

UNIVERSIDADE FEDERAL DE SÃO CARLOS  
CENTRO DE CIÊNCIAS E TECNOLOGIAS PARA A SUSTENTABILIDADE  
CAMPUS DE SOROCABA  
PROGRAMA DE PÓS-GRADUAÇÃO EM PLANEJAMENTO E USO DE RECURSOS  
RENOVÁVEIS

LÍVIA LANZI ALÓ

**USO DE COMPONENTES DE IMAGENS DE SATÉLITES NA  
MODELAGEM ESPACIAL DO VOLUME EM POVOAMENTOS DE  
*Eucalyptus sp.***

Sorocaba  
2016





UNIVERSIDADE FEDERAL DE SÃO CARLOS  
CENTRO DE CIÊNCIAS E TECNOLOGIAS PARA A SUSTENTABILIDADE  
CAMPUS DE SOROCABA  
PROGRAMA DE PÓS-GRADUAÇÃO EM PLANEJAMENTO E USO DE RECURSOS  
RENOVÁVEIS

LÍVIA LANZI ALÓ

**USO DE COMPONENTES DE IMAGENS DE SATÉLITES NA  
MODELAGEM ESPACIAL DO VOLUME EM POVOAMENTOS DE  
*Eucalyptus sp.***

Dissertação apresentada ao Programa de Pós-Graduação em Planejamento e Uso de Recursos Renováveis, para obtenção do título de mestre em Planejamento e Uso de Recursos Renováveis.

Orientação: Prof. Dr. Cláudio Roberto Thiersch

Coorientação: Prof. Dr. Antônio César Germano Martins

Sorocaba  
2016

Lanzi Aló, Livia

Uso de componentes de imagens de satélites na modelagem espacial do volume em povoamento de Eucalyptus sp. / Livia Lanzi Aló. -- 2016.

112 f. : 30 cm.

Dissertação (mestrado)-Universidade Federal de São Carlos, campus Sorocaba, Sorocaba

Orientador: Cláudio Roberto Thiersch

Banca examinadora: Rafael Mingoti, Mônica Fabiana Bento Moreira Thiersch

Bibliografia

1. Geoestatística. 2. Krigagem com deriva externa. 3. Inventário florestal.  
I. Orientador. II. Universidade Federal de São Carlos. III. Título.

Ficha catalográfica elaborada pelo Programa de Geração Automática da Secretaria Geral de Informática (SIn).

DADOS FORNECIDOS PELO(A) AUTOR(A)

LÍVIA LANZI ALÓ

**USO DE COMPONENTES DE IMAGENS DE SATÉLITES NA  
MODELAGEM ESPACIAL DO VOLUME EM POVOAMENTO  
DE *Eucalyptus* sp.**

Dissertação apresentada ao Programa de Pós-Graduação para obtenção do título de  
mestre em Planejamento e Uso de Recursos Renováveis.  
Universidade Federal de São Carlos.  
Sorocaba, 17 de maio de 2016.

Orientador:



Prof. Dr. Cláudio Roberto Thiersch  
Universidade Federal de São Carlos – UFSCar *Campus* Sorocaba

Examinadores:



Dr. Rafael Mingoti  
Embrapa



Profa. Dra. Monica Fabiana Bento Moreira Thiersch  
Universidade Federal de São Carlos – UFSCar *Campus* Sorocaba



## **DEDICATÓRIA**

À minha mãe Vera,

Ao meu esposo Rafael por seu amor.

Às minhas irmãs, cunhados e sobrinhos,  
Lia, Laís, Thiago's, Laura, Lis e Benício  
que sempre me apoiaram nesta trajetória.



## AGRADECIMENTOS

A Deus, por mais uma conquista maravilhosa em minha vida.

Ao amor da minha vida, meu esposo Rafael, por sua compreensão, incentivos, ajudas, amor, e por fazer cada dia da minha vida mais feliz e melhor.

À minha mãe Vera, às minhas irmãs Lia e Laís e aos meus cunhados Thiago's, pelo ensinamento, pela ajuda nos momentos difíceis e pelo amor incondicional.

Aos meus sobrinhos queridos, Laura, Lis e Benício, por sua inocência e alegria de vida.

Ao meu professor orientador, Cláudio Roberto Thiersch, pelos ensinamentos, paciência e dedicação.

Ao professor coorientador, Antônio César Germano Martins, pelas contribuições a este estudo e ensinamentos.

Ao meu pai Valdir, Yuri, aos meus familiares e padrinhos pelo carinho, ajuda e por acreditarem em mim, em especial a minha avó Lélia.

À minha sogra Sevane e minha cunhada Livia, pelo apoio e ajuda.

Aos meus grandes amigos pelo incentivo, amor e compreensão pela distância física entre nós, em especial Carla, Rodrigo, Diogo (Shek), Luciano (Ximu), Laís, Carol, Lari, Karen, Thaís, Eltinho, Augusto, Adriana, Naldo, Paulinha, Michel e Cléber.

Aos meus amigos de curso, pelos momentos incríveis compartilhados ao longo desses dois anos, pela ajuda e carinho, em especial Juliette, Gabi, Thiago e Ana Larissa.

Aos professores da Universidade Federal de São Carlos- *campus* Sorocaba e do Programa de Pós Graduação em Planejamento e Uso de Recursos Renováveis (PPGPUR) por seus ensinamentos, apoios e orientações.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pela concessão da bolsa de estudos.

À todos que, direta ou indiretamente, contribuíram para a realização deste trabalho, o meu muito obrigado.





“O sucesso nasce do querer, da determinação e persistência em se chegar a um objetivo. Mesmo não atingindo o alvo, quem busca e vence obstáculos, no mínimo fará coisas admiráveis”.

José de Alencar



## RESUMO

ALÓ, Livia Lanzi. *Uso de componentes de imagens de satélites na modelagem espacial do volume em povoamento de Eucalyptus sp.* 2016. 112 f. Dissertação (Mestrado em “Planejamento e uso de recursos renováveis”) – Centro de Ciências e Tecnologias para Sustentabilidade, Universidade Federal de São Carlos, Sorocaba, 2016.

O inventário florestal é uma importante ferramenta utilizada para estimar a produção dos povoamentos florestais. Contudo, algumas metodologias utilizadas no inventário são embasadas na Estatística Clássica, que desconsidera qualquer continuidade espacial que possa existir entre as unidades amostrais. Alguns interpoladores geoestatísticos, tais como a krigagem ordinária (KO) e a krigagem de deriva externa (KDE), permitem avaliar essa estrutura espacial. Além disso, interpoladores como a KDE utilizam, além da variável espacial, uma ou mais variáveis auxiliares. As imagens de satélites possuem diferentes componentes que se correlacionam com as variáveis dendrométricas podendo ser usados como variáveis auxiliares, visando o aumento do grau de precisão das estimativas. O objetivo deste estudo foi avaliar o desempenho da KDE na estimativa do volume de povoamentos florestais de *Eucalyptus sp.*, utilizando os componentes de imagens de satélites como variáveis auxiliares e compará-la com o desempenho da KO. Com esse propósito, processou-se um inventário florestal de 210 parcelas circulares de 500 m<sup>2</sup>, a fim de estimar o volume (m<sup>3</sup> ha<sup>-1</sup>) por parcela. As imagens obtidas da área do estudo continham as bandas azul, verde, vermelho e infravermelho próximo. A partir destas, foram extraídos o nível de cinza (NC) de cada banda, da razão simples entre as bandas, índices de vegetação (NDVI, SAVI e ARVI), medidas de textura e índices gerados a partir das texturas referentes à área da parcela. Para a geoestatística KDE, foi feito o ajuste do modelo de covariância através do método *Stepwise* e a seleção pelo método AIC (Critério de Informação de Akaike). Os semivariogramas da KDE e da KO foram ajustados por diferentes modelos teóricos por meio do método dos Mínimos Quadrados Ordinários (MQO) e a escolha do melhor modelo se deu pelo menor valor do erro padrão residual. Nas análises das estatísticas das imagens e da matriz de correlação geradas, foi possível observar a correlação das variáveis com o volume e também a autocorrelação existente entre as variáveis. O melhor modelo de covariância selecionado foi composto por banda 2, medida de textura COR (correlação) da banda 2, índice de textura MULCOR (correlação multiplicado pela banda) da banda 1 e pela idade. Nos dois semivariogramas, o modelo que melhor se ajustou foi o exponencial. Nas análises dos resultados, as estimativas de volume geradas pela KDE produziram melhores resultados que as estimativas da KO, obtendo o menor valor de erro padrão residual e a melhor área sob a curva (AUC) na análise da curva ROC (Receiver Operating Characteristic).

**Palavras-chave:** Inventário florestal. Geoestatística. Krigagem com deriva externa.



## ABSTRACT

ALÓ, Livia Lanzi. *Use of satellite imagery components in spatial modeling of volume in Eucalyptus sp. stands*. 2016. 112 f. Dissertation (Master in “Planning and Use of Renewable Resources”) – Science and Technology Center of Sustainability, Federal University of São Carlos, Sorocaba, 2016.

Forest inventory is an important tool used to estimate forest wood production. However, some methodologies used in forest inventory are based in Classical Statistics, which disregards any spatial continuity that may exist between sample unities. Some geostatistic interpolators such as ordinary kriging (OK) and external drift kriging (EDK) allow us to assess this spatial structure. Furthermore, besides spatial variability, interpolators as EDK use one or more auxiliary variables. Satellite images have different components that interrelate with dendrometric variables and that can be used as auxiliary variables in order to increase the degree of precision of estimates. The aim of this study was to assess EDK performance on the volume estimation of *Eucalyptus* sp. stands using satellite image components as secondary variables and to compare it with OK performance. With this purpose, a forest inventory of 210 circular plots of 500 m<sup>2</sup> was carried out in order to estimate the volume (m<sup>3</sup> ha<sup>-1</sup>) in each plot. Images obtained of studied area had blue, green, red and near infrared band. From these bands, it were extracted: gray level in each band, the ratio between bands, vegetation index (NDVI, SAVI e ARVI), texture measures and index generated from textures related to plot area. Covariance model adjustment throughout Stepwise method and selection by AIC (Akaike Information Criterion) method were made to EDK geostatistic. EDK and OK semivariograms were adjusted by different theoretical models through Ordinary Least Squares (OLS) method and the choice of the best model was given by the lowest value of residual standard error. From statistic analysis of images and correlation matrix, it was observed a correlation of variables with volume and also autocorrelation between these variables. The best covariance model selected was composed by band 2, measure of COR texture of band 2, MULCOR texture index of band 1 and by age. In the two semivariograms, the best model adjusted was the exponential one. Analysing the results, volume estimates generated by EDK produced better results than OK estimates and had the lowest value of residual standard error and the best area under curve (AUC) in receiver operating characteristic (ROC) curve analysis.

**Keywords:** Forest inventory. Geostatistic. External drift kriging.



## LISTA DE FIGURAS

Figura 1 - Espectro eletromagnético.....	32
Figura 2 - Efeito da posição do pixel no terreno com relação a radiância da cultura.....	35
Figura 3 – Variações angulares utilizadas no cálculo da matriz de coocorrência. ....	42
Figura 4 – Parâmetros do semivariograma.....	47
Figura 5 – Três curvas ROC associadas a conjunto de pontos distintos. ....	54
Figura 6 – Área sob a curva ROC de dois conjunto de pontos distintos. ....	56
Figura 7 – Povoamento onde foram coletados os dados usados no estudo. ....	57
Figura 8 – Distribuição das parcelas nas áreas de plantio da fazenda.....	58
Figura 9 – Seleção das árvores no interior das parcelas.....	59
Figura 10 – Valores de nível de cinza das quatro bandas.....	74
Figura 11 – Distribuição dos volumes observados pelos NC da: (A) banda 1; (B) banda 2; (C) banda 3; e (D) banda 4.....	76
Figura 12 – Distribuição dos volumes observados pela razão simples dos NC da: (A) banda 1/banda 2; (B) banda 1/banda 3; (C) banda 1/banda 4; (D) banda 2/banda 3; (E) banda 2/banda 4; e (F) banda 3/banda 4. ....	77
Figura 13 – Distribuição dos volumes observados pelos valores do índice: (A) NDVI; (B) SAVI; e (C) ARVI.....	78
Figura 14 – Distribuição dos volumes observados pelas medidas de textura: (A) CON banda 1; (B) COR banda 1; (C) ENE banda 1; (D) HOM banda 1; (E) CON banda 2; (F) COR banda 2; (G) ENE banda 2; e (H) HOM banda 2.....	80
Figura 15 – Distribuição dos volumes observados pelas medidas de textura: (A) CON banda 3; (B) COR banda 3; (C) ENE banda 3; (D) HOM banda 3; (E) CON banda 4; (F) COR banda 4; (G) ENE banda 4; e (H) HOM banda 4.....	81
Figura 16 – Distribuição dos volumes observados pelos índices obtidos da textura: (A) MULCON banda 1; (B) MULCOR banda 1; (C) MULENE banda 1; (D) MULHOM banda 1; (E) MULCON banda 2; (F) MULCOR banda 2; (G) MULENE banda 2; e (H) MULHOM banda 2.....	82
Figura 17 – Distribuição dos volumes observados pelos índices obtidos da textura: (A) MULCON banda 3; (B) MULCOR banda 3; (C) MULENE banda 3; (D) MULHOM banda 3; (E) MULCON banda 4; (F) MULCOR banda 4; (G) MULENE banda 4; e (H) MULHOM banda 4.....	83
Figura 18 – Distribuição dos volumes observados pelos índices obtidos da textura: (A) INVCON banda 1; (B) INVCOR banda 1; (C) INVENE banda 1; (D) INVHOM banda 1; (E) INVCON banda 2; (F) INVCOR banda 2; (G) INVENE banda 2; e (H) INVHOM banda 2..	84
Figura 19 – Distribuição dos volumes observados pelos índices obtidos da textura: (A) INVCON banda 3; (B) INVCOR banda 3; (C) INVENE banda 3; (D) INVHOM banda 3; (E) INVCON banda 4; (F) INVCOR banda 4; (G) INVENE banda 4; e (H) INVHOM banda 4..	85
Figura 20 – Correlação entre as variáveis selecionadas para o modelo. ....	87
Figura 21 – Semivariograma ajustado ao modelo exponencial para KDE.....	90
Figura 22 – Semivariograma ajustado ao modelo exponencial para KO. ....	91
Figura 23 –Distribuição dos volumes observados pelos volumes estimados do ajuste: (A) KDE e (B) KO.....	94
Figura 24 – Erro padrão residual por classe de idade.....	95
Figura 25 – Curva ROC geradas para cada limiar da KDE e da KO. ....	96





## LISTA DE TABELAS

Tabela 1 - Tabela de contingência para modelos de classificação. ....	52
Tabela 2 - Resolução espectral das imagens obtidas pelo satélite SPOT-6.....	62
Tabela 3 - Modelos geoestatísticos ajustados ao semivariograma. ....	68
Tabela 4 - Estatísticas descritivas dos volumes totais ( $\text{m}^3 \text{ha}^{-1}$ ) por classe de idade.....	73
Tabela 5 - Estatísticas aplicadas aos valores de nível de cinza de cada banda.....	73
Tabela 6 - Parâmetros do modelo selecionado para KDE. ....	87
Tabela 7 - Parâmetros do modelo exponencial para KDE e KO. ....	91
Tabela 8 - Estatísticas dos volumes por classe de idade estimados pela KDE e KO. ....	93



## LISTA DE ABREVIATURAS, SIGLAS E SÍMBOLOS

AIC	Critério de Informação de Akaike
ARVI	Atmosphere Resistant Vegetation Index
AUC	Área sob a curva
CAP	Circunferência à altura do peito
COM	Contraste da textura
COR	Correlação da textura
DAP	Diâmetro à altura do peito
ENE	Energia da textura
GLCM	Matriz de coocorrência de níveis de cinza
HDOM	Altura dominante
HOM	Homogeneidade da textura
HT	Altura total
INVCON	Inverso do contraste da textura multiplicada pela banda
INVCOR	Inverso da correlação da textura multiplicada pela banda
INVENE	Inverso da energia da textura multiplicada pela banda
INVHOM	Inverso da homogeneidade da textura multiplicada pela banda
KO	krigagem ordinária
KDE	krigagem de deriva externa
MQO	Mínimos Quadrados Ordinários
MULCON	Contraste da textura multiplicada pela banda
MULCOR	Correlação da textura multiplicada pela banda
MULENE	Energia da textura multiplicada pela banda
MULHOM	Homogeneidade da textura multiplicada pela banda
NASA	National Aeronautics and Space Administration
NC	Nível de cinza
ND	Números digitais
NDVI	Normalized Difference Vegetation Index
ROC	Receiver Operating Characteristic
SAVI	Soil Adjusted Vegetation Index
TVP	Taxa de verdadeiros positivos
TFP	Taxa de falsos positivos



## SUMÁRIO

<b>1. INTRODUÇÃO</b> .....	23
<b>2. FUNDAMENTAÇÃO TEÓRICA</b> .....	27
2.1. INVENTÁRIO FLORESTAL .....	27
2.2. IMAGENS DE SENSORIAMENTO REMOTO .....	30
2.2.1. Aplicação na área florestal e na geoestatística .....	35
2.2.2. Segmentação e classificação .....	37
2.2.3. Índices de vegetação .....	38
2.2.4. Texturas .....	40
2.3. GEOESTATÍSTICA .....	43
2.3.1. Semivariograma .....	45
2.3.2. Modelos de semivariograma, métodos de ajuste e seleção de modelo .....	47
2.3.3. Krigagem .....	49
2.4. ANÁLISE ROC .....	50
2.4.1. Aplicações .....	51
2.4.2. Classificação .....	52
2.4.3. Curva ROC .....	53
2.4.4. Área sob a curva ROC (AUC) .....	54
<b>3. MATERIAL E MÉTODOS</b> .....	57
3.1. ÁREA DE ESTUDO E PROCESSAMENTO DA BASE DE DADOS .....	57
3.2. CARACTERIZAÇÃO E PROCESSAMENTO DAS IMAGENS .....	61
3.3. GEOESTATÍSTICA .....	67
3.4. ANÁLISE DAS ESTIMATIVAS .....	71
<b>4. RESULTADOS E DISCUSSÃO</b> .....	73
4.1. VOLUME POR PARCELA .....	73
4.2. CARACTERIZAÇÃO E PROCESSAMENTO DAS IMAGENS .....	73
4.3. GEOESTATÍSTICA .....	85
4.4. ANÁLISE DAS ESTIMATIVAS .....	92
<b>5. CONCLUSÕES</b> .....	99
REFERÊNCIAS .....	101



## 1. INTRODUÇÃO

O destaque que as florestas plantadas vêm adquirindo no Brasil e no mundo é incontestável, uma vez que existem muitos produtos oriundos dessa matéria-prima, tais como papel, celulose, cavaco, madeira maciça, óleos, entre outros. Dessa forma, o conhecimento dos recursos existentes em uma floresta plantada é essencial para uma boa gestão e manejo, possibilitando a otimização dessa matéria-prima.

Dentro deste contexto, uma das ferramentas mais utilizadas para obter informações quantitativas e qualitativas de uma floresta e suas potencialidades é o inventário florestal. De acordo com Mello et al. (2005), os métodos estatísticos aplicados ao inventário florestal tradicional são embasados na Teoria de Amostragem Clássica, porém esse método analisa todos os eventos de forma independente, isto é, a variação da característica de interesse é aleatória desconsiderando, assim, a dependência espacial.

O conhecimento do comportamento espacial de uma floresta é muito importante na análise do inventário florestal, pois esta informação pode contribuir de forma positiva na precisão dos resultados do inventário. Segundo Bognola et al. (2008), uma maneira de se conhecer as variações de um plantio florestal é a utilização de métodos estatísticos que explorem as relações existentes entre as unidades amostrais, isto é, um interpolador geoestatístico.

A geoestatística é uma técnica baseada na Teoria de Variáveis Regionalizadas, que permite descrever a variabilidade espacial de uma dada característica, isto é, a autocorrelação espacial entre as unidades amostrais e, assim, estimar valores da característica de interesse em locais não amostrados (MELLO, 2004; RUFINO et al., 2006; BOGNOLA et al., 2008).

Muitos trabalhos vêm sendo desenvolvidos com o intuito de incluir a informação da dependência espacial no inventário florestal através de um interpolador geoestatístico, aplicando esta técnica na estimativa de volume, na formação de estratos, na estimativa de índices de sítio e na continuidade espacial da variável de interesse (BIONDI et al., 1994; HOCK et al., 1994; MELLO et al., 2006; RUFINO et al., 2006; KANEGAE JUNIOR et al., 2007; ASSIS et al., 2009; MELLO et al., 2009a; RÉ et al., 2009; SANTANA, 2011; GUEDES et al., 2012).

Quando as variáveis estão relacionadas espacialmente, o uso da geoestatística permite a obtenção de um modelo mais adequado para a descrição dessas variáveis, por considerar a relação de dependência espacial. Dessa forma, é possível ter um modelo que possa descrever melhor os dados, proporcionando estimativas mais precisas da variável em localizações não medidas.

Segundo Bivand et al. (2008), a krigagem conhecida como krigagem com deriva externa (KDE), interpolador geoestatístico empregado neste estudo, utiliza, além da variável espacial, uma ou mais variáveis auxiliares capazes de explicar parte da variação observada nos dados. Porém, as variáveis auxiliares utilizadas devem possuir correlação com a variável principal, sendo que, quanto maior for a correlação maior será a contribuição da variável auxiliar na predição da característica de interesse.

Uma opção para o uso de variáveis auxiliares na KDE é a utilização de componentes de imagens de satélites. As imagens multiespectrais adquiridas por satélites têm sido muito utilizadas como auxiliares no monitoramento e mapeamento de áreas florestais. Contudo, as informações armazenadas nas imagens vão além da sua localização geográfica, possuindo informações como a quantidade de radiação refletida e emitida pela floresta com suas diferentes características. Sendo assim, a utilização de componentes dessas imagens como a radiação refletida, índices de vegetação e textura da imagem podem contribuir significativamente para a obtenção de informações sobre a variabilidade do plantio de uma determinada área.

Portanto, a inclusão da estrutura de continuidade espacial das características dendrométricas juntamente com as componentes de imagens usadas como covariável na krigagem na análise do inventário florestal, podem proporcionar resultados mais precisos e acurados, que contribuirão de maneira significativa com o manejo e com o planejamento florestal.

Dentro destas circunstâncias, o objetivo deste estudo é avaliar o desempenho da krigagem com deriva externa na estimativa do volume total com casca de povoamentos florestais de *Eucalyptus* sp., utilizando componentes de imagens de satélites como variáveis auxiliares e compará-la com o desempenho da krigagem ordinária. Especificamente, pretende-se:



- Verificar a melhor forma de extrair informações das imagens de satélites para serem utilizadas como segunda variável na geoestatística;
- Ajustar e selecionar modelos de função de semivariância para a estimativa volumétrica;
- Avaliar o uso de componentes de imagens de satélites como covariável na predição do volume de madeira para um povoamento de *Eucalyptus* sp.;
- Avaliar as estimativas dos volumes geradas pela KDE e pela KO com os valores de volumes observados.



## 2. FUNDAMENTAÇÃO TEÓRICA

### 2.1. INVENTÁRIO FLORESTAL

A alta produtividade das florestas plantadas justifica cada vez mais a importância de conhecer e quantificar os recursos florestais existentes em uma determinada área. O conhecimento e as informações sobre os recursos de uma floresta contribuem significativamente para o planejamento e a gestão das florestas plantadas, isto é, no planejamento do corte, na comercialização do produto final, na tomada de decisão acerca da viabilidade do empreendimento e na otimização de todo o processo da cadeia florestal (LEITE; ANDRADE, 2002; ASSIS et al., 2009; ALVARENGA, 2012).

As informações do estoque de madeira, como o volume de uma floresta ou outra variável dendrométrica, são obtidas por meio do inventário florestal (SCOLFORO; MELLO, 2006; MELLO et al., 2009b). Para Péllico Netto e Brena (1997), o inventário florestal tem a função de determinar e/ou estimar, a partir de uma amostra, uma característica de interesse, seja ela quantitativa ou qualitativa.

Segundo Leite e Andrade (2002), o inventário florestal pode ser definido como uma atividade que visa obter informações sobre os recursos florestais existentes em uma área pré-estabelecida, onde são feitas medições em unidades amostrais para gerar, posteriormente, estimativas para a área total. Além disso, possui o propósito de aplicar e avaliar sistemas de amostragem capazes de gerar estimativas precisas da população amostrada (SÉ et al., 2013).

O inventário florestal tem como base a teoria de amostragem clássica, onde as variações espaciais de uma determinada característica são aleatórias, isto é, independentes (MELLO, 2004; ALVARENGA, 2012). De acordo com Bolfarine e Bussad (2007), na teoria de amostragem, a população é um conjunto de indivíduos que possuem características semelhantes no qual será feito o estudo, e a amostra é o conjunto de subpopulações retiradas da população de interesse, onde as amostras devem ser representativas em relação às características que se quer estimar dessa população. Portanto, as estimativas de uma característica de interesse são obtidas a partir das unidades amostrais que seguem um delineamento amostral (MELLO, 2004).

Nos procedimentos de amostragem da teoria clássica, um valor médio de uma determinada característica, em uma região amostrada, é igual ao valor da esperança desta característica em qualquer ponto dentro desta região, com um erro de estimativa correspondente à variância dos dados amostrados (COCHRAN, 1977; TRANGMAR et al., 1985).

De acordo com Guedes (2009), a Teoria de Amostragem Clássica consiste em duas componentes, conforme a equação apresentada a seguir:

$$Z(x_i) = M(x) + \varepsilon_i \quad (1)$$

onde  $Z(x_i)$  é a variável avaliada,  $M(x)$  é uma função determinística que descreve a componente estrutural de  $Z$  e  $\varepsilon_i$  é o erro aleatório. A análise dos resíduos é fundamental, pois o que se deseja é que o erro seja aleatório, não apresentando qualquer comportamento estruturado. Portanto, a formulação desta teoria distingue somente o tamanho da variabilidade, por meio do desvio padrão, mas não distingue a estrutura da variabilidade (continuidade espacial) (MELLO, 2004; GUEDES, 2009).

Segundo Kanegae Junior (2004), devido ao princípio da casualização em que se baseia a estatística clássica, as posições das unidades amostrais no campo podem ser ignoradas, pois as variáveis são consideradas independentes e, portanto, não possuem continuidade espacial. Assim, esse método pressupõe que as variações que ocorrem numa dada característica de um local para o outro são aleatórias.

Os métodos tradicionais de estatística utilizados na análise dos inventários florestais, não consideram as possíveis correlações entre observações vizinhas, pois utilizam a média e a variância para descrever uma variável dendrométrica. Portanto, eles não exploram suficientemente as relações que possam existir entre as unidades amostrais (MELLO, 2004; MELLO et al., 2005).

A amostragem utilizada em um inventário florestal depende de quais características se quer estimar, da composição florestal da área e de muitas outras variáveis. Assim, as unidades amostrais são selecionadas conforme o procedimento de um delineamento amostral. Existem diferentes métodos de amostragem nos inventários florestais, sendo os mais comuns: amostragem casual simples, amostragem sistemática e amostragem casual estratificada. É

possível encontrar as descrições detalhadas destes procedimentos de amostragem em vários trabalhos como Cochran (1977), Péllico Netto e Brena (1997) e Scolforo e Mello (2006).

Existem diversos estudos na área florestal visando avaliar o desempenho de diferentes procedimentos de amostragem procurando otimizar a relação custo-precisão, dentre eles: Cesaro (1994), Mello e Scolforo (2000), Soares et al. (2004), Costa e Regazzi (2010), Druszcz et al. (2010) e Fick (2011). A precisão de uma estimativa está associada a sua amostragem, pois em toda amostragem existe um erro devido à fração da população não amostrada (BOLFARINE; BUSSAD, 2007). Quanto menor o erro da amostragem, melhor será a precisão das estimativas (ASSIS et al., 2009).

Segundo Mello et al. (2006), os métodos usados no inventário florestal procuram obter o menor erro para as estimativas juntamente com o menor custo possível. Assim, torna-se importante investigar métodos e processos de amostragem que permitam reduzir o custo do inventário (CESARO, 1994). Entretanto, encontrar o melhor método de amostragem para um inventário florestal é complexo, já que em cada plantio há uma variedade de fatores que não podem ser controlados e influenciam o crescimento das árvores. Portanto, para garantir o sucesso do planejamento florestal, é preciso utilizar técnicas de amostragem que consigam capturar essa variedade existente.

Em áreas reflorestadas é possível observar variações significativas ao longo do plantio (RUFINO et al., 2006). Como na teoria de amostragem clássica a caracterização do estado atual e temporal de uma variável de interesse dentro da floresta usualmente é adotada por um valor médio obtido de uma amostra, essa variabilidade espacial pode, muitas vezes, afetar de forma significativa a qualidade das estimativas feitas pelo inventário (MELLO, 2004; MELLO et al., 2005). Para Sé et al. (2013), se as correlações entre as unidades amostrais forem desprezadas, a precisão advinda desses inventários florestais será distorcida em termos de amplitude do intervalo de confiança da variável estudada. Isto é, os intervalos de confiança resultantes serão superestimados ou subestimados, dependendo da intensidade da correlação (MINGOTI; FIDELIS, 2001).

Em vários estudos, autores como Kanegae Junior (2004), Mello (2004), Assis (2005), Mello et al. (2005) e Rufino et al. (2006), identificaram a presença de uma estrutura de continuidade espacial nas características dendrométricas avaliadas. Logo, a informação do comportamento espacial das características dendrométricas é fundamental para a escolha de

um método estatístico que otimize recursos sem perda de precisão no processo de estimativa (MELLO, 2004; MELLO et al., 2005).

Uma alternativa capaz de efetuar a predição de variáveis dendrométricas com eficácia e sem acréscimo de custos em relação aos métodos tradicionais é a utilização do interpolador geoestatístico, pois este possui a capacidade de considerar a autocorrelação existente entre as parcelas a fim de aumentar a eficiência da estimativa (MELLO et al., 2009b).

## 2.2. IMAGENS DE SENSORIAMENTO REMOTO

As imagens de sensoriamento remoto são adquiridas através de sensores multiespectrais instalados em satélite. Cada sensor é projetado para fornecer dados a um determinado nível de detalhe espacial. Os sensores possuem características diferentes quanto a sua resolução espacial, radiométrica, espectral e temporal (ALVARENGA, 2012).

A resolução espacial está relacionada diretamente às dimensões e às formas dos alvos e à capacidade dos sensores de detectar estas características (PONZONI; DISPARATI, 1994). A resolução espectral determina o quanto o sensor consegue capturar do comprimento de onda do espectro da superfície terrestre ou do objeto (SILVA, 2009). A resolução temporal se refere ao período em que um sensor consegue imagear uma mesma área, ou seja, é dada pelo tempo em que o sensor consegue repetir as informações referentes ao objeto (ALVARENGA, 2012). Já a radiométrica representa o NC, usados para expressar os dados coletados pelo sensor (GONZALEZ; WOODS, 2007).

Quando a radiação solar (eletromagnética) incide sobre a superfície da atmosfera, parte dessa radiação é espalhada e/ou refletida pelas partículas atmosféricas, outra parte é absorvida e parte pode ser transmitida, caso a matéria possua alguma transparência. A partir daí, podemos observar os três fenômenos que expressam as propriedades radiométricas dos objetos ou superfície: Absortância, é a radiação absorvida pelo objeto dividida pela radiação incidente; Reflectância, é o quociente entre a radiação refletida pelo objeto dividida pela radiação incidente; e Transmitância, é a radiação transmitida pelo objeto dividida pela radiação incidente (SOUSA, 1997; MENESES, 2001; ROSA, 2003; MARCUSSI et al., 2010).

Pelo princípio da conservação da energia, segundo Sousa (1997) e Meneses (2001), a soma desses três componentes (Reflectância, Absortância e Transmitância) pode ser expressa pela equação:

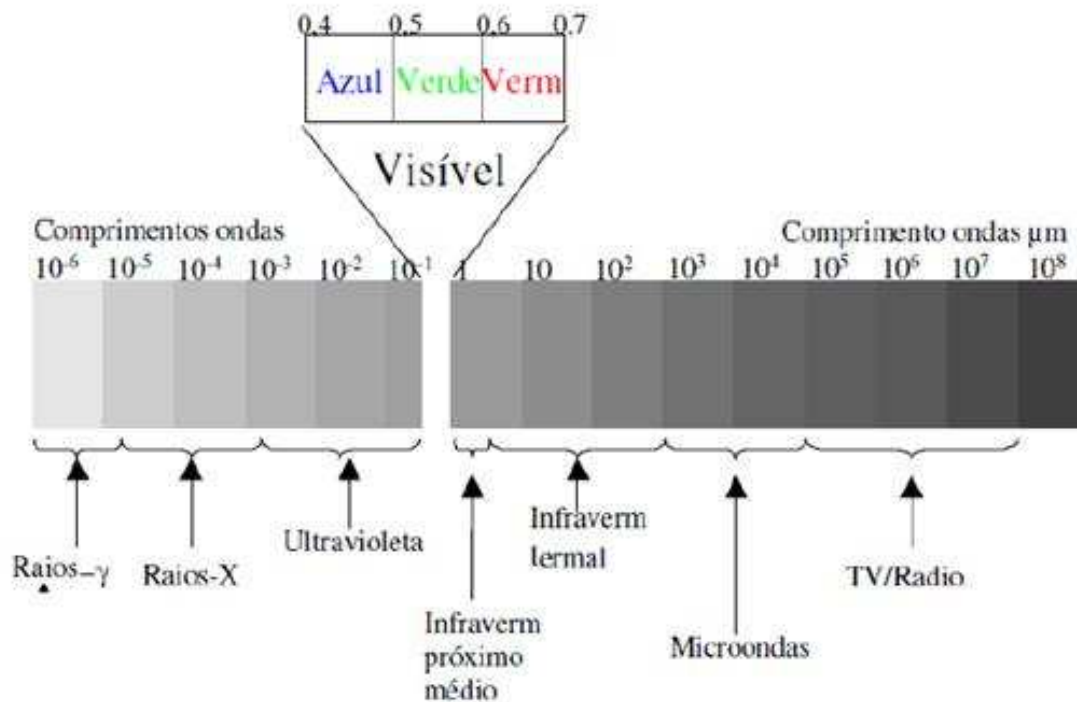
$$A + R + T = 1 \quad (2)$$

Onde as proporções de cada um dos componentes variam de uma feição para outra, dependendo do material, das condições e do comprimento de onda, permitindo a distinção das diferentes feições terrestres (LILLESAND; KIEFER, 2000). A reflectância é o componente mais utilizado e explorado em estudos com enfoque nos recursos naturais, pois seus sensores operam em regiões onde a energia refletida predomina (0,4 a 3  $\mu\text{m}$ ) (PONZONI; DISPARATI, 1994; LILLESAND; KIEFER, 2000).

A completa faixa de comprimentos de onda e de frequência da radiação eletromagnética é chamada de espectro eletromagnético (Figura 1) (ALVARENGA, 2012). De acordo com Rosendo (2005), o espectro pode ser subdividido em faixas, onde cada faixa possui características próprias em função de seus processos físicos de detecção de energia, conforme pode ser observado a seguir:

- ✓ Raios Cósmicos, raios  $\gamma$ : ( $< 0,003 - 0,4 \mu\text{m}$ );
- ✓ Raios X: ( $0,03 - 3,0 \mu\text{m}$ );
- ✓ Ultravioleta: ( $0,003 - 0,4 \mu\text{m}$ );
- ✓ Visível: ( $0,4 - 0,72 \mu\text{m}$ ), sendo o azul ( $0,45 - 0,50 \mu\text{m}$ ), o verde ( $0,50 - 0,54 \mu\text{m}$ ) e o vermelho ( $0,65 - 0,72 \mu\text{m}$ );
- ✓ Infravermelho Próximo: ( $0,72 - 1,3 \mu\text{m}$ );
- ✓ Infravermelho Médio: ( $1,3 - 4,0 \mu\text{m}$ );
- ✓ Infravermelho Distante: ( $4,0 - 300 \mu\text{m}$ );
- ✓ Microondas: ( $1,0 - 100 \text{ cm}$ );
- ✓ Ondas de Rádio: ( $> 100 \text{ cm}$ ).

Figura 1 - Espectro eletromagnético.



Fonte: Moraes (2002).

Nas imagens de sensores remotos, os dados estão distribuídos na forma de uma malha ou grade, onde cada ponto dessa grade obtido pelos sensores corresponde a uma área mínima conhecida como pixel (picture cell). Cada pixel tem sua localização definida em um sistema de coordenadas cartesianas (X,Y), no qual são registrados números digitais (ND) que indicam o nível de cinza (NC) desses pontos, ou seja, representam a intensidade de energia refletida ou emitida pelos objetos em faixas (bandas), definidas no espectro eletromagnético (CRÓSTA, 1993; LORENÇO; LANDIM, 2003; 2004).

Segundo Campbell (2002), os ND são registrados como uma sequência de dígitos binários chamados de bits. O número de bits determina a resolução radiométrica, isto é, o número de valores de brilho em uma imagem digital. Uma vez que cada bit admite apenas dois valores, e sendo os bits independentes, a resolução radiométrica é igual a  $2^b$ . Assim, sendo 0 o menor valor de ND, o seu valor máximo será  $2^b - 1$ . As resoluções radiométricas mais comuns são de 8 e 11 bits/pixel. Quanto maior é a resolução radiométrica, maior é o poder de contraste e de discriminação das imagens (BAUERMANN, 2008).

Desse modo, uma imagem digital pode ser definida por uma função bidimensional, na qual se considera a intensidade de luz refletida ou emitida em cada coordenada espacial (x,y),



onde essa intensidade é identificada por um valor inteiro, não negativo e finito, chamado de NC (GONZALEZ; WOODS, 2007; LORENÇO; LANDIM, 2003; 2004). Vale ressaltar que o ND de um pixel corresponde sempre à média da intensidade da energia refletida ou emitida pelos diferentes materiais presentes nesse pixel. A imagem digital, então, pode ser vista como uma matriz de dimensões X e Y com seus respectivos ND. No caso das imagens de sensoriamento remoto, essas matrizes possuem dimensões de até alguns milhares de linhas e de colunas (CRÓSTA, 1993).

Portanto, as imagens obtidas por satélites são compostas por intensidades de radiação eletromagnética refletida e emitida pelos objetos imageados em diferentes intervalos de comprimento de onda, onde em cada banda é adquirida uma imagem em tons de cinza. O conjunto dessas imagens em várias bandas forma uma imagem multiespectral (FLORENZANO, 2002).

Como existem diversos tipos de sensores para diferentes aplicações, cada um deles gera distintos produtos com diferentes qualidades. Por exemplo o satélite Landsat, que de acordo com Moreira (2007), possuem dois sensores: o MSS, que opera com quatro bandas, sendo duas na região do visível, uma na região do infravermelho próximo e na região termal; e o sensor TM, que opera com sete bandas, sendo três na região do visível, três na região do infravermelho refletido e uma na região do termal.

O satélite SPOT-5 também possui dois sensores, um do tipo HRG que operam na região do visível e no infravermelho próximo (bandas 1,2 e 3, comprimentos de onda centrais de 0,540  $\mu\text{m}$ , 0,650  $\mu\text{m}$  e 0,830  $\mu\text{m}$ , respectivamente) e na faixa do infravermelho (banda 4, comprimento central de 1,630  $\mu\text{m}$ ) (LISITA et al., 2013). Já o Ikonos possui um sensor que opera em cinco faixas espectrais, gerando uma imagem multiespectral (composição colorida com quatro bandas) e uma imagem pancromática (composição em nível de cinza e uma única banda) (CINTRA, 2007).

As faixas espectrais do visível, infravermelho próximo, infravermelho médio, infravermelho distante e microondas são as faixas mais utilizadas em sensoriamento remoto (MORAES, 2002). Cada objeto se comporta de maneira diferente com relação à quantidade refletida de frequência da radiação eletromagnética, isto é, cada um deles apresenta comportamentos únicos de comprimentos de onda chamados de curvas de assinatura espectral (ALVARENGA, 2012).

Em relação à vegetação, segundo os autores, na região do visível (0,4 a 0,7  $\mu\text{m}$ ) é possível captar a reflectância espectral das folhas que são influenciadas pela ação dos pigmentos fotossintetizantes (SZEKIELDA, 1988). No processo de fotossíntese, as radiações que correspondem à faixa azul (0,45 a 0,50  $\mu\text{m}$ ) e à faixa vermelha (0,65 a 0,72  $\mu\text{m}$ ) são absorvidas pela clorofila e por carotenos apresentando baixa reflectância, enquanto na faixa verde (0,50 a 0,54  $\mu\text{m}$ ) boa parte é refletida (ROSENDO, 2005; BAUERMANN, 2008).

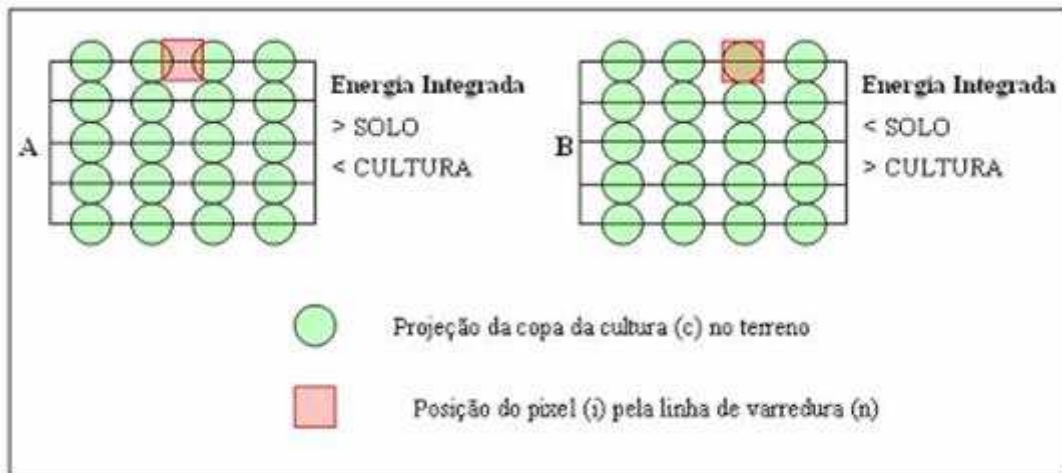
Já a faixa do infravermelho próximo (0,74 a 1,3  $\mu\text{m}$ ) é pouco afetada pela clorofila, mas muito influenciada pela estrutura foliar da vegetação alvo (BAUERMANN, 2008). Nesta região, as folhas verdes absorvem pouca energia, apresentando uma alta reflectância causada pelo espalhamento resultante dos diferentes índices de refração do líquido intracelular e dos espaços intercelulares do mesófilo. No intervalo entre 1,3 e 4  $\mu\text{m}$ , região do infravermelho médio, a reflectância das folhas é dominada por uma forte absorção pela água (ROSENDO, 2005).

As curvas de reflectância espectral dos solos são menos complexas que as da vegetação, e a característica mais marcante são os valores crescentes de reflectância com o aumento do comprimento de onda. Mesmo o solo sendo composto por uma mistura complexa de materiais, os diferentes tipos de solos possuem suas curvas espectrais com formatos semelhantes, mas com diferenças na amplitude (LILLESAND; KIEFER, 2000).

Outro fato interessante apresentado por Rosendo (2005) é com relação às análises feitas em plantios dispostos em linhas (Figura 2). No exemplo A, o pixel da linha está posicionado de forma a incorporar 50% de biomassa e 50% de solo; portanto, o valor da radiância registrada pelo sensor será uma junção da radiância do solo e da vegetação. Assim, como na faixa do vermelho a radiância do solo é maior que a radiância da vegetação, o sinal referente ao pixel será dominado pelo comportamento espectral do solo, podendo comprometer a análise do pixel.

Já no exemplo B, o pixel está praticamente com sua extensão sobre a cultura, de modo que a resposta registrada pelo sensor referente àquele pixel será dominada pelo comportamento espectral da vegetação. Esse exemplo mostra a diversidade de registros de radiância que poderia ser encontrada em decorrência da gama de situações intermediárias entre pixel, vegetação e solo.

Figura 2 - Efeito da posição do pixel no terreno com relação a radiância da cultura.



Fonte: Rosendo (2005) adaptado de Novo (1989).

Portanto, o que é efetivamente medido por um sensor, oriundo de determinada vegetação, não pode ser explicado somente pelas características intrínsecas dessa vegetação, pois inclui também a interferência de vários outros parâmetros e fatores tais como: a fonte de radiação, o espalhamento atmosférico, as características tanto das folhas quanto do dossel, os teores de umidade do solo, a interferência da reflectância do solo, sombra, entre outros (BORATTO; GOMIDE, 2013).

### 2.2.1. Aplicação na área florestal e na geostatística

As primeiras imagens voltadas para a área de recursos naturais da Terra foram geradas pela NASA através do programa LANDSAT na década de 70. Isso permitiu o desenvolvimento de estudos sobre delimitações de terras, planejamento de áreas urbanas, coberturas vegetais, informações meteorológicas, entre outras (CINTRA, 2007; MOREIRA, 2007).

Um novo sensor orbital foi desenvolvido no final da década de 90, o IKONOS-2, iniciando uma nova era em sensoriamento remoto devido a sua alta resolução temporal em escala global e ao seu baixo custo (SAWAYA et al., 2003). Isso permitiu uma compreensão mais detalhada tanto do padrão de cobertura do espaço como do nível de alteração deste padrão ao longo do tempo (ANTUNES, 2003).

Nos últimos anos, é notável o desenvolvimento tecnológico que os satélites têm proporcionado aos estudos na área ambiental, especialmente no que se refere à resolução espacial, definida em termos da mínima distância entre dois objetos que um sensor pode registrar como sendo distintos (ROSENDO, 2005). Pode-se destacar a utilização das imagens para detecção e monitoramento de focos de incêndios e áreas queimadas, desmatamento, erosão, estágio sucessional, previsão do tempo, conteúdo de biomassa de uma floresta, dentre outros (CINTRA, 2009; WATZLAWICK et al., 2009; FERRAZ et al., 2014).

No caso das florestas, as imagens multiespectrais adquiridas por satélites podem ser utilizadas para o monitoramento e a avaliação da cobertura e da dinâmica dessas florestas. Embora o seu uso, na maioria das vezes, esteja associado ao mapeamento das áreas florestais, as informações que podem ser extraídas dessas imagens vão além da localização geográfica (BAUERMANN, 2008).

Diversos estudos vêm apresentando uma gama de informações extraídas das imagens de satélites, como a correlação existente entre volume de madeira em florestas de coníferas e as bandas TM/Landsat (NILSON; OLSSON, 1995), a utilização de índices de vegetação, estatística descritiva e texturas extraídas de imagens de sensores remotos em regressão simples e múltipla para o volume e idade da floresta (BAUERMANN, 2008) e também o emprego de métodos geoestatísticos com imagens digitais de satélite para estratificação de florestas em cerrado e em plantios de eucaliptos (GUEDES, 2009; ALVARENGA, 2012). Sousa (1997), utilizou imagens transformadas em imagens índice de vegetação e imagens proporção na detecção de alteração do volume de madeira de *Pinus* em uma área de desbaste. Já no trabalho de Sousa e Ponzoni (1998), utilizaram as reflectâncias das bandas TM/Landsat5, nove índices de vegetação e a razão simples entre as bandas na detecção do volume de *Pinus* após o desbaste.

No trabalho dos autores Alencar-Silva e Maillard (2011), foram utilizadas informações das bandas da imagem Ikonos2 e cinco medidas de textura de segunda ordem das bandas vermelho e infravermelho para modelos de regressão múltipla na estimativa de várias medidas alométricas como altura, volume, e outras. Gomes e Maillard (2013), estudaram a utilização das bandas de uma imagem RapidEye, índice de vegetação e cinco medidas de textura em modelos de regressão múltipla para selecionar o melhor conjunto de variáveis que explica-se a idade e características estruturais da vegetação do cerrado em diferentes estágios sucessionais. Segundo Lorenço e Landim (2003, 2004) há muitos trabalhos nacionais e

internacionais que também utilizam a geoestatística em imagens digitais para caracterizar a cobertura de solo, a distribuição de vegetação, a fitossanidade das culturas e muitos outros.

### **2.2.2. Segmentação e classificação**

Antes de qualquer tipo de processamento de imagens é preciso verificar o registro e a normalização radiométrica, a fim de garantir a consistência geométrica e radiométrica, respectivamente (LIMA et al., 2007), pois as imagens de satélite, ao expressar a realidade de uma região, muitas vezes o fazem de forma plana, o que causa algumas distorções (CINTRA, 2007).

A segmentação nada mais é do que a subdivisão da imagem em porções menores de acordo com as características da imagem, como: textura, NC, uniformidade do pixel, reflectância e outros (WOODCOCK et al., 1994). O agrupamento dos pixels na segmentação dependerá do grau de homogeneidade calculado, onde este delimita a máxima heterogeneidade dos objetos de interesse (DEFINIENS, 2003).

O procedimento de classificação de imagem visa categorizar todos os pixels da imagem em classes ou temas (LILLESAND; KIEFER, 2000). Segundo Bauermann (2008), normalmente as características são agrupadas em quatro grandes categorias: i) características morfológicas, que são medidas da forma dos objetos que compõem a imagem e não levam em consideração a intensidade dos pixels; ii) características cromáticas, as quais descrevem a cor ou a composição espectral da radiação emitida ou refletida dos objetos, quantificada pela intensidade dos pixels em diferentes bandas espectrais; iii) características texturais, que são medidas que caracterizam a variabilidade local das intensidades dos pixels; e iv) características estruturais ou contextuais, as quais descrevem a relação entre um ou mais objetos que compõem a imagem.

As características que aparecem com mais frequência no sensoriamento remoto são as cromáticas, utilizando os índices de vegetação (BAUERMANN, 2008). Entretanto, com o avanço tecnológico nos processamentos, a textura tem sido utilizada no emprego de técnicas geoestatísticas para análise de dependência espacial dos números digitais (LORENÇO; LANDIM, 2003; 2004; BARTELS et al., 2005).

De acordo com Alvarenga (2012), a classificação de imagens pode ser visual, baseada pixel a pixel ou baseada em objetos, podendo ser diferenciada em classificação

supervisionada, na qual o usuário define os primeiros pixels e deixa que o algoritmo classifique os outros, ou em não supervisionada, na qual o computador realiza toda a classificação. As técnicas computacionais empregadas durante a classificação tentam simular a capacidade de reconhecer padrões do sistema visual humano. Portanto, a vantagem do uso computacional para esse fim reside na capacidade que estes têm de processar um grande volume de informações presentes na imagem (MOREIRA, 2007).

### **2.2.3. Índices de vegetação**

Os índices de vegetação são um tipo de operação aritmética entre bandas espectrais de uma imagem de sensoriamento remoto, que consistem na razão de valores digitais, de reflectância ou de radiância de duas ou mais bandas, referentes à mesma área, para formar uma nova imagem, denominada imagem índice de vegetação (BAUERMAN, 2008).

Inicialmente, esses índices foram desenvolvidos com o intuito de acompanhar o desenvolvimento de culturas agrícolas, porém tiveram seu uso estendido também para a área de vegetação visando avaliar a cobertura vegetal, diagnosticar o índice de área foliar, a biomassa, a porcentagem de cobertura do solo, a atividade fotossintética, a produtividade e a colheita (SOUSA, 1997; PONZONI, 2001; ROSENDO, 2005; MARCUSSI et al., 2010).

Segundo Boratto e Gomide (2013), a modelagem dos índices de vegetação baseia-se no comportamento oposto da reflectância da vegetação na região do visível. Sendo assim, com o aumento quantitativo da vegetação, menor será a reflectância em função da absorção da radiação pelos pigmentos fotossintetizantes, e com a diminuição da densidade vegetal, maior a reflectância devido ao espalhamento nas diferentes camadas das folhas. No entanto, de acordo com Silva (2004), a modelagem não se baseia apenas no fato de que a energia refletida no vermelho e no infravermelho próximo é diretamente relacionada à atividade fotossintética da vegetação, mas também na suposição de que a utilização de duas ou mais bandas espectrais pode minimizar as principais fontes de “ruídos” que afetam a resposta da vegetação.

Os índices de vegetação são recomendados em estudos de crescimento e vigor da vegetação e também podem ser utilizados no diagnóstico de vários parâmetros biofísicos com os quais apresentam altas correlações (PONZONI, 2001). Segundo vários autores, as correlações foram encontradas com índice de área foliar, porcentagem de cobertura do solo, capacidade fotossintética, estresses, DAP, altura e volume (SOUSA, 1997; PONZONI, 2001;

ROSENDO, 2005; BAUERMANN, 2008; BORATTO; GOMIDE, 2013; FERRAZ et al., 2014).

Devido aos crescentes estudos envolvendo a aplicação dos índices de vegetação e a sua simples implementação, vários índices têm sido criados e modificados ao longo dos anos da década de 90 (SOUSA, 1997). De acordo com Rosendo (2005) e Marcussi (2010), existe diversos índices de vegetação que podem ser utilizados como: CTVI (*Corrected Transformed Vegetation Index*); NDVI (*Normalized Difference Vegetation Index*) - um dos mais utilizados; RATIO (*Ratio Vegetation Index*); RVI (*Ratio Vegetation Index*); TVI (*Transformed Vegetation Index*); e TTVI (*Thiam's Transformed Vegetation Index*).

Ultimamente, vários índices vêm sendo desenvolvidos ou aperfeiçoados com a finalidade de amenizar ou eliminar o efeito do brilho do solo quando a vegetação não é adensada. Os índices utilizados para este objetivo são: AVI (*Ashburn Vegetation Index*); DVI (*Difference Vegetation Index*); MSAVI (*Modified Soil Adjusted Vegetation Index*); NRVI (*Normalized Ratio Vegetation Index*); PVI (*Perpendicular Vegetation Index*); SAVI (*Soil Adjusted Vegetation Index*); TSAVI (*Transformed Soil Adjusted Vegetation Index*); e WDV ( *Weighted Difference Vegetation Index*) (ROSENDO, 2005; MARCUSSI et al., 2010).

Há também os índices desenvolvidos especificamente para a correção do efeito atmosférico, como fumaça, poeira e partículas de poluição: ARVI (*Atmosphere Resistant Vegetation Index*); EVI (*Enhanced Vegetation Index*); GEMI (*Global Environment Monitoring Index*); GVI (*Green Vegetation Index*); e MGVI (*Misra's Green Vegetation Index*) (ROSENDO, 2005; MARCUSSI et al., 2010).

Embora existam muitos índices de vegetação, o mais frequentemente utilizado para a caracterização da vegetação tem sido o NDVI, por oferecer bons contrastes em relação a outros elementos da paisagem. O NDVI foi desenvolvido por Rouse et al. (1974) com a finalidade de explorar o comportamento espectral da vegetação nas regiões do vermelho e do infravermelho próximo, onde a vegetação na região do vermelho possui baixa reflectância e na região do infravermelho próximo, ela é relativamente mais alta do que as outras superfícies (ROSENDO, 2005; BAUERMANN, 2008; MARCUSSI et al., 2010).

O desempenho deste índice é citado em alguns estudos na literatura. Amaral et al. (1996) encontraram uma alta correlação do NDVI com a regeneração secundária da floresta

Amazônica e também uma alta correlação com a área basal. Sims e Gamon (2003) também encontraram uma boa correlação do NDVI com o teor de umidade da planta.

O índice SAVI também vem sendo bastante utilizado para minimizar o efeito do solo sobre a resposta da vegetação. Ele foi proposto por Huete (1998) com base no NDVI, onde minimiza o brilho do solo ao incorporar uma constante de ajuste de solo, o fator  $L$  (ROSENDO, 2005; BAUERMANN, 2008). Segundo Xavier (1998), há uma relação linear inversa entre o SAVI e o diâmetro à altura do peito (DAP) e também com a altura em alguns clones de eucalipto.

Outros índices também têm sido estudados com menor frequência, como o ARVI, que em estudo de vegetação vem demonstrando melhores resultados em relação a sensibilidade à atmosfera que o NDVI assim, sendo mais indicado em estudos que envolvam áreas predominantemente ocupadas pela vegetação (BELTRAME et al., 2007).

Portanto, a escolha do índice de vegetação deve ser baseada na capacidade de informação que este consegue transmitir, isto é, a particularidade que se tem em realçar ou considerar um determinado fator de interesse por meio da imagem (SOUSA, 1997).

#### **2.2.4. Texturas**

A textura é uma importante característica para a identificação de objetos ou regiões de interesse em uma imagem, pois suas propriedades representam o arranjo estrutural das superfícies e seus relacionamentos com o ambiente ao redor, isto é, a distribuição espacial e dependência espacial entre o NC (HARALICK et al., 1973; HARALICK, 1979).

Embora a textura seja um fenômeno fácil de reconhecer e entender, é algo difícil de conceituar, não existindo uma definição única (ILEA; WHELAN, 2011; GONÇAVES, 2014). Segundo Mather (2004), na área do sensoriamento remoto a textura pode ser definida como a variabilidade de um tom em uma vizinhança, ou o padrão das relações espaciais entre o NC de cinza de pixels vizinhos. Outra definição descreve a textura como sendo um conjunto de características estatísticas ou outras propriedades locais da imagem, que sejam constantes, com pouca variação ou aproximadamente periódicas (GONZALEZ; WOODS, 2007).

A textura de imagem possui uma grande variedade de aplicações e tem sido alvo de muitas pesquisas ao longo dos anos (CHEN et al., 2016). Para Vieira (2013), os esforços têm



se concentrado no desenvolvimento de métodos computacionais que obtenham as características de imagens e interpretem seu conteúdo a partir das suas propriedades texturais.

As aplicações mais conhecidas para a textura são: análise de imagens aéreas ou de satélite (NONATO, 2010; MIYAMOTO, 2013), processamento de imagens médicas (MEDEIROS, 2001; SOLTANIAN-ZADEH et al., 2004), inspeção automática (MARTINS, 2010, PAULA FILHO, 2012), classificação de documentação (GONÇALVES, 2014), entre outras.

Na área florestal, têm sido utilizadas as texturas extraídas das imagens de sensoriamento remoto como informação auxiliar para a classificação de estágios sucessionais de vegetação (SETTE, 2009), classificação de cobertura da vegetação (JOHANSEN; PHINN, 2006; LUZ, 2009; SILVEIRA, 2013) e para a estimação de parâmetros da vegetação (HALL; SKAKUN, 2007; GOMES; MAILLARD, 2013).

Segundo Gonzalez e Woods (2007), as principais abordagens para descrever uma determinada textura são: espectral, estrutural e estatística. A abordagem espectral refere-se às propriedades baseadas no espectro de Fourier, que permite detectar a existência de padrões periódicos em imagens através da identificação de pontos máximos de espectro (GONZALEZ; WOODS, 2007). As técnicas estruturais, por outro lado, lidam com o arranjo espacial dos tons estruturais, os quais são ordenados a partir de regras sintáticas, formando padrões (GONZALEZ; WOODS, 2007; AMARAL, 2010). Já na abordagem estatística, considera-se a distribuição dos tons de cinza e o inter-relacionamento entre eles, utilizando um conjunto de medidas estatísticas locais para representar uma imagem (GONZALEZ; WOODS, 2007; AMARAL, 2010).

A abordagem estatística é uma das abordagens mais utilizadas para descrever a textura de uma imagem. De acordo com Gonçalves (2014), essa abordagem pode realizar uma contagem da ocorrência de NC relativos aos pixels da imagem e verificar o modo como os pixels com diferentes intensidades se relacionam na imagem. Inúmeras medidas podem ser extraídas da distribuição do NC em imagens, das quais podem-se citar o cálculo de média, variância, desvio padrão e energia ou entropia (MEDEIROS, 2001; MARTINS, 2010).

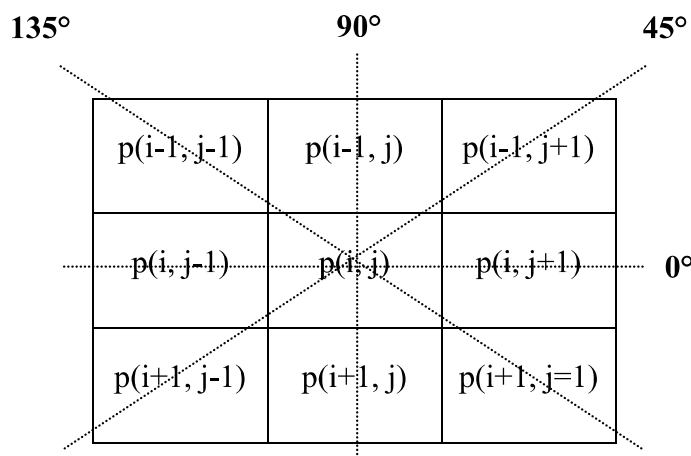
Essa abordagem pode ser dividida em duas importantes vertentes: a abordagem de primeira ordem, onde se pode extrair características a partir de histogramas de primeira

ordem, e a abordagem de segunda ordem, que considera o posicionamento relativo da ocorrência do NC (MEDEIROS, 2001; MARTINS, 2010). Segundo os mesmos autores, a principal diferença entre as duas abordagens é que, dadas duas imagens que contenham a mesma quantidade de NC, mas com posicionamento diferente, na abordagem de primeira ordem apresentarão um mesmo histograma, enquanto na abordagem de segunda ordem os histogramas serão diferentes.

Atualmente, a literatura apresenta diversas técnicas de extração de características de textura, sendo o método proposto por Haralick et al. (1973) o mais utilizado. As primeiras aplicações de medidas de texturas em imagens digitais feitas por esses autores foram embasadas em estatística de segunda ordem através da matriz de coocorrência de NC (GLCM).

Uma GLCM é uma matriz quadrada onde cada elemento  $(i, j)$  é medida da probabilidade de uma ocorrência entre dois tons de cinza separados por uma determinada distância ( $d$ ) e direção angular ( $\theta$ ) (LUZ, 2009; SETTE, 2009; MIYAMOTO, 2013; SILVEIRA, 2013). Normalmente utilizam-se distâncias entre 1 a 5, e os quatros ângulos:  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$  e  $135^\circ$  (SILVEIRA, 2013; GONÇAVES, 2014). A Figura 3 mostra a variação do pixel  $(i, j)$  com distância igual a um em relação aos seus vizinhos de acordo com as quatro direções.

Figura 3 – Variações angulares utilizadas no cálculo da matriz de coocorrência.



Fonte: Adaptado de Martins (2010).

Os quatro direcionamentos dão origem a matrizes de coocorrência distintas onde, após a construção de cada matriz, cada uma delas é então normalizada pela divisão de cada célula

pelo número total de pares (LUZ, 2009). A partir da GLCM definida é possível extrair um conjunto de 14 características de textura como: homogeneidade (ou momento diferencial inverso), contraste, segundo momento angular, energia, correlação, soma da média, variância, soma da variância, diferença da variância, entropia, soma da entropia, diferença da entropia, informação de medidas de correlação e coeficiente de correlação máxima (HARALICK et al., 1973).

### 2.3. GEOESTATÍSTICA

A geoestatística é um ramo da estatística que visa controlar parte do erro aleatório associado à localização da unidade amostral, isto é, considera que o comportamento da variável de interesse sofre influência da sua localização no espaço (GUEDES, 2009). Ao analisar a influência que essa variável sofre, é possível construir um modelo estatístico que estima o seu valor em locais não amostrados, levando em conta o comportamento espacial da variável amostrada. Existem várias técnicas para predizer esses valores, entre elas a krigagem, que utiliza a função chamada semivariograma para conhecer o comportamento dessa variação (LIMA et al., 2007; FERNANDES, 2009; ALVARENGA, 2012; RODRIGUES, 2013).

O início dos estudos da estatística espacial se deu com Krige (1951) que, trabalhando com dados de mineração de ouro em jazidas na África do Sul, concluiu que as informações dadas pela variância obtidas através da amostragem clássica eram insuficientes para explicar o fenômeno em estudo se as distâncias entre as amostras não fosse levadas em consideração (MELLO, 2004; GUEDES, 2009).

Matheron (1963), baseado nas observações de Krige (1951), desenvolveu a teoria das variáveis regionalizadas, definindo-a como uma função espacial numérica, que varia de um ponto a outro com continuidade aparente e cuja variação não pode ser representada por uma simples função matemática. A aplicação dessa teoria a problemas voltados para a geologia e mineração recebeu o nome de geoestatística (MELLO, 2004; MELLO et al., 2005; GUEDES, 2009). De acordo com Diniz (2007), a ideia desenvolvida foi a de que há influência mútua entre dois ou mais pontos, localizados numa determinada área e cuja diferença quadrática (variância) pode ser estimada apenas com a distância entre eles. Contudo, esta influência é ditada pelo comportamento da distância, deixando de existir a partir de um valor específico, conhecido como alcance.

Segundo Burrough et al. (1994), na teoria das variáveis regionalizadas a variação de uma variável qualquer pode ser dada pela soma de três componentes, sendo eles: uma componente estrutural, agregada a um valor médio constante ou a uma tendência; uma componente aleatória, espacialmente correlacionada; e um ruído aleatório ou erro residual. Se  $x$  representa uma posição em uma, duas ou três dimensões, então o valor da variável  $Z$ , em  $x$ , é dado por:

$$Z(x_i) = M(x) + \varepsilon'_x + \varepsilon'_i, \quad (3)$$

onde  $M(x)$  é a função determinística que descreve a componente estrutural  $Z$  em  $x$ ;  $\varepsilon'_x$  é o termo estocástico, que varia localmente e depende espacialmente de  $M(x)$ ; e  $\varepsilon'_i$ , é o ruído aleatório não correlacionado, com distribuição normal, média zero e variância  $\sigma^2$ . A geoestatística atua no termo estocástico  $\varepsilon'_x$  e é considerada como um grande ramo da estatística espacial, visto que estuda a relação de uma variável regionalizada com ela mesma numa outra posição (CRESSIE, 1993).

Embora a Geoestatística tenha sido muito aplicada na Geologia e na Ciência do Solo, ela apresentou grande potencial na aplicação em outras áreas da ciência, tais como agricultura, meio ambiente, indústria petrolífera, dentre outras (MELLO, 2004; RODRIGUES, 2013).

Os primeiros relatos na área florestal a respeito da continuidade espacial foram descritos por Matérn (1960), que trabalhou com a continuidade espacial de volumes de madeira oriunda de parcelas distribuídas sistematicamente na área. Já nos anos 80, o uso de métodos geoestatísticos em estudos envolvendo características dendrométricas aumentou significativamente. De acordo com Mello (2004), no estudo de Samra et al. (1989), os autores procuraram modelar o crescimento de altura para *Melia azedarach* na Índia fazendo uso de técnicas da geoestatística. A associação de técnicas geoestatísticas com o sensoriamento remoto foi trabalhada por Cochen et al. (1990), os quais procuraram caracterizar a estrutura de copa de coníferas a partir de imagens de satélite.

Atualmente, estudos envolvendo a geoestatística e a área florestal, como o de Mello (2004), constataram que as variáveis dendrométricas obtidas por meio de dois procedimentos de amostragem estão estruturadas espacialmente, não importando o procedimento adotado. Kanegae Junior (2004), trabalhando com plantios clonais de *Eucalyptus* sp., analisou a eficiência do uso da krigagem e do inverso da distância como estratificadores, em relação ao

tipo de amostragem, idade e espaçamento. No trabalho realizado por Rufino et al. (2006), foi encontrada uma dependência espacial entre as variáveis dendrométricas de povoamentos de *Eucalyptus* sp. com os atributos de solo.

A krigagem de bloco foi utilizada por Mello et al. (2006) para predição do volume de madeira por talhão em plantio de *Eucalyptus grandis*. Diniz (2007) avaliou a estrutura da dependência espacial do número de fustes e volume em eucalipto em três intensidades amostrais. No estudo conduzido por Lundgren (2011), encontrou-se uma forte correlação entre o DAP e o volume, além da já verificada dependência espacial dos dois atributos.

Segundo Bognola et al. (2008) e Rodrigues (2013), a dependência espacial encontrada nos trabalhos citados acima permite, através da modelagem, a obtenção quantitativa da variabilidade espacial da variável de interesse sobre uma determinada área e ainda estimar, sem tendência, a variância mínima em locais não amostrados. Na área florestal, o uso de técnicas geoestatísticas juntamente com as informações geográficas possibilita aprimorar o processamento dos inventários florestais, contribuindo assim para o planejamento e a otimização da produção florestal (DINIZ, 2007; BOGNOLA et al., 2008; RODRIGUES, 2013).

### **2.3.1. Semivariograma**

Na aplicação da geoestatística é necessário avaliar o grau de dependência espacial entre as amostras, sendo o semivariograma ou simplesmente variograma o método aplicado para diagnosticar a presença da correlação espacial entre as unidades amostradas (MELLO, 2004; GUEDES, 2009). Segundo Ribeiro Júnior (1995), os semivariogramas são preferidos para caracterizar a estrutura de continuidade espacial por exigirem hipótese de estacionaridade menos restritiva (Hipótese Intrínseca).

O semivariograma mostra o grau de dependência entre amostras ao longo de um suporte específico para as chamadas variáveis regionalizadas, representadas por  $Z(x)$ , em que  $x$  indica a localização da amostra, coletada em diversos pontos que podem estar regularmente distribuídos ou não. Caso haja a dependência espacial, o valor de cada ponto deverá apresentar, de algum modo, uma relação com valores obtidos no entorno deste ponto e a influência dos mesmos será tanto maior quanto menor forem as distâncias que os separam (SANTOS et al., 2011). Espera-se que as amostras mais próximas geograficamente tenham

comportamentos mais semelhantes entre si com relação a uma característica de interesse àquelas amostras mais distantes entre si (RIBEIRO JÚNIOR, 1995).

De acordo com Rodrigues (2013), a função variograma  $2\gamma(h)$  é definida como a esperança matemática do quadrado da diferença entre as variáveis  $Z(x + h)$  e  $Z(x)$  separadas por uma distância  $h$ , conforme a seguinte expressão:

$$2\gamma(h) = E\{[Z(x) - Z(x + h)]^2\} \quad (4)$$

Quando há correlação entre os pontos amostrados, é possível obter a semivariância das diferenças em cada ponto  $h$ . Segundo Mello (2004) e Guedes (2009), a função semivariograma é definida como a metade da função variograma, que representa uma função de semivariâncias em relação às respectivas distâncias. O estimador da semivariância  $\hat{\gamma}(h)$  é igual a média aritmética das diferenças entre pares de valores experimentais ao quadrado, em todos os pontos separados pela distância  $h$ .

$$\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(x) - Z(x + h)]^2 \quad (5)$$

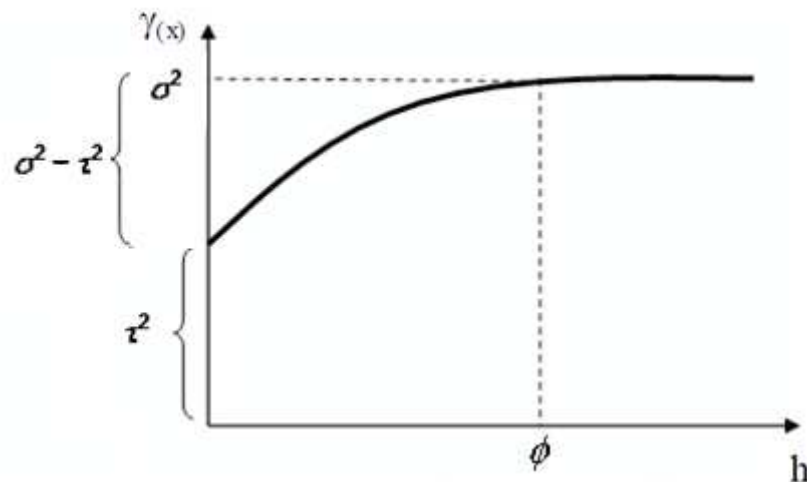
em que  $\hat{\gamma}(h)$  é a semivariância estimada para cada distância ( $h$ ),  $N(h)$  é o número de pares de pontos separados pela distância  $h$ ,  $Z(x)$  é o valor da variável regionalizada no ponto  $x$  e  $Z(x+h)$  é o valor no ponto  $x + h$ .

A função semivariograma permite gerar o gráfico da semivariância em função da distância ( $h$ ), denominado de semivariograma experimental, o qual permite interpretar a continuidade espacial da variável regionalizada (MELLO, 2004; GUEDES, 2009). Portanto, após o ajuste do modelo teórico ao semivariograma, são definidos alguns parâmetros, como o efeito pepita, o alcance e a contribuição, os quais podem ser utilizados em métodos de interpolação, a fim de gerar uma superfície contínua que represente o fenômeno de interesse (LIMA et al., 2007). Segundo Mello (2004), Diniz (2007) e Oda-Souza et al. (2008), os parâmetros do semivariograma podem ser observados na Figura 4, cujas definições são:

- ✓ Alcance ou *range* ( $\phi$ ): mostra a distância dentro da qual as amostras apresentam-se correlacionadas espacialmente.
- ✓ Patamar ( $\sigma^2$ ) ou *sill*: é o ponto em que os dados deixam de apresentar dependência espacial.

- ✓ Efeito Pepita ou *nugget* ( $\tau^2$ ): é a variação aleatória ou não estruturada do processo. Quanto menor o efeito pepita, melhor será a estimativa no processo de inferência.
- ✓ Contribuição ou *sill partial* ( $\sigma^2 - \tau^2$ ): representa o quanto de variação total pode ser explicada pelo componente espacial. É a diferença entre o patamar ( $\sigma^2$ ) e o efeito pepita ( $\tau^2$ ).

Figura 4 – Parâmetros do semivariograma.



Fonte: Adaptado de Oda-Souza (2008).

A partir da análise do semivariograma juntamente com a confirmação de dependência espacial entre as amostras, é possível interpolar valores em qualquer posição no campo de estudo, sem tendência e com variação mínima (BOGNOLA et al., 2008).

### 2.3.2. Modelos de semivariograma, métodos de ajuste e seleção de modelo

Segundo Mello (2004), há vários modelos de ajuste aplicáveis a diferentes fenômenos com continuidade espacial. Os modelos são chamados de modelos teóricos e devem fornecer soluções estáveis para o estimador linear – krigagem (geoestatístico). Para isso, as covariâncias tem que ser definidas positivamente, ou seja, a matriz de covariância utilizada na krigagem pode ser invertida. A condição de não negatividade do modelo limita o conjunto de funções usadas na modelagem do semivariograma experimental.

Os modelos básicos permissíveis de semivariogramas mais utilizados em geoestatística são: o esférico, o exponencial, o gaussiano e a família de funções de Matérn (MELLO, 2004). Por meio do comportamento do semivariograma experimental, o

pesquisador é capaz de definir o modelo que melhor descreve o comportamento dos dados (GUEDES, 2009).

Após a definição do modelo, tem-se o ajuste da função matemática ao semivariograma experimental ou aos dados. Até a década de 80, o ajuste do modelo espacial ao semivariograma experimental era chamado de ajuste a *sentimento*, o qual era feito de forma visual, sem nenhum procedimento matemático. Por meio deste, escolhe-se um modelo que será sobreposto aos pontos estimados e, a partir de uma nuvem de pontos, obtém-se a estimativa dos parâmetros do semivariograma por meio de uma avaliação visual (RIBEIRO JÚNIOR, 1995).

Posteriormente, com o aumento da capacidade dos recursos computacionais, outros métodos de ajuste, sem subjetividade, foram e continuam sendo estudados. Os ajustes podem ser feitos com funções matemáticas ao semivariograma experimental, ou ajustando-se um modelo diretamente aos dados. Dentre estes métodos, destacam-se os Mínimos Quadrados Ordinários (MQO), os Mínimos Quadrados Ponderados e a Máxima Verossimilhança (MELLO, 2004; GUEDES, 2009). Segundo Cressie (1993), esses métodos foram desenvolvidos com o intuito de retirar ou atenuar o caráter de subjetividade na estimação dos parâmetros do semivariograma de forma visual.

A avaliação do desempenho de cada modelo se dá através de várias técnicas que podem ser quantitativas, como por exemplo o critério de Informação de Akaike (AIC). O critério de AIC foi desenvolvido por Akaike (1983), que relacionou a distância de Kulback e Leibler com a Máxima Verossimilhança. Esta distância é uma medida de discrepância relativa entre o modelo verdadeiro, que descreve a relação entre a variável dependente e as diversas variáveis explanatórias, e o modelo aproximado (MELLO, 2004).

Outras técnicas também utilizadas são as de validações cruzadas e preditivas, que consistem em um meio efetivo para avaliar a qualidade das estimativas de krigagem, bem como para selecionar modelos de semivariograma (MELLO, 2004). O Teste de Razão de Verossimilhanças compara os modelos com e sem a estrutura de dependência espacial, onde a razão de verossimilhanças somente é válida para modelos sequencialmente reduzidos (ODA-SOUZA et al., 2008) ou a escolha pode ser feita simplesmente pelo modelo que obtiver o maior coeficiente de determinação ( $R^2$ ) (LUNDGREN, 2011).



Deste modo, a sensibilidade dos semivariogramas para detectar a variabilidade espacial das amostras está diretamente ligada ao melhor ajuste entre os dados experimentais e os modelos teóricos do semivariograma (ODA-SOUZA et al., 2006).

### 2.3.3. Krigagem

A krigagem é um interpolador geoestatístico que estima os valores de pontos não amostrados com base nos pontos amostrados, considerando a estrutura de dependência espacial da característica em estudo (MELLO, 2004; GUEDES, 2009; ALVARENGA et al., 2012). Ou seja, ele usa a dependência espacial entre amostras vizinhas, expressas no semivariograma, para estimar valores em qualquer posição dentro do campo, sem tendência e com variância mínima (BOGNOLA et al., 2008).

Este método foi desenvolvido por Matheron na década de 60, que deu este nome em homenagem ao pesquisador Daniel Krige (MELLO, 2004; ALVARENGA, 2012). Ele utiliza funções matemáticas para acrescentar pesos maiores nas posições mais próximas aos pontos amostrais e pesos menores nas posições mais distantes e criar, assim, os novos pontos interpolados com base nas combinações lineares de dados (BOGNOLA et al., 2008).

Segundo Landim e Sturaro (2002), a krigagem utiliza informações a partir do variograma para encontrar os pesos ótimos a serem associados às amostras com valores conhecidos, que irão estimar pontos desconhecidos. Ou seja, utiliza técnicas de análise de regressão que procura minimizar a variância estimada a partir de um modelo, levando em conta a dependência entre os dados. Nessa situação, o método fornece, além dos valores estimados, o erro associado a tal estimativa, o que o distingue dos demais algoritmos de interpolação.

As principais características da krigagem são: poder fornecer valores maiores ou menores do que os valores amostrados; usar a distância e a geometria entre todas as unidades amostrais; levar em conta a minimização do erro da variância (estimativas sem tendências com erro mínimo); determinar os pesos a partir de uma análise espacial, baseada no semivariograma; e apresentar somatório dos pesos igual a 1 (sem viés) (ZIMBACK, 2003).

A diferença da krigagem com os outros métodos de interpolação está no modo de como é feita a atribuição dos pesos nas diferentes amostras. Na interpolação linear simples, os pesos são todos iguais a  $1/N$  ( $N$  = número de amostras); no caso da interpolação baseada no

inverso do quadrado das distâncias, os pesos são definidos como: o inverso do quadrado da distância que separa o valor interpolado dos valores observados. Já na krigagem, o procedimento é parecido com a interpolação feita por média móvel ponderada porém, na krigagem os pesos são atribuídos a partir da análise do semivariograma experimental, isto é, a uma análise espacial. Além disso a krigagem buscar fornecer estimativas não tendenciosas e com variância mínima (ALVARENGA, 2012).

Existem vários tipos de krigagem, segundo Isaaks & Srivastava (1989). Dentre eles, é possível citar a krigagem simples, pontual, ordinária, preditiva, universal, de blocos, com deriva externa, cokrigagem e outras. A krigagem ordinária (KO) é uma das mais difundidas, onde a estimativa da variável de interesse em um local dependerá apenas da distância e do peso de seus vizinhos mais próximos (JOURNEL; HUIJBREGTS, 1978). Segundo Yamamoto (2001), a krigagem pontual é a estimativa ponto a ponto na área avaliada. A krigagem de bloco, de acordo com Burgess e Webster (1980) apud Mello (2004), envolve estimativas de valores da variável regionalizada para uma subárea da área total e é apropriada para situações em que a informação média é mais útil do que o valor exato no ponto (BURGESS;WEBSTER, 1980). Já a krigagem com deriva externa (KDE) é um método multivariado que utiliza uma variável auxiliar, que tenha boa correlação linear com a principal, para estimar a variável principal, porém existe a necessidade da multicolocalização dos dados (FERNANDES, 2009).

Portanto, os métodos de interpolação geoestatísticos, como a krigagem, proporcionam excelentes estimativas a partir de dados independentes. Para isso, é necessário referenciar todas as informações juntamente com a amostragem, onde cada característica terá suas informações quantitativas e/ou qualitativas associadas a um ponto no espaço (BOGNOLA et al., 2008).

#### 2.4. ANÁLISE ROC

A análise da curva *Receiver Operating Characteristic* (ROC) é uma importante ferramenta para selecionar possíveis modelos ótimos entre os modelos testados independentemente da distribuição da classe (SAMPAIO, 2009). Na análise ROC pondera-se essencialmente duas variáveis: a) uma dicotômica, que representa um evento que se encontra presente ou ausente, como por exemplo o estado de saúde; e b) uma variável contínua ou

ordinal, com o determinado ponto de corte que categoriza os indivíduos identificados como portadores ou como não portadores da doença (PEPE, 2003 apud PEREIRA, 2013).

Segundo Fawcett (2006), o gráfico ROC é um método para avaliação, organização e seleção de classificadores e/ou predições. Esse gráfico é baseado em duas quantidades que contêm toda a informação: a taxa de verdadeiros positivos (TVP) e a taxa de falsos positivos (TFP) (BRAGA, 2000; OLIVARES, 2009), retratando as compensações relativas entre benefícios (verdadeiros positivos) e custos (falsos positivos) (FAWCETT, 2006).

#### **2.4.1. Aplicações**

A ideia inicial com relação à análise ROC se deu no campo da psicologia experimental, onde um matemático, Fechner, queria avaliar a capacidade de indivíduos de distinguirem entre estímulos e não estímulos, formando assim o critério de escolha forçada (BRAGA, 2000; OLIVARES, 2009).

Contudo, foi durante a Segunda Guerra Mundial que a utilização prática da curva ROC realmente apareceu. A análise ROC foi utilizada para quantificar a habilidade dos operadores de radares (chamados originalmente de *receiver operators*) na detecção de sinal em um canal com ruído, isto é, na detecção correta de aviões através dos sinais de radar, representando o equilíbrio entre as taxas de sucesso e as taxas de falso alarme de classificadores (SWETS et al., 2000; OLIVARES, 2009; SAMPAIO, 2009; BEZERRA 2012).

Nos anos 70 a aplicação da curva ROC ganhou força na área da medicina e engenharia biomédica, onde é uma importante ferramenta para calibrar e validar modelos de diagnóstico e detecção precoce de doenças (AZEVEDO; PEREIRA, 2010; OLIVARES, 2009). Desde então, a curva ROC tem sido utilizada nas mais diversas aplicações práticas: em economia, é utilizada para validação de modelos de risco de crédito (BRITO; ASSAF NETO, 2008); em previsão do tempo, na avaliação de predições de eventos raros; e atualmente foi inserida na aprendizagem de máquinas e mineração de dados com a finalidade de se tornar uma ferramenta útil e eficiente na avaliação de modelos de classificação (OLIVARES, 2009).

### 2.4.2. Classificação

A classificação do resultado fornecido pelo modelo de precisão é necessária para a obtenção da TVP e da TFP. No caso de resultados contínuos, é preciso classificá-los de forma binária, isto é, rotulá-los como classe positiva (p) ou negativa (n) (SAMPAIO, 2009).

De acordo com Prati et al. (2008), para criar uma classificação é necessário determinar um limiar, isto é, selecionar um ponto de corte onde os valores de interesse se tornem um resultado binário, sendo as predições acima desse limiar atribuídas a uma classe e as predições abaixo desse limiar atribuídas à outra classe. A determinação do ponto de corte é feito de forma arbitrária e normalmente baseado na taxa de erro. Uma opção na determinação um limiar é a utilização de amostras de casos para as quais se conhece a classificação verdadeira (OLIVARES, 2009).

A partir da escolha do ponto de corte, é possível avaliar um modelo de classificação de maneira simples através da apresentação das estatísticas pela tabela de contingência, conhecida também como matriz de confusão. A tabela de contingência utiliza uma tabela de tabulação cruzada entre a classe prevista pelo modelo e a classe real dos exemplos, como pode ser observada na Tabela 1 (OLIVARES, 2009).

Tabela 1 - Tabela de contingência para modelos de classificação.

	Predito		
Real	VP	FN	POS
	FP	VN	NEG
	PP	PN	N

Fonte: Olivares (2009).

Na tabela de contingência, quando um exemplo positivo é classificado como positivo, ele é chamado de verdadeiro positivo (VP). Quando um exemplo negativo é classificado como positivo, ele é chamado de falso positivo (FP). Nomenclatura parecida é utilizada no caso dos exemplos classificados como negativos. Nessa tabela, *PP* e *PN* correspondem ao número de exemplos preditos como positivos e negativos e *POS* e *NEG* ao número real de exemplos positivos e negativos na amostra. *N* é o número de elementos da amostra (FAWCETT, 2006; PRATI et al., 2008; OLIVARES, 2009).

Para Bezerra (2012), as principais estatísticas obtidas pela tabela de contingência são a taxa de verdadeiros positivos, também chamada de sensibilidade, ou seja, a probabilidade do resultado ser positivo quando o evento é verdadeiro, e a taxa de verdadeiros negativos ou especificidade, que designa a probabilidade de um teste ser negativo quando o evento é negativo. Os cálculos dessas estatísticas são as seguintes:

$$\text{Sensibilidade} = TVP = \frac{TP}{PP} \quad (6)$$

$$\text{Especificidade} = 1 - TFP = 1 - \frac{FP}{PN} \quad (7)$$

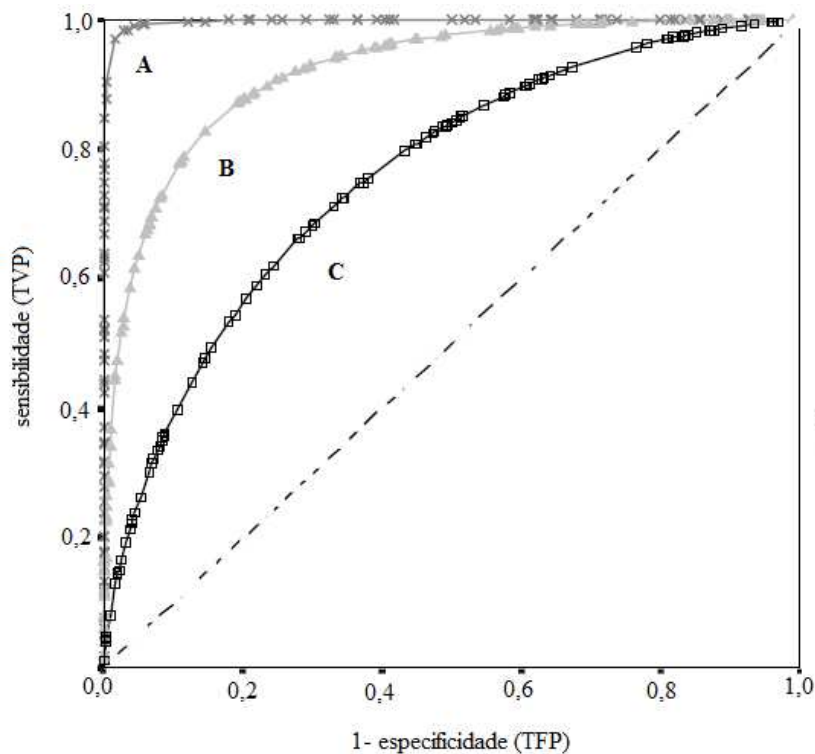
Nota-se que a sensibilidade e especificidade são dependentes e variam em função do ponto de corte. Portanto, o aumento de um dos índices implica geralmente numa diminuição do outro índice (PEREIRA, 2013).

### 2.4.3. Curva ROC

A curva ROC é representada no gráfico por pares de 1- especificidade (abscissas) e sensibilidade (ordenadas), resultantes do ponto de corte (BRAGA, 2000; PEREIRA, 2013), onde as coordenadas representam medidas de probabilidade e por esse motivo variam entre zero e um (MARINI, 2007).

Segundo Fawcett (2006) e Silva (2006), alguns classificadores produzem uma única saída discreta que indica a classe para cada elemento, gerando assim apenas um ponto no espaço ROC. Outros classificadores produzem uma probabilidade para cada elemento que representa o grau de pertinência do elemento à determinada classe. A partir de um limiar será produzido um conjunto distinto de pontos no espaço ROC. A Figura 5 mostra exemplos de curvas ROC produzidas a partir de um conjunto distinto de pontos.

Figura 5 – Três curvas ROC associadas a conjunto de pontos distintos.



Fonte: Adaptado Braga (2000).

A linha diagonal pontilhada representa um modelo de comportamento aleatório. Os pontos localizados no triângulo superior a essa diagonal representam modelos que desempenham uma previsão melhor que a aleatória; já os pontos localizados no triângulo inferior representam modelos piores que o aleatório (OLIVARES, 2009).

Ao analisar a Figura 5, de acordo com Schiavoni (2010), é fácil perceber que uma curva no espaço ROC é melhor que a outra quando pontos em que as linhas de um preditor são superiores aos pontos de outro preditor, como por exemplo a curva B em relação à C. No caso da curva A, o modelo que se encontra mais próximo à ligação da linha (0,0) para (0,1) e (0,1) para (1,1) é um modelo considerado ótimo (RESMINI et al., 2012). Já qualquer modelo que aparece no triângulo inferior direito executa pior do que a previsão aleatória (FAWCETT, 2006).

#### 2.4.4. Área sob a curva ROC (AUC)

A curva ROC permite comparar dois ou mais modelos de previsão, sendo esta uma de suas maiores virtudes (MARGOTTO, 2010). Da mesma forma, ao invés de escolhermos um

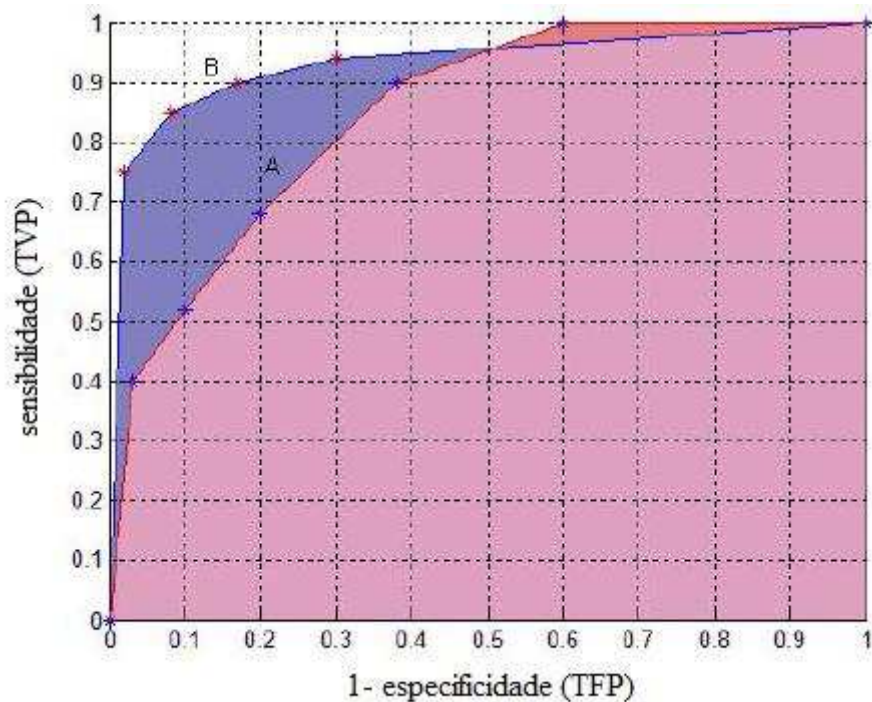
limiar e representar o desempenho desse modelo no espaço ROC, podemos simular o comportamento desse modelo para vários limiares (MARINI, 2007; OLIVARES, 2009).

Segundo Margotto (2010) e Schiavoni (2010), para verificar se há diferença significativa entre os modelos testados é preciso reduzir a curva ROC a um único valor escalar representando desempenho esperado. Um método comum para realizar esta redução é calcular a área sob a curva ROC (AUC). A AUC é um dos índices mais importante da análise ROC, já que ela é estimada levando-se em consideração todas as TVP e TFP relativas a cada um dos valores estimados (SOLTANIAN-ZADEH et al., 2004; BEZERRA 2012).

Quanto mais a curva se aproximar do canto superior esquerdo (área = 1), maior será a sua capacidade de classificação ou predição. Quando a curva se aproximar da diagonal (área = 0.5), a capacidade de classificação é nula, isto é, não é melhor que a classificação aleatória. Já quando assume valores inferiores à diagonal, esta curva apresenta capacidade de classificação contrária, não realista (SOLTANIAN-ZADEH et al., 2004; FAWCETT, 2006; MARGOTTO, 2010; PEREIRA, 2013).

A Figura 6 mostra a área sob a curva ROC de dois classificadores. O classificador B possui uma área maior e, portanto, tem um melhor desempenho médio se comparado com o classificador A. É possível que em algumas regiões do espaço ROC um classificador seja melhor que outro, como acontece com o classificador A no ponto  $TFP > 0,6$ , onde possui uma ligeira vantagem em relação ao B (FAWCETT, 2006; SILVA, 2006).

Figura 6 – Área sob a curva ROC de dois conjunto de pontos distintos.



Fonte: Adaptado Silva (2006).

Para Fawcett (2006) e Silva (2006), a AUC não pode ser o único índice a ser avaliado para selecionar os melhores classificadores, pois sem uma medida de variância não é possível comparar os classificadores. Segundo os mesmos autores, a média de curvas ROC é uma operação simples, onde é preciso projetar as curvas que estão num espaço bidimensional para o espaço unidimensional. Contudo, é preciso escolher qual projeção é a mais apropriada para se preservar a característica de interesse. As projeções mais utilizadas são: média vertical e limiar.

O método da média vertical utiliza amostras do eixo TFP e calcula a média entre as curvas ROC, dos respectivos valores do eixo TVP portanto, ela é baseada na posição dos pontos no espaço ROC. Esse método deve ser utilizado quando TFP pode ser fixado pelo pesquisador, ou quando uma simples medida de variância é desejada. Já o método limiar é baseado nas amostras geradas por diferentes limiares, onde para cada limiar acha-se o ponto correspondente em cada curva ROC e assim faz-se a média desses pontos. Mais detalhes das formas de projeção para a média da AUC se encontram nos trabalhos desenvolvidos por Fawcett (2006) e Silva (2006).

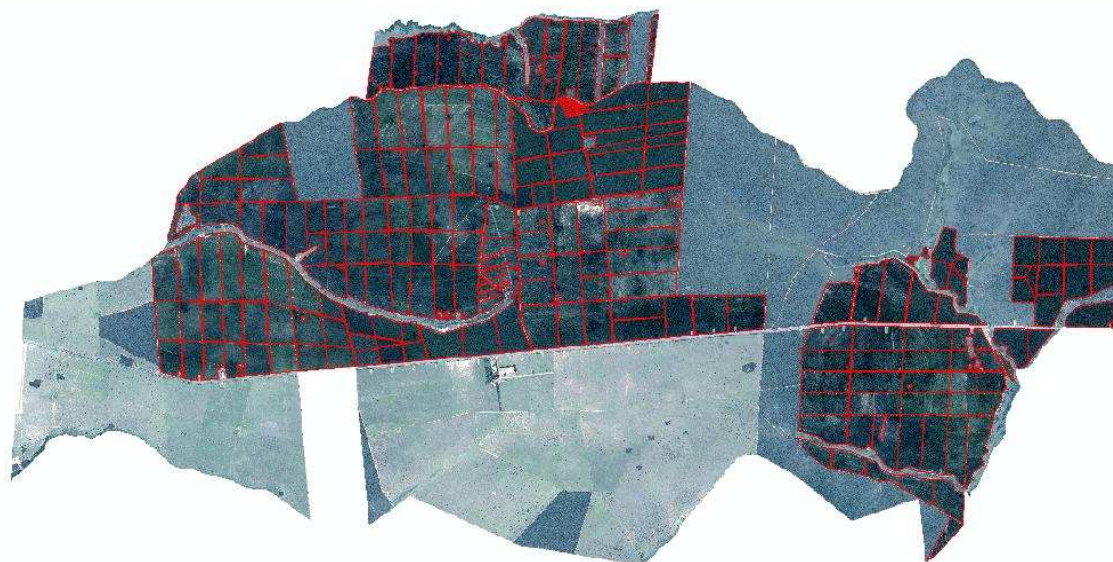


### 3. MATERIAL E MÉTODOS

#### 3.1. ÁREA DE ESTUDO E PROCESSAMENTO DA BASE DE DADOS

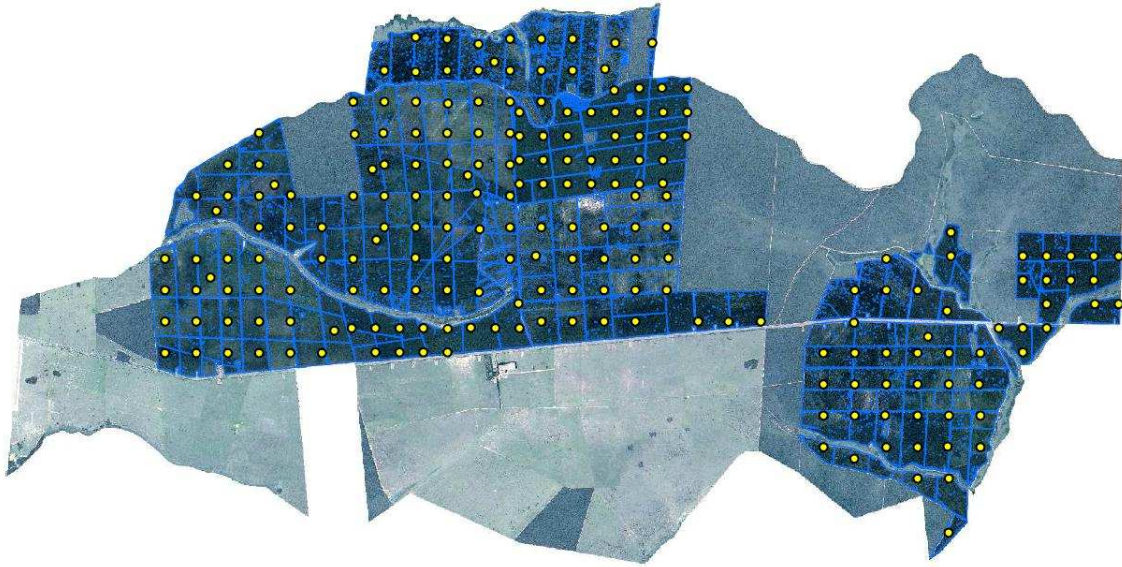
Os dados dendrométricos foram obtidos através do inventário florestal realizado em um povoamento de *Eucalyptus* sp. situado na região sul do estado do Mato Grosso do Sul (Figura 7). A altitude média da região é de 374 metros e o clima, segundo a classificação de Köppen, é do tipo *Aw*, caracterizado por um clima tropical com estação seca no inverno.

Figura 7 – Povoamento onde foram coletados os dados usados no estudo.



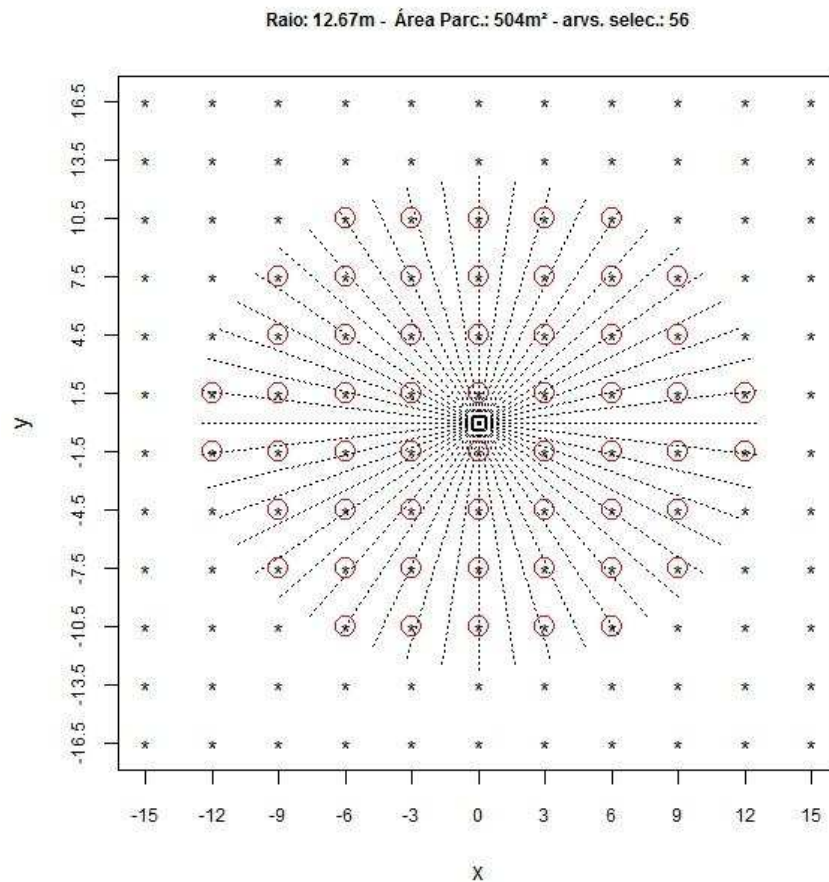
O povoamento possui uma área de 9.431,21 ha com 11 cultivares e idades que variam de 1,5 a 4 anos. De forma sistemática, foram distribuídas 210 unidades amostrais (parcelas), onde a malha da amostragem alocada na fazenda pode ser observada na Figura 8. Todas as parcelas foram georreferenciadas com a utilização de GPS de navegação a partir do ponto central da parcela. O georreferenciamento é necessário na geoestatística, tanto para conhecer a distribuição espacial das unidades amostrais como também a sua localização em relação às imagens orbitais.

Figura 8 – Distribuição das parcelas nas áreas de plantio da fazenda.



Como os plantios possuem um espaçamento de 3 x 3m, a ideia era de que as parcelas circulares possuísem em torno de 60 covas para as medições das variáveis. Foram adotadas, portanto, parcelas de 500m<sup>2</sup>, isto é, um raio de 12,61m com centro entre duas árvores da mesma linha (Figura 9).

Figura 9 – Seleção das árvores no interior das parcelas.



Em cada parcela, foi realizada a mensuração da circunferência à altura do peito (CAP), isto é, a 1,30 m do solo, de todas as árvores. Além disso, foi medida a altura total (HT) de 25% das árvores, iniciando-se pela linha central. A altura dominante (HDOM) foi coletada conforme o conceito de Assmann (1970) e todas as árvores foram categorizadas, sendo as árvores com CAP menores que 10 cm classificadas como dominadas e, assim, desconsideradas.

Além da mensuração das parcelas, também foi realizada a cubagem, sendo que as árvores a serem cubadas foram selecionadas de acordo com as classes de idade (2, 3 e 4 anos). Após a derrubada das árvores selecionadas, realizou-se medições ao longo do fuste em duas angulações, perpendiculares entre si, com uma suta nas posições longitudinais 0,10 m, 0,70 m, 1,3 m, 2 m, 4 m e assim por diante, a cada 2 m até o diâmetro mínimo de 2 cm. Foram medidos os CAPs e também as alturas totais de cada árvore.

Após a mensuração das variáveis, os dados foram analisados de acordo com sua distribuição, forma e tendência. Para a estimativa das alturas totais não mensuradas, foi realizado um ajuste por parcela do modelo hipsométrico proposto por Curtis (1967), cuja expressão é:

$$\ln(HT) = \beta_0 + \beta_1 \left( \frac{1}{DAP} \right) \pm \varepsilon_i \quad (8)$$

em que  $HT$  é a altura total da árvore (m),  $DAP$  é o diâmetro medido a 1,30 m do solo (cm),  $\ln$  é o logaritmo neperiano,  $\beta_0$  e  $\beta_1$  os parâmetros estimados e  $\varepsilon_i$  é o erro aleatório.

As estimativas dos volumes das árvores foram dadas pelo ajuste por classe de idade do polinômio de quinto grau proposto por Schopfer (1966), cuja expressão é:

$$\frac{d_i}{DAP} = \beta_0 + \beta_1 \left( \frac{h_i}{HT} \right) + \beta_2 \left( \frac{h_i}{HT} \right)^2 + \beta_3 \left( \frac{h_i}{HT} \right)^3 + \beta_4 \left( \frac{h_i}{HT} \right)^4 + \beta_5 \left( \frac{h_i}{HT} \right)^5 \pm \varepsilon_i \quad (9)$$

em que  $d_i$  é o diâmetro do fuste medido à uma altura  $h_i$  ao longo do fuste (cm),  $DAP$  é o diâmetro à altura de 1,30 m,  $h_i$  é a altura da secção,  $HT$  é a altura total,  $\beta_{0,1...5}$  são os coeficientes encontrados nos dados da cubagem e  $\varepsilon_i$  é o erro aleatório da estimativa.

Isolando ( $d_i$ ), obtém-se a função de afilamento através da qual pode-se estimar o diâmetro correspondente a qualquer altura na árvore, desde que sejam fornecidos o  $DAP$  e a altura total.

$$d_i = DAP \left[ \beta_0 + \beta_1 \left( \frac{h_i}{HT} \right) + \beta_2 \left( \frac{h_i}{HT} \right)^2 + \beta_3 \left( \frac{h_i}{HT} \right)^3 + \beta_4 \left( \frac{h_i}{HT} \right)^4 + \beta_5 \left( \frac{h_i}{HT} \right)^5 \right] \quad (10)$$

Para integrar a função e obter a expressão que permite a estimativa dos volumes, fez-se a seguinte simplificação:

$$c_0 = \beta_0; c_1 = \beta_1/HT; c_2 = \beta_2/HT^2; \dots; c_5 = \beta_5/HT^5$$

Feita a simplificação, a expressão a ser integrada assume a seguinte forma:

$$d_i = DAP * [c_0 + c_1 * h + c_2 * h^2 + c_3 * h^3 + c_4 * h^4 + c_5 * h^5] \quad (11)$$

O volume ( $V$ ) de um sólido é obtido pela integração de suas áreas seccionais ( $g_i$ ) entre o limite inferior ( $h_1$ ) e o superior ( $h_2$ ) que se deseja estabelecer. Para a determinação do volume de uma árvore, onde o interesse é o volume total, então  $h_1=0$  e  $h_2=HT$ . A representação da integral é mostrada a seguir:

$$V = \int_{h_1}^{h_2} g_i \delta h \quad \rightarrow \quad V = \int_{h_1}^{h_2} \frac{\pi * d_i^2}{40000} \delta h$$

$$V = K \int_{h_1}^{h_2} d_i^2 \delta h$$

Onde:  $K = \pi/40000$  e  $d_i$  = diâmetro correspondente a qualquer altura  $h_i$  ao longo do fuste da árvore. Então, substituindo as expressões acima, tem-se:

$$V = K * DAP^2 * \int_{h_1}^{h_2} (c_0 + c_1 h_i + c_2 h_i^2 + c_3 h_i^3 + c_4 h_i^4 + c_5 h_i^5)^2 \delta h_i \quad (12)$$

Integrando esta função, obtêm-se a expressão que fornece os sortimentos ou volumes correspondentes a qualquer porção da árvore, além do volume total.

$$V = K * DAP^2 * \left[ c_0^2 h_i + c_0 c_1 h_i^2 + \left( \frac{2}{3} c_0 c_2 + \frac{1}{3} c_1^2 \right) h_i^3 + \left( \frac{1}{2} c_0 c_3 + \frac{1}{2} c_1 c_2 \right) h_i^4 + \right. \\ \left. \left( \frac{2}{5} c_0 c_4 + \frac{2}{5} c_1 c_3 + \frac{1}{5} c_2^2 \right) h_i^5 + \left( \frac{1}{3} c_0 c_5 + \frac{1}{3} c_1 c_4 + \frac{1}{3} c_2 c_3 \right) h_i^6 + \right. \\ \left. \left( \frac{2}{7} c_1 c_5 + \frac{2}{7} c_2 c_4 + \frac{1}{7} c_3^2 \right) h_i^7 + \left( \frac{1}{4} c_2 c_5 + \frac{1}{4} c_3 c_4 \right) h_i^8 + \right. \\ \left. \left( \frac{2}{9} c_3 c_5 + \frac{1}{9} c_4^2 \right) h_i^9 + \frac{1}{5} c_4 c_5 h_i^{10} + \frac{1}{11} c_5^2 h_i^{11} \right]_{h_1}^{h_2} \quad (13)$$

### 3.2. CARACTERIZAÇÃO E PROCESSAMENTO DAS IMAGENS

Neste estudo, as imagens utilizadas como variáveis auxiliares na KDE foram obtidas no ano de 2015 a partir de um satélite comercial de observação denominado SPOT-6. As imagens do satélite SPOT-6 fusionadas com a banda Pancromática possuem uma resolução espacial de 1,5 m e uma resolução radiométrica de 12 bits por pixel, isto é, possui 4.096 variações de NC.

A resolução espectral é composta por 4 bandas que englobam os intervalos espectrais (Tabela 2). A banda pancromática possui o intervalo espectral de 450 a 745 nm.

Tabela 2 - Resolução espectral das imagens obtidas pelo satélite SPOT-6.

<b>Bandas</b>	<b>Bandas do sensor</b>	<b>Intervalo espectral (nm)</b>
<b>1</b>	Azul	455 - 525
<b>2</b>	Verde	530 - 590
<b>3</b>	Vermelho	625 - 695
<b>4</b>	Infravermelho próximo	760 - 890

A projeção adotada foi a UTM - *Universal Transverso de Mercator* - e o DATUM SIRGAS 2000 - *Sistema de Referência Geocêntrico para as Américas* -, zona 22 sul.

Devido ao tamanho da imagem e em função da capacidade computacional disponível, visando o processamento destas a resolução espacial foi reduzida, onde o tamanho do novo pixel ficou em 12 por 12 m.

Após a alteração da resolução espacial, os *shapes* das imagens foram recortados de acordo com os talhões da fazenda, onde apenas as informações das áreas de efetivo plantio constam no arquivo das imagens. A partir disso, um *raster* foi gerado contendo as informações de cada pixel, isto é, suas coordenadas (X, Y) e seus respectivos NC de cada uma das quatro bandas espectrais.

Com as informações de cada pixel, os valores médios dos NC de cada banda foi analisado estatisticamente visando conhecer o comportamento destas. Também foi calculado os índices de vegetação e a razão simples entre as bandas, isto é, banda 1 dividida pela banda 2, banda 1 pela banda 3 e assim sucessivamente (TERRA et al., 2010). Os índices de vegetação calculados neste trabalho foram:

- *NDVI*: foi definido de modo a explorar o fato de que o dossel vegetativo absorve fortemente a radiação espectral da região do vermelho e reflete intensamente na região do infravermelho próximo (BAUERMANN, 2008; MARCUSSI et al., 2010). Ele é expresso

como a razão entre a diferença e a soma das medidas nos canais infravermelho próximo e vermelho, como mostra sua expressão:

$$NDVI = \frac{IVP-VM}{IVP+VM} \quad (14)$$

Onde  $IVP$  é a energia refletida na região do infravermelho próximo e  $VM$  a energia refletida na região do vermelho do espectro eletromagnético.

- *SAVI*: proposto por Huete (1988) citado por Rosendo (2005), possui a propriedade de minimizar os efeitos do solo de fundo no sinal da vegetação ao incorporar uma constante de ajuste de solo, o fator  $L$  no denominador da equação NDVI. Para a vegetação com baixa densidade, sugere-se utilizar o fator  $L = 1.0$ ; para vegetação intermediária,  $L = 0,5$ ; para altas densidades,  $L = 0,25$ . Neste trabalho foi utilizado o fator  $L = 0,5$ , sendo a equação do SAVI é escrita da seguinte maneira:

$$SAVI = \frac{IVP-VM}{IVP+VM+L} \cdot (1 + L) \quad (15)$$

Em que  $IVP$  é a energia refletida na região do infravermelho próximo,  $VM$  é a energia refletida na região do vermelho do espectro eletromagnético e  $L$  é o fator de cobertura do solo.

- *ARVI*: neste índice, os efeitos atmosféricos do espalhamento atmosférico por aerossóis, principalmente evidenciados no vermelho, podem ser corrigidos, pixel a pixel, por meio da diferença entre as reflectâncias azul e vermelho (ROSENDO, 2005; MARCUSSI et al., 2010). A expressão é dada por:

$$ARVI = \frac{IVP+\gamma(AZ-VM)-VM}{IVP+\gamma(AZ-VM)+VM} \quad (16)$$

Em que  $IVP$  é a energia refletida na região do infravermelho próximo,  $AZ$  é a energia refletida na região do azul,  $VM$  é a energia refletida na região do vermelho e  $\gamma$  é o fator de correção, em que na falta de um modelo pra efeito de aerossol, podemos usar o valor como 1.

Com as informações das coordenadas dos pixels, os valores de NC de cada banda, razão simples entre as banda e os valores dos índices, foram extraídos os pixels para cada parcela amostral em campo através da utilização de *buffer* de 24 por 24 m, isto é, amostrando



uma média entre os pixels de 576 m<sup>2</sup>. Resumindo assim, as informações de coordenadas, o NC de cada banda e os valores dos índices para cada parcela amostrada.

As informações das imagens e dos índices de vegetação foram obtidos através do software estatístico R (R CORE TEAM, 2016), com o auxílio dos pacotes '*sp*', '*raster*' e '*mapproj*' (PEBESMA; BIVAND, 2005; HIJMANS et al., 2015; BIVAND et al., 2016).

Para a extração de características da textura, a abordagem estatística de segunda ordem foi escolhida para ser utilizada neste trabalho. Segundo Amaral (2010), essa abordagem é sugerida porque se baseia na análise de imagens e na extração das suas características através da relação espacial entre os pixels. Ela também é uma das mais utilizadas para analisar imagens naturais, pois estas imagens possuem texturas com um certo grau de aleatoriedade e são mais facilmente modeladas nessa abordagem (PAULA FILHO, 2012).

O método usado para a extração das medidas de textura foi a matriz de coocorrência (GLCM). Esse método foi proposto por Haralick et al. (1973), e de acordo com Sette (2009) e Paula Filho (2012), é o método mais utilizado para a extração de textura em imagens. A GLCM foi desenvolvida com o intuito de ser a mais genérica possível, tornando-se assim um método simples e com baixo custo computacional, o que permite que possa ser utilizado em uma vasta gama de situações. Além disso, o método geralmente apresenta bons resultados nos mais diversos tipos de problemas (SILVA, 2013).

Para selecionar as amostras de onde foram extraídas as medidas de textura, primeiramente recortaram-se as áreas referentes às parcelas nas imagens de cada uma das quatro bandas originais (1,5 x 1,5 m), isto é, para a imagem da banda 1 foram recortadas 210 imagens referentes às parcelas, e o mesmo foi feito para as outras bandas. As imagens recortadas de cada parcela possuem uma janela de 17 x 18 pixels.

O tamanho da janela é um gargalo para a construção da GLCM, pois ao contrário de outras características (brilho, cor), a textura não pode ser definida num pixel, mas sim através de uma região ou conjunto de pixels (AMARAL, 2010; PAULA FILHO, 2012). Segundo Silva (2013), quanto maior for essa janela, maior será a quantidade de probabilidades de coocorrências utilizadas nos cálculos e maior será o processamento realizado. Assim, a definição do tamanho da janela utilizada neste trabalho foi baseada na área da parcela, sendo



essa mesma definição utilizada em outros trabalhos como em Sette (2009), Alencar-Silva e Maillard (2011), Gomes e Maillard (2013).

Após a obtenção das imagens de cada parcela, o processamento para obtenção dos descritores texturais foi realizado utilizando funções do *toolbox* desenvolvidas para o software MATLAB 7.14 (R2012a). As funções utilizadas foram a *graycomatrix*, que gera a matriz GLCM, e a *graycoprops*, que gera o valor de cada uma das características para a GLCM. Para a utilização destas funções, foi necessário transformar a resolução radiométrica das imagens de 16 para 8 bits, resultando em 256 NC.

As GLCM foram calculadas para cada imagem utilizando a distância  $d=1$  e os quatro ângulos de orientação do pixel vizinho ( $0^\circ$ ,  $45^\circ$ ,  $90^\circ$  e  $135^\circ$ ) para abranger os variados tipos de relacionamentos espaciais entre os pixels, sendo por fim calculada a média dos quatro direcionamentos.

No seu estudo, Haralick et al. (1973) definiram 14 diferentes medidas de textura, mas a maioria dos trabalhos na literatura utilizam um número menor de características, pois, segundo Silva (2013), nem sempre deve-se utilizar todas as medidas de textura, o que na verdade pode piorar o desempenho do método ao invés de melhorar. As medidas de textura calculadas nesse trabalho foram:

- *Contraste (CON)*: é uma medida que representa a variação local da intensidade de um pixel e seu vizinho, ou seja, mede a presença de uma transição abrupta de NC na imagem (HARALICK et al., 1973; MEDEIROS, 2001; LUZ, 2009; AMARAL, 2010; NONATO, 2010; MIYAMOTO, 2013; SILVA, 2013; GONÇAVES, 2014). A fórmula do contraste é:

$$CON = \sum_{i=0}^{Ng} \sum_{j=0}^{Ng} (i - j)^2 P(i, j, d, \theta) \quad (17)$$

Em que  $P(i, j)$  é o conteúdo da coordenada  $(i, j)$ ,  $Ng$  é o número de NC da imagem,  $d$  é a distância e  $\theta$  é o ângulo.

- *Correlação (COR)*: representa a dependência linear no NC entre um pixel com o seu vizinho sobre toda a imagem. A comparação é realizada em todos os pixels da imagem (HARALICK et al., 1973; MEDEIROS, 2001; LUZ, 2009; AMARAL, 2010; MIYAMOTO, 2013; SILVA, 2013; GONÇAVES, 2014). Seu cálculo é dado por:

$$COR = \frac{\sum_{i=0}^{Ng} \sum_{j=0}^{Ng} ij P(i,j,d,\theta) - \mu_x \mu_y}{\sigma_x \sigma_y} \quad (18)$$

Em que  $P(i, j)$  é o conteúdo da coordenada  $(i, j)$ ,  $Ng$  é o número de NC da imagem,  $d$  é a distância,  $\theta$  é o ângulo,  $\mu_x, \mu_y$  são as médias nas direções x e y, e  $\sigma_x, \sigma_y$  são os desvios-padrão.

- *Energia (ENE)*: também é conhecida como segundo momento angular e avalia a uniformidade da textura em uma imagem, ou seja, uma medida da suavidade do local (HARALICK et al., 1973; MEDEIROS, 2001; LUZ, 2009; AMARAL, 2010; NONATO, 2010; MIYAMOTO, 2013; GONÇAVES, 2014). A energia pode ser dada por:

$$ENE = \sum_{i=0}^{Ng} \sum_{j=0}^{Ng} [P(i, j, d, \theta)]^2 \quad (19)$$

Em que  $P(i, j)$  é o conteúdo da coordenada  $(i, j)$ ,  $Ng$  é o número de NC da imagem,  $d$  é a distância e  $\theta$  é o ângulo.

- *Homogeneidade (HOM)*: conhecida também como momento diferencial inverso, ela representa a proximidade da distribuição dos elementos na GLCM com a sua diagonal, isto é, mede a regularidade presente na imagem (HARALICK et al., 1973; MEDEIROS, 2001; LUZ, 2009; AMARAL, 2010; NONATO, 2010; MIYAMOTO, 2013; GONÇAVES, 2014). A expressão é dada por:

$$HOM = \sum_{i=0}^{Ng} \sum_{j=0}^{Ng} \frac{P(i,j,d,\theta)}{1+(i-j)^2} \quad (20)$$

Em que  $P(i, j)$  é o conteúdo da coordenada  $(i, j)$ ,  $Ng$  é o número de NC da imagem,  $d$  é a distância e  $\theta$  é o ângulo.

As quatro medidas acima mencionadas estão entre as mais utilizadas e são consideradas as mais relevantes nos estudos de sensoriamento remoto (SETTE, 2009; MARTINS, 2010; SILVA, 2013; SILVEIRA, 2013). Também foram criados índices a partir das medidas de textura obtidas. Os índices MUL, são medidas de textura multiplicadas pela banda a qual foi extraída a textura exemplo, o MULCON da banda 1, onde os valores da textura COM obtidos da banda 1 foram multiplicados pelos valores da banda 1, o mesmo foi feito para todas as medidas de textura de cada banda gerando um total de 16 índices MUL. Também criou-se os índices INV, onde o inverso das medidas de textura foram multiplicadas

pelas bandas que compõem a textura exemplo, INVCON da banda 1, onde o inverso da textura COM da banda 1 foi multiplicado pela banda 1, o mesmo foi feito para todas as medidas de textura de cada banda produzindo 16 índices INV.

As médias de cada textura, os índices MUL e INV de cada textura e banda calculados para cada parcela foram tabuladas, gerando 48 novas variáveis utilizadas para a seleção do modelo geoestatístico KDE.

### 3.3. GEOESTATÍSTICA

Para as estimativas do volume da área de interesse utilizou-se a técnica de geoestatística multivariada KDE e a KO. A KDE caracteriza-se como uma metodologia capaz de integrar as informações do inventário florestal amostral obtidas em campo (parcelas) com informações de sensoriamento remoto obtidas em área total (imagens de satélite). Já a KO utiliza apenas as informações de espacialidade, atribuindo pesos distintos conforme a distância do ponto de interesse com os seus vizinhos.

Para a função do semivariograma da KDE, foi necessário ajustar um modelo de covariância através do método *Stepwise*, o qual analisou a construção de um modelo através de várias combinações entre variáveis como idade, NC de cada uma das bandas, índices de vegetação, medidas de textura e índices de textura para o volume.

O método de *Stepwise* avalia os efeitos das variáveis através da adição ou remoção de apenas uma variável em cada etapa. Para a escolha do melhor modelo de covariância da KDE encontrado pelo método de *Stepwise* utilizou-se o método de AIC. O método AIC foi desenvolvido por Akaike (1983), sendo sua equação:

$$AIC = -2L + 2p \quad (21)$$

onde  $L$  é o logaritmo neperiano do máximo da função de verossimilhança e  $p$  é o número de parâmetros deste modelo. De acordo com esse critério, o melhor modelo é aquele que apresenta o menor valor de AIC (ODA-SOUZA et al., 2008).

Após a seleção do modelo de covariância para KDE, ajustou-se aos semivariogramas da KDE e da KO, para a característica volume, os modelos teóricos mais utilizados como:

Exponencial, Esférico, Cúbico, Martén e Gaussiano, cujas formulações podem ser visualizadas na Tabela 3.

O semivariograma construído para KDE permitiu verificar a estrutura da dependência espacial do volume (m<sup>3</sup>) com as informações das bandas, dos índices de vegetação, das medidas de textura, índices de textura e com a idade.

Tabela 3 - Modelos geoestatísticos ajustados ao semivariograma.

Modelos	Forma de Ajuste
Exponencial	$\gamma(h) = \exp\{-h/\phi\} \pi/3,$
Esférico	$\gamma(h) = \exp\begin{cases} 1 - \frac{3}{2}(h/\phi) + \frac{1}{2}(h/\phi)^3, & 0 < h \leq \phi, \\ 0, & h > \phi. \end{cases}$
Cúbico	$\gamma(h) = \exp\begin{cases} 7(h/\phi)^2 - 8,75(h/\phi)^3 + 3,5(h/\phi)^5 - 0,75(h/\phi)^7, & 0 < h \leq \phi, \\ 0, & h > \phi. \end{cases}$
Matérn	$\gamma(h) = \{2^{k-1}\Gamma(k)\}^{-1}(h/\phi)^k K_k(h/\phi),$
Gaussiano	$\gamma(h) = \exp\{(-h/\phi)^2\},$

Nestas formulações,  $h$  é a distância em metros,  $\phi$  é o parâmetro alcance,  $\Gamma(k)$  é a função gama e  $K_{(k)}$  é a função Bessel do terceiro tipo de ordem  $K$ .

O método utilizado para ajustar os modelos da Tabela 3h foi o MQO. De acordo com Guedes (2009), o método dos MQO consistem em obter valores dos parâmetros de um modelo que minimizam a soma do quadrado da diferença entre os valores observados e os valores estimados. A estimativa dos parâmetros do semivariograma pelo método dos MQO foi obtida pela seguinte expressão:

$$Q(\theta) = \sum_{i=1}^k [\hat{g}(h_i) - \gamma(h_i; \theta)]^2 \quad (22)$$

em que  $\theta$  é o vetor de parâmetros estimados que define completamente o semivariograma, sendo cada estimativa denotada por  $\hat{g}(h_i)$  e  $k$  refere-se ao número de lags de distância do

semivariograma experimental. A quantidade  $\gamma(h_i; \theta)$  é a semivariância calculada pelo modelo e depende dos parâmetros estimados  $\theta$ ; após a minimização, a função  $\hat{g}(h_i)$  se torna a estimativa  $\hat{\gamma}(h_i)$  de quadrados mínimos (BUSSAB; MORETTIN, 2002).

O modelo que melhor se ajustou aos dados de acordo com a característica volume foi selecionado pelo menor valor do erro padrão residual encontrado através do ajuste do modelo não centrado para média da semivariância, onde a equação (23) representa o ajuste da regressão e a equação (24) o erro padrão residual.

$$S \sim (\beta_i \cdot \hat{S}) \pm \mathcal{E}_i \quad (23)$$

$$S_{xy} = \sqrt{\frac{\sum_{i=1}^n (S_i - \hat{S}_i)^2}{n-1}} \quad (24)$$

onde, na equação (23),  $S$  é a semivariância estimada,  $\beta_i$  é o parâmetro estimado,  $\hat{S}$  é a semivariância utilizada e  $\mathcal{E}_i$  é o erro aleatório. Na equação (24), o  $S_{xy}$  é o erro padrão residual e  $n$  é o número de amostras. O erro padrão residual mede a dispersão média entre os valores observados e estimados ao longo da linha da regressão. Quanto menor os valores do erro melhor será o ajuste (FIGUEIREDO, 2005a).

Com a escolha do melhor modelo teórico para o semivariograma, foi possível obter as predições espaciais da variável através do procedimento da KDE e da KO. O procedimento da KDE se deu segundo Wackernagel (2003), em que o estimador resulta da somatória dos valores da variável regionalizada  $Z$  nos pontos  $x_i$  multiplicados por um peso  $\lambda_i$ :

$$\hat{z}^{KED}(x_0) = \sum_{i=1}^n \lambda_i Z(x_i) \quad (25)$$

Os pesos  $\lambda_i$  provêm da resolução de um sistema de equações lineares:

$$\begin{cases} \sum_{j=1}^n \lambda_j C_r(x_1 - x_j) - \mu_1 - \mu_2 s(x_i) = C_r(x_1 - x_0) & \text{para } i = 1, n \\ \sum_{j=1}^n \lambda_j = 1 \\ \sum_{j=1}^n \lambda_j s(x_j) = s(x_0) \end{cases} \quad (26)$$

Onde:  $\hat{Z}^{KDE}$  é o estimador do valor na posição  $x_0$ ,  $n$  o número de pontos amostrais vizinhos,  $\lambda_i$  é o  $i$ -ésimo peso atribuído a cada  $i$ -ésima observação da variável na posição  $x_i$ ,  $Z(x_i)$  definido pelo semivariograma,  $\mu_1$  e  $\mu_2$  são os multiplicadores de Lagrange e  $C_r(x_i - x_j)$  é a covariância centrada dos resíduos, entre a variável primária e a secundária, da  $j$ -ésima amostra em relação à  $i$ -ésima.

A aplicação deste método considera a função variograma residual como modelo de correlação espacial. O cálculo do variograma residual exige que as variáveis primária e secundária sejam conhecidas em todos os pontos amostrais e correlacionadas entre si para que se possa obter primeiramente a reta de regressão  $V_p^* = a + bV_s$  e, posteriormente, os resíduos  $V_p^* - V_p$  (WATANABE et al., 2009).

Segundo Fernandes (2009), a KDE é aplicada no caso em que a variável principal apresenta relação de dependência com a variável auxiliar externa. Este método utiliza variáveis auxiliares para estimar um modelo de tendência. Estas variáveis devem ser conhecidas em todo o domínio onde a variável principal será estimada, ou seja, as variáveis auxiliares devem ser multicolocalizadas.

O procedimento da KO utiliza médias locais estimadas a partir dos elementos amostrais vizinhos, ou seja, o resultado da estimativa de um ponto não amostrado se dará a partir da combinação linear dos valores encontrados na vizinhança próxima (YAMAMOTO; LANDIN, 2013). O estimador da KO é dado, segundo Rocha et al. 2008, pela seguinte equação:

$$\hat{Z}^{KO}(x_0) = \sum_{i=1}^n \lambda_i Z(x_i) \quad (27)$$

Onde:  $\hat{Z}^{KO}$  é o estimador do valor na posição  $x_0$ ,  $n$  o número de pontos amostrais vizinhos,  $\lambda_i$  é o  $i$ -ésimo peso atribuído a cada  $i$ -ésima observação da variável na posição  $x_i$  e  $Z(x_i)$  definido pelo semivariograma.

Para os cálculos geoestatísticos foi empregado o software estatístico R (R CORE TEAM, 2016), juntamente com os seguintes pacotes: 'automap', 'gstat', 'geostatsp' e 'MASS' (HIEMSTRA et al., 2009; PEBESMA, 2004; BROWN, 2015; VENABES; RIPLEY, 2002).

### 3.4. ANÁLISE DAS ESTIMATIVAS

Após as predições espaciais para o volume por parcela, feitas tanto pela KDE quanto pela KO, foi analisado os volumes gerados pelos dois métodos testados. Outra análise feita para verificar a eficiência do modelo foi a análise ROC, na qual os valores estimados pela KDE e pela KO foram comparados com os valores observados, por intermédio da construção de uma curva ROC, a fim de avaliar o desempenho do procedimento de krigagem. Segundo Resmini et al. (2012), a análise da curva ROC é uma ferramenta útil para avaliar o desempenho de modelos em geral.

Para a construção da curva ROC para cada krigagem, foi necessário classificar os valores de volume estimado de forma binária. Para isso, estabeleceu-se um limiar a partir de um intervalo para cada valor de volume por parcela, onde as predições dentro desse intervalo foram atribuídas a uma classe e as predições abaixo e acima desse intervalo foram atribuídas à uma outra classe. Foram criados três intervalos a partir de uma margem de 5%, 10% e 15% em relação ao volume observado de cada parcela. As atribuições das classes foram feitas para cada intervalo tanto para as predições da KDE quanto para as predições da KO gerando, assim, três curvas com limiares distintos para cada tipo de krigagem.

As curvas ROC foram construídas utilizando o pacote *pROC* gerado pelos autores Robin et al. (2016) para o software estatístico R (R CORE TEAM, 2016). Estas curvas foram construídas colocando-se os valores da sensibilidade no eixo vertical, e o complemento da especificidade,  $1 - \text{especificidade}$ , no eixo horizontal (FORTES et al., 2004). Com o objetivo de avaliar o desempenho dos modelos utilizados, foi calculada a área sob a curva (AUC) das seis curvas construídas. A AUC avalia a habilidade do modelo de prever corretamente os valores do volume, sendo que quanto mais próximo de 1 e mais distante de 0,5, melhor será o modelo usado (FORTES et al., 2004; OLIVERA et al., 2012).

Segundo Olivares (2009), existem várias maneiras de calcular a área sob a curva, como: regra do trapézio; estimação de máxima verossimilhança; e aproximação à estatística U de Wilcoxon-Mann-Whitney. O pacote *pROC* utilizado calcula a AUC pelo método da regra do trapézio. Também foram calculados pelo mesmo pacote os intervalos de confiança da curva a 95% através do Bootstrap com 2.000 replicações.





## 4. RESULTADOS E DISCUSSÃO

### 4.1. VOLUME POR PARCELA

Por somatório, os volumes totais por parcela foram obtidos a partir das estimativas dos volumes totais das 13.407 árvores mensuradas no inventário florestal. Um resumo das estimativas por classe de idade pode ser observado na Tabela 4.

Tabela 4 - Estatísticas descritivas dos volumes totais ( $\text{m}^3 \text{ha}^{-1}$ ) por classe de idade.

<b>Classe idade</b>	<b>Mín.</b>	<b>Máx.</b>	<b>Média</b>	<b>Mediana</b>	<b>1° Q.</b>	<b>3° Q.</b>	<b>Var.</b>	<b>Desvio Pad.</b>
<b>2</b>	1,16	76,50	25,76	22,13	14,85	32,38	231,32	15,21
<b>3</b>	82,41	174,80	121,20	119,50	109,20	136,90	475,08	21,80
<b>4</b>	117,70	249,80	180,40	177,50	166,60	192,80	1060,24	32,56

Onde: Mín.= mínimo, Máx. = máximo, Q. = quartil, Var. = variância e Pad. = padrão.

### 4.2. CARACTERIZAÇÃO E PROCESSAMENTO DAS IMAGENS

As análises estatísticas realizadas com os valores das quatro bandas visam conhecer o comportamento do NC das bandas (Tabela 5). É possível observar que 25% dos maiores valores de NC da banda 1 se sobrepõe aos valores de até 50% (mediana) da banda 2. O mesmo acontece com os valores da banda 2 comparados aos valores da banda 3. A sobreposição indica que há correlação entre as bandas.

As médias encontradas mostram as diferenças entre os valores médios de NC de cada banda, que correspondem ao brilho da imagem (Tabela 5). A variância da banda 4 (73,13) é superior à variância das demais bandas, possuindo assim um contraste maior que as outras bandas. A variância é uma medida que expressa a variação dos dados em torno da média. Assim, quanto maior for a variância, maior será o contraste da imagem, ou seja, maior será o espalhamento dos NC na imagem (FIGUEIREDO, 2005b; PERROTTA, 2005).

Tabela 5 - Estatísticas aplicadas aos valores de nível de cinza de cada banda.

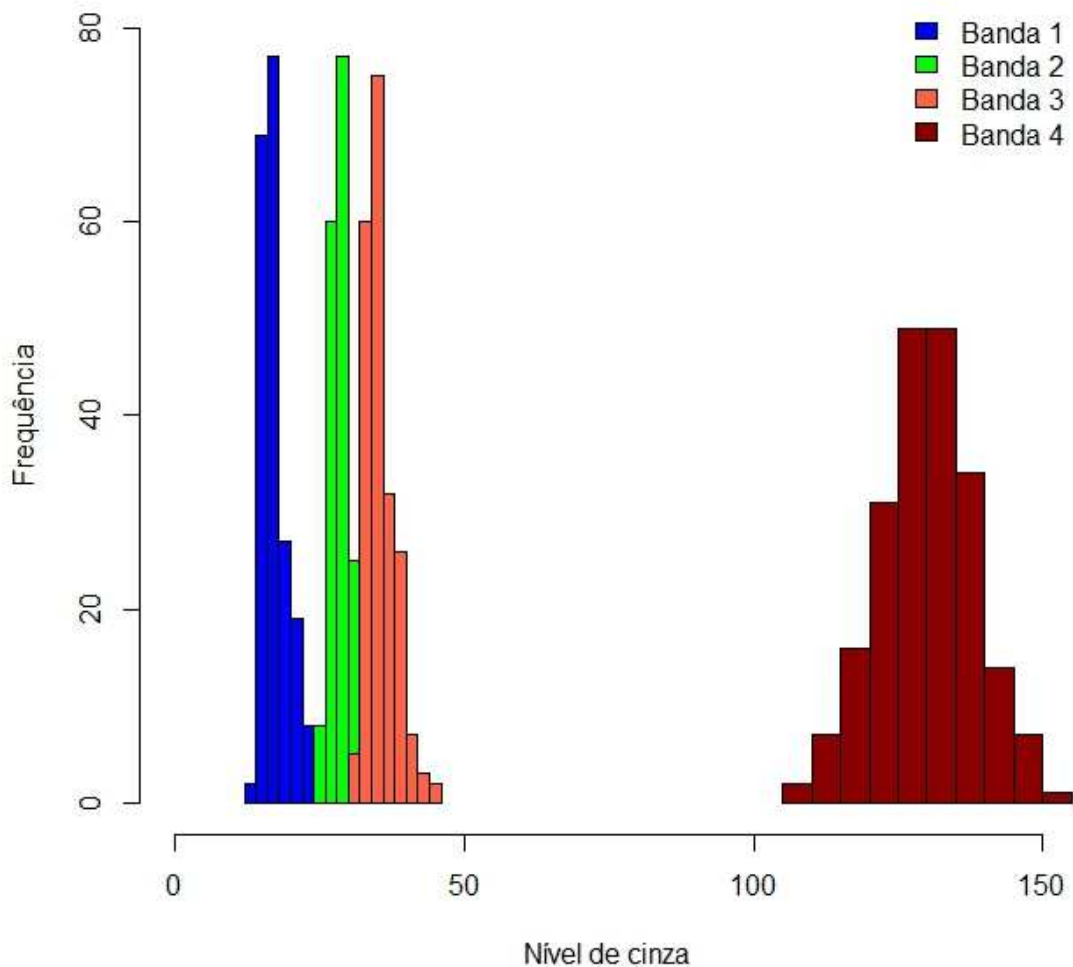
<b>Bandas</b>	<b>Mín.</b>	<b>Máx.</b>	<b>Média</b>	<b>Mediana</b>	<b>1° Q</b>	<b>3° Q</b>	<b>Var.</b>	<b>Desvio Pad.</b>
<b>B1</b>	13,89	31,96	17,49	16,58	15,72	18,53	7,720	2,78
<b>B2</b>	25,01	40,58	29,48	28,57	27,58	30,95	7,930	2,82
<b>B3</b>	31,61	44,74	35,50	34,77	33,80	36,87	6,480	2,55
<b>B4</b>	105,87	150,79	129,88	139,04	124,42	135,78	73,130	8,55

Onde: Mín.= mínimo, Máx.= máximo, Q.= quartil, Var.= variância e Pad.= padrão.

O histograma apresentado na Figura 10 representa a ocorrência dos valores de NC das quatro bandas, onde também é possível observar a ocorrência de sobreposição dos valores de NC entre as bandas encontradas. Os histogramas das bandas 1, 2 e 3 estão concentrados nos baixos valores de NC e com distribuição estreita, ou seja, baixo brilho e baixo contraste. Já a banda 4 possui uma distribuição mais ampla e está concentrada na média dos NC que variam de 0 a 256. Portanto, a banda 4 possui brilho mediano e maior contraste quando comparada com as outras bandas.

A importância de conhecer as características do histograma, em termos de processamento de imagens, é que elas ajudam a conhecer o comportamento do NC e a selecionar o tipo mais apropriado para a área específica (MENESES; ALMEIDA, 2012).

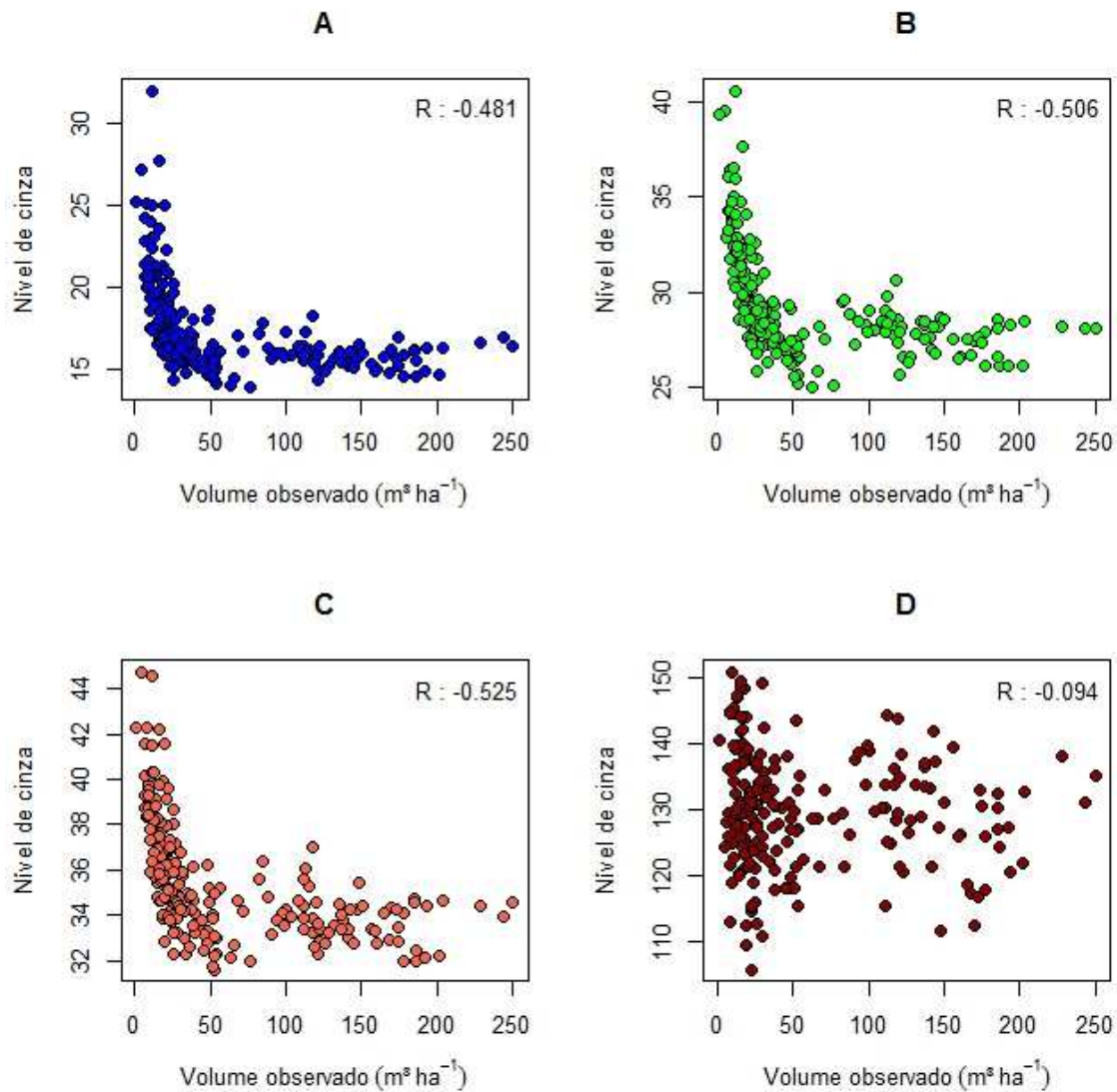
Figura 10 – Valores de nível de cinza das quatro bandas.



As distribuições dos valores de volume observado ( $\text{m}^3 \text{ha}^{-1}$ ) pelos valores de NC de cada uma das quatro bandas, pela razão simples entre as bandas e pelos índices de vegetação podem ser visualizadas nas Figuras 11, 12 e 13, respectivamente. É possível observar nos gráficos A, B e C da Figura 11, o comportamento das variáveis e a existência de correlação inversa dos NC das bandas 1, 2 e 3 com os volumes. Já no gráfico D, a correlação entre a banda 4 e o volume é praticamente nula.

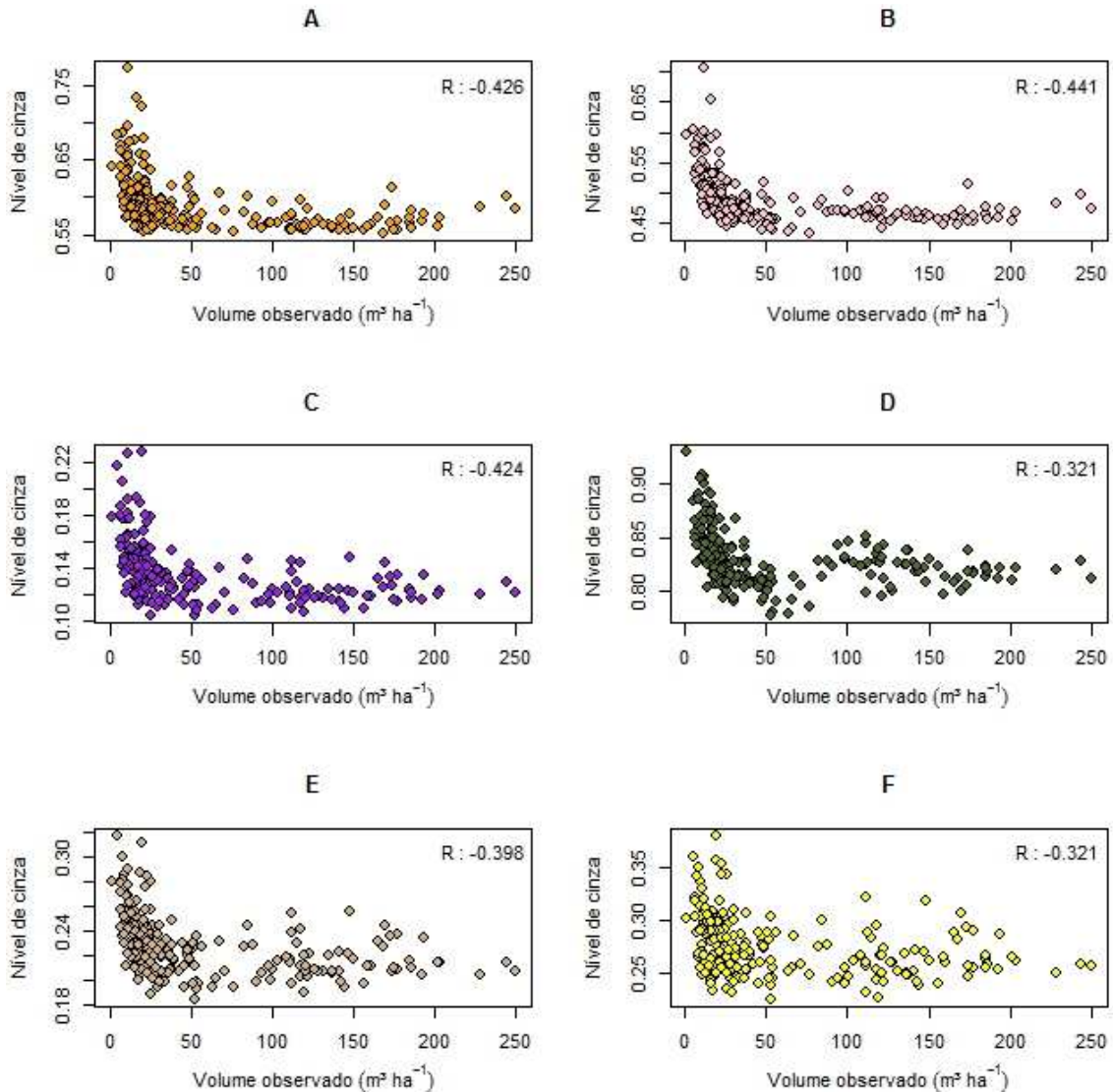
Em relação aos valores de correlação, valores próximos também foram encontrados para as bandas do satélite RapidEye's quando correlacionadas com o volume em povoamentos clonais de *Eucalyptus* sp. localizados no Estado do Espírito Santo, onde para a banda 1 a correlação foi de -0,47, banda 2 de -0,61, banda 3 de 0,56 e banda 4 de 0,26, sendo esta última a mais diferente (GASPAROTO, 2013). No trabalho de Ponzoni e Rezende (2004), as correlações também foram próximas, com exceção da banda 4. Porém, as correlações utilizaram a variável DAP em vegetação nativa localizada no Pará com imagens TM/Landsat 5. As correlações encontradas foram -0,50, -0,57, -0,42 e -0,30 para as bandas TM1, TM2, TM3 e TM4, respectivamente.

Figura 11 – Distribuição dos volumes observados pelos NC da: (A) banda 1; (B) banda 2; (C) banda 3; e (D) banda 4.



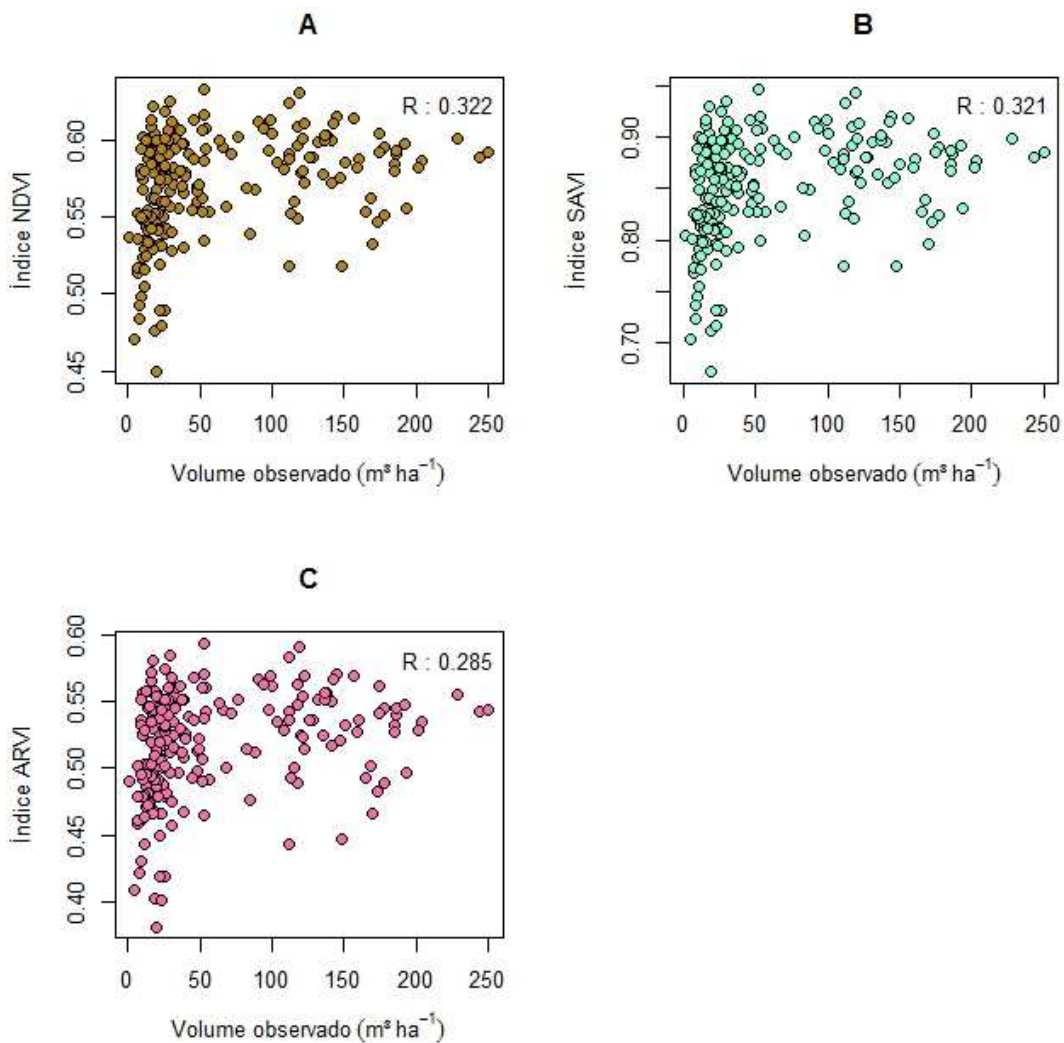
Observa-se que existe uma correlação inversa entre os valores de NC da razão simples entre as bandas com o volume observado ( $\text{m}^3 \text{ha}^{-1}$ ) em todas as razões estudadas (Figura 12). Porém, todas as correlações encontradas entre a razão simples das bandas com o volume foram menores que as correlações das bandas originais com o volume, com exceção da banda 4. Para Perrotta (2005), a razão entre as bandas é capaz de realçar intensamente as diferenças espectrais de um par de bandas.

Figura 12 – Distribuição dos volumes observados pela razão simples dos NC da: (A) banda 1/banda 2; (B) banda 1/banda 3; (C) banda 1/banda 4; (D) banda 2/banda 3; (E) banda 2/banda 4; e (F) banda 3/banda 4.



Na Figura 13, verifica-se que as correlações encontradas entre os índices e os volumes são menores que as encontradas nas Figuras 11 e 12, com exceção da banda 4. No trabalho de Gasparoto (2013), com imagens do satélite RapidEye's, foi encontrado uma correlação de 0,52 entre o NDVI e o volume de clones de eucaliptos. De acordo com Watzlawick et al. (2009), as correlações encontradas entre a biomassa de floresta de araucária com os índices NDVI e SAVI foram -0,74 e -0,77, respectivamente. Os índices foram obtidos da imagem do satélite IKONOS II. Já as correlações encontradas entre o volume de pinus com os índices obtidos de imagens TM/Landsat 5 foram -0,19 para o NDVI, -0,18 para o SAVI e 0,01 para o ARVI (SOUSA; PONZONI, 1998).

Figura 13 – Distribuição dos volumes observados pelos valores do índice: (A) NDVI; (B) SAVI; e (C) ARVI.



Os valores encontrados variaram de 0,45 a 0,63 para o NDVI, de 0,67 a 0,95 para o SAVI e de 0,38 a 0,59 para ARVI com  $L = 0,5$  (Figura 13). Segundo Lorenço e Landim (2004), o NDVI varia de 0,10 a 0,60 em cenas com vegetação em que os valores mais altos são associados com o maior vigor e a densidade da cobertura vegetal. Ele é empregado com frequência para o monitoramento da vegetação em grandes escalas porque compensa parcialmente a variação nas condições de iluminação, inclinação da superfície terrestre e aspectos das visadas do sensor em função da ampla largura da órbita.

No trabalho de Boratto e Gomide (2013), que abrangeu medições nos períodos úmido e seco, foi encontrado um NDVI variando de 0 a 0,94 na mesma imagem, onde os valores muito próximos a zero indicavam uma vegetação seca, pouca vegetação ou solo exposto. Já

em áreas irrigadas ou próximas de água com vegetação preservada, o NDVI variou de 0,48 a 0,81 em período seco e de 0,70 a 0,94 em período úmido, o qual proporcionou melhores condições no desenvolvimento da vegetação. De acordo com os mesmo autores, os valores encontrados para o índice SAVI ajustado com  $L=0,5$  foram similares ao NDVI, destacando-se a influência do solo exposto, principalmente no início da retomada do desenvolvimento da vegetação nas áreas estudadas.

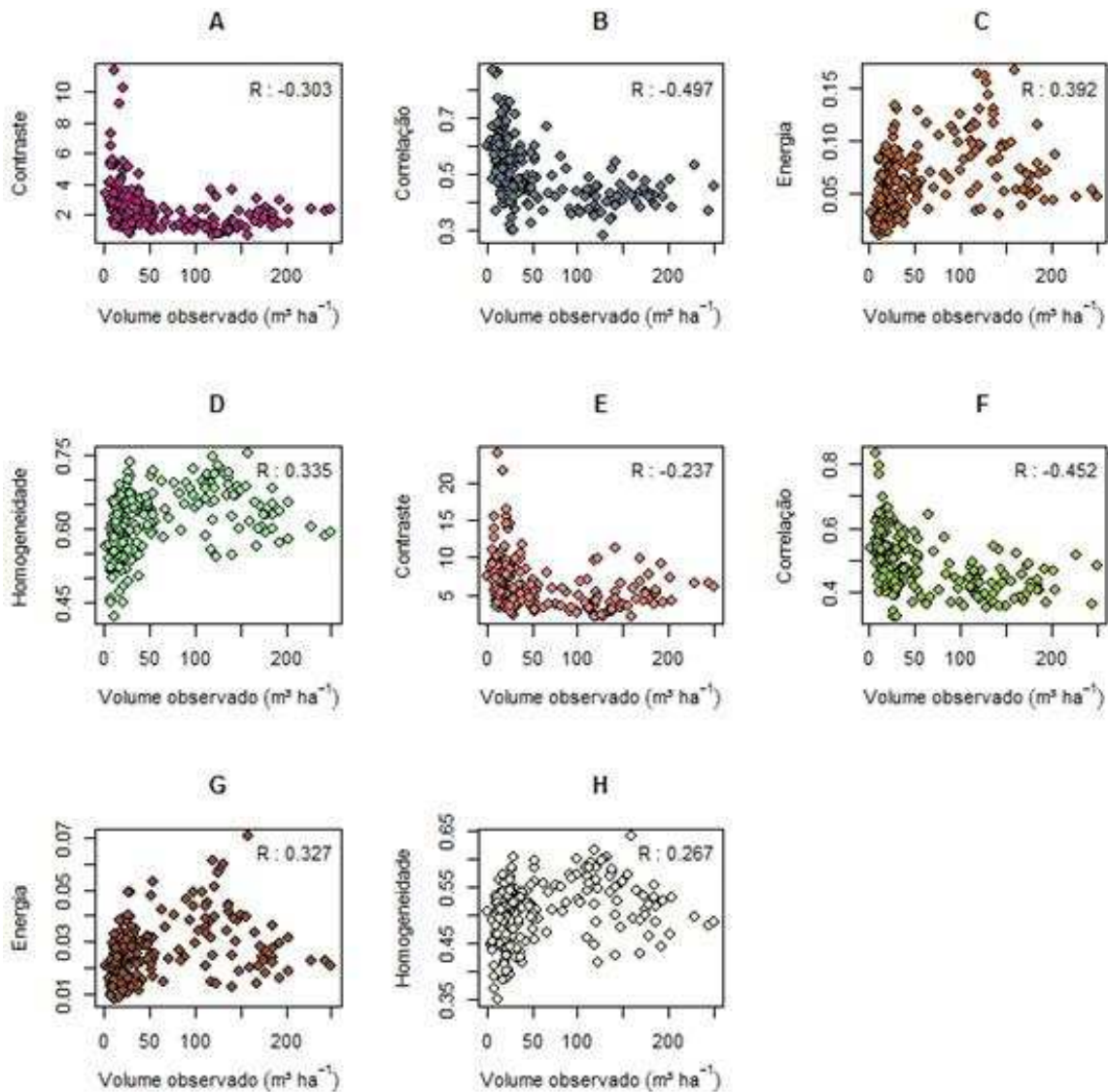
Segundo Rosendo (2005), o índice SAVI também apresentou similaridade em relação aos valores do índice NDVI, ou seja, estes índices apresentaram praticamente os mesmos resultados na avaliação do índice correlacionado com a cobertura vegetal em um bioma do cerrado localizado no estado de Minas Gerais.

A análise das imagens é algo complexo, pois a vegetação apresenta uma variedade de NC. Por sua vez, há o valor de NC referente ao solo, à água e à outras superfícies que são próprias e que contribuem no processo de modelagem das propriedades de NC da vegetação (EPIPHANIO et al., 1996).

As distribuições dos valores de volume observado ( $m^3 ha^{-1}$ ) pelas medidas de texturas (contraste, correlação, energia e homogeneidade) de cada uma das quatro bandas podem ser visualizadas nas Figuras 14 e 15. O menor valor de correlação encontrado entre as medidas de texturas com o volume foi o da homogeneidade (HOM) da banda 4 com 0,095 (Figura 15H). Já o maior valor foi o da correlação (COR) da banda 1 com -0,497 (Figura 14B). Estes valores foram diferentes dos encontrados por Sette (2009) em seu estudo sobre estágios sucessionais da Mata Atlântica no sul da Bahia, onde encontrou, entre as medidas de texturas obtidas pela banda 2 (verde) pelo satélite FORMOSAT - 2 e o volume do estágio avançado de regeneração, uma correlação de 0,571 para o COM, 0,029 para COR, 0,539 para ENE e 0,776 para HOM.



Figura 14 – Distribuição dos volumes observados pelas medidas de textura: (A) CON banda 1; (B) COR banda 1; (C) ENE banda 1; (D) HOM banda 1; (E) CON banda 2; (F) COR banda 2; (G) ENE banda 2; e (H) HOM banda 2.



Com relação às medidas de contraste (CON) obtidas pelas texturas de cada banda, a variação de contraste foi de: 0,70 - 11,33 para a banda 1; 1,09 - 24,05 para a banda 2; 2,80 - 31,98 para a banda 3; e 42,15 - 334,12 para a banda 4, a qual teve a maior variância de CON. Segundo Luz (2009), uma imagem com maior quantidade de variações locais apresentará valores de contraste mais elevados do que uma imagem com menores variações locais.

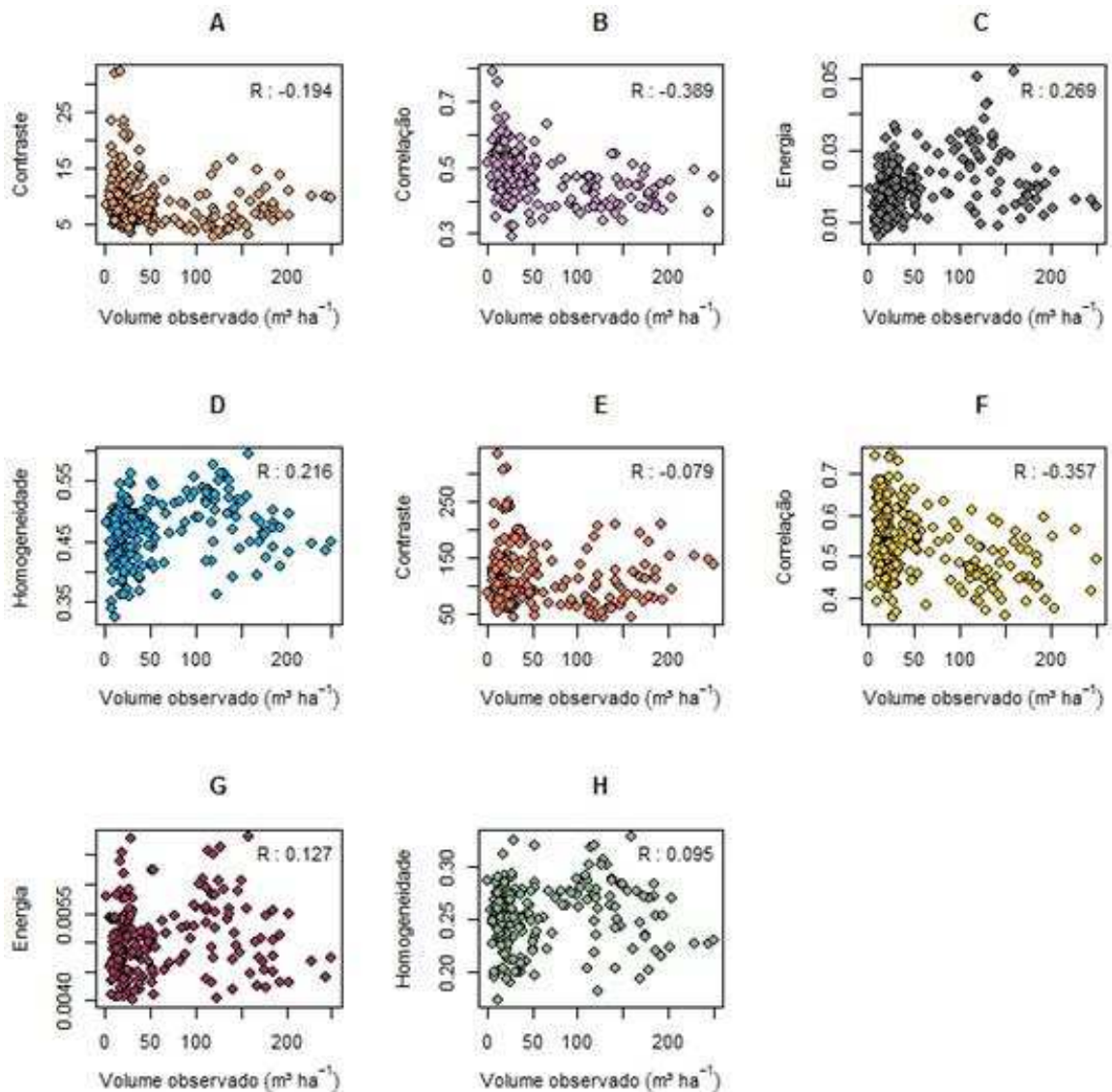
Os valores de COR encontrados para as bandas 1, 2, 3 e 4 variaram de 0,28 - 0,87, 0,32 - 0,83, 0,29 - 0,79 e 0,35 - 0,75, respectivamente. Os valores de COR podem variar de -1 a 1, sendo que quanto mais próximo de 1, maior é a dependência entre os tons de cinza da imagem; caso contrário, está descorrelacionada (HARALICK et al., 1973).



Para as bandas 1, 2, 3 e 4, os valores de ENE variaram de: 0,01 - 0,17; 0,01- 0,07; 0,01 - 0,05; e 0,00 - 0,01, respectivamente. De acordo com Haralick et al. (1973), a faixa de possíveis valores de ENE vai de 0 a 1, onde a imagem com energia próxima a 1 é uma imagem homogênea, isto é, que possui o mesmo tom de cinza em toda a sua extensão. Portanto, as imagens das parcelas possuem muitas transições de NC.

As medidas de textura HOM variaram de 0,42 - 0,75 para a banda 1, 0,35 - 0,64 para a banda 2, 0,32 - 0,60 para a banda 3 e 0,17 - 0,33 para a banda 4. A faixa de variação da HOM é de 0 a 1, onde quanto mais próximo de 1, os valores mais se aproximam da diagonal da matriz de coocorrência de tons de cinza (HARALICK et al., 1973).

Figura 15 – Distribuição dos volumes observados pelas medidas de textura: (A) CON banda 3; (B) COR banda 3; (C) ENE banda 3; (D) HOM banda 3; (E) CON banda 4; (F) COR banda 4; (G) ENE banda 4; e (H) HOM banda 4.



As distribuições dos índices MUL obtidos através das quatro medidas de texturas e das quatro bandas pelos valores de volume observado ( $\text{m}^3 \text{ha}^{-1}$ ) podem ser observadas nas Figuras 16 e 17. Os maiores valores de correlação encontrados foram de 0,520 para MULCOR banda 1 (Figura 16 B) e -0,522 MULCOR banda 2 (Figura 16 F). Já o menor valor de correlação foi de 0,04 para MULHOM banda 4 (Figura 17 H).

Figura 16 – Distribuição dos volumes observados pelos índices obtidos da textura: (A) MULCON banda 1; (B) MULCOR banda 1; (C) MULENE banda 1; (D) MULHOM banda 1; (E) MULCON banda 2; (F) MULCOR banda 2; (G) MULENE banda 2; e (H) MULHOM banda 2.

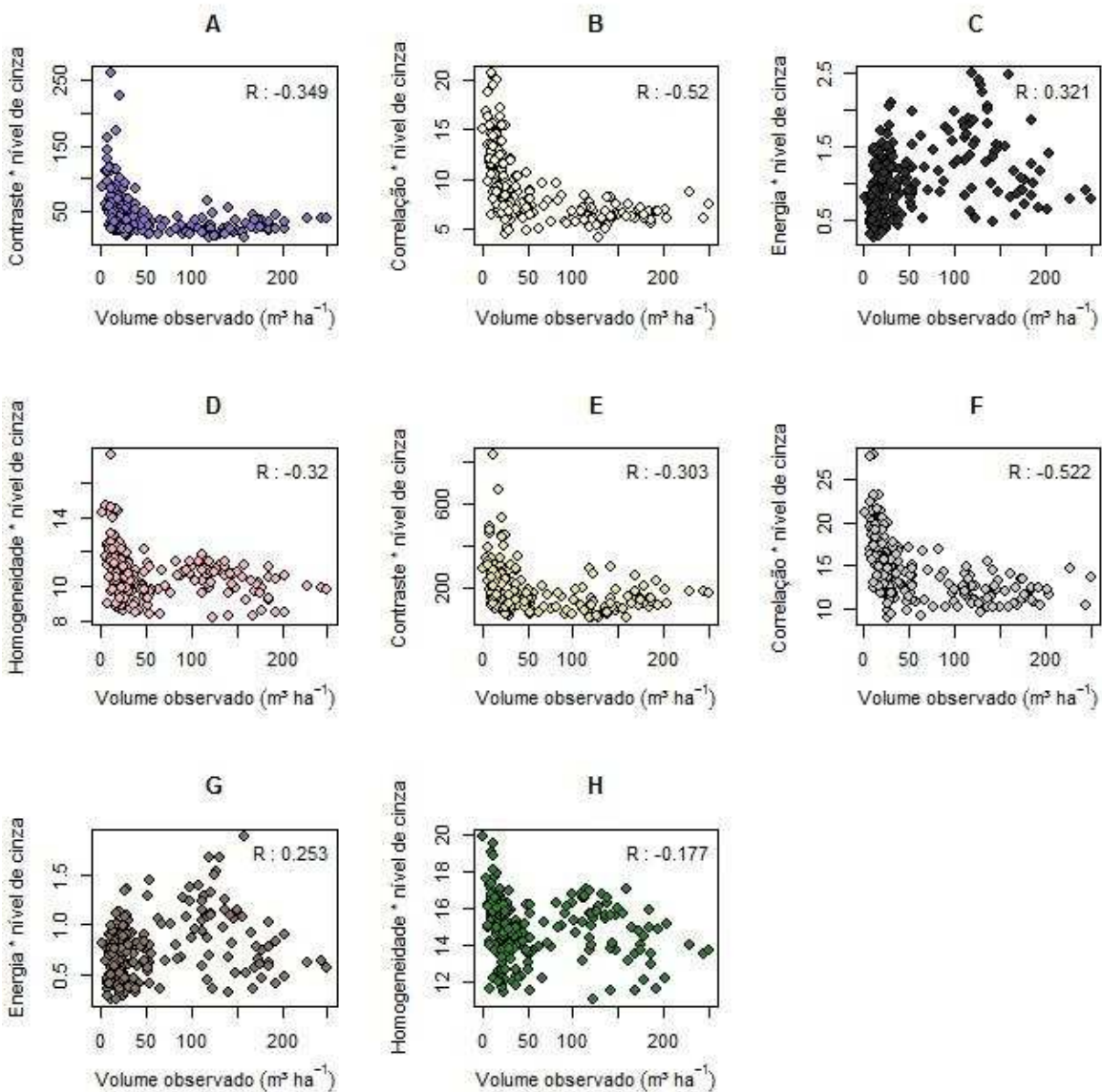
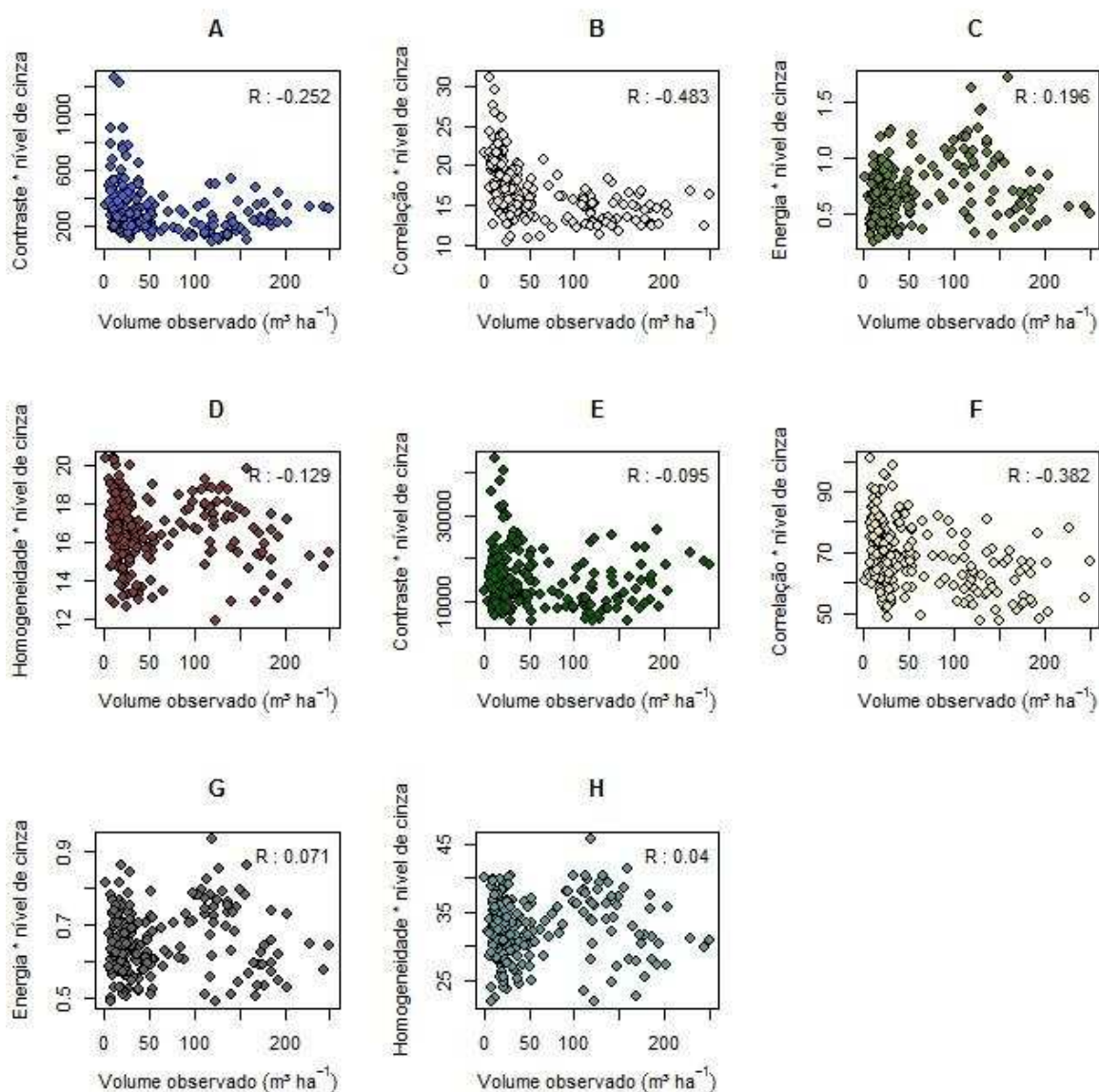


Figura 17 – Distribuição dos volumes observados pelos índices obtidos da textura: (A) MULCON banda 3; (B) MULCOR banda 3; (C) MULENE banda 3; (D) MULHOM banda 3; (E) MULCON banda 4; (F) MULCOR banda 4; (G) MULENE banda 4; e (H) MULHOM banda 4.



As distribuições dos valores de volume observado ( $\text{m}^3 \text{ha}^{-1}$ ) pelos valores dos índices INV obtidos através das quatro medidas de texturas para as quatro bandas podem ser visualizadas nas Figuras 18 e 19. Em relação às correlações, o maior valor encontrado foi de -0,464 para INVHOM banda 1 (Figura 18 D) e o menor valor foi de 0,08 para INVCON banda 4 (Figura 19 E).



Figura 18 – Distribuição dos volumes observados pelos índices obtidos da textura: (A) INVCON banda 1; (B) INVCOR banda 1; (C) INVENE banda 1; (D) INVHOM banda 1; (E) INVCON banda 2; (F) INVCOR banda 2; (G) INVENE banda 2; e (H) INVHOM banda 2.

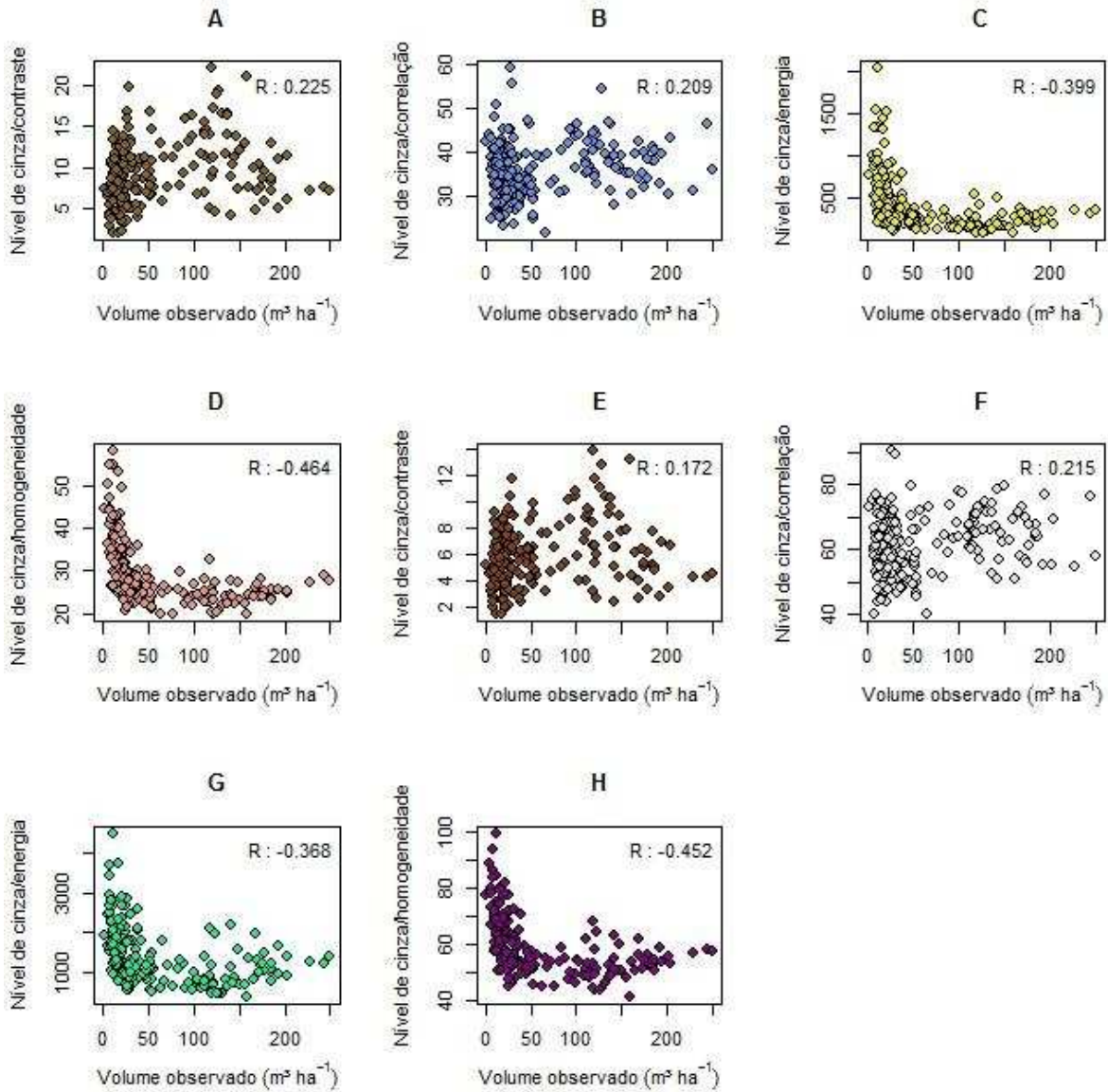
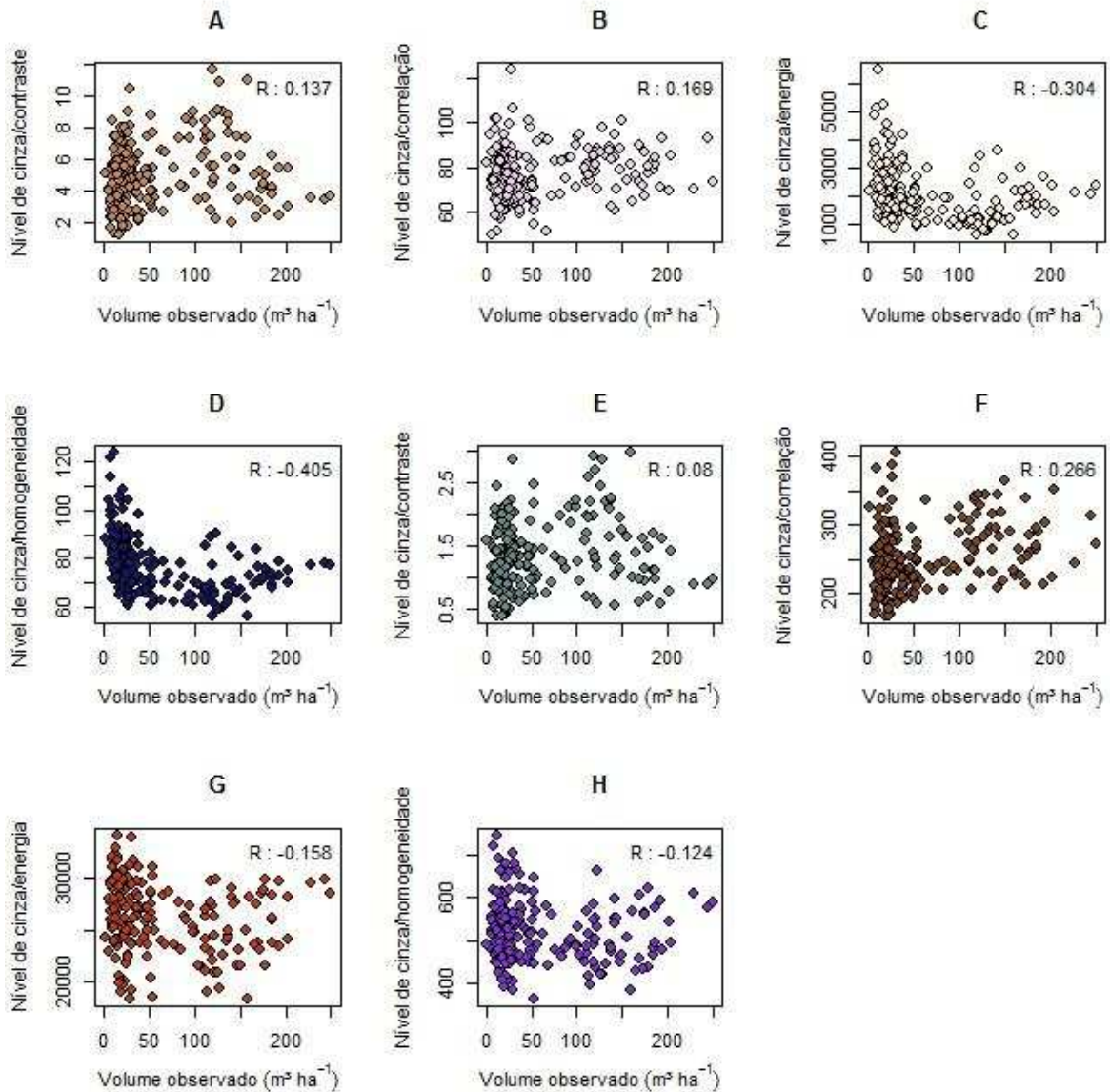


Figura 19 – Distribuição dos volumes observados pelos índices obtidos da textura: (A) INVCON banda 3; (B) INVCOR banda 3; (C) INVENE banda 3; (D) INVHOM banda 3; (E) INVCON banda 4; (F) INVCOR banda 4; (G) INVENE banda 4; e (H) INVHOM banda 4.



### 4.3. GEOESTATÍSTICA

A matriz de correlação gerada para seleção de covariáveis foi feita com as 63 variáveis obtidas, sendo uma delas a variável principal (volume observado) e as outras 62 como variável auxiliar, que se dividem em: idade (variável cadastral); quatro bandas originais do satélite SPOT-6; seis razões entre bandas; três índices de vegetação (NDVI, SAVI e ARVI), 16 medidas de textura (CON, COR, ENE e HOM para cada banda); 32 índices de textura (16 MUL e 16 INV). Pela matriz, foi possível verificar que muitas das covariáveis são altamente correlacionadas entre si, como por exemplo, entre as bandas da faixa do visível (bandas 1, 2 e

3), evidenciando que muitas das informações extraídas de uma banda também são encontradas nas outras duas bandas. O mesmo ocorre em relação à razão simples entre as bandas, os índices de vegetação, as medidas de textura e os índices de textura. Há também alta correlação entre uma variável e índices derivados da mesma variável.

A alta correlação deve-se à combinação de dois fatores: muitos objetos (alvos) possuem valores de NC muito próximos na região do visível ou na região do infravermelho, resultando em valores de NC pouco diferenciados entre si; e o sombreamento topográfico, que é o mesmo em todas as bandas de um mesmo sensor, sendo que em áreas de relevo acidentado com baixo ângulo de iluminação solar faz com que o grau de correlação entre as bandas aumente (MENESES; ALMEIDA, 2012).

Outro ponto importante na análise da matriz gerada é a correlação existente entre as variáveis auxiliares com a variável volume ( $\text{m}^3 \text{ha}^{-1}$ ). É possível observar que a idade possui a maior correlação com o volume (0,95) e que, entre as variáveis obtidas pelas imagens, as maiores correlações foram de -0,52 para as quatro variáveis (banda 3, COR banda 1 e MULCOR das bandas 1 e 2).

As 62 variáveis auxiliares juntamente com a variável principal foram utilizadas, por meio da análise do *Stepwise*, para a construção do modelo de covariância, cujo modelo gerado indica quais variáveis espectrais e de textura melhor explicam o volume. O modelo selecionado pelo menor AIC (-495,38) foi:

$$\widehat{\text{Volume}} = 93,65659 - 4,7001 B2 - 93,04338 \text{COR}_{B2} + 3,6777 \text{MULCOR}_{B1} + 56,7391 Id \quad (28)$$

Onde:  $B2$  é a banda 2,  $\text{COR}_{B2}$  é a medida de textura COR da banda 2,  $\text{MULCOR}_{B1}$  é a multiplicação da COR da banda pela banda 1 e  $Id$  é a idade.

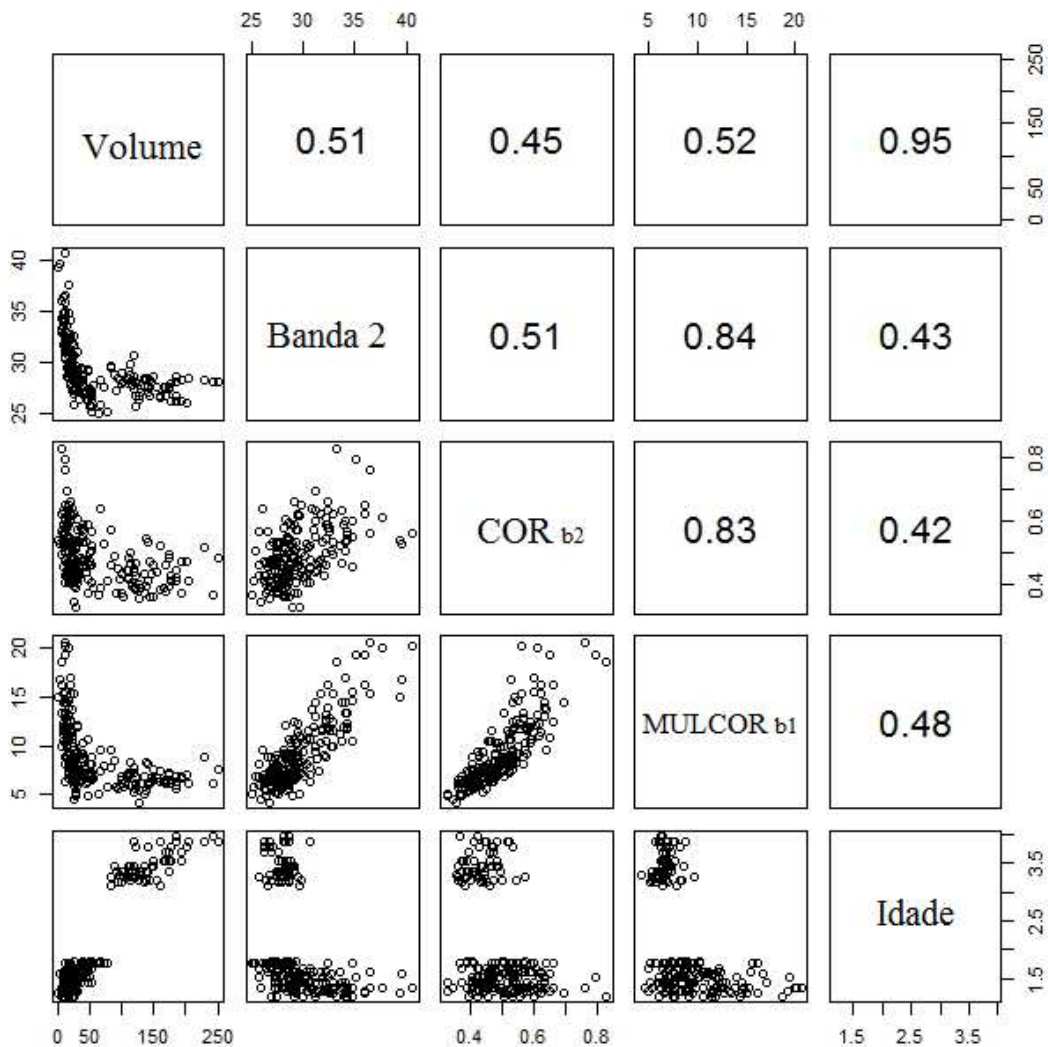
As análises dos parâmetros das variáveis utilizadas pelo modelo podem ser observadas na Tabela 6. Onde o erro padrão residual foi de  $17,86 \text{ m}^3$  e o coeficiente de determinação ajustado ( $R^2_{\text{ajd}}$ ) foi de 0,9081, isso significa que o modelo selecionado para a KDE conseguiu explicar 90,81% do volume a partir das variáveis selecionadas. Todas as variáveis foram significantes ao nível de 1%.

Tabela 6 - Parâmetros do modelo selecionado para KDE.

Variáveis	Coefficientes	Erro padrão	Valor de t	Pr (> t )
Intercepto	93,65659	31,196	3,002	0,003010
Banda 2	-4,70010	1,015	-4,629	0,000006
COR banda 2	93,04338	31,944	-2,913	0,003980
MULCOR banda 1	3,67772	1,328	2,769	0,006140
Idade	56,73910	1,532	37,027	0,000000

As correlações entre as variáveis selecionadas para o modelo são apresentadas na Figura 20.

Figura 20 – Correlação entre as variáveis selecionadas para o modelo.



No estudo apresentado por Gasparoto (2016) para predição volumétrica de eucalipto utilizando vários tipos de imagens, as variáveis que compuseram o modelo foram: bandas 2 e 3 e razão simples entre bandas 3 e 2 para imagem GEOEYE; banda 1 e razão simples entre

bandas 1 e 2 para imagem RapidEye's; e razão simples entre bandas 4 e 3, razão simples entre bandas 5 e 4 e razão simples entre bandas 9 e 3 para imagem Landsat8, juntamente com a idade e o material genético para todos os modelos. Segundo Gomes e Maillard (2013), as bandas e as texturas selecionadas para o modelo de predição de volume em uma vegetação de cerrado com imagens de satélite RapidEye's foram as bandas 3 e 4 e o desvio da correlação da banda 3. Já a textura que melhor explicou a variável volume em estágio avançado sucessional da Mata Atlântica foi a ENE (energia), seguida pela COR (correlação), ambas extraídas da banda 3 do satélite FORMOSAT-2 (SETTE, 2009).

A seleção da banda 2 corrobora com as informações extraídas do seu intervalo espectral referente à vegetação. Isto é, a banda verde, com intervalo 530-590 nm, reflete boa parte da radiação eletromagnética (luz), sendo a responsável pela coloração esverdeada da maioria das folhas (SOUSA, 1997; MENEZES, 2001; BAUERMANN, 2008).

Com relação à razão simples entre as bandas e os índices testados neste trabalho, nenhum deles foi selecionado para compor o modelo de predição de volume. O NDVI utiliza bandas nas regiões do vermelho (625 - 695nm), onde a clorofila é refletida, e nas regiões do infravermelho próximo (760 - 890), onde a radiação é pouco afetada pela clorofila e muito influenciada pela estrutura foliar (BAUERMANN, 2008). Na região do infravermelho próximo, as folhas verdes absorvem pouca energia, apresentando uma alta reflectância causada pelo espalhamento resultante do líquido intracelular e dos espaços intercelulares do mesófilo (ROSENDO, 2005).

Levando em consideração o forte contraste existente entre as duas bandas, o NDVI foi definido de modo a realçar a contribuição espectral da vegetação, enquanto minimiza a contribuição espectral do solo, do ângulo solar, da vegetação senescente e da atmosfera (WATZLAWICK, 2003). Contudo, há limitações no uso do NDVI, o que pode ter limitado a inclusão deste na composição do modelo escolhido. Uma limitação é a rápida saturação do índice que se torna insensível ao aumento da densidade do dossel, onde os valores estabilizam-se em um patamar independente do aumento da densidade do dossel. A outra limitação é a influência de outros fatores no resultado, como características dos sensores, propriedades da atmosfera e do solo (BAUERMANN, 2008).

De acordo com Sousa (1997), o uso do NDVI como identificador nas modificações da vegetação após desbastes de florestas de pinus não surtiu efeito, ou seja, seus valores



permaneceram estagnados, concluindo que por estar trabalhando com florestas adultas o índice saturou. Para Rosendo (2005) e Bauermann (2008), tanto o índice NDVI como o SAVI são mais apropriados para realizar o acompanhamento da evolução temporal de uma vegetação de uma mesma área, onde é possível comparar o desempenho dos índices em relação à mudança na cobertura vegetal e à diferença sazonal, pois assim espera-se uma menor influência das condições atmosféricas. O índice ARVI tem apresentado bons resultados nos estudos de vegetação, mostrando-se mais sensível à atmosfera que o NDVI (MARCUSSE et al., 2010). No entanto, esse índice também não foi selecionado para o modelo de predição de volume.

Das 16 medidas e dos 32 índices de textura, apenas duas variáveis foram selecionadas para o modelo de predição de volume, sendo a medida de textura COR da banda 2 e o índice de textura MULCOR da banda 1. Entretanto, de acordo com Haralick et al. (1973), não existe uma regra para a escolha das melhores medidas de textura, pois um dado conjunto pode funcionar bem para um caso, mas pode não ser muito útil para outro caso, sendo recomendado não utilizar uma quantidade muito grande de características para descrever as texturas.

A medida de textura COR representa a ideia de linearidade de dependências de tons de cinza em uma imagem. Portanto, em uma imagem onde exista uma certa ordenação local de NC, o valor da correlação é alto (HARALICK et al., 1973).

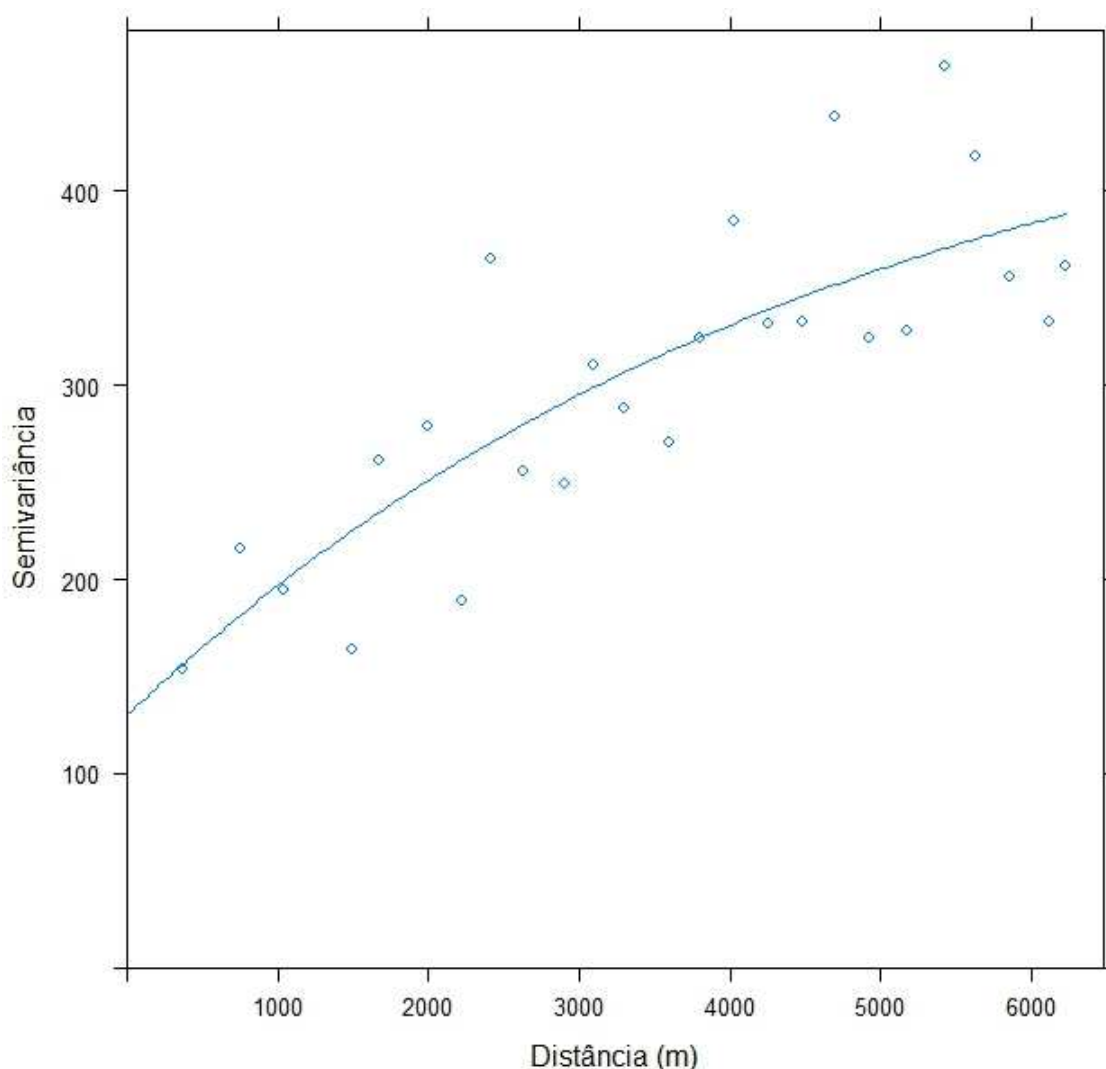
Para Sette (2009), quando a classificação com as três bandas da imagem mais as cinco bandas de textura foi realizada, houve um aumento de 37,1% na acurácia, chegando a 83,8% de acurácia da classificação. Observou-se que para a textura, a banda verde foi a mais eficiente para a análise de vegetação. Também foi observado que em florestas jovens há uma melhor correlação entre a altura e as medidas de textura e que nas florestas mais velhas a correlação é melhor com o DAP. Isso se deve ao fato de que nas jovens há um maior crescimento em altura, provocando uma irregularidade e uma diferenciação textural, enquanto nas mais velhas há uma estabilização da altura e um aumento no crescimento secundário (DAP), deixando a textura mais homogênea.

No estudo feito por Gomes e Maillard (2013), em que foi feita a estimativa da idade da vegetação do cerrado, a união das bandas espectrais e da textura dos dados RapidEye trouxe um incremento de mais de 15% no poder preditivo do modelo. Porém, para os outros parâmetros da vegetação como volume, DAP, densidade e área basal, as texturas não

apresentaram altos coeficientes de determinação. Outros autores também utilizaram as medidas de textura relacionadas à vegetação florestal (JOHANSEN; PHINN, 2006; HALL; SKAKUN, 2007; ALENCAR-SILVA; MAILLARD, 2011).

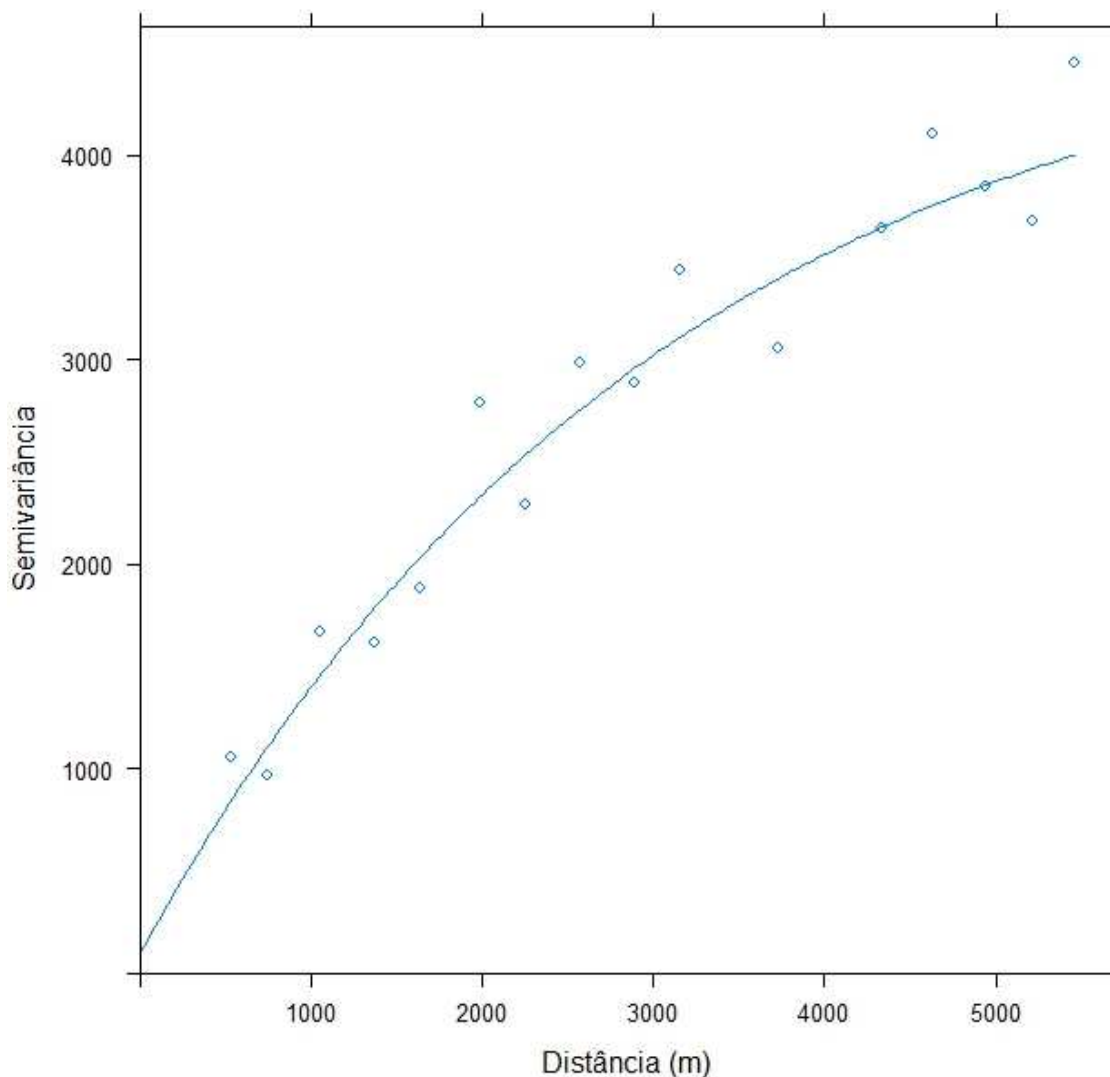
A Figura 21 representa a estrutura de dependência espacial para a característica volume, obtida através do semivariograma ajustado ao modelo exponencial pelo método MQO para a KDE. A seleção do modelo exponencial se deu pelo menor valor do erro padrão residual de 25,58%.

Figura 21 – Semivariograma ajustado ao modelo exponencial para KDE.



Para o semivariograma da KO, entre os modelos avaliados, o exponencial foi selecionado com o menor valor do erro padrão residual (50,25%), sendo possível observar a estrutura espacial dos dados na Figura 22.

Figura 22 – Semivariograma ajustado ao modelo exponencial para KO.



Ambos os semivariogramas (Figuras 21 e 22) mostraram que a característica estudada apresentou-se estruturada espacialmente, ou seja, existe uma função estrutural, com semivariância de comportamento modelável. Isto mostra que as estimativas de volume com o modelo exponencial, considerando a componente espacial, podem gerar estimativas mais precisas no processo da krigagem. Os parâmetros dos semivariogramas encontrados para KDE e KO podem ser visualizados na Tabela 7.

Tabela 7 - Parâmetros do modelo exponencial para KDE e KO.

<b>Krigagem</b>	<b>Efeito pepita</b>	<b>Patamar</b>	<b>Alcance (m)</b>
KDE	130,20	354,30	4778,90
KO	102,30	4698,60	3083,80

Este resultado é análogo ao de Alvarenga (2012), cujo estudo mostrou que o menor efeito pepita encontrado na avaliação de dependência espacial para o volume foi o do modelo exponencial (112,48), se comparado com o modelo esférico (293,97), ambos ajustados pelo método MQO. De acordo com o mesmo autor, o efeito pepita, sendo diferente de zero, indica que nem toda a variação pode ser explicada pela componente espacial. O valor do efeito pepita encontrado no presente estudo também foi próximo ao valor encontrado por Mello et al. (2005), que foi de 261,34 para a variável volume em *Eucalyptus grandis* com 7 anos de idade no ajuste do modelo exponencial. Gasparoto (2016) encontrou um efeito pepita de 100,7 para KO com modelo exponencial e de 209 para KDE com o modelo geoestatístico de potência e componentes de imagens RapidEye. Com relação ao patamar, é possível observar na Tabela 6 que o KDE foi inferior ao da KO, isso se deve a variância das variáveis auxiliares utilizadas pela KDE.

Ainda no mesmo estudo, Alvarenga (2012) encontrou um alcance de 2.254,80 metros, indicando que quando a distância entre as amostras forem menores que o valor do alcance, a estrutura de continuidade espacial deve ser considerada nas análises. Mello et al. (2005) encontraram um valor de alcance médio de 645 metros para a variável volume em *Eucalyptus grandis* com 7 anos de idade no ajuste do modelo exponencial. Kanegae Júnior et al. (2007) encontraram, para povoamentos clonais de *Eucalyptus* sp., alcances médios de: 1.083,06 metros aos 2 anos de idade; 1.177,66 metros aos 3 anos; 1.265,88 metros aos 4 anos; e 786,49 metros aos 5 anos. Gasparoto (2016) encontrou um alcance de 664,5 para KO e de 109,1 para KDE. Este parâmetro reflete o grau de homogeneização entre unidades amostrais, isto é, quanto maior o alcance, maior será a homogeneidade entre os pontos amostrais (YAMAMOTO, 2001). O valor do alcance encontrado neste estudo é superior aos valores encontrados nos trabalhos acima citados, o que indica uma homogeneidade entre as amostras.

#### 4.4. ANÁLISE DAS ESTIMATIVAS

Os volumes ( $\text{m}^3 \text{ha}^{-1}$ ) por parcela obtidos após a seleção do modelo exponencial para ambos os métodos de krigagem (KDE e KO) podem ser verificados na Tabela 8, juntamente com o volume observado e suas estatísticas. Observa-se que os valores médios preditos pela KDE estão mais próximos do volume médio observado por classe de idade.

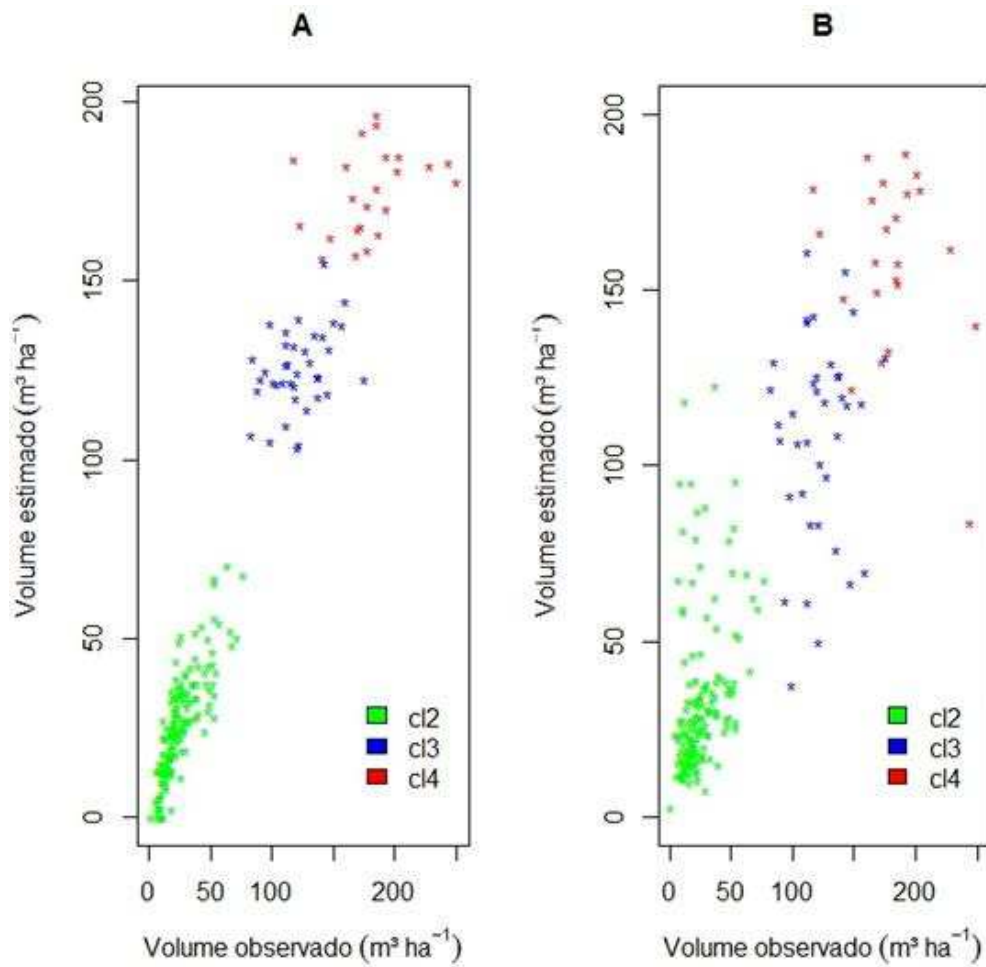
Tabela 8 - Estatísticas dos volumes por classe de idade estimados pela KDE e KO.

<b>Volume</b>	<b>Cl. idade</b>	<b>Mín.</b>	<b>Máx.</b>	<b>Média</b>	<b>Mediana</b>	<b>Var.</b>	<b>Desvio Pad.</b>
<b>Observado</b>	<b>2</b>	1,16	76,50	25,76	22,13	231,32	15,21
	<b>3</b>	82,41	174,80	121,20	119,50	475,08	21,80
	<b>4</b>	117,70	249,80	180,40	177,50	1060,24	32,56
<b>KDE</b>	<b>2</b>	0,00	70,23	25,97	24,75	224,79	14,99
	<b>3</b>	103,20	154,90	125,20	123,80	128,62	11,34
	<b>4</b>	156,30	196,60	175,00	175,90	144,95	12,03
<b>KO</b>	<b>2</b>	2,80	122,60	33,40	26,64	506,90	22,51
	<b>3</b>	37,70	161,10	108,50	116,30	867,98	29,46
	<b>4</b>	83,94	188,90	158,50	161,70	624,45	24,99

Onde: Cl.= classe, Mín.= mínimo, Máx.= máximo, Var.= variância e Pad.= padrão.

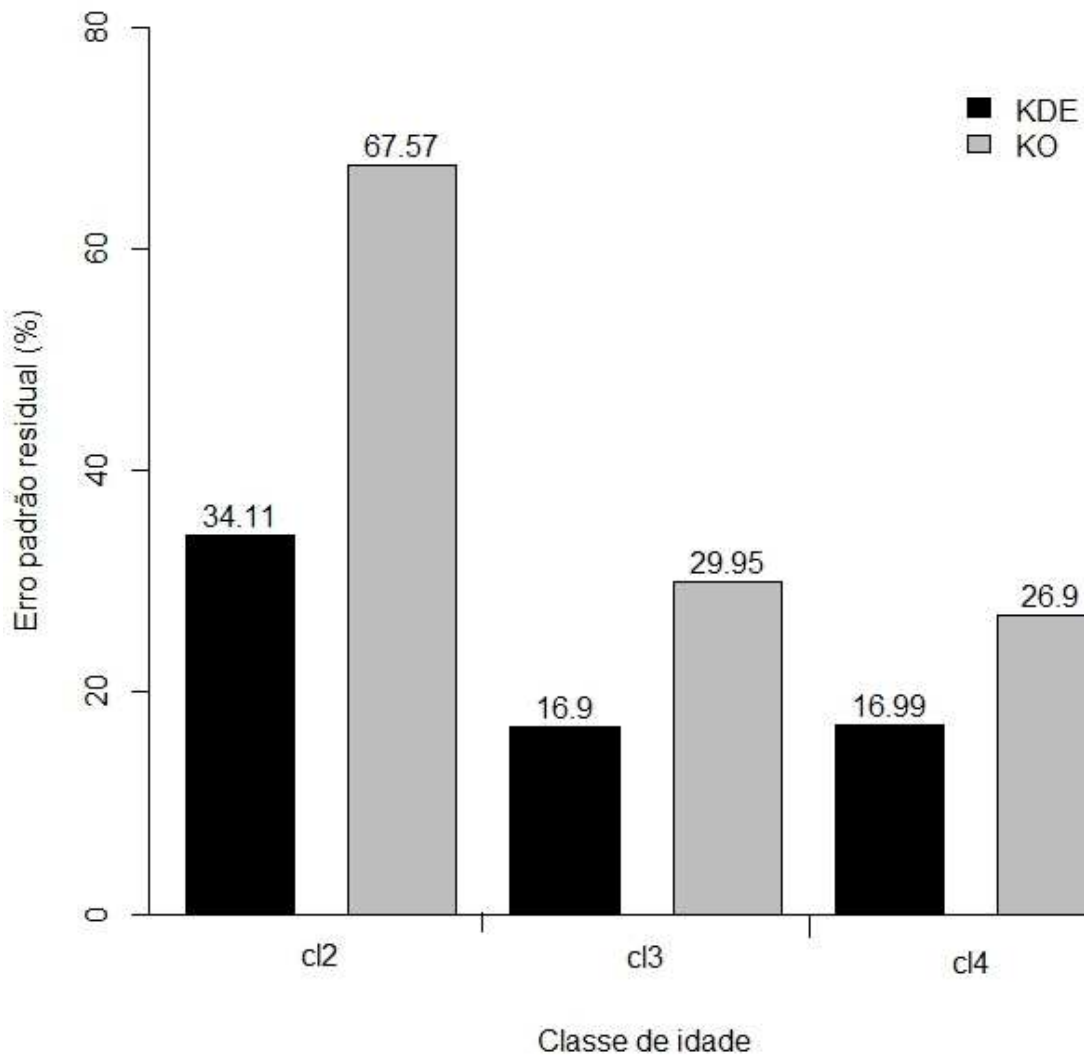
A Figura 23 demonstra a distribuição dos volumes observados pelos volumes ajustados pela KDE e KO. Comparando-se os dois métodos de krigagem é possível observar que o comportamento dos dados gerados pela KDE com o volume observado possui uma linearidade, diferentemente da distribuição dos dados da KO pelo volume observado.

Figura 23 –Distribuição dos volumes observados pelos volumes estimados do ajuste: (A) KDE e (B) KO.



Na Figura 24, é possível observar os valores do erro padrão residual para cada classe de idade de acordo com os métodos de krigagem. Os valores dos erros da KO são maiores que os valores dos erros da KDE em todas as classes.

Figura 24 – Erro padrão residual por classe de idade.

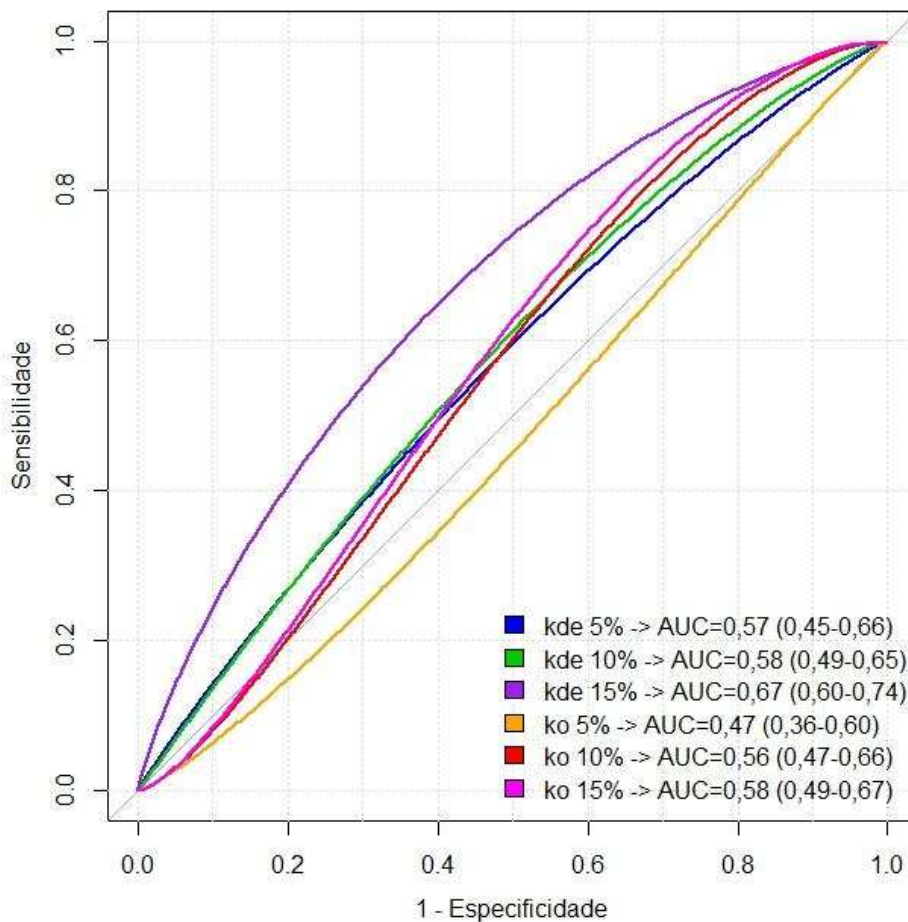


A outra análise utilizada para verificar o desempenho dos modelos é a da curva ROC. As curvas produzidas para cada limiar podem ser observadas na Figura 25, demonstrando as diferenças entre os limiares e os tipos de krigagem. Segundo Margotto (2010), um modelo perfeito corresponderia a uma linha horizontal no topo do gráfico, porém, na prática, as curvas entre a linha diagonal e a linha perfeita são consideradas boas, sendo que quanto maior a distância da linha diagonal, melhor o modelo.

A avaliação dos modelos é feita pela AUC de cada curva ROC, estes valores podem ser visualizados no Figura 25. Quando comparado entre os métodos de krigagem, a AUC da KDE foi melhor que a KO em todos os limiares. A KDE apresentou uma AUC de 0,67 e um intervalo de confiança variando entre 0,60 a 0,74. Isto significa que o modelo gerado pela KDE com limiar de 15% apresentou uma acurácia de 67% na predição da variável de

interesse, um valor estatisticamente significativo, pois descarta a hipótese nula que seria AUC de 0,50, em que se encontra um modelo aleatório, ou seja, sem informação. O valor de 67% não é considerado um desempenho muito bom, segundo Francois et al. (2008), um modelo com AUC a partir de 0,80 é considerado um modelo adequado. Já para Margotto (2010) e Bezerra (2012), modelos com valores de AUC acima de 0,70 possuem um desempenho satisfatório e contribuem na avaliação da adequação do modelo ao propósito ao qual ele se destina. Assim, com a análise ROC é possível ter uma avaliação simples do desempenho de um modelo, contribuindo com mais uma medida na avaliação de modelo.

Figura 25 – Curva ROC geradas para cada limiar da KDE e da KO.



Com base nos resultados obtidos, nota-se que a KDE apresentou melhores resultados que a KO, possuindo um menor valor do erro padrão residual e uma maior AUC na análise da curva ROC. Isso indica que a utilização de preditores como medidas espectrais e texturais proporcionou estimativas mais precisas do que as geradas pela KO.



Lundgren et al. (2015) avaliaram a estimação de volume de eucalipto por Cokrigagem, krigagem e regressão, sendo a Cokrigagem melhor que a krigagem na estimação do volume. Em outro trabalho também houve a comparação entre as técnicas de regressão linear múltipla, KO e KDE para a predição volumétrica de plantios de eucaliptos, sendo que o melhor desempenho foi da KDE, que permitiu reduzir um erro amostral de 3,27% para 2,34% (GASPAROTO, 2016).

Segundo Assis et al. (2009), a partir da confirmação de dependência espacial e da escolha do modelo, é possível gerar mapas que correspondem à característica de interesse sem apresentar tendência ou enviesamento, isto é, mapas mais precisos em relação ao detalhamento espacial da variável de interesse. Assim, as distribuições dos volumes de acordo com os talhões é uma ferramenta importante para o planejamento da colheita. Neste trabalho, o volume para todos os talhões não foi estimado e nem foi gerado um mapa com a distribuição do volume para a fazenda, pois é preciso obter informações das variáveis auxiliares que compõem o modelo para a área toda de interesse. Dessa forma, um ponto para um próximo estudo seria a obtenção das medidas de textura da área toda a partir da sua divisão em subáreas com tamanho de uma parcela.



## 5. CONCLUSÕES

Para o modelo de predição de volume gerado para KDE, dos 61 componentes de imagem obtidos, três componentes de imagens foram selecionados, sendo a banda 2, a medida de textura COR (correlação) da banda 2 e o índice MULCOR da banda 1, que é a medida de textura da correlação da banda 1 multiplicada pelos valores de NC da mesma banda. Outra variável que também compôs o modelo da KDE foi a idade, por apresentar uma alta correlação com o volume.

As correlações obtidas pelos três índices de vegetação, os quais são muito usados para informações de vegetação, foram baixas com relação ao volume quando comparadas com os outros componentes. O contrário ocorreu com alguns outros componentes de imagem, em que a união de bandas com as medidas de textura apresentaram-se como importantes variáveis auxiliares na predição de volume de madeira de eucalipto.

Com relação ao ajuste do semivariograma, o volume de madeira apresentou uma dependência espacial, sendo que em ambos os métodos de krigagem (KDE e KO), o modelo teórico que obteve o melhor ajuste foi o exponencial, pelo método MQO. Assim, a geoestatística aplicada pelos interpoladores de krigagem teve êxito ao estimar os volumes por parcela.

Através da análise dos volumes estimados pelos métodos de krigagem e das curvas ROC, a KDE apresentou melhores resultados do que a KO. Nos diferentes limiares testados para as curvas ROC dos dois tipos de krigagem, os valores de AUC variaram de acordo com os limiares estabelecidos, sendo que em todos os casos a KDE com AUC foi superior à KO. O valor de AUC da KDE não foi alto, porém, o valor ainda foi superior a um desempenho aleatório, mostrando que o modelo possibilitou agregar informações úteis à predição do volume.

Para um futuro trabalho de pesquisa nesta área, é recomendado utilizar o modelo selecionado pela KDE para estimar o volume da área total do estudo e gerar mapas da distribuição do volume para observar o comportamento do volume nas diferentes regiões da área. Um outro ponto seria extrair outras informações das imagens de satélites, tais como bandas com um intervalo espectral acima de 890 nm, ou mesmo a composição de bandas de outras imagens.



## REFERÊNCIAS

- AKAIKE, H. Information measures and model selection. Bulletin of the International Statistical Institute, v. 50, n. 1, p. 277-291, 1983.
- ALENCAR-SILVA, T.; MAILLARD, P. Determinação da estrutura biofísica da vegetação ribeirinha através de técnica de segmentação, análise de textura de imagem e modelagem estatística. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 15., Curitiba, 2011. Anais... Curitiba: INPE, 2011, p. 7828-7835.
- ALVARENGA, L. H. V. Imagens de alta resolução e geoestatística na estratificação da fisionomia Cerrado para inventários florestais. 2012. 91 f. Dissertação (Mestrado em Ciências Florestais) - Universidade Federal de Lavras, Lavras, 2012.
- AMARAL, S. et al. Relações entre Índice de Área Foliar (LAI), Área Basal e Índice de Vegetação (NDVI) em relação a diferentes estágios de crescimento secundário na Floresta Amazônica em Rondônia. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 8., Salvador, 1996. Anais... Salvador: INPE, 1996. p. 485-489.
- AMARAL, P. M. J. Análise de imagem: medição de área florestal. 2010. 86 f. Monografia (Graduação em Engenharia Informática) - Universidade da Beira do Interior, Covilhã, 2010.
- ANTUNES, A. F. B. Classificação de ambiente ciliar baseada em orientação a objeto em imagens de alta resolução espacial. 2008. 146f. Tese (Doutorado em Ciências Geodésicas) - Universidade Federal do Paraná, Curitiba, 2003.
- ASSIS, A. L. et al. Development of a sampling strategy for young stands of *Eucalyptus* sp. using geostatistics. Cerne, Lavras, v. 15, n. 2, p. 166-173, abr./jun. 2009.
- ASSMANN, E. The principles of forest yield study. New York: Pergamon, 1970. 506 p.
- AZEVEDO, L. F.; PEREIRA, A. da C. Avaliação crítica e implementação prática de estudo sobre a variabilidade de teste de diagnóstico - Parte II. Revista Nascer e Crescer, Porto-Portugal, v. 19, n. 4, p. 265-277, 2010.
- BAUERMAN, G. C. Uso de imagens de sensores remotos na estimativa de características dendrométricas de povoamentos de eucalipto. 2008. 78 f. Dissertação (Mestrado em Engenharia de Produção) – Universidade Federal de Santa Maria. Santa Maria, 2008.
- BARTELS, M.; WEI, H.; MASON, D. C. Wavelet packets and co-occurrence matrices for texture-based image segmentation. In: ADVANCED VIDEO AND SIGNAL BASED SURVEILLANCE, 2005. AVSS 2005. IEEE Conference on. IEEE, 2005. p. 428-433.
- BELTRAME, A. M. K. et al. Análise da vegetação em faixas de servidão de linhas de transmissão de energia elétrica. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 13., 2007, Florianópolis. Anais... Florianópolis: INPE, 2007. p. 5579 – 5586.

BEZERRA, G. K. A. Modelo de regressão logística para previsão do óbito na unidade de terapia intensiva. 2012. 90 f. Dissertação (Mestrado em Modelos de Decisão e Saúde) - Universidade Federal da Paraíba, João Pessoa-PB, 2012.

BIONDI, F.; MYERS, D. E.; AVERY, C. C. Geostatistically modeling stem size and increment in an old-growth forest. Canadian Journal of Forest Research-Revue Canadienne de Recherche Forestiere, v. 24, n. 7, p. 1354-1368, 1994.

BIVAND, R. S.; PEBESMA, E. J.; GÓMEZ-RUBIO, V. Applied spatial data analysis with R. New York: Springer, 2008.

BIVAND, R. et al. Maptools: tools for reading and handling spatial objects. R package, version 0.8-39, 81 p., 2016.

BOLFE, E. L.; GOMES, J. B. V. Geoestatística como subsídio à implantação de agricultura de precisão. Pesquisadores da Embrapa Tabuleiros Costeiros. 2005.

BOGNOLA, I. A. et al. Modelagem uni e bivariada da variabilidade espacial de rendimento de *Pinus taeda* L. Floresta, Curitiba-PR, v. 38, n. 2, p. 373-385, abr/jun. 2008.

BOLFARINE, H.; BUSSAB, W. O. Elementos de amostragem. São Paulo: Edgard Blücher, 2007. 274 p.

BORATTO, I. M.; GOMIDE, R. L. Aplicação dos índices de vegetação NDVI, SAVI e IAF na caracterização da cobertura vegetativa da região Norte de Minas Gerais. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 16., Foz do Iguaçu, 2013. Anais... São José dos Campos: INPE, 2013. p. 7345-7352.

BRAGA, A. C. Curvas ROC: aspectos funcionais e aplicações. 2000. 267f. Tese (Doutorado em Engenharia de Produção e Sistemas) -Universidade do Minho, Braga- Portugal, 2000.

BRITO, G. A. S.; ASSAF NETO, A. Modelo de classificação de risco de crédito de empresas. Revista Contabilidade & Finanças, São Paulo-SP, v. 19, n. 46, p. 18-29, 2008.

BROWN, P. E. Model-Based Geostatistics the Easy Way. Journal of Statistical Software, v. 63, n. 12, p. 1-24, 2015.

BURROUGH, P. A.; BOUMA, J.; YATES, S. R. The state of the art in pedometrics. Geoderma, Amsterdam, v. 62, n. 1, p. 311-326, 1994.

BUSSAB, W. de O.; MORETTIN, P. A. Estatística básica. 5. ed. São Paulo: Saraiva, 2002.

CAMPBELL, J. B. Introduction to Remote Sensing. 3 ed. New York: The Guilford Press, 2002. 621 p.

CESARO, A. et al. Comparação dos métodos de amostragem de área fixa, relascopia, e de seis árvores, quanto a eficiência, no inventário florestal de um povoamento de *Pinus* sp. Ciência Florestal, Santa Maria v. 4, n. 1, p. 97-108, dez. 1994.

CHEN, C.; PAU, L.; WANG, P.; WANG, S. Handbook of Pattern Recognition & Computer Vision. 2 ed. World Scientific, 1999. 1019 p. Disponível em: <<http://books.google.com.br/books?id=pCghq7-i1oMC>> Acesso em: 12 fev. 2016.

CINTRA, D. P. Classificação de estágios sucessionais florestais por meio de imagens de alta resolução (IKONOS) no Parque Estadual da Pedra Branca, RJ. 2007. 64 f. Dissertação (Mestrado em Ciências) – Universidade Federal de Lavras, Lavras, 2007.

COCHRAN, W. G. Sampling techniques. 3. ed. New York: J. Wiley, 1977. 555 p.

COHEN, W.B.; SPIES, T.A.; BRADSHAW, G.A. Semivariograms of digital imagery for analysis of conifer canopy structure. Remote Sensing of Environment, New York, v.34, n.3, p.167-178, Dec. 1990.

COSTA, T. C. C. da; REGAZZI, A. J. Amostragem para inventário florestal com probabilidade de superposição de parcelas circulares. Revista Árvore, Viçosa-MG, v. 34, n. 1, p. 137-145, 2010.

CRESSIE, N. Statistics for spatial data. New York: J. Wiley, 1993. 900p.

CRÓSTA, A. P. Processamento Digital de Imagens de Sensoriamento Remoto. Campinas-SP, IG, 170 p. 1993.

CUNHA, U. S. Dendrometria e inventário florestal. Série técnica: curso técnico em manejo florestal. Escola Agrotécnica Federal de Manaus, Manaus, 2004. 58 p.

CURTIS, R. O. Height-diameter and height-diameter-age equations for second-growth Douglas-Fir. Forest Science, Amsterdam, v. 13, n. 4, p. 365-375, 1967.

DEFINIENS, I. Ecognition user guide 3. München: Definiens Imaging, 2003. 413 p.

DINIZ, F. S. Métodos de Amostragem e Geoestatística Aplicada ao Inventário Florestal. 2007.101 f. Dissertação (Mestrado em Engenharia Florestal) – Universidade Federal de Lavras, Lavras, 2007.

DRUSZCZ, J. P et al. Comparação entre os métodos de amostragem de Bitterlich de área fixa com parcela circular em plantações de *Pinus taeda*. Floresta, Curitiba, v. 40, n. 4, p. 739-754, out./dez. 2010.

EIPHANIO, J. C. N. et al. Índices de vegetação no sensoriamento remoto da cultura do feijão. Pesquisa Agropecuária Brasileira, Brasília, v. 31, n. 6, p. 445-454, jun. 1996.

FACCHIN, S. (2005). Técnicas de análise multivariável aplicadas ao desenvolvimento de analisadores virtuais. 2005. 140 f. Dissertação (Mestrado em Engenharia) - Universidade Federal do Rio Grande do Sul, Porto Alegre, 2005.

FAWCETT, T. An introduction to ROC graphs. Pattern Recognition Letters, Amsterdam, v. 27, n. 8, p. 861-874, 2006.

FERNANDES, J. A. B. Krigagem com deriva externa aplicada à avaliação de recursos minerais de calcário e de minério laterítico. 2009. 66 f. Dissertação (Mestrado) - Instituto de Geociências, Universidade de São Paulo, São Paulo, 2009.

FERRAZ, A. S. et al. Estimativa do estoque de biomassa em um fragmento florestal usando imagens orbitais. Floresta e Ambiente, v. 21, n. 3, p. 286-296, jul./set. 2014.

FICK, T. A. Amostragem para inventário florestal em sistemas silvipastoris. Revista Árvore, Viçosa-MG, v. 35, n. 5, p. 1033-1038, 2011.

FIGUEIREDO, E. O. Método da variável fictícia para ajuste de modelos volumétricos estáveis e compatíveis em povoamentos florestais. Rio Branco: Embrapa Acre. 2005a. 28p.

FIGUEIREDO, D. Conceitos Básicos de Sensoriamento Remoto. Brasília: Companhia Nacional de Abastecimento - CONAB. 2005b. 32p.

FLORENZANO, T. G. Imagens de Satélite para Estudos Ambientais. São Paulo: Oficina de Textos. 2002. 97 p.

FORTES, B. de P. M. D. et al. Modelagem geoestatística da infecção por *Ascaris lumbricoides*. Cad. Saúde Pública, Rio de Janeiro, v. 20, n. 3, p. 727-734, 2004.

FRANCAIS, A.; VESIN, A.; TIMSIT, J. F. Como realizar ensaios clínicos em terapia intensiva utilizando base de dados de alta qualidade. Rev Bras Ter Intensiva, São Paulo, v. 20, n. 3, p. 296-304, 2008.

GASPAROTO, E. A. G. O uso de imagens de satélite como covariável na modelagem geoestatística volumétrica de florestas de *Eucalyptus* spp. 2013. 30 f. Monografia (Graduação em Engenharia Florestal) - Universidade Federal de São Carlos, Sorocaba, 2013.

GASPAROTO, E. A. G. Estatística espacial e sensoriamento remoto para a predição volumétrica em florestas de *Eucalyptus* spp. 2016. 120 f. Dissertação (Mestrado em Ciências de Recursos Florestais) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, 2016.

GOME, M. F; MAILLARD, P. O uso de feições de textura em imagens RapidEye para estimativas da idade e de parâmetros estruturais da vegetação do cerrado. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 16., Foz do Iguaçu, 2013. Anais... Foz do Iguaçu: INPE, 2013, p. 3260-3267.

GONÇALVES, D. B. Identificação e verificação de descritores usando características texturais e dissimilaridade. 2014. 117 f. Tese (Doutorado em Ciência da Computação) - Universidade Federal do Paraná, Curitiba, 2014.

GONZALEZ, R. C.; WOODS, R. E. Processamento de Imagens Digitais, [Digital image processing]. Roberto Marcondes Cesar Junior (Trad.); Luciano da Fontoura Costa (Trad.). São Paulo: Blucher, 2007. 509 p.



GUEDES, I. C. de L. Técnicas geoestatísticas e interpoladores espaciais na estratificação de povoamentos de *Eucalyptus* sp. 2009. 84 f. Dissertação (Mestrado em Engenharia Florestal) – Universidade Federal de Lavras, Lavras, 2009.

GUEDES, L. C. I. et al. Técnicas geoestatísticas e interpoladores espaciais na estratificação de povoamentos de *Eucalyptus* sp. Ciência Florestal, Santa Maria, v. 22, n. 3, p. 541-550, jul./set. 2012.

HALL, R. J.; SKAKUN, R. S. Mapping forest inventory attributes across coniferous, deciduous and mixedwood stand types in the northwest territories from high spatial resolution QuickBird satellite imagery. In: Proceedings of the specialty conference: our common Borders-Safety, security, and the environment through remote sensing, edited: American Society for Photogrammetry and Remote Sensing (ASPRS). Unnumbered. Ottawa, Ontario: CRSS/ASPRS, 2007.

HARALICK, R. M.; SHANMUGAM, K.; DINSTEN, I. Textural features for image classification. IEEE Transactions on Systems, Man and Cybernetics, v. 3, n. 6, p. 610-621, 1973.

HARALICK, R. M. Statistical and structural approaches to texture. Proceedings of the IEEE, v. 67, n. 5, p. 786-804, 1979.

HIEMSTRA, P.H. et al. Real-time automatic interpolation of ambient gamma dose rates from the Dutch Radioactivity Monitoring Network. Computers & Geosciences, v. 35, n. 8, p. 1711-1721, 2009.

HIJMANS, R. J. et al. raster: Geographic data analysis and modeling. R package, version 2.5-2, 242 p., 2015.

HOCK, B. K.; PAYN, T. W.; SHIRLEY, J. W. Using a geographic information system and geostatistics to estimate site index of *Pinus radiata* for Kaingaroa forest, New Zealand. NZJ For. Sci., New Zealand, v. 23, n. 3, p. 264-277, 1993.

HUETE, A. R. A soil-adjusted vegetation index (SAVI). Remote sensing of environment, New York, v. 25, n. 3, p. 295-309, aug. 1988.

ILEA, D. E.; WHELAN, P. F. Image segmentation based on the integration of colour texture descriptors: A review. Pattern Recognition, v. 44, n. 10, p. 2479-2501, 2011.

ISAAKS, E. H.; SRIVASTAVA, R. M. An introduction to applied geostatistics. New York: Oxford University, 1989. 560 p.

JI, L.; PETERS, A. J. Performance evaluation of spectral vegetation indices using a statistical sensitivity function. Remote Sensing of Environment, v. 106, n. 1, p. 59-65, 2007.

JOHANSEN, K.; PHINN, S. Mapping structural parameters and species composition of riparian vegetation using IKONOS and Landsat ETM+ data in Australian tropical savannahs. Photogrammetric Engineering & Remote Sensing, v. 72, n. 1, p. 71-80, 2006.

JOURNEL, A. G.; HUIJBREGTS, C. J. Mining geostatistics. London: Academic Press, 1978. 600 p.

KANEGAE JUNIOR, H. Avaliação de sistemas de inventários sucessivos e dos interpoladores espaciais como instrumento de estratificação de povoamentos clonais de *Eucalyptus* sp. Lavras: UFLA, 2004, 131p.

KANEGAE JUNIOR, H. et al. Avaliação da continuidade espacial de características dendrométricas em diferentes idades de povoamentos clonais de *Eucalyptus* sp. Revista Árvore, Viçosa- MG, v. 31, n. 5, p. 859-866, 2007.

KRIGE, D.G. A statistical approach to some basic mine valuation problems on the Witwatersrand. Journal of the Chemical, Metallurgical and Mining Society of South Africa, v.52, p.119-139, dec. 1951.

LANDIM, P . M. B.; STURARO, J. R. Krigagem Indicativa aplicada à elaboração de mapas probabilísticos de riscos. Rio Claro: DGA, IGCE, UNESP, 2002. (Geomatemática, Texto Didático 6).

LEITE, H. G.; ANDRADE, V. C. L. Um método para condução de inventários florestais sem o uso de equações volumétricas. R. Árvore, Viçosa-MG, v.26, n.3, p.321-328, 2002.

LILLESAND, T. M.; KIEFER, R. W. Remote sensing and image interpretation. 4. ed. Nova York: John Wiley and Sons, Inc., 2000. 724 p.

LIMA, D. L. et al. Caracterização da dispersão espacial de plantas aquáticas usando sensoriamento remoto e geoestatística. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 13. 2007, Florianópolis. Anais... São José dos Campos: INPE, 2007. p. 5879-5886.

LISITA, A.; SANO, E. E.; DURIEUX, L. Detecção de plantios de Cannabis sativa no semi-árido pernambucano a partir da análise das anomalias de dinâmica temporal dos índices de vegetação derivados do satélite spot-5 HRG. Bol. Ciênc. Geod., Curitiba, v. 19, n. 1, p. 45-64, jan./mar. 2013.

LOURENÇO, R. W.; LANDIM, P. M. B. Tratamento de imagem de satélite por meio de metodologia geoestatística. Revista Geociências, Guarulhos, v. 8, n. 6, p. 5-14, dez. 2003.

LOURENÇO, R. W.; LANDIM, P. M. B. Estudo da variabilidade do “Índice de Vegetação por Diferença Normalizada/NDVI” utilizando krigagem indicativa. Holos Environment, v. 4, n. 1, p. 38-55. 2004.

LUNDGREN, W. J. C.; SILVA, J. A. A.; FERREIRA, R. L. C. Estimção de volume de madeira de eucalipto por cokrigagem, krigagem e regressão. Cerne, Lavras, v.21, n. 2, p. 243-250, abr./jun. 2015.

LUZ, J. da Imagens ALOS para o mapeamento da vegetação arbórea e outros usos do solo em área de floresta ombrófila mista. 2009. 122 f. Dissertação (Mestrado em Engenharia Florestal) - Universidade Federal do Paraná, Curitiba, 2009.

MARCUSSI, A. B. et al. Utilização de índices de vegetação para os sistemas de informação geográfica. Caminhos de Geografia, Uberlândia, v. 11, n. 35, p. 41-53, set. 2010.

MARGOTTO, P. R. Curva ROC como fazer e interpretar no SPSS. Escola Superior de Ciências da Saúde, Distrito Federal, 2010.

MARINI, A. Análise de estratégias de rejeição para problemas com múltiplas classes utilizando curvas ROC. 2007. 107 f. Dissertação (Mestrado em Informática) - Pontifícia Universidade Católica do Paraná, Curitiba-PR, 2007.

MARTINS, L. A. de O. Sistema de inspeção visual automática aplicado à detecção de defeitos em aços laminados. 2010. 124 f. Dissertação (Mestrado em Modelagem Matemática e Computacional) - Centro Federal de Educação Tecnológica de Monas Gerais, Belo Horizonte, 2010.

MATÉRN, B. Spatial variation: stochastic models and their application to some problems in forest surveys and other sampling investigations. Meddelanden fran statens skogsforskningsinstitut, Stockholm, v. 49, n. 5, 144p., may. 1960.

MATHER, P. M. Computer processing of remotely-sensed images: an introduction. 3 ed. London: Wiley and Sons, 2004. 350 p.

MATHERON, G. Principles of geostatistics. Economic geology, v. 58, n. 8, p. 1246-1266, dec. 1963.

MATHWORKS Matlab for Windows. The MathWorks Inc., version 7.14, 2012.

MEDEIROS, G. C. F. Uso de texturas para o acompanhamento da evolução do tratamento de úlceras dermatológicas. 2001. 93 f. Dissertação (Mestrado em Engenharia Elétrica) - Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2001.

MELLO, J. M.; SCOLFORO, J. R. S. Análise comparativa de procedimentos de amostragem em um remanescente de Floresta Estacional Semidecídua Montana. Revista Árvore, Viçosa, v. 24, n. 1, p. 55-62, 2000.

MELLO, J. M. Geoestatística aplicada ao inventário florestal. 2004. 111 f. Tese (Doutorado em Recursos Florestais) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, 2004.

MELLO, J. M. et al. Estudo da dependência espacial de características dendrométricas para *Eucalyptus grandis*. CERNE, Lavras, v. 11, n. 2, p. 113-126, jul./dez. 2005.

MELLO, J. M. et al. Uso do estimador geoestatístico para predição volumétrica por talhão. Floresta, Curitiba, v. 36, n. 2, p. 251-260, mai./ago.2006.

MELLO, J. M. et al. Continuidade espacial para características dendrométricas (número de fustes e volume) em plantios de *Eucalyptus grandis*. Revista Árvore, Viçosa-MG, v. 33, n. 1, p. 185-194, 2009a.

MELLO, J. M. et al. Métodos de amostragem e geoestatística para estimativa do número de fustes e volume em plantios de *Eucalyptus grandis*. Floresta, Curitiba, v. 39, n. 1, p. 157-166, jan./mar. 2009b.

MENESES, P. R. Fundamentos de Radiometria Óptica Espectral. In: MENESES, P. R.; NETTO, J. S. M. Sensoriamento Remoto: Reflectância dos alvos naturais. Brasília, DF: UnB; Planaltina: Embrapa Cerrados. 2001.

MENESES, P. R.; ALMEIDA, T. de Introdução ao processamento de imagens de sensoriamento remoto. Brasília, DF: Universidade de Brasília - CNPq, 2012. 266 p.

MINGOTI, S. A.; FIDELIS, M. T. Aplicando a geoestatística no controle estatístico de processos. Produto & Produção, Porto Alegre, v. 5, n. 2, p. 55-70, 2001.

MIYAMOTO, B. S. Método de processamento digital de imagens para inferências da qualidade de materiais preparados com fibras vegetais. 2013. 98 f. Dissertação (Mestrado em Ciência da Computação) - Universidade Federal de São Carlos, São Carlos, 2013.

MORAES, E. C. Fundamento de Sensoriamento Remoto. Instituto Nacional de Pesquisa Espacial. DSR/INPE. São José dos Campos. 2002. 191 p.

MOREIRA, M. A. Fundamentos do Sensoriamento Remoto e metodologias de aplicação. 3. ed. Viçosa: UFV, 2007.

MOREIRA, M. A. et al. Tecnologia de informação: imagens de satélite para o mapeamento de áreas de café de Minas Gerais. Informe Agropecuário, Belo Horizonte-MG, v. 28, n. 241, p. 27-37, nov./dez. 2007.

NILSON, T.; OLSSON, H. Effect of thinning cutting on boreal forest reflectance: a comparison of simulations and Landsat TM estimates. International Journal of Remote Sensing, v. 16, n. 15, p. 2963-2968, Oct. 1995.

NONATO, R. T. Aplicação de mineração de dados na identificação de áreas cultivadas com cana-de-açúcar em imagens de sensoriamento remoto no Estado de São Paulo. 2010. 147 f. Dissertação (Mestrado em Engenharia Agrícola) - Faculdade de Engenharia Agrícola, Universidade Estadual de Campinas, Campinas, 2010.

ODA-SOUZA, M. et al. Aplicação de métodos geoestatísticos para identificação de dependência espacial na análise de dados de um ensaio de espaçamento florestal em delineamento sistemático tipo leque. Revista Árvore, Viçosa-MG, v. 32, n. 3, p. 499-509, 2008.

OLIVARES, V. da S. A curva ROC e suas aplicações. 2013. 84 f. Monografia (Graduação em Matemática Aplicada Computacional) - Universidade de São Paulo, São Paulo, 2009.

OLIVERA, G. O. S. et al. Técnicas de mineração de dados aplicadas a imagens térmicas mastológicas. In: CONGRESSO NACIONAL DE ENGENHARIA MECÂNICA, 7., 2012. Anais... São Luís, 2012, p. 1-10.

PAULA FILHO, P. L. de Reconhecimento de espécies florestais através de imagens macroscópicas. 2012. 116 f. Tese (Doutorado em Informática) - Universidade Federal do Paraná, Curitiba, 2012.

PELLICO-NETTO, S.; BRENA, D. A. Inventário florestal. Curitiba: [s.n.], 1997. 316p.

PEBESMA, E. J. Multivariable geostatistics in S: the gstat package. Computers & Geosciences, v. 30, n. 7, p. 683-691, 2004.

PEBESMA, E. J., BIVAND, R.S. Classes and methods for spatial data in R. R News, v.5, n. 2, 2005.

PEREIRA, M. Estimação da sensibilidade, da especificidade e da curva ROC. 2013. 140f. Dissertação (Mestrado em Estatística) -Universidade do Minho, Braga- Portugal, 2013.

PERROTTA, M. M. Processamento digital básico de imagens de sensores remotos ópticos para uso em mapeamento geológico. Tutorial: Divisão de Sensoriamento Remoto - DISERE, 2005. 43 p.

PONZONI, F. J.; DISPERATI, A. A. Aplicação do modelo SAIL no estudo de reflectância espectral de dosséis de mudas de *Eucalyptus* sp. e de *Aspidosperma* sp. Floresta, Curitiba, v. 21, n. 1/2, p. 45-69, 1994.

PONZONI, F. J. Comportamento espectral da vegetação. In: MENESES, P. R.; MADEIRA NETTO, J. S. (Org.). Sensoriamento remoto: reflectância de alvos naturais. 1. ed. Brasília, DF: UNB, 2001. v. 1, p. 157-199

PONZONI, F. J.; REZENDE, A. C. P. Caracterização espectral de estágios sucessionais de vegetação secundária arbórea em Altamira (PA), através de dados orbitais. Revista Árvore, v. 28, n. 4, p. 535-545, 2004.

PRATI, R. C.; BATISTA, G. E. A. P. A.; MONARD, M. C. Curvas ROC para avaliação de classificadores. Revista IEEE América Latina, v. 6, n. 2, p. 215-222, 2008.

R Development Core Team., R: A language and environment for statistical computing. R Foundation for Statistica Computing, Vienna, Austria, 2016.

RÉ, D. S.; THIERSCH, C. R.; ANDRADE, C. M. Avaliação do uso da Krigagem na definição de estratos para o inventário florestal. In: SIMPÓSIO DE GEOESTATÍSTICA APLICADA EM CIÊNCIAS AGRÁRIAS, Botucatu, 2009. Anais... Botucatu: Universidade Estadual de São Paulo, 2009.

RESMINI, R. et al. Diagnóstico precoce de doenças mamárias usando imagens térmicas e aprendizado de máquina. REAVI-Revista Eletrônica do Alto Vale do Itajaí, Itajaí - SC, v. 1, n. 1, p. 55-67, 2012.

RIBEIRO JÚNIOR, P. J. Métodos geoestatísticos no estudo da variabilidade espacial de parâmetros do solo. 1995. 99 f. Dissertação (Mestrado) – Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, Piracicaba, 1995.

ROBIN, X. et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics, v.12, p. 1-77, 2011. Disponível em: < <https://cran.r-project.org/web/packages/pROC/index.html>> Acesso em: 04 mar. 2016.

ROCHA, G. F. et al. Detecção de desmatamentos no bioma cerrado entre 2002 e 2009: Padrões, tendências e impactos. Revista Brasileira de Cartografia, v. 63, n. 3, p. 341-349, 2011.

RODRIGUES, J. N. da CRUZ. Análise geoestatística multi-pontos. 2013. 87 f. Tese (Doutorado em Ciências) – Universidade de São Paulo, São Paulo, 2013.

ROSA, R. Introdução ao Sensoriamento Remoto. 5ª Ed. Uberlândia: EDUFU, 2003. 238 p.

ROSENDO, J. S. Índices de vegetação e monitoramento do uso do solo e cobertura vegetal na Bacia do Rio Araguari -MG- Utilizando Dados do Sensor Modis. 2005. 130 f. Dissertação (Mestrado em Geografia) – Universidade Federal de Uberlândia, Uberlândia-MG, 2005.

RUFINO, T. M. C. et al. Uso da Geoestatística no estudo da relação entre variáveis dentrométricas de povoamentos de *Eucalyptus* sp. e atributos do solo. Ambiência, Guarapuava-PR, v. 2, n. 3, p. 83-93, edição especial 1,2006.

SAMPAIO, W. B. de Detecção de massas em imagens mamográficas usando redes neurais celulares, funções geoestatísticas e máquinas de vetores de suporte. 2009. 120 f. Dissertação (Mestrado em Engenharia de Eletricidade) - Universidade Federal do Maranhão, São Luís-MA, 2009.

SANTANA, R. A. Avaliação de técnicas geoestatísticas no inventário de povoamento de *Tectona grandis* L.f. 2011. 43 f. Dissertação (Mestrado em Ciências) - Universidade Federal de Viçosa, Viçosa-MG, 2011.

SANTOS, P. C. et al. O emprego da geoestatística na determinação do tamanho "ótimo" de amostras aleatórias com vistas à obtenção de estimativas dos volumes dos fustes de espécies florestais em Paragominas, estado do Pará. Acta Amazônica, v.41, n. 2, p. 213-222, 2011.

SAWAYA, K. E. et al. Extending satellite remote sensing to local scales: land and water resource monitoring using high-resolution imagery. Remote Sensing of Environment, v. 88, n. 1, p. 144-156, Nov. 2003.

SCHIAVONI, A. S. Um estudo comparativo de métodos para balanceamento do conjunto de treinamento em aprendizado de redes neurais artificiais. 2010. 69 f. Monografia (Graduação em Ciências da Computação) - Universidade Federal de Lavras, Lavras - MG, 2010.

SCHOEPFER, W. Autimatisierung des massen, sorten and wertberechnung stenender waaldbestande schriftenreihe bad. Berlim: Wurt-Forstl., 1966.

SCOLFORO, J. R. S.; MELLO, J. M. de. Inventário florestal. Lavras: UFLA/FAEPE, 2006. 561 p.

SÉ, D. C. et al. Use of the correlation coefficient between plots in order to improve the accuracy of forest inventories. Cerne, Lavras-MG, v.19, n.4, p. 575-580. Oct./Dec. 2013.

SETTE, P. G. C. Análise de textura de imagem de alta resolução para a classificação dos estágios sucessionais de Mata Atlântica do Sul da Bahia. 2009. 73 f. Dissertação (Mestrado em Análise e Modelagem de Sistemas Ambientais) - Universidade Federal de Minas Gerais, Belo Horizonte, 2009.

SILVA, E. T. J. B. Utilização dos índices de Vegetação do Sensor MODIS para Detecção de Desmatamentos no Cerrado: Investigação de Parâmetros e Estratégias. 2004, 146 f. Dissertação (Mestrado em Geologia) – Universidade de Brasília, Brasília, 2004.

SILVA, F. C. Análise ROC. Instituto Nacional de Pesquisa Espacial. DPI/INPE. São José dos Campos- SP, 2006. 9 p.

SILVA, M. A. T. B. Uso de matrizes de co-ocorrências para classificação automática de imagens de cana-de-açúcar. 2013. 74 f. Monografia (Graduação em Ciência da Computação) - Universidade Federal de Lavras, Lavras- MG, 2013.

SILVA, S. T. Estratificação da fitofisionomia Cerrado para inventário florestal utilizando geotecnologia. 2009. 68 f. Dissertação (Mestrado em Engenharia Florestal) - Universidade Federal de Lavras, Lavras, 2009.

SILVEIRA, L. S. da Identificação de áreas cultivadas com café por meio de descritores texturais. 2013. 74 f. Dissertação (Mestrado em Engenharia Agrícola) - Universidade Federal de Viçosa, Viçosa-MG, 2013.

SIMS, D. A.; GAMON, J. A. Estimation of vegetation water content and photosynthetic tissue area from spectral reflectance: a comparison of indices based on liquid water and chlorophyll absorption features. Remote Sensing of Environment, v. 84, n. 4, p. 526-537, 2003.

SOARES, T. S. et al. Uso de diferentes alternativas para viabilizar a relação hipsométrica no povoamento florestal. Revista Árvore, Viçosa, v. 28, n. 6, p. 845-854, nov./dez. 2004.

SOLTANIAN-ZADEH, H.; HAFIEE-RAD, F.; POURABDOLLAH-NEJAD, S. Comparison of multiwavelet, wavelet, Haralick, and shape features for microcalcification classification in mammograms. Pattern Recognition, v. 37, n. 10, p. 1973-1986, 2004.

SOUSA, C. L. Uso de imagens-índice e de imagens-proporção para avaliar a quantidade de madeira em povoamentos de *Pinus* spp. 1997. 133f. Dissertação (Mestrado em Sensoriamento Remoto) - INPE, São José dos Campos, 1997.

SOUSA, C. L.; PONZONI, F. J. Avaliação de índices de vegetação e de bandas TM/Landsat para estimativa de volume de madeira em floresta implantada de *Pinus* spp. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 9. 1998, Santos. Anais... Santos: INPE, 1998. p. 1537-1547.

SWETS, J. A.; DAWES, R. M.; MONAHAN, J. Better DECISIONS through science. Scientific American, v. 283, p. 82-87, 2000.

SZEKIELDA, K. Satellite monitoring of the earth. New York, John Wiley, 1988. 326p.

TERRA, F. da S.; FONSECA, E. L. da; SALDANHA, D. L. Contribuições do solo e dossel em modelo de estimativa de biomassa aérea no Bioma Pampa. Pesquisa Agropecuária Brasileira, v. 45, n. 4, p. 384-390, 2010.

TRANGMAR, B. B.; YOST, R. S.; UEHARA, G.. Application of geostatistics to spatial studies of soil properties. Advances in agronomy, New York, v. 38, p. 45-94, 1985.

VENABLES, W. N.; RIPLEY, B. D. Modern Applied Statistics with S. Fourth Edition. Springer, New York, 2002.

VIEIRA, R. T. Análise de micropadrões em imagens digitais baseada em números fuzzy. 2013. 155 f. Dissertação (Mestrado em Engenharia Elétrica) - Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2013.

WACKERNAGEL, H. Multivariate Geostatistics. Berlin: Springer-Verlag, 2003. 387 p.

WATANABE, J. et al. Estudo da influência da correlação inicial entre variáveis nos resultados de co-estimativas Geociências, São Paulo, v. 28, n. 4, p. 467-484, 2009.

WATZLAWICK L. F; KIRCHNER F. F; SANGUETTA C. R. Estimativa de biomassa e carbono em floresta com araucária utilizando imagens do satélite IKONOS II. Ciência Florestal, Santa Maria v. 19, n. 2, p. 169-181, abr./jun. 2009.

WOODCOCK, C. E. et al. Mapping forest vegetation using Landsat TM imagery and a canopy reflectance model. Remote Sensing of Environment, New York, v. 50, n. 3, p. 240-254, 1994.

YAMAMOTO, J. K. Avaliação e classificação de reservas minerais. São Paulo: Editora da Universidade de São Paulo, 2001. 226p.

YAMAMOTO, J. K.; LANDIM, P. M. B. Geoestatística: conceitos e aplicações. São Paulo: Oficina de Textos, 2015. 215 p.