9,927
76	PALSAR	II - RBF	1,550	40 (33)	0,747	8,974	0,521	11,129
76	AV+P	II - RBF	2,250	25 (13)	0,992	2,887	0,860	6,421
47	AVNIR-2	II - RBF	0,250	23 (13)	0,822	7,155	0,622	10,077
47	PALSAR	II - RBF	1,850	10 (7)	0,612	10,430	0,376	15,499
47	AV+P	II - RBF	1,850	30(16)	0,997	1,556	0,820	7,358
19	AVNIR-2	II - RBF	0,250	9 (2)	0,743	8,758	0,351	19,876
19	PALSAR	II - RBF	1,300	14 (8)	0,884	8,001	0,134	24,110
19	AV+P	II - RBF	2,200	10(4)	0,974	3,911	0,466	11,581

em que: AV+P: combinação dos dados dos sensores AVNIR-2 e PALSAR; II-RBF: combinação da configuração regressão tipo II e função *kernel* Gaussiana Radial Basis function. Vet.: Vetores de margem máxima e margem suave (parênteses).

4.2.2.1 Teste de validação e aplicação

Procedeu-se a análise das MVS com a integração das variáveis de entrada e com redução da amostragem. A Tabela 19, mostra as estatísticas de qualidade de ajuste das generalizações das MVS treinadas com a redução do número de parcelas amostrais para treinamento. Assim como ocorrido nos treinamentos para estimativas de volume de madeira, conforme o número de amostras foi reduzido, diminuiu-se também o número de vetores responsáveis por determinar as margens máxima e suave, o que leva a uma

generalização com menor desempenho. De modo geral, todos os treinamentos apresentaram forte correlação, mesmo com um menor número de amostras, porém o desempenho das generalizações caiu, com o número de amostras reduzido.

Tabela 19 – Melhores MVS selecionadas para estimar biomassa aérea de povoamentos de eucalipto a partir de dados dos sensores AVNIR-2 e PALSAR, com a redução do número de parcelas amostrais para treinamento.

N	MVS	r_{yy}	r_{yy}	Vet.	RQEM (teste)		Bias (teste)		MDA (teste)	
		(treino)	(teste)		(t/ha)	(%)	(t/ha)	(%)	(t/ha)	(%)
142	II-RBF	0,981	0,922	102 (50)	5,670	4,456	-0,979	-0,769	4,346	3,416
114	II-RBF	0,988	0,904	54 (27)	5,502	4,343	0,045	0,036	4,915	0,036
95	II-RBF	0,992	0,901	45 (4)	5,851	4,566	-0,394	-0,307	5,201	4,059
76	II-RBF	0,992	0,860	47 (21)	6,421	5,024	0,553	0,432	5,403	4,227
47	II-RBF	0,997	0,820	41 (22)	7,358	5,879	1,155	0,901	6,969	5,435
19	II-RBF	0,974	0,466	10 (0)	11,581	9,039	2,093	1,633	11,784	9,197

em que: N equivale ao número de parcelas amostrais para treinamento das MVS. II-RBF é a combinação da configuração regressão tipo II e função *kernel* Gaussiana Radial Basis function. Vet.: Vetores de margem máxima e margem suave, entre parênteses.

A RQEM, o Bias e o MDA apresentaram resultados fora do esperado para os treinamentos com 114 parcelas, se considerado a suficiência amostral para explicar o comportamento dos dados. O treinamento com 114 parcelas apresentou os menores valores residuais em comparação com 142 e 95 parcelas. Este resultado ressalta a atenção que deve ser tomada na busca exaustiva por um valor ótimo do parâmetro ajustável pelo usuário, no caso, o parâmetro γ da função *kernel*.

A Tabela 20 mostra a eficiência do ajuste das MVS, com índices de correlação alto e treinamento aceitável para todas as reduções do número de parcelas para estimar biomassa aérea. Resultado similar ao encontrado para estimar volume do povoamento. Porém, mesmo validados, nem todos os treinos foram considerados aplicáveis. Somente os treinamentos com número mínimo de 95 parcelas, também, foram aceitáveis para generalização dos dados, isto é, aceitáveis para aplicação. Estas generalizações aceitas pelo teste, apresentaram RQEM menor que 5,851 t/ha (4,566% da média) e MDA abaixo de 4,059% da média das parcelas testadas (Tabela 19). O Bias pode ser melhor analisado com os gráficos residuais da figura a seguir.

A Figura 20 mostra os gráficos residuais das generalizações das MVS, reduzindo-se o número de amostras de treinamento para estimar biomassa aérea de povoamentos de eucalipto. Por esta figura, nota-se que os resíduos apresentam uma dispersão maior do que as estimativas de volume. Reduzindo-se a 47 parcelas amostrais (Figuras 20m, 20n e

20o) os resíduos passam a apresentarem maior amplitude, e maior tendência para as estimativas com menores e maiores valores, fenômeno que passa ser considerável com 19 parcelas (Figuras 20p e 20q).

Tabela 20 - Teste de validação e aplicação das melhores MVS selecionadas, com a redução do número de parcelas amostrais para treinamento, na estimativa de biomassa aérea de povoamentos de eucalipto.

Est.	142	114	95	76	47	19
Treino						
F(H ₀)	0,9348 ^{ns}	0,767 ^{ns}	0,045 ^{ns}	4,269*	2,258*	19,338*
r _{yŷ}	0,981	0,988	0,992	0,992	0,997	0,974
Erro(%)	0,001	0,002	0,001	0,003	0,004	0,009
Situação	Aceitável	Aceitável	Aceitável	Aceitável	Aceitável	Aceitável
Generaliz.						
F(H ₀)	34,362*	12,927*	9,332*	5,9087*	62,160*	749,491
r _{yŷ}	0,922	0,904	0,901	0,860	0,820	0,466
Erro(%)	-0,005	0,364	-0,038	0,646	1,548	3,037
Situação	Aceitável	Aceitável	Aceitável	Inaceitável	Inaceitável	Inaceitável

* Teste F(H₀) significativo 5% de probabilidade; ^{ns} Teste F(H₀) não significativo a 5% de probabilidade.

Assim como a análise para estimativas de volume de madeira, o número mínimo de 95 parcelas amostrais foi o suficiente para que as MVS gerassem estimativas consistentes de biomassa aérea de povoamentos de eucalipto, configuradas com a regressão tipo II, função *kernel* RBF e a combinação de todos os dados dos sensores AVNIR-2 e PALSAR como variáveis de entrada.

4.2.3 Redução do número de variáveis de entrada

Com base nos resultados já encontrados até aqui para predição de biomassa aérea, com relação as melhores configurações e o mínimo de parcelas para treinamento, testou-se as MVS com o mínimo de variáveis de entrada. O treino das MVS consistiu na configuração MVS regressão tipo II, função RBF como *kernel*, usando a combinação dos sensores e 95 parcelas amostrais tanto para treino quanto para teste de generalização dos ajustes.

Os treinamentos e generalizações com o menor número possível de variáveis de entrada que retornassem ajuste ótimo são apresentados na Tabela 21, bem como suas estatísticas de ajuste.

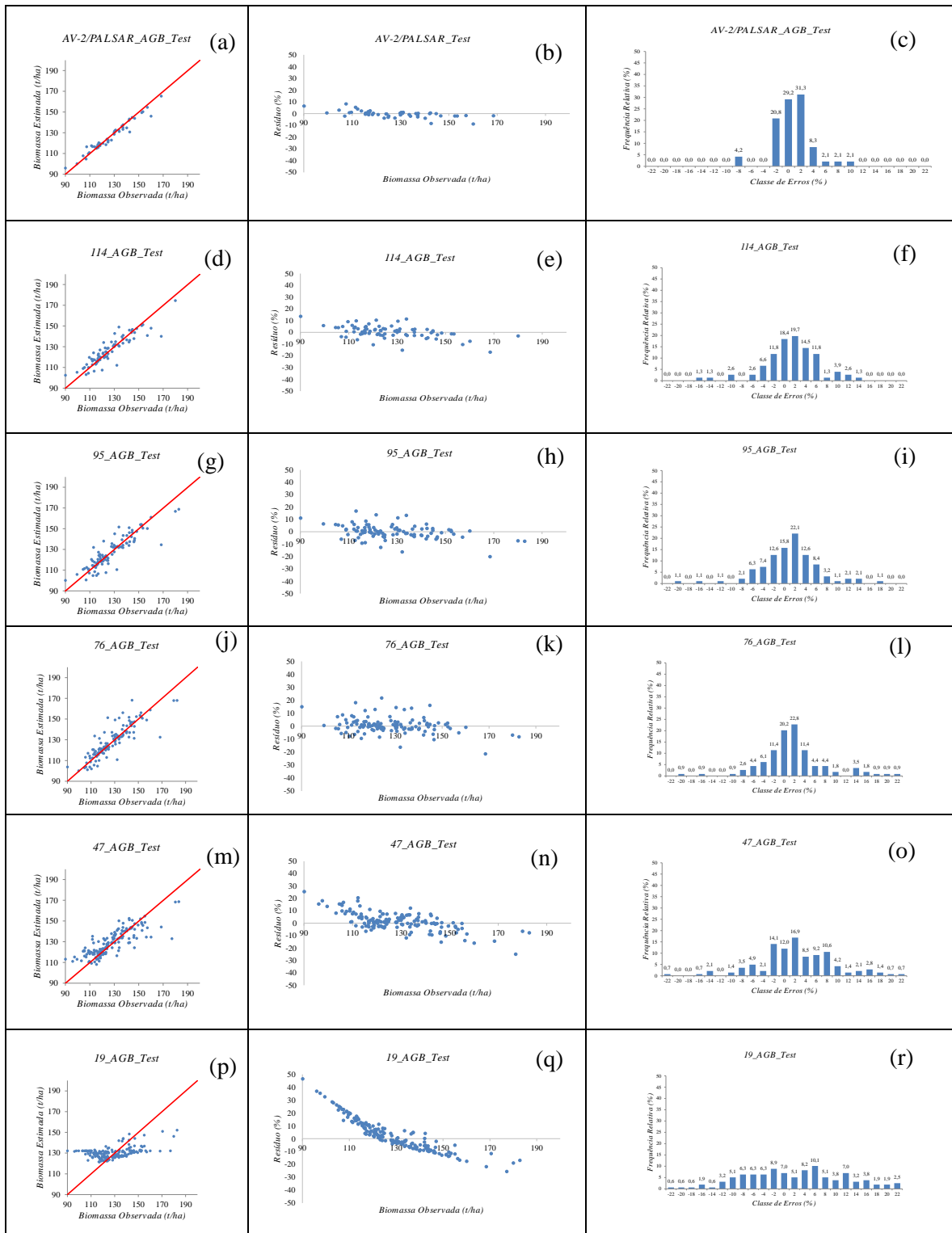


Figura 20 – Gráficos de resíduos das generalizações das MVS para as previsões de biomassa aérea utilizando dados dos sensores AVNIR-2 e PALSAR com a redução do número de parcelas amostrais para treinamento. As Figuras 20a, 20d, 20g, 20j, 20m e 20p representam os gráficos da relação dos valores observados e estimados pelas MVS treinadas, respectivamente, com: 142, 114, 95, 76, 47 e 19 parcelas. As Figuras 20b, 20e, 20h, 20k, 20n e 20q, representam os gráficos de dispersão dos resíduos percentuais das estimativas das MVS treinadas, respectivamente, com: 142, 114, 95, 76, 47 e 19 parcelas. As Figuras 20c, 20f, 20i, 20l, 20o e 20r representam os gráficos de frequência de classes de erros das estimativas das MVS treinadas, respectivamente, com: 142, 114, 95, 76, 47 e 19 parcelas.

Tabela 21 – MVS com duas variáveis combinadas dos sensores AVNIR-2 e PALSAR, utilizando 95 parcelas amostrais para treinamento, para estimativas de biomassa aérea de povoamentos de eucalipto

Variáveis de Entrada	MVS		$r_{\hat{y}y}$	$r_{y\hat{y}}$	RQEM		Bias (teste)	MDA (teste)
	Kernel	(γ)			(teste)			
			(t/ha)	(%)	(%)	(%)		
AV1,L _{HH}	II-RBF	15,000	0,901	0,833	7,037	2,766	-1,624	3,544
AV1,L _{HV}	II-RBF	10,000	0,976	0,938	5,954	2,337	-0,103	2,982
AV1,L _{VV}	II-RBF	18,000	0,985	0,941	6,054	2,376	-0,102	3,209
AV1,L _{VH}	II-RBF	10,000	0,952	0,941	6,027	2,366	-1,310	3,270
AV2,L _{HH}	II-RBF	10,000	0,950	0,921	6,506	2,554	-0,073	3,252
AV2,L _{HV}	II-RBF	8,000	0,970	0,960	4,894	1,921	0,241	2,876
AV2,L _{VV}	II-RBF	10,000	0,973	0,942	6,077	2,385	1,286	3,542
AV2,L _{VH}	II-RBF	9,750	0,966	0,959	5,020	1,971	-0,429	2,761
AV3,L _{HH}	II-RBF	15,000	0,971	0,910	6,912	2,713	-0,152	3,523
AV3,L _{HV}	II-RBF	18,000	0,982	0,920	6,660	2,614	-0,840	3,230
AV3,L _{VV}	II-RBF	10,000	0,971	0,927	6,660	2,614	0,146	3,362
AV3,L _{VH}	II-RBF	40,000	0,982	0,940	5,892	2,313	-0,263	2,937
AV4,L _{HH}	II-RBF	17,000	0,979	0,942	5,622	2,207	0,332	2,949
AV4,L _{HV}	II-RBF	14,250	0,980	0,949	5,312	2,085	-0,424	2,859
AV4,L _{VV}	II-RBF	16,000	0,979	0,951	5,338	2,095	0,201	3,068
AV4,L _{VH}	II-RBF	21,000	0,982	0,944	5,340	2,098	-0,195	3,073

Onde: γ é o parâmetro ajustável da função *kernel* RBF; * Teste F(H₀) significativo 5% de probabilidade; ^{ns} Teste F(H₀) não significativo a 5% de probabilidade. II-RBF: combinação da configuração regressão tipo II e função *kernel* Gaussiana *Radial Basis function*.

Com duas variáveis de entrada, os treinamentos apresentaram bom desempenho retornando estimativas com alta correlação. A RQEM foi menor que 3%, os maiores valores absolutos de Bias relativo não ultrapassaram 2% e o MDA foi abaixo de 3,5% da média das parcelas testadas.

4.2.3.1 Teste de validação e aplicação

A Tabela 22 mostra que segundo o teste de validação e aplicação de Guimarães (1994) modificado, o treinamento com pelo menos uma banda de cada sensor resultou em MVS válidas e aplicáveis. Somente uma combinação não apresentou o mesmo resultado, a banda AV1 com a polarização L_{HH}. Como foi verificado para as predições de volume, tanto os treinos quanto as generalizações das MVS com um menor número de variáveis

de entrada resultaram em estimativas com maior correlação, comparadas ao uso das 8 variáveis integradas dos sensores AVNIR-2 e PALSAR.

Tabela 22 – Teste de validação e aplicação para as melhores e mais simples MVS ajustadas, com duas variáveis de entrada para a estimativa de biomassa aérea de povoamentos de eucalipto.

Variáveis de Entrada	MSV (γ)	Treinamento				Generalização			
		r_{yy}	F (H_0)	Erro %	Situação	r_{yy}	F (H_0)	Erro %	Situação
AV1,L _{HH}	15,000	0,901	9,881*	0,008	Aceitável	0,833	44,784*	0,010	Inaceitável
AV1,L _{HV}	10,000	0,976	9,240*	0,001	Aceitável	0,938	28,307*	0,002	Aceitável
AV1,L _{VV}	18,000	0,985	4,706*	0,001	Aceitável	0,941	42,212*	0,003	Aceitável
AV1,L _{VH}	10,000	0,952	4,884*	0,008	Aceitável	0,941	33,327*	0,010	Aceitável
AV2,L _{HH}	10,000	0,950	7,542*	0,004	Aceitável	0,921	21,825*	0,003	Aceitável
AV2,L _{HV}	8,000	0,970	6,619*	0,002	Aceitável	0,960	26,611*	0,005	Aceitável
AV2,L _{VV}	10,000	0,973	7,657*	0,003	Aceitável	0,942	39,178*	0,016	Aceitável
AV2,L _{VH}	9,750	0,966	8,741*	0,001	Aceitável	0,959	30,324*	0,001	Aceitável
AV3,L _{HH}	15,000	0,971	4,491*	0,002	Aceitável	0,910	22,382*	0,002	Aceitável
AV3,L _{HV}	18,000	0,982	6,368*	0,001	Aceitável	0,920	24,715*	0,005	Aceitável
AV3,L _{VV}	10,000	0,971	5,375*	0,003	Aceitável	0,927	46,145*	0,006	Aceitável
AV3,L _{VH}	40,000	0,982	4,365*	0,002	Aceitável	0,940	30,590*	0,001	Aceitável
AV4,L _{HH}	17,000	0,979	4,510*	0,001	Aceitável	0,942	15,921*	0,006	Aceitável
AV4,L _{HV}	14,250	0,980	3,232*	0,001	Aceitável	0,949	17,198*	0,002	Aceitável
AV4,L _{VV}	16,000	0,979	4,958*	0,001	Aceitável	0,951	24,068*	0,002	Aceitável
AV4,L _{VH}	21,000	0,982	4,537*	0,001	Aceitável	0,944	25,138*	0,006	Aceitável

em que: γ é o parâmetro ajustável da função *kernel* RBF; * Teste F(H_0) significativo 5% de probabilidade; ^{ns} Teste F(H_0) não significativo a 5% de probabilidade.

Os gráficos residuais dos ajustes mostraram que apesar das estimativas de generalizações serem aceitas pelo teste de aplicação proposto, algumas MVS selecionadas utilizando 2 variáveis de entrada apresentaram viés ora para as parcelas com menor valor de biomassa, ora para as parcelas com maior valor (Figura 21). Entretanto, os treinamentos com as bandas AV-4 quando combinadas com variáveis das polarizações do tipo L_{HV} e L_{VV} apresentaram valores residuais mais uniformes para todas as parcelas testadas, gerando estimativas de generalização acuradas.

A análise gráfica, nesta etapa da avaliação, mais uma vez se mostrou essencial para a escolha de MVS que gerassem estimativas com menores Bias, ainda que o teste de validação e aplicação indicassem a viabilidade de aplicação.

Neste estudo, as MVS mais simples capazes de gerar estimativas com menores bias para biomassa aérea de povoamentos de eucalipto, foram aquelas que utilizaram a MVS regressão tipo II, função *kernel* RBF, 95 parcelas para treinamento, e duas variáveis de entrada, combinando a banda AV4 do sensor AVNIR-2 com a polarização L_{HV} ou L_{VV} do sensor PALSAR. O melhor desempenho para estimar biomassa aérea, com a combinação AV4,L_{VV}, obteve um r_{yy} de 0,951 e RQEM de 5,338 t.ha⁻¹, o equivalente a 2,095 % da média das parcelas testadas.

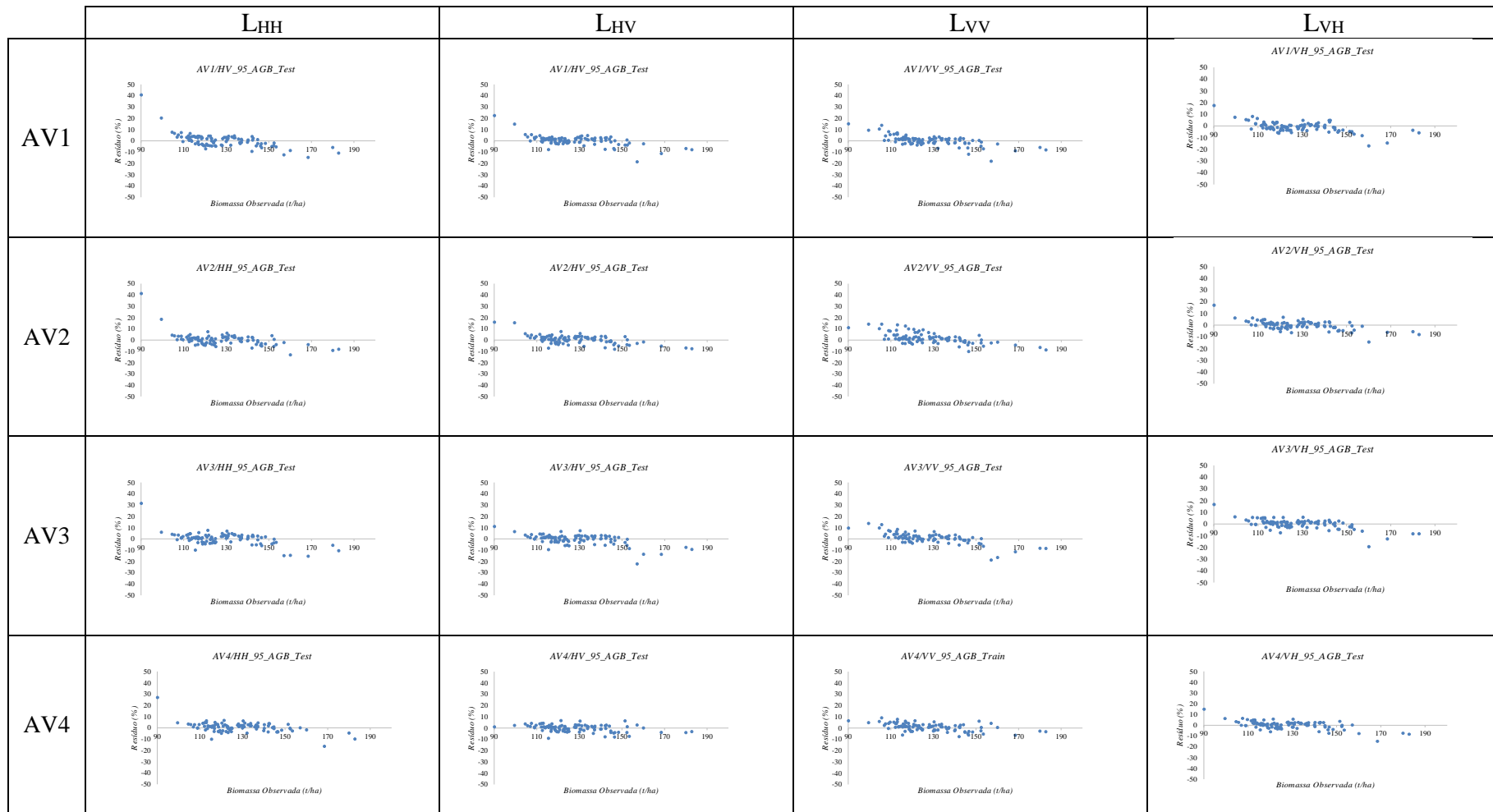


Figura 21 - Gráficos de resíduos (%) das generalizações das MVS para as predições de biomassa aérea utilizando a combinação de dados dos sensores AVNIR-2 e PALSAR, com 95 parcelas amostrais para treinamento.

5 DISCUSSÃO

O presente trabalho teve como objetivo avaliar o desempenho das MVS com foco na simplicidade dos ajustes mantendo os níveis de exatidão a fim de estimar o volume de madeira e biomassa aérea de parcelas inventariadas por meio de imagens provenientes dos sensores AVNIR-2 e PALSAR, ambos a bordo do satélite ALOS. Dados do sensor AVNIR-2 compreenderam 4 bandas multiespectrais da região óptica do espectro que puderam ser usadas como variáveis de entrada. Da mesma forma, dados de entrada do sensor PALSAR compreenderam 4 polarizações da banda L.

Os resultados do presente estudo mostraram que as MVS utilizando a combinação de dados ópticos e radarmétricos forneceram estimativas satisfatórias de volume e biomassa aérea para fins operacionais. As MVS alcançaram um $r_{\hat{y}y}$ de 0,991 para estimativas de volume, com um RQEM de 4,120 m³, equivalente a 1,617% da média das parcelas testadas. Já para biomassa aérea, as MVS alcançaram um $r_{\hat{y}y}$ de 0,951 e um RQEM de 5,338 t.ha⁻¹. Dube et al. (2015) combinou bandas ópticas do satélite SPOT-4 e dados polarimétricos do sensor ALOS-PALSAR para estimar volume de madeira de povoamentos de eucalipto na África do Sul. Com um número amostral superior ao do presente estudo e aplicando regressão múltipla, os autores chegaram a uma acurácia inferior, com r^2 de 0,83 e RQEM de 17,63%. Ismail et al. (2015) utilizando regressão múltipla, também encontrou boas estimativas de volume com a combinação de dados ópticos e de RADAR, porém foram consideradas insatisfatórias ($r^2 = 0,68$). O algoritmo utilizado no presente trabalho mostrou um aprimoramento na acurácia das estimativas quando comparado ao trabalho prévio realizado por Oliveira (2011). O estudo de Oliveira (2011) empregou a regressão *Stepwise*, alcançando um r^2 de 0,69 e 0,71 respectivamente para volume e biomassa aérea, ainda que com uma número de parcelas maior – 204. Trabalhos como Chen, Hay e Zhou (2010), Meng et al. (2016) e Shataee (2013) mostram a eficiência das MVS para estimativas de parâmetros biofísicos de povoamentos florestais naturais e mistos.

A combinação AV4,L_{HV} apresentou estimativas com alta correlação e precisão para as variáveis alvo - volume e biomassa aérea dos povoamentos florestais. Este fato pode ser explicado pela interação dos tipos de ondas eletromagnéticas com os parâmetros biofísicos do povoamento. Na região do infravermelho próximo (NIR) a reflectância pelas folhas é alta, assim também o dossel florestal (TUCKER; SELLERS, 1986). Logo, ondas

desta região conseguem detectar o fechamento de copa pela área foliar e fornecer informações dos parâmetros biofísicos da floresta como o volume e biomassa aérea (DAHLNERG, 2001; HEISKANEN, 2006; SPANNER et al., 1990). Alguns estudos como Gibbs et al. (2007) e Kronseder et al. (2012) reportam as limitações de se trabalhar com dados NIR como a saturação do sinal. No caso do presente estudo, utilizando dados NIR, foram observados dificuldades de estimativas mais precisas para as parcelas com menores valores de volume ou biomassa aérea. Uma alternativa para contornar esse problema é a utilização de dados de sensores ativos SAR, no caso deste trabalho ALOS-PALSAR (ATTARCHI; GLOAGUEN, 2014). A integração de dados aumentou a acurácia das estimativas, com destaque para a polarização L_{HV} . Alguns estudos como os de Harrell et al. (1997) e Behera et al. (2016) mostram uma maior sensibilidade de micro-ondas com polarização cruzada (L_{HV} ou L_{VH}) do que as co-polarizadas (L_{HH} ou L_{VV}), o que pode ser explicado pela interação das ondas com os componentes de mesma dimensão do comprimento de onda, e também componentes verticais (troncos) que estão diretamente relacionadas com o volume e biomassa aérea estocada (WOLTER; TOWNSEND, 2011).

Em contraste com os desempenhos com os dados NIR e polarizações cruzadas, a MVS treinada com a banda AV1 combinada com a polarização L_{HH} não apresentou desempenho satisfatório para estimar a biomassa aérea, evidenciando menor sensibilidade destes comprimentos de onda aos componentes do povoamento florestal. A mesma polarização, L_{HH} , foi a que gerou as piores estimativas no estudo de Behera et al. (2016) para povoamentos florestais de *Tectona grandis* e *Shorea robusta* na Índia.

As MVS treinadas somente com dados ópticos não conseguiram gerar estimativas consistentes. Ainda que com $r_{\hat{y}y} > 0,8$, a metodologia proposta refutou a aplicação destas MVS. O resultado apresentado para estimativas de volume foi superior ao encontrado por Trotter, Dymond e Goulding (1997), que utilizaram regressão múltipla, imagens Landsat-TM para estimar volume de madeira de povoamentos florestais de coníferas, e alcançaram, com 188 parcelas, um r^2 de 0,3. As limitações para estimar volume ou biomassa pelos sensores ópticos pode ser explicada pela interação das ondas eletromagnéticas que se limita a superfície do dossel, diminuindo a sensibilidade da interação com componentes verticais (WULDER, 1998). Bauermann (2008) conseguiu aumentar a acurácia e precisão das estimativas de volume de povoamentos de eucalipto, por meio de imagens de bandas ópticas do satélite CBERS, com a inclusão de variáveis texturais e idade do povoamento, chegando ao um r^2 de 0,99. O presente estudo não

utilizou operações de bandas e informações de inventário povoamento como *inputs*, mas como já discutido, alcançou também uma forte correlação das estimativas integrando os dados ópticos e dados de RADAR.

As informações das polarizações do PALSAR, por si só também, não foram suficientes para que as MVS gerassem estimativas consistentes. As predições de volume chegaram a um índice de correlação fraca de 0,730. Resultado semelhante fora encontrado por Amrutkar et al. (2012), que utilizando o mesmo sensor, estimaram o volume de madeira para povoamentos florestais mistos, incluindo povoamentos de eucalipto, com estimativas que não ultrapassaram um r^2 de 0,573. Amrutkar et al. (2012), estimou o volume de povoamentos florestais mistos, usando somente dados do sensor PALSAR e chegou a estimativas com correlação semelhante aos do presente estudo ($r_{\hat{y}y}$ de 0,786). Baghdadi et al. (2015) estimando biomassa aérea de povoamentos de eucalipto com o PALSAR, e aplicando a regressão RF, também apresentou estimativas com fraca correlação ($r_{\hat{y}y} < 0,700$).

É importante destacar também, que a baixa correlação entre as informações provenientes dos dados AVNIR-2 e PALSAR podem explicar os melhores resultados das estimativas, em vista de suas complementariedades (OLIVEIRA, 2011). Santoro, Eriksson e Fransson (2015) também encontraram resultado semelhante para povoamentos florestais naturais boreais na Suécia, onde a polarização L_{HV} foi a que apresentou estimativas mais acuradas.

As MVS configuradas pela regressão tipo II foram de modo geral superiores as MVS regressão tipo I. Isso porque o parâmetro v na regressão tipo II, torna a determinação de ϵ (margem máxima) menos subjetiva no processo de treinamento dos dados. Com relação as funções *kernel*, houve um claro destaque para os resultados com função RBF. Segundo Boswell (2002), não existe uma regra para a escolha da função *kernel*. A escolha deve se atentar na capacidade de transformações dos pontos para uma dimensão maior, garantindo a capacidade de generalização de novos dados.

O critério de escolha das melhores MVS com o teste de validação e aplicação, Guimarães (1994) modificado, foi utilizado devido a trabalhos prévios realizados com os dados do presente estudo. Entretanto a avaliação baseada por si só no teste proposto não foi suficiente para selecionar os melhores desempenhos. A análise gráfica foi primordial para a seleção de MVS que gerassem estimativas menos enviesadas, assim como proposto por Campos; Leite (2013).

Segundo Soares, Paula Neto e Souza (2011), o bom planejamento de um inventário florestal deve conter, dentre outros, informações iniciais do local, bem como a compilação e procedimento de cálculos. Com o presente estudo, fica claro a importância dos subsídios que o SR pode prover para a atividade de manejo florestal, no tocante ao seu planejamento. Com o uso de imagens do satélite ALOS pôde-se estimar o volume de madeira e biomassa aérea de povoamentos de eucalipto com um mínimo de 95 parcelas inventariadas. A redução do total de amostras inicial (142 parcelas) significa um menor custo para as atividades de inventário, em um manejo baseado no SR.

No manejo florestal, a aplicação de algoritmos de aprendizagem de máquina e a busca pela redução do número de medições em campo não é exclusiva do SR. Alguns trabalhos mostram a eficiência das redes neurais artificiais para mensuração florestal quantificando e estimando variáveis com uma amostragem menor do que a abordagem tradicional e, conseqüentemente, reduzindo custos com inventário (BINOTI, M.; 2012; BINOTI, M.; BINOTI, D.; LEITE, 2013; BINOTI, D. et al., 2015).

A simplicidade de ajuste das MVS atrelado ao poder de generalização dos dados mostrou o potencial desse algoritmo no aprimoramento das facilidades operacionais para a aplicação do SR nas atividades de quantificação e mapeamento dos recursos florestais. Segundo Beaudoin et al. (1994), povoamentos florestais homogêneos são os locais mais propícios para aplicação de modelagem por SR, visto suas homogeneidades de condições ambientais, principalmente para estudos que utilizam SAR para estimar biomassa aérea. Além disso, o autor também salienta que a generalização da modelagem por SR, com dados empíricos, limita a aplicação dos modelos às mesmas condições ambientais de origem dos dados de treinamento.

Uma sugestão para posteriores estudos de SR, aplicando algoritmos de aprendizagem de máquina, é a avaliação do desempenho das MSV para estimar parâmetros florestais com a inclusão de variáveis de povoamento, como a idade e variáveis texturais, assim como incluídos por Bauermann (2008) com ótimos resultados com regressão múltipla, a nível de talhão. Entretanto, Ismail et al. (2015) salienta que a enumeração dos povoamentos com um ano base, e a utilização da variável idade, sugere o uso de dados empíricos para quantificação dos recursos atuais e futuros por modelos calibrados com informações históricas. No contexto de mudanças climáticas, períodos anormais de chuva ou seca podem influenciar significativamente os valores de índice de sítio, inviabilizando técnica e financeiramente o procedimento proposto. Assim, o mais sensato seria a implementação de modelos que incluíssem os efeitos climáticos recentes.

6 CONCLUSÕES

- As MVS se mostraram eficientes para estimar o volume de madeira e biomassa aérea de povoamentos de eucalipto.
- A MVS regressão tipo II (v-SVR) configurada com a função Gaussiana RBF como *kernel*, e como variáveis de entrada a combinação dos dados dos sensores AVNIR-2 e PALSAR apresentaram os melhores desempenhos.
- Uma quantidade de 95 parcelas é suficiente para uma boa modelagem e generalização dos dados de volume de madeira e biomassa aérea de povoamentos de eucalipto pelas MVS a partir dos sensores AVNIR-2 e PALSAR do satélite ALOS.
- Para estimar volume de madeira de povoamentos de eucalipto, a MVS mais simples foi obtida combinando as bandas ópticas AV3 ou AV4 do sensor AVNIR-2 com uma polarização cruzada (L_{HV} ou L_{VH}) do PALSAR. Já para biomassa aérea, a MVS mais simples foi obtida combinando a banda AV4 do sensor AVNIR-2 com a polarização L_{HV} ou L_{VV} do sensor PALSAR. A combinação AV4, L_{HV} apresentou estimativas consistentes tanto para volume do povoamento, quanto biomassa aérea de eucalipto.

7 BIBLIOGRAFIA

AIZERMAN, M. A.; BRAVERMAN, E. M.; ROZONER, L. I. Theoretical foundations of the potential function method in pattern recognition learning. *Automation and Remote Control*, v. 25, p. 821-837.

ALI, I.; GREIFENEDER, F.; STAMENKOVIC, J.; NEUMANN, M.; NOTARNICOLA, C. Review of machine learning approaches for biomass and soil moisture retrievals from remote sensing data. *Remote Sensing*, v. 7, n. 12, p. 16398–16421, 2015. Disponível em: <<http://www.mdpi.com/2072-4292/7/12/15841/htm>>.

AMRUTKAR, R. P.; KUMAR, S.; KUSHUWAHA, S. P. S.; BHATT, G. D. Forest Biophysical Parameter Retrieval Using PolSAR Technique. In: International Conference on Microwaves, Antena, Propagation and Remote Sensing, 8th, Jodhpur, India, 2012. *Anais...* 2012. n. 1970, p. 128–133.

ATTARCHI, S.; GLOAGUEN, R. Improving the Estimation of Above Ground Biomass Using Dual Polarimetric PALSAR and ETM+ Data in the Hyrcanian Mountain Forest (Iran). *Remote Sensing*, v.6, p. 3693-3715, 2014.

ATTEMA, E. P. W.; ULABY, F. T. Vegetation modeled as a water cloud. *Radio Science*, v. 13, n. 2, p. 357–364, 1978. Disponível em: <<http://doi.wiley.com/10.1029/RS013i002p00357>>. Acesso em: 30/3/2016.

AVTAR, R.; SUZUKI, R.; SAWADA, H. Natural forest biomass estimation based on plantation information using PALSAR data. *PLoS ONE*, v. 9, n. 1, p. 1–15, 2014.

BAGHDADI, N.; MAIRE, G. LE; BAILLY, J.-S.; OSÉ, K.; NOUVELLON, Y.; ZRIBI, M.; LEMOS, C.; HAKAMADA, R. Evaluation of ALOS PALSAR-L Band for the Estimation of Eucalyptus Plantation Aboveground Biomass in Brazil. *IEEE Transactions on Geoscience and Remote Sensing*, v. 8, n. 8, p. 721-724. 2015. Disponível em: <<http://ieeexplore.ieee.org/articleDetails.jsp?arnumber=6946525>>. Acesso em: 25/09/2016.

BAUERMAN, G. C. *Uso de imagens de sensores remotos na estimativa de características dendrométricas de povoamentos de eucalipto*, 2008. 78f. Dissertação (Mestrado em Engenharia de Produção) – Universidade Federal de Santa Maria, Santa Maria, RS, 2008.

BEAUDOIN, A.; LE TOAN, T.; GOZE, S.; NEZRY, E.; LOPES, A.; MOUGIN, E.; HSU, C. C.; HAN, H. C.; KONG, J. A.; SHIN, R. T. Retrieval of forest biomass from SAR data. *International Journal of Remote Sensing*, v. 15, n. 14, p. 2777–2796, 1994. Disponível em: <<http://www.tandfonline.com/doi/abs/10.1080/01431169408954284>>. Acesso em: 20/4/2016.

BEHERA, M. D.; TRIPATHI, P.; MISHRA, B.; KUMAR, S.; CHITALE, V. S.; BEHERA, S. K. Above-ground biomass and carbon estimates of *Shorea robusta* and *Tectona grandis* forests using QuadPOL ALOS PALSAR data. *Advances in Space Research*, v. 57, n. 2, p. 552–561, 2016. Disponível em:

<<http://www.sciencedirect.com/science/article/pii/S0273117715008078>>. Acesso em: 31/3/2016.

BINOTI, M. L. M. S. *Emprego de redes neurais artificiais em mensuração e manejo florestal*. 2012. 130f. Tese (Doutorado em Ciência Florestal) – Universidade Federal de Viçosa, Viçosa, MG, 2012. Disponível em: <<http://www.bibliotecaflorestal.ufv.br:80/handle/123456789/3343>>.

BINOTI, M. L. M. S.; BINOTI, D. H. B.; LEITE, H. G. Aplicação de redes neurais artificiais para estimação da altura de povoamentos equiâneos de eucalipto. *Revista Árvore*, v. 37, n. 4, p. 639–645, 2013. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-67622013000400007&lng=en&nrm=iso&tlng=pt>. Acesso em: 20/4/2016.

BINOTI, M. L. M. S.; LEITE, H. G.; BINOTI, D. H. B.; GLERIANI, J. M. Prognose em nível de povoamento de clones de eucalipto empregando redes neurais artificiais. *CERNE*, v. 21, n. 1, p. 97–105, 2015. UFPA - Universidade Federal de Lavras. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0104-77602015000100097&lng=en&nrm=iso&tlng=pt>. Acesso em: 20/4/2016.

BOSWELL, D. *Introduction to Support Vector Machines*. 2002. Disponível em: <<http://dustwell.com/PastWork/IntroToSVM.pdf> >. Acesso em: 23/09/2016.

CAMPBELL, J. B.; WYNNE, R. H. *Introduction to Remote Sensing*. 5. ed. New York: Guilford Press, 2011.

CAMPOS, J. C. C.; LEITE, H. G. *Mensuração Florestal - Perguntas e respostas*. 4. ed. Viçosa: Editora UFV, 2013. 605p.

CCRS. *Fundamentals of Remote Sensing Tutorial*. Canada Center of Remote Sensing, 2007. Disponível em: <http://www.nrcan.gc.ca/sites/www.nrcan.gc.ca/files/earthsciences/pdf/resource/tutor/fundamentals_e.pdf >. Acesso em: 23/09/2016.

CHAVEZ, P.S. Atmospheric, solar and MTF corrections for ERTS digital imagery. In: *American Society of Photogrammetric Fall Technical Meeting*, Phoenix. 1975. p. 69.

CHEHADE, N. H.; BOUREAU, J.-G.; VIDAL, C.; ZERUBIA, J. Multi-class SVM for forestry classification. IEEE International Conference on Image Processing (ICIP), 16th, Cairo, Egito, 2009. *Anais... IEEE*, 2009. p.1673–1676. Disponível em: <<http://ieeexplore.ieee.org/articleDetails.jsp?arnumber=5413395>>. Acesso em: 22/3/2016.

CHEN, G.; HAY, G. J.; ZHOU, Y. Estimation of forest height, biomass and volume using support vector regression and segmentation from LiDAR transects and Quickbird imagery. In: International Conference on Geoinformatics, 18th, Galway, Ireland. *Anais... IEEE*, 2010. p.1–4. Disponível em: <<http://ieeexplore.ieee.org/articleDetails.jsp?arnumber=5567501>>. Acesso em: 20/4/2016.

CHUVIECO, E.; HUETE, A. *Fundamentals of Satellite Remote Sensing*. CRC Press, 2009.

COOPS, N. C. Eucalypt forest structure and synthetic aperture radar backscatter: a theoretical analysis. *Trees*, v. 16, n. 1, p. 28–46, 2002. Disponível em: <<http://link.springer.com/10.1007/s00468-001-0137-3>>. Acesso em: 23/3/2016.

COSENZA, D. N.; LEITE, H. G.; MARCATTI, G. E.; BINOTI, D. H. B.; ALCÂNTRA, A. E. M.; RODE, R. Classificação da capacidade produtiva de sítios florestais utilizando máquina de vetor de suporte e rede neural artificial. *Scientia Forestalis*, v. 43, n. 108, p. 955–963, 2015. Disponível em: <<http://www.ipef.br/publicacoes/scientia/leitura.asp?Article=19&Number=108&p=n>>. Acesso em: 23/09/2016.

DEL FRATE, F.; SOLIMINI, D. On Neural Network Algorithms for Retrieving Forest Biomass from SAR Data. *IEEE Transactions on Geoscience and Remote Sensing*, v. 42, n. 1, p. 24-34, 2004.

DENÈGRE, J. *Thematic Mapping from Satellite Imagery: a guidebook*. Elsevier, 2013.

DOBSON, M. C. Forest Information from Synthetic Aperture Radar. *Journal of Forestry*, v. 98, n. 6, p. 41–43, 2000. Disponível em: <<http://www.ingentaconnect.com/content/saf/jof/2000/00000098/00000006/art00014>>. Acesso em: 20/4/2016.

DOBSON, M. C.; ULABY, F. T.; LE TOAN, T.; BEAUDOIN, A.; KASISCHKE, E.; CHRISTENSEN, N. Dependence of radar backscatter on coniferous forest biomass. *IEEE Transactions on Geoscience and Remote Sensing*, v. 30, n. 2, p. 412–415, 1992. Disponível em: <<http://ieeexplore.ieee.org/articleDetails.jsp?arnumber=134090>>. Acesso em: 22/3/2016.

DOMINGUES, G. F. *Estimação de Biomassa Aérea e Volume em Povoamentos de Eucaliptos através de Imagens do Satélite ALOS*. 2015. 56f. Dissertação (Mestrado em Ciência Florestal) – Universidade Federal de Viçosa, Viçosa, MG, 2015.

DUBE, T.; MUTANGA, O.; ELHADI, A.; ISMAIL, R. Intra-and-inter species biomass prediction in a plantation forest: testing the utility of high spatial resolution spaceborne multispectral RapidEye sensor and advanced machine learning algorithms. *Sensors*, Basel (Suíça), v. 14, n. 8, p. 15348–70, 2014.

DUBE, T.; MUTANGA, O.; ABDEL-RAHMAN, E. M.; ISMAIL, R.; SLOTOW, R. Predicting Eucalyptus spp. stand volume in Zululand, South Africa: an analysis using a stochastic gradient boosting regression ensemble with multi-source data sets. *International Journal of Remote Sensing*, v. 36, n. 14, p. 3751–3772, 2015.

DUDLEY, R. G. A little REDD model to quickly compare possible baseline and policy scenarios for reducing emissions from deforestation and forest degradation. *Mitigation*

and *Adaptation Strategies for Global Change*, v. 15, n. 1, p. 53–69, 2009. Disponível em: <<http://link.springer.com/10.1007/s11027-009-9204-7>>. Acesso em: 22/3/2016.

DURBHA, S. S.; KING, R. L.; YOUNAN, N. H. Support vector machines regression for retrieval of leaf area index from multiangle imaging spectroradiometer. *Remote Sensing of Environment*, v. 107, n. 1-2, p. 348–361, 2007. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0034425706004159>>. Acesso em: 22/3/2016.

FERNANDEZ-ORDÓÑEZ, Y.; SORIA-RUIZ, J.; LEBLON, B. Forest Inventory using Optical and Radar Remote Sensing. In: JEDLOVEC, G. (Ed.). *Geoscience and Remote Sensing*. InTech, 2009. cap. 26. Disponível em: <<http://www.intechopen.com/books/advances-in-geoscience-and-remote-sensing/forest-inventory-using-optical-and-radar-remote-sensing>>. Acesso em: 23/09/2016.

FERRAZ, A. S.; SOARES, V. P.; SOARES, C. P. B.; RIBEIRO, C. A. A. S.; BINOTI, D. H. B.; LEITE, H.G. Estimativa do estoque de biomassa em um fragmento florestal usando imagens orbitais. *Floresta e Ambiente*, v. 21, n. 3, p. 286–296, 2014. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S2179-80872014000300002&lng=en&nrm=iso&tlng=pt>. Acesso em: 22/3/2016.

FLETCHER, T. *Support Vector Machines Explained*. 2009. Disponível em: <<http://www.tristanfletcher.co.uk/SVM%20Explained.pdf>>. Acesso em: 23/09/2016.

FREEMAN, T. **What is imaging radar?** Disponível em: <<http://southport.jpl.nasa.gov/>>. Acesso em: 20/4/2016.

GAL-EL-HAK, M. *Large-Scale Disasters: Prediction, Control, and Mitigation*. Londres: Cambridge University Press, 2008.

GAMA, F. F.; SANTOS, J. R.; MURA, J. C. Eucalyptus biomass and volume estimation using interferometric and polarimetric SAR Data. *Remote Sensing*, v. 2, n. 4, p. 939–956, 2010.

GAMA, F. F.; SANTOS, J. R.; MURA, J. C.; RENNÓ, C. D. Estimativa de Parâmetros Biofísicos de Povoamentos de Eucalyptus Através de Dados SAR Estimation of Biophysical Parameters in the Eucalyptus Stands by SAR Data. *Ambiência*, v. 2, p. 29–42, 2006.

GEBRESLASIE, M. T.; AHMED, F. B.; AARDT, J. A. N. VAN. Extracting structural attributes from IKONOS imagery for Eucalyptus plantation forests in KwaZulu-Natal, South Africa, using image texture analysis and artificial neural networks. *International Journal of Remote Sensing*, v. 32, n. 22, p. 7677–7701, 2011. Disponível em: <<http://www.tandfonline.com/doi/abs/10.1080/01431161.2010.527392>>. Acesso em: 17/3/2016.

GIBBS, H. K.; BROWN, S.; NILES, J. O.; FOLEY, J. A. Monitoring and estimating tropical forest carbon stocks: making REDD a reality. *Environmental Research Letters*, v. 2, 2007. Disponível

em:<http://redd.unfccc.int/uploads/2_112_redd_20081022_tfg.pdf>. Acesso em: 25/09/2016.

GLEASON, C. J.; IM, J. Forest biomass estimation from airborne LiDAR data using machine learning approaches. *Remote Sensing Environment*, v.125, p. 80-91, 2012.

GONG, Y. X.; HE, C.; YAN, F.; FENG, Z. K.; CAO, M. L.; GAO, Y.; MIAO, J. L. Study on artificial neural network combined with multispectral remote sensing imagery for forest site evaluation. *Guang pu xue yu guang pu fen xi = Guang pu*, v. 33, n. 10, p. 2815–22, 2013. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/24409742>>. Acesso em: 22/3/2016.

GUIMARÃES, D. P. *Desenvolvimento de um modelo de distribuição diamétrica de passo invariante para prognose e projeção da estrutura de povoamentos de eucalipto*. 1994. 160f. Tese (Doutorado em Ciência Florestal) – Universidade Federal de Viçosa, Viçosa, MG, 1994.

HARRELL, P. A.; KASISCHKE, E. S.; BOURGEOU-CHAVEZ, L. L.; HANEY, E. M.; CHRISTENSEN, N. L. Evaluation of approaches to estimating aboveground biomass in Southern pine forests using SIR-C data. *Remote Sensing of Environment*, v. 59, n. 2, p. 223–233, 1997. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0034425796001551>>. Acesso em: 20/4/2016.

HEDLEY, J. D.; ROELFSEMA, C. M.; PHINN, S. R.; MUMBY, P. J. Environmental and Sensor Limitations in Optical Remote Sensing of Coral Reefs: Implications for Monitoring and Sensor Design. *Remote Sensing*, v. 4, n. 12, p. 271–302, 2012. Molecular Diversity Preservation International. Disponível em: <<http://www.mdpi.com/2072-4292/4/1/271/htm>>. Acesso em: 22/3/2016.

HEISKANEN, J. Estimating aboveground tree biomass and leaf area index in a mountain birch forest using ASTER satellite data. *International Journal of Remote Sensing*, v. 27, n. 6, p. 1135-1158, 2006.

HENDERSON, F.; LEWIS, A. J. Radar Fundamentals: The Geoscience Perspective. In: ENDERSON, F. M.; LEWIS, A. J. (Eds.). *Principles and Application of Imaging Radar: Manual of Remote Sensing*. 3. ed, v. 2. New York: John Wiley & Sons Inc., 1998. p.567–629.

HUANG, C.; SONG, K.; KIM, S.; TOWNSHEND, J. R. G.; DAVIS, P.; MASEK, J.G.; GOWARD, S. N. Use of a dark object concept and support vector machines to automate forest cover change analysis. *Remote Sensing of Environment*, v. 112, n. 3, p. 970–985, 2008. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0034425707003951>>. Acesso em: 22/12/2015.

HUANG, G. B.; LEARNED-MILLER, E. L. Labeled faces in the wild : Updates and new reporting procedures. *University of Massachusetts Amherst Technical Report*, Amherst, n.3, 2014.

HYDE, P.; DUBAYAH, R.; WALKER W.; BLAIR, B.; HOFTON, M.; HUNSAKER, C. Mapping forest structure for wildlife habitat analysis using multi-sensor (LiDAR, SAR/InSAR, ETM+, Quickbird) synergy. *Remote Sensing Environment*, v. 102, p. 63-73, 2006.

IGARASHI, T. ALOS mission requirement and sensor specifications. *Advances in Space Research*, v. 28, n. 1, p. 127–131, 2001. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0273117701003167>>. Acesso em: 22/3/2016.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA – IBGE. *Mapa de solos do Brasil*. Rio de Janeiro: Ministério do Planejamento, Orçamento e Gestão, 2001. 1 mapa, color., Escala 1: 5.000.000.

ISMAIL, R.; KASSIER, H.; CHAUKE, M.; HOLECZ, F.; HATTINGH, N. Assessing the utility of ALOS PALSAR and SPOT 4 to predict timber volumes in even-aged Eucalyptus plantations located in Zululand, South Africa. *Southern Forests: a Journal of Forest Science*, v. 77, n. 3, p. 203–211, 2015. Disponível em: <<http://www.tandfonline.com/doi/abs/10.2989/20702620.2014.1001681>>. Acesso em: 31/3/2016.

JAXA. Agência de Exploração Aeroespacial Japonesa. *Advanced Land Observing Satellite – DAICHI-2 – DAICHI*. 2016. Disponível em: <http://www.eorc.jaxa.jp/ALOS/en/top/about_top.htm> Acesso em: 24/09/2016.

JENSEN, J. R. *Remote Sensing of the Environment: An Earth Resource Perspective*. New Jersey: Prentice-Hall, 2000.

KRONSEDER, K.; BALLHORN, U.; BÖHM, V.; SIERGERT, F. Above ground biomass estimation across forest types at different degradation levels in Central Kalimantan using LiDAR data. *International Journal of Applied Earth Observation and Geoinformation*, v. 18, p. 37-48, 2012.

LE TOAN, T.; BEAUDOIN, A.; RIOM, J.; GUYON, D. Relating forest biomass to SAR data. *IEEE Transactions on Geoscience and Remote Sensing*, v. 30, n. 2, p. 403–411, 1992. Disponível em: <<http://ieeexplore.ieee.org/articleDetails.jsp?arnumber=134089>>. Acesso em: 22/3/2016.

LECKIE, D.; RANSON, J. K. Forestry Applications Using Imaging Radar. In: ENDERSON, F. M.; LEWIS, A. J. (Eds.). *Principles and Application of Imaging Radar: Manual of Remote Sensing*. 3. ed, v. 2. New York: John Wiley & Sons Inc., 1998. p. 435-501.

LI, F.; WANG, J.; NIE, L.; ZHANG, W. Computational Methods to Interpret and Integrate Metabolomic Data. In: ROESSNER, U. (Ed.). *Metabolomics*. InTech, p. 99-131, 2012, cap 5.

LI, Z.; XU, D.; GUO, X. Remote sensing of ecosystem health: opportunities, challenges, and future perspectives. *Sensors*, Basel (Suíça), v. 14, n. 11, p. 21117–39, 2014.

Disponível em: <<http://www.mdpi.com/1424-8220/14/11/21117/htm>>. Acesso em: 12/2/2016.

LILLESAND, T.; KIEFER, R. W.; CHIPMAN, J. *Remote Sensing and Image Interpretation*. 7. ed. Wiley, 2015. 768p.

LIM, K.; TREITZ, P.; WULDER, M.; ST-ONGE, B.; FLOOD, M. LiDAR remote sensing of forest structure. *Progress in Physical Geography*, v. 27, n. 1; p. 88-106, 2003.

LORENA, A. C. *Investigação de estratégias para a geração de máquinas de vetores suporte multiclassas*. 2006. 229f. Tese (Doutorado em Ciências de Computação e Matemática Computacional) – Universidade de São Paulo, São Carlos, SP, 2006.

LORENA, A. C.; CARVALHO, A. C. P. L. F. Uma Introdução às Support Vector Machines. *Revista de Informática Teórica e Aplicada*, v. 14, n. 2, p. 43–67, 2007.

LUCAS, R. M.; CRONIN, N.; LEE, A.; MOGHADDAM, M.; WITTE, C.; TICKLE, P. Empirical relationships between AIRSAR backscatter and LiDAR-derived forest biomass, Queensland, Australia. *Remote Sensing Environment*, v. 100, p. 407-425, 2006.

MAGALHÃES, J. G. DE S. *Prognose e Regulação Florestal de Eucalipto Utilizando Modelo de Crescimento e Inteligência Computacional*. 2014. 66f. Dissertação (Mestrado em Ciência Florestal) – Universidade Federal de Viçosa, Viçosa, MG, 2014.

MENG, S.; PANG, Y.; ZHANG, Z.; JIA, W.; LI, Z. Mapping Aboveground Biomass using Texture Indices from Aerial Photos in a Temperate Forest of Northeastern China. *Remote Sensing*, v. 8, n. 3, p. 230, 2016. Disponível em: <<http://www.mdpi.com/2072-4292/8/3/230/htm>>. Acesso em: 24/3/2016.

MOREIRA, M. A. *Fundamentos do Sensoriamento Remoto*. 4. ed. Viçosa: Editora UFV, 2011.

MOREL, A. C.; SAATCHI, S. S.; MALHI, Y.; BERRY, N. J.; BANIN, L.; BURSLEM, D.; NILUS, R; ONG, R. Estimating aboveground biomass in forest and oil palm plantation in Sabah, Malaysian Borneo using ALOS PALSAR data. *Forest Ecology and Management*, v. 262, n. 9, p. 1786–1798, 2011. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0378112711004361>>. Acesso em: 20/4/2016.

MOSES, W. J.; BOWLES, J. H.; CORSON, M. R. Expected improvements in the quantitative remote sensing of optically complex waters with the use of an optically fast hyperspectral spectrometer-a modeling study. *Sensors*, Basel (Suíça), v. 15, n. 3, p. 6152–73, 2015. Multidisciplinary Digital Publishing Institute. Disponível em: <<http://www.mdpi.com/1424-8220/15/3/6152/htm>>. Acesso em: 22/3/2016.

NILSSON, N. J. *Learning Machines*. Nova York: McGraw-Hill, 1965.

NOBLE, W. S. What is a support vector machine? *Nature biotechnology*, v. 24, n. 12, p. 1565–7, 2006. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/17160063>>. Acesso em: 7/3/2016.

OLIVEIRA, F. S. DE. *Uso de imagens do satélite ALOS para estimativa do volume de madeira em plantações de Eucaliptos*. 2011. 146f. Tese (Doutorado em Ciência Florestal) – Universidade Federal de Viçosa, Viçosa, MG, 2011. Disponível em: <[http://alexandria.cpd.ufv.br:8000/teses/ciencia florestal/2011/245688f.pdf](http://alexandria.cpd.ufv.br:8000/teses/ciencia%20florestal/2011/245688f.pdf)>.

OLIVEIRA, T. C. A.; CARVALHO, L. M. T.; OLIVEIRA, L. T.; MARTINHAGO, A. Z.; ACERBI JÚNIOR, L. T.; LIMA, M. P. Mapping deciduous forests by using time series of filtered MODIS NDVI and neural networks. *CERNE*, v. 16, n. 2, p. 123–130, 2010. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0104-77602010000200002&lng=en&nrm=iso&tlng=en>. Acesso em: 22/3/2016.

PAX-LENNEY, M.; WOODCOCK, C. E.; MACOMBER, S. A.; GOPAL, S.; SONG, C. Forest mapping with a generalized classifier and Landsat TM data. *Remote Sensing of Environment*, v. 77, n. 3, p. 241–250, 2001. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0034425701002085>>. Acesso em: 22/3/2016.

PUTZ, F. E.; ZUIDEMA, P. A.; PINARD, M. A.; BOOT, R. G. A.; SAYER, J. A.; SHEIL, D.; SIST, P.; ELIAS; VANKLAY, J.K. Improved tropical forest management for carbon retention. *PLoS biology*, v. 6, n. 7, p. e166, 2008. Disponível em: <<http://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.0060166>>. Acesso em: 22/3/2016.

RICHARDS, J. A. *Remote Sensing with Imaging Radar*. Berlim: Springer, 2009.

ROSENQVIST, A.; SHIMADA, M.; ITO, N.; WATANABE, M. ALOS PALSAR: A Pathfinder Mission for Global-Scale Monitoring of the Environment. *IEEE Transactions on Geoscience and Remote Sensing*, v. 45, n. 11, p. 3307–3316, 2007. Disponível em: <<http://ieeexplore.ieee.org/articleDetails.jsp?arnumber=4358862>>. Acesso em: 12/2/2016.

ROY, D. P.; WULDER, M. A.; LOVELAND, T. R.; WOODCOCK, C. E.; ALLEN, R. G.; ANDERSON, M. C.; HELDER, D.; IRONS, J. R.; JOHNSON, D. M.; KENNEDY, R.; SCAMBOS, T. A.; SCHAAF, C. B.; SCHOTT, J. R.; SHENG Y.; VERMOTE, E. F.; BELWARD, A. S.; BLINDSCHADLER, R.; COHEN, W.B.; GAO, F.; HIPPLE, J. D.; HOSTERT, P.; HUNTINGTON, J.; JUSTICE, C. O.; KILIC, A.; KOVALSKYY, V.; LEE, Z. P.; LYMBURNER, L.; MASEK, J. G.; MACCORKEL, J.; SHUAI, Y.; TREZZA, R.; VOGELMANN, J.; WYNNE, R. H.; ZHU, Z. Landsat-8: Science and product vision for terrestrial global change research. *Remote Sensing of Environment*, v.145, p. 154- 172, 2013.

SAMMUT, C.; WEBB, G. I. *Encyclopedia of Machine Learning*. Boston: Springer US, 2010.

SANTORO, M.; ERIKSSON, L.; FRANSSON, J. Reviewing ALOS PALSAR Backscatter Observations for Stem Volume Retrieval in Swedish Forest. *Remote*

Sensing, v. 7, n. 4, p. 4290–4317, 2015. Disponível em: <<http://www.mdpi.com/2072-4292/7/4/4290/htm>>. Acesso em: 17/4/2016.

SCHÖLKOPF, B.; BARTLETT, P.; SMOLA, A. J.; WILLIAMSON, R. Shrinking the tube: A new support vector regression algorithm. In: KEARNS, M. S.; SOLLA, S. A.; COHN, D. A. (Eds.). *Advances in Neural Information Processing Systems*, v. 11. Cambridge: MIT Press, 1999. Disponível em: <<http://papers.nips.cc/paper/1563-shrinking-the-tube-a-new-support-vector-regression-algorithm.pdf>>. Acesso em: 25/09/2016.

SCHÖLKOPF, B.; SMOLA, A. J.; WILLIAMSON, R. C.; BARTLETT, P. L. New support vector algorithms. *Neural computation*, v. 12, n. 5, p. 1207–1245, 2000.

SCHÖLKOPF, B.; SMOLA, A. J. *Support Vector Machines and Kernel Algorithms*. 2002. 22p. Disponível em: <<http://alex.smola.org/papers/2002/SchSmo02b.pdf>>. Acesso em: 25/09/2016

SHATAEE, S. Forest Attributes Estimation Using Aerial Laser Scanner and TM Data. *Forest Systems*, v. 22, n. 3, p. 484-496, 2013. Disponível em: <<http://revistas.inia.es/index.php/fs/article/view/3874>>. Acesso em: 25/3/2016.

SMOLA, A. J.; SCHÖLKOPF, B. A Tutorial on Support Vector Regression. *Statistics and Computing*, v. 14, n. 3, p. 199–222, 2004.

SOARES, C. P. B.; PAULA NETO, F.; SOUZA, A. L. *Dendrometria e Inventário Florestal*. 2. ed. Viçosa: Editora UFV, 2011.

SPANNER, M.A.; PIERCE, L.L.; PETERSON, D.L.; RUNNING, S.W. Remote sensing of temperate coniferous forest leaf area index - The influence of canopy closure, understory vegetation and background reflectance. *International Journal of Remote Sensing*, v. 11, n. 1, p. 95–111, 1990.

TREITZ, P. M.; HOWARTH, P. J. Hyperspectral remote sensing for estimating biophysical parameters of forest ecosystems. *Progress in Physical Geography*, v. 23, n. 3, p. 359-390, 1999.

TROTTER, C. M.; DYMOND, J. R.; GOULDING, C. J. Estimation of timber volume in a coniferous plantation forest using Landsat TM. *International Journal of Remote Sensing*, v. 18, n. 10, p. 2209–2223, 1997. Disponível em: <<http://www.tandfonline.com/doi/abs/10.1080/014311697217846>>. Acesso em: 17/3/2016.

TUCKER, C. J.; SELLERS, P. J. Satellite Remote Sensing of primary production. *International Journal of Remote Sensing*, v. 7, n. 11, p. 1395-1416, 1986. Disponível em: <<http://www.tandfonline.com/doi/pdf/10.1080/01431168608948944>>. Acesso em: 25/09/2016.

TYMOTHY, D.; ONISIMO, M.; CLETAH, S.; ADELABU, S.; TSITSI, B. Remote Sensing of aboveground forest biomass: A review. *Tropical Ecology*, v. 57, n. 2, p. 125-132, 2016.

VAPNIK, V. N.; CHERVONENSK A. Y. *Theory of Patterns recognition*. Moscou: Nauka, 1974.

VAPNIK, V. N. *The nature of statistical learning theory*. Nova York: Springer-Verlag, 1995. 188p. Disponível em: <<http://dl.acm.org/citation.cfm?id=211359>>. Acesso em: 22/3/2016.

WANG, K.; FRANKLIN, S. E.; GUO, X.; CATTET, M. Remote sensing of ecology, biodiversity and conservation: a review from the perspective of remote sensing specialists. *Sensors*, Basel (Suíça), v. 10, n. 11, p. 9647–67, 2010. Disponível em: <<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3231003&tool=pmcentrez&rendertype=abstract>>. Acesso em: 2/2/2016.

WATZLAWICK, L. FARINHA; KOEHLER, H. S.; KIRCHNER, F. F. Estimativa de Biomassa e carbono em Plantios de Pinus taeda Utilizando Imagens de Satélite Ikonos II. *Ciência e Natura*, v. 28, n. 1, p. 45–60, 2006. Disponível em: <<http://cascavel.ufsm.br/revistas/ojs2.2.2/index.php/cienciaenatura/article/download/9708/5816>>.

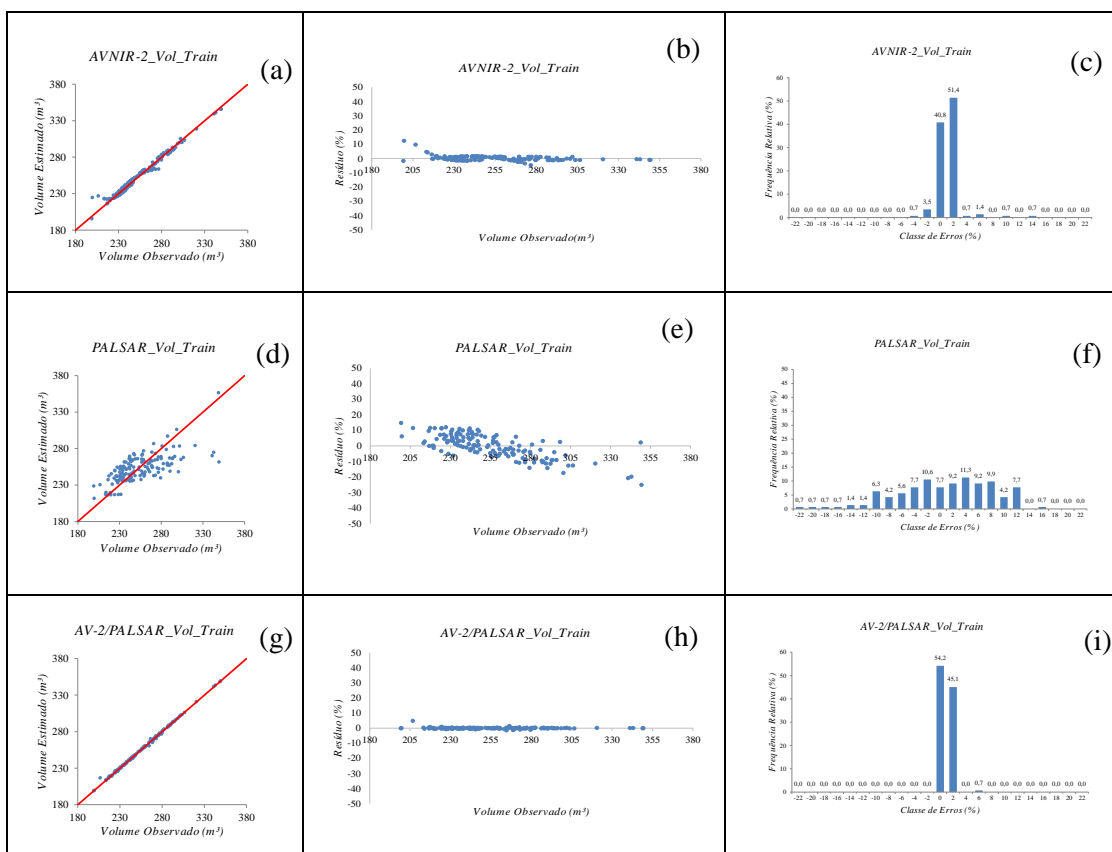
WOLTER, P. T.; TOWNSEND, P. A. Estimating forest species composition using a multi-sensor fusion approach. *Remote Sensing of Environment*, v. 115, p. 671–691, 2011.

WULDER, M. Optical remote-sensing techniques for the assessment of forest inventory and biophysical parameters. *Progress in Physical Geography*, v. 22, n. 4, p. 449–476, 1998. Disponível em: <<http://ppg.sagepub.com/content/22/4/449.abstract>>. Acesso em: 1/4/2016.

ZHENG, G.; MOSKAL, L. M. Retrieving Leaf Area Index (LAI) Using Remote Sensing: Theories, Methods and Sensors. *Sensors*, Basel (Suíça), v. 9, n. 4, p. 2719–45, 2009. Molecular Diversity Preservation International. Disponível em: <<http://www.mdpi.com/1424-8220/9/4/2719/htm>>. Acesso em: 29/11/2015.

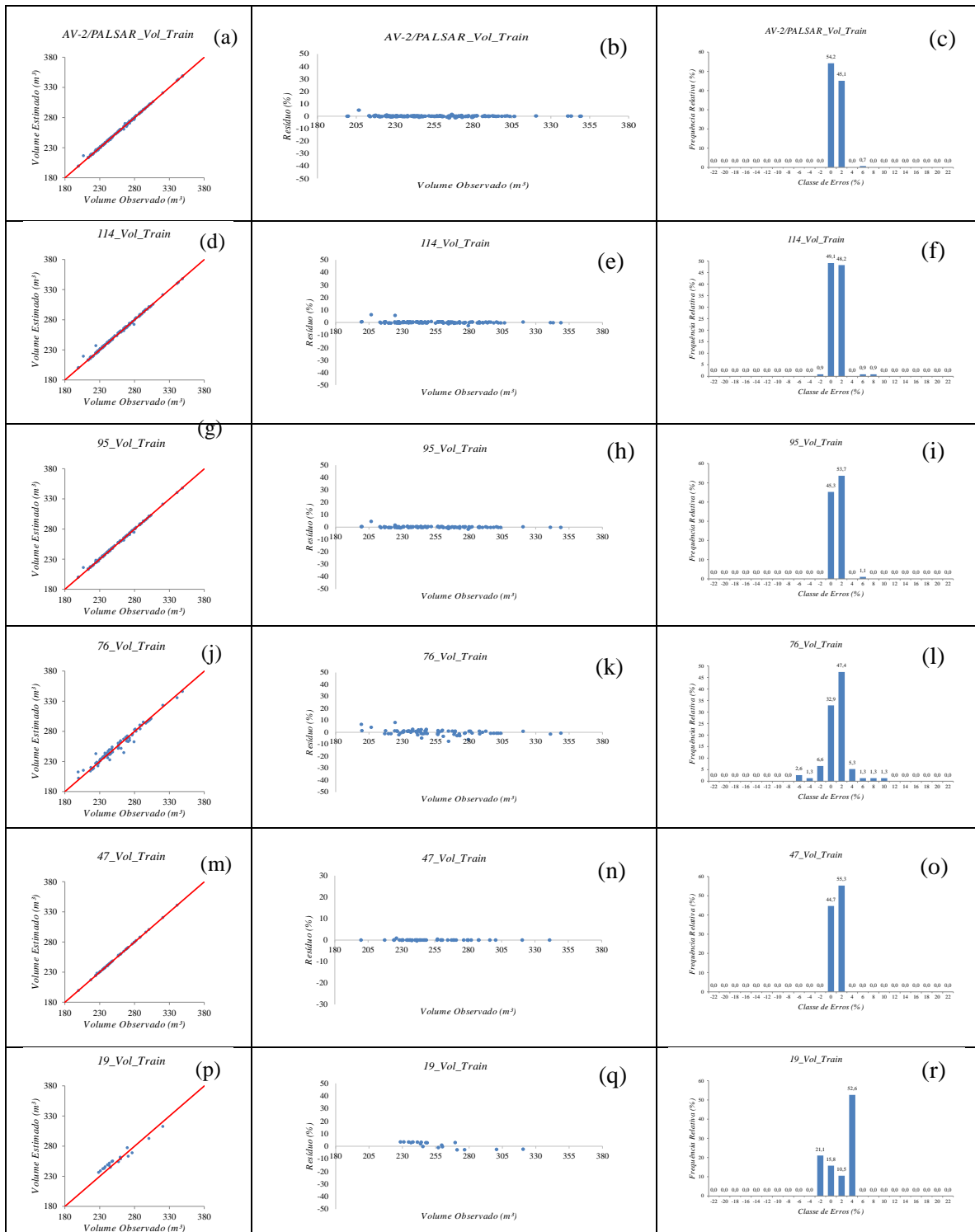
APÊNDICES

APÊNDICE A



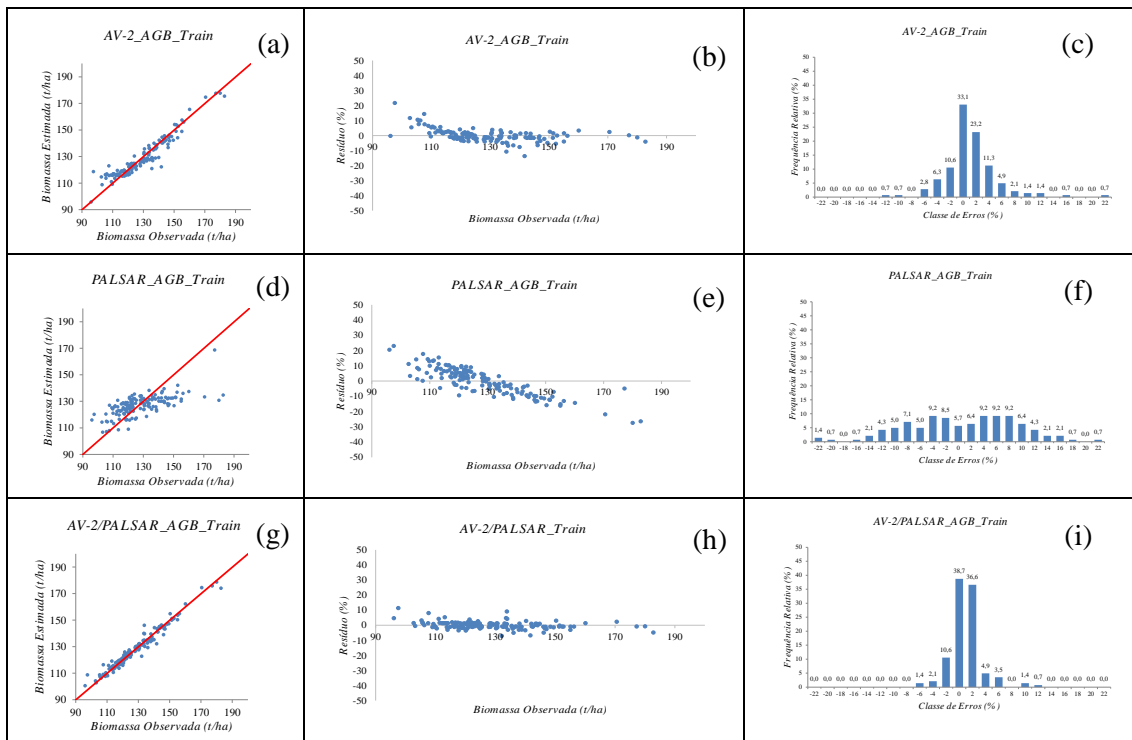
Gráficos de resíduos dos treinamentos das MVS para as predições de volume utilizando dados dos sensores AVNIR-2 e PALSAR e 142 parcelas para treinamento. As Figuras a, d e g representam os gráficos da relação dos valores observados e estimados pelas MVS, utilizando como variáveis de entrada, respectivamente, os dados dos sensores: AVNIR-2, PALSAR, AVNIR-2+PALSAR. As Figuras b, f e h representam os gráficos de dispersão dos resíduos percentuais das estimativas das MVS utilizando como variáveis de entrada, respectivamente, os dados dos sensores: AVNIR-2, PALSAR, AVNIR-2+PALSAR. As Figuras c, f e i representam os gráficos de frequência de classes de erros das estimativas das MVS utilizando como variáveis de entrada, respectivamente, os dados dos sensores: AVNIR-2, PALSAR, AVNIR-2+PALSAR

APÊNDICE B



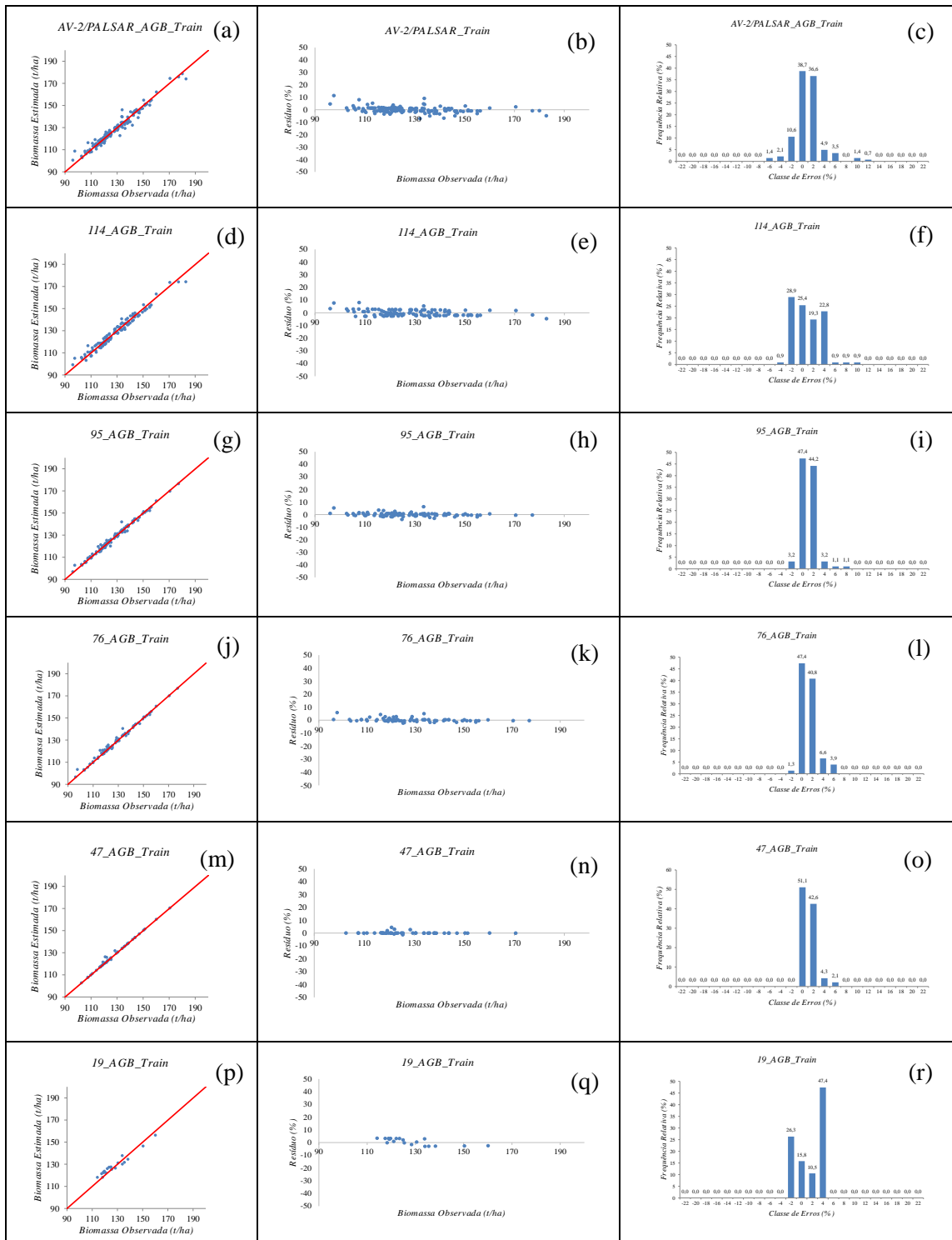
Gráficos de resíduos dos treinamentos das MVS para as previsões de volume utilizando dados dos sensores AVNIR-2 e PALSAR com a redução do número de parcelas amostrais para treinamento. As Figuras a, d, g, j, m e p representam os gráficos da relação dos valores observados e estimados pelas MVS treinadas, respectivamente, com: 142, 114, 95, 76, 47 e 19 parcelas. As Figuras b, e, h, k, n e q, representam os gráficos de dispersão dos resíduos percentuais das estimativas das MVS treinadas, respectivamente, com: 142, 114, 95, 76, 47 e 19 parcelas. As Figuras c, f, i, l, o e r representam os gráficos de frequência de classes de erros das estimativas das MVS treinadas, respectivamente, com: 142, 114, 95, 76, 47 e 19 parcelas.

APÊNDICE C



Gráficos de resíduos dos treinamentos das MVS para as previsões de biomassa aérea utilizando dados dos sensores AVNIR-2 e PALSAR e 142 parcelas para treinamento. As Figuras a, d e g representam os gráficos da relação dos valores observados e estimados pelas MVS, utilizando como variáveis de entrada, respectivamente, os dados dos sensores: AVNIR-2, PALSAR, AVNIR-2+PALSAR. As Figuras b, f e h representam os gráficos de dispersão dos resíduos percentuais das estimativas das MVS utilizando como variáveis de entrada, respectivamente, os dados dos sensores: AVNIR-2, PALSAR, AVNIR-2+PALSAR. As Figuras c, f e i representam os gráficos de frequência de classes de erros das estimativas das MVS utilizando como variáveis de entrada, respectivamente, os dados dos sensores: AVNIR-2, PALSAR, AVNIR-2+PALSAR

APÊNDICE D



Gráficos de resíduos dos treinamentos das MVS para as predições de biomassa a utilizando dados dos sensores AVNIR-2 e PALSAR com a redução do número de parcelas amostrais para treinamento. As Figuras a, d, g, j, m e p representam os gráficos da relação dos valores observados e estimados pelas MVS treinadas, respectivamente, com: 142, 114, 95, 76, 47 e 19 parcelas. As Figuras b, e, h, k, n e q, representam os gráficos de dispersão dos resíduos percentuais das estimativas das MVS treinadas, respectivamente, com: 142, 114, 95, 76, 47 e 19 parcelas. As Figuras c, f, i, l, o e r representam os gráficos de frequência de classes de erros das estimativas das MVS treinadas, respectivamente, com: 142, 114, 95, 76, 47 e 19 parcelas.