

**APLICAÇÃO DE MÉTODOS GEOESTATÍSTICOS PARA
IDENTIFICAÇÃO DE DEPENDÊNCIA ESPACIAL NA
ANÁLISE DE DADOS DE UM EXPERIMENTO EM
DELINEAMENTO SISTEMÁTICO TIPO “LEQUE”**

MELISSA LOMBARDI ODA

Dissertação apresentada à Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, para obtenção do título de Mestre em Agronomia, Área de Concentração: Estatística e Experimentação Agronômica.

P I R A C I C A B A
Estado de São Paulo - Brasil
Março - 2005

**APLICAÇÃO DE MÉTODOS GEOESTATÍSTICOS PARA
IDENTIFICAÇÃO DE DEPENDÊNCIA ESPACIAL NA
ANÁLISE DE DADOS DE UM EXPERIMENTO EM
DELINEAMENTO SISTEMÁTICO TIPO “LEQUE”**

MELISSA LOMBARDI ODA

Engenheiro Agrônomo

Orientador: Prof^o Dr^o DÉCIO BARBIN

Dissertação apresentada à Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, para obtenção do título de Mestre em Agronomia, Área de Concentração: Estatística e Experimentação Agronômica.

PIRACICABA
Estado de São Paulo - Brasil
Março - 2005

**Dados Internacionais de Catalogação na Publicação (CIP)
DIVISÃO DE BIBLIOTECA E DOCUMENTAÇÃO - ESALQ/USP**

Oda, Melissa Lombardi

Aplicação de métodos geoestatísticos para identificação de dependência espacial na análise de dados de um experimento em delineamento sistemático tipo "Leque" / Melissa Lombardi Oda. - - Piracicaba, 2005.
72 p.

Dissertação (mestrado) - - Escola Superior de Agricultura Luiz de Queiroz, 2005.
Bibliografia.

1. Corte (Plantas) 2. Espaçamento 3. Estatística aplicada 4. Estatística espacial
5. Eucalipto 6. Geoestatística 7. Verossimilhança I. Título

CDD 634.9734

"Permitida a cópia total ou parcial deste documento, desde que citada a fonte – O autor"

DEDICATÓRIA

A

DEUS

Senhor todo poderoso.

Autor da minha vida e senhor do meu destino.

Aos meus pais,

Lucy Ney Lombardi Oda e Milton Isao Oda,
tias,

Elza, Laura, Lira, Luiza,

Mery, Tereza

e vós,

Kaor (*in memorian*) e Zélia,

pelo amor incondicional que foi fundamental nesta conquista, pelo apoio, dedicação e confiança, que nortearam minha vida e me fizeram progredir.

Ao meu companheiro

Bruno Almeida de Souza

o amor, compreensão, incentivo e, principalmente, por se manter sempre forte nos momentos decisivos.

AGRADECIMENTOS

Ao Prof^o Dr^o Décio Barbin a orientação, a amizade, ensinamentos, apoio e crédito ao meu trabalho. Por acreditar e me fazer acreditar que chegar até aqui seria possível, o meu agradecimento especial.

Ao Prof^o Dr^o Paulo Justiniano Ribeiro Júnior do Departamento de Estatística da Universidade Federal do Paraná a amizade, a paciência, os ensinamentos, as incansáveis discussões e sugestões ao meu trabalho.

Ao Prof^o Dr^o José Luiz Stape do Departamento de Ciências Florestais ESALQ/USP a paciência, os ensinamentos, as sugestões e que gentilmente cedeu dados para a realização deste trabalho.

Aos amigos do mestrado e doutorado, em especial a Ana Alice Pilon, Denise Viola, Juliana Garcia Cespedes, David Miquelluti, Adriano Ferreti Borgatto, Antônio Willians Moita, Willian Saconato, Ana Maria Araújo, João Mauricio Araújo Mota, Elisabeth Strapasson e aos amigos Tammy Kiihl, Gustavo Bastos Lyra, Alailson Santiago, Evandro Zanini Righi, Magda Fidelis de Lima, Daniela Talora, Antônio Batista Sampaio, Natacha Lima, Wyratan da Silva Santos a amizade, o apoio e a compreensão.

Aos amigos da Universidade Federal do Paraná Elias Teixeira Krainski, Marcos Aurélio Carrero e Pedro Ribeiro de Andrade Neto, a amizade e ajuda.

À Solange Sabadin e Luciane Brajão e aos demais funcionários do Departamento de Ciências Exatas ESALQ/USP a paciência, atenção e respeito.

Ao responsável pelos laboratórios de informática, Jorge Alexandre Wiendl,

e aos funcionários e estagiários da Estação Experimental de Itatinga, em especial Rildo Moreira e Moreira e Estevão Araújo, a atenção e ajuda.

Aos professores do Departamento de Ciências Exatas da ESALQ/USP, em especial a Prof^a Dr^a Clarice Garcia Borges Demétrio, que me propiciaram condições para a realização deste trabalho.

Ao CNPQ, o apoio financeiro.

À bibliotecária Silvia Zinsly a revisão das normas.

A todos que de forma direta ou indireta contribuíram para realização deste novo desafio em minha vida agradeço.

SUMÁRIO

	Página
LISTA DE FIGURAS	viii
LISTA DE TABELAS	x
RESUMO	xii
SUMMARY	xiv
1 INTRODUÇÃO	1
2 REVISÃO DE LITERATURA	3
2.1 Delineamentos sistemáticos	3
2.1.1 Delineamento sistemático tipo “leque”	7
2.2 Geoestatística	8
2.2.1 Geoestatística aplicada em experimentos de campo	10
2.2.2 Semivariograma	11
2.3 Modelos geoestatísticos	14
2.3.1 Funções de correlação	16
2.3.2 Modelos teóricos de semivariogramas	17
2.3.3 Métodos de estimação dos parâmetros do semivariograma	18
2.3.4 Critério de Akaike na seleção de modelos	20
3 MATERIAL E MÉTODOS	21
3.1 Material	21
3.1.1 Localização do experimento	21
3.1.2 Avaliações dendrométricas	21
3.1.3 Variável de interesse	21
3.2 Métodos	22
3.2.1 Instalação e caracterização do experimento sistemático tipo “leque”	22
3.2.2 Falhas e bifurcações	24
3.2.3 Análise Clássica	25

3.2.3.1	Modelo considerando a eliminação das plantas vizinhas	26
3.2.4	Análise de Covariância	26
3.2.4.1	Modelo com a covariável morte da planta	27
3.2.4.2	Modelo com as covariáveis morte da planta e tempo	28
3.2.5	Construção das covariáveis	28
3.2.6	Transformação de dados	28
3.2.7	Análise exploratória espacial	29
3.2.8	Modelo geoestatístico	30
3.2.9	Construção do semivariograma empírico	31
3.2.10	Ajuste de um modelo de semivariograma teórico para as semivariâncias	32
3.2.11	Estimação de máxima verossimilhança	32
3.2.12	Critério de Akaike na seleção de modelos	33
3.2.13	Conjunto de dados e programas	33
4	RESULTADOS E DISCUSSÃO	34
4.1	Transformação de dados	34
4.2	Análise exploratória espacial	35
4.3	Resultados para os modelos estudados	37
5	CONCLUSÕES	41
	ANEXOS	42
	REFERÊNCIAS BIBLIOGRÁFICAS	67

LISTA DE FIGURAS

		Página
1	Valores do raio inicial (r_0), raios dos tratamentos (r_1 a r_n), ângulo entre os raios (θ) e área associada a cada planta (A_1 a A_n) no delineamento sistemático em leque	8
2	Semivariograma com efeito pepita	13
3	Croqui de instalação do experimento sistemático “leque” com 10 tratamentos e 36 repetições com <i>Eucalyptus dunnii</i>	23
4	Gráfico da média x variância para o modelo I (a), e modelos II e III (b)	36
5	Localização geográfica dos resíduos da variável volume considerando as variáveis explanatórias para os modelos I (a), II (b) e III (c)	37
6	Ajuste dos semivariogramas teóricos aos valores do semivariograma empírico e envelope simulado para o Modelo I. Parâmetros ajustados por máximo verossimilhança considerando-se as estruturas de covariância exponencial, Matérn com $\kappa = 1$ e Matérn com $\kappa = 2$	38
7	Ajuste dos semivariogramas teóricos aos valores do semivariograma empírico e envelope simulado para o Modelo II. Parâmetros ajustados por máximo verossimilhança considerando-se as estruturas de covariância exponencial, Matérn com $\kappa = 1$ e Matérn com $\kappa = 2$	39
8	Ajuste dos semivariogramas teóricos aos valores do semivariograma empírico e envelope simulado para o Modelo III. Parâmetros ajustados por máximo verossimilhança considerando-se as estruturas de covariância exponencial, Matérn com $\kappa = 1$ e Matérn com $\kappa = 2$	39
9	Resíduos vs valores observados (a), “box-plot” (b), gráfico de probabilidade normal (c) e distribuição de freqüências (d) para o modelo I	65
10	Resíduos vs valores observados (a), “box-plot” (b), gráfico de probabilidade normal (c) e distribuição de freqüências (d) para o modelo II	65

11 Resíduos vs valores observados (a), “box-plot” (b), gráfico de probabilidade normal (c) e distribuição de frequências (d) para o modelo III 66

LISTA DE TABELAS

		Página
1	Valores da distância radial, área por planta e densidade de plantas para 10 tratamentos (espaçamentos) no delineamento sistemático tipo “leque”	23
2	Número do raio das plantas devido à morte no delineamento sistemático tipo “leque”, aos 6 anos	25
3	Números dos raios das plantas que serão eliminadas, e suas vizinhas devido a morte no delineamento sistemático tipo “leque”, aos 6 anos	25
4	Resultados do teste de Shapiro-Wilk para os resíduos da variável volume (m^3) nos modelos I, II e III	34
5	Resultados dos testes de Bartlett e Hartley para a variável volume (m^3) dos modelos I, II e III	35
6	Resultados das estimativas dos parâmetros, $(\hat{\sigma}^2, \hat{\tau}^2, \hat{\phi})$, logaritmo da função de verossimilhança (Log), número de parâmetros do modelo considerado (Npars) e critério de Akaike (AIC) para os modelos I, II e III, obtidas pelo método de estimação de parâmetros do modelo geoestatístico utilizando diferentes modelos de função de correlação	40
7	Conjunto de dados para o tratamento 1, aos 6 anos. Ap - Área por Planta, N° Raio - Número do raio das plantas, X1 e X2 (m) - Coordenadas, Trat - Tratamento, H (m) - Altura, Dap (cm) - Diâmetro altura do peito, Vol (m^3) - Volume, Mp - Covariável morte da planta e T - Covariável tempo	43
8	Conjunto de dados para o tratamento 2, aos 6 anos. Ap - Área por Planta, N° Raio - Número do raio das plantas, X1 e X2 (m) - Coordenadas, TRAT - Tratamento, H (m) - Altura, Dap (cm) - Diâmetro altura do peito, Vol (m^3) - Volume, Mp - Covariável morte da planta e T - Covariável tempo	44

9	Conjunto de dados para o tratamento 3, aos 6 anos. Ap - Área por Planta, N° Raio - Número do raio das plantas, X1 e X2 (m) - Coordenadas, Trat - Tratamento, H (m) - Altura, Dap (cm) - Diâmetro altura do peito, Vol (m^3) - Volume, Mp - Covariável morte da planta e T - Covariável tempo	45
10	Conjunto de dados para o tratamento 4, aos 6 anos. Ap - Área por Planta, N° Raio - Número do raio das plantas, X1 e X2 (m) - Coordenadas, Trat - Tratamento, H (m) - Altura, Dap (cm) - Diâmetro altura do peito, Vol (m^3) - Volume, Mp - Covariável morte da planta e T - Covariável tempo	46
11	Conjunto de dados para o tratamento 5, aos 6 anos Ap - Área por Planta, N° Raio - Número do raio das plantas, X1 e X2 (m) - Coordenadas, Trat - Tratamento, H (m) - Altura, Dap (cm) - Diâmetro altura do peito, Vol (m^3) - Volume, Mp - Covariável morte da planta e T - Covariável tempo	47
12	Conjunto de dados para o tratamento 6, aos 6 anos. Ap - Área por Planta, N° Raio - Número do raio das plantas, X1 e X2 (m) - Coordenadas, Trat - Tratamento, H (m) - Altura, Dap (cm) - Diâmetro altura do peito, Vol (m^3) - Volume, Mp - Covariável morte da planta e T - Covariável tempo	48
13	Conjunto de dados para o tratamento 7, aos 6 anos. Ap - Área por Planta, N° Raio - Número do raio das plantas, X1 e X2 (m) - Coordenadas, Trat - Tratamento, H (m) - Altura, Dap (cm) - Diâmetro altura do peito, Vol (m^3) - Volume, Mp - Covariável morte da planta e T - Covariável tempo	49
14	Conjunto de dados para o tratamento 8, aos 6 anos. Ap - Área por Planta, N° Raio - Número do raio das plantas, X1 e X2 (m) - Coordenadas, Trat - Tratamento, H (m) - Altura, Dap (cm) - Diâmetro altura do peito, Vol (m^3) - Volume, Mp - Covariável morte da planta e T - Covariável tempo	50
15	Conjunto de dados para o tratamento 9, aos 6 anos. Ap - Área por Planta, N° Raio - Número do raio das plantas, X1 e X2 (m) - Coordenadas, Trat - Tratamento, H (m) - Altura, Dap (cm) - Diâmetro altura do peito, Vol (m^3) - Volume, Mp - Covariável morte da planta e T - Covariável tempo	51
16	Conjunto de dados para o tratamento 10, aos 6 anos. Ap - Área por Planta, N° Raio - Número do raio das plantas, X1 e X2 (m) - Coordenadas, Trat - Tratamento, H (m) - Altura, Dap (cm) - Diâmetro altura do peito, Vol (m^3) - Volume, Mp - Covariável morte da planta e T - Covariável tempo	52

APLICAÇÃO DE MÉTODOS GEOESTATÍSTICOS PARA IDENTIFICAÇÃO DE DEPENDÊNCIA ESPACIAL NA ANÁLISE DE DADOS DE UM EXPERIMENTO EM DELINEAMENTO SISTEMÁTICO TIPO “LEQUE”

Autora: MELISSA LOMBARDI ODA

Orientador: Prof^o Dr^o DÉCIO BARBIN

RESUMO

Os delineamentos sistemáticos são usados nas mais diversas áreas, como: florestal, horticultura, solos, etc. Na área florestal, os delineamentos sistemáticos são frequentemente usados para estudos preliminares e têm o objetivo de testar o maior número de espaçamentos possíveis. No entanto, existem algumas limitações para a sua utilização. A primeira é o arranjo sistemático (não casualizado) das plantas, que não permite o uso das análises convencionais. A segunda é a alta sensibilidade para valores perdidos. Quando uma planta é perdida, o espaçamento das plantas vizinhas é alterado, assim esses valores não podem ser incluídos no conjunto de dados e informações consideráveis são excluídas das análises. O objetivo deste trabalho foi aplicar a metodologia geoestatística para identificação de dependência espacial em um experimento em delineamento sistemático tipo “leque”, levando-se em consideração: a eliminação dos dados das plantas vizinhas aos valores perdidos e as informações de ocorrência de parcelas perdidas e o tempo que ocorreram. Os dados de volume sólido por planta utilizados neste trabalho são provenientes de um experimento de espaçamento

de *Eucalyptus dunnii* em delineamento sistemático tipo “leque”. Neste trabalho foram utilizados os dados referentes ao sexto ano, idade comercial de corte da espécie, com os seguintes procedimentos: eliminação dos dados das plantas vizinhas às plantas mortas (Modelo I); as informações de mortes das plantas foram consideradas como uma covariável no modelo (Modelo II) e além da covariável morte das plantas, também foi levado em consideração o tempo da ocorrência da morte (Modelo III). Os parâmetros do semivariograma foram estimados pelo método de máxima verossimilhança e para seleção de modelos, utilizou-se o Critério de Akaike (AIC). Os resultados obtidos permitem concluir que se identificou uma fraca dependência espacial, o que não justificaria considerá-la com a aplicação de um modelo geoestatístico. A função de correlação que apresentou melhor desempenho foi a Matérn com $\kappa = 2$ para os três modelos considerados. Comparando-se esses modelos e seguindo o critério de Akaike, o modelo mais adequado foi o II, pois apresentou menor valor de AIC.

APPLICATION OF GEOSTATISTICAL METHODS TO IDENTIFY THE SPATIAL DEPENDENCE IN THE DATA ANALYSIS OF A FAN SYSTEMATIC DESIGN EXPERIMENT

Author: MELISSA LOMBARDI ODA

Adviser: Prof^o Dr^o DÉCIO BARBIN

SUMMARY

Systematic designs are utilized in many areas, such as: forestry, horticulture, soils, etc. In forestry, the systematic designs are frequently used for preliminary studies and they aim at evaluating the largest number of possible spacings. However, there are some limitations on their use. The first limitation is the systematic design (non-randomized) of plants, which does not allow the use of conventional analyses. The second is the high sensitivity to lost values. When a plant is lost, the neighboring plant spacings are altered, so these values cannot be added to the data collection, and a great sum of information is excluded from the analyses. This study aimed at applying geostatistical methods to identify the spatial dependence in the data analysis of a fan systematic design experiment, taking into account: the exclusion of neighboring plant data to the lost values and the information regarding the occurrence of lost parcels as well as the time of their occurrence. The plant solid volume data utilized in this study were taken from a fan systematic design *Eucalyptus dunnii* spacing study. The data utilized were referent to the sixth year, commercial age for cutting of

the specie, with the following procedures exclusion of the data from a neighboring plant next to a dead tree (Model I); the information of tree mortality as covariable in the model (Model II); and the time of occurrence of tree mortality, besides the tree mortality covariable (Model III). The semivariogram parameters were estimated by the maximum likelihood method, and the model selection was done by the utilization of the Akaike's Information Criterion (AIC). It was possible to conclude from the result analyses that there is a weak spatial dependence, which does not justify neither taking it into account nor the utilization of a geostatistical model. The correlation function that showed the best performance was the Matérn, with $\kappa=2$ for the three models considered. By the comparison of these three models and the utilization of the Akaike's Information Criterion, the most suitable model was Model II, as it showed lower AIC value.

1 INTRODUÇÃO

O principal idealizador dos métodos estatísticos utilizados nas ciências experimentais modernas foi o matemático inglês Ronald A. Fisher. Ao ser nomeado, em 1919, por Sir John Russel para trabalhar na estação experimental em Rothamsted, desenvolveu métodos para o planejamento e a análise de experimentos de campo (Pontes, 2002). Estes métodos são baseados em três princípios: repetição, casualização e controle local (Barbin, 2003).

A primeira contribuição de Fisher para a experimentação agrícola, de acordo com Greenberg (1951), foi o emprego da técnica de análise de variância na avaliação de resultados de experimentos agrícolas. Nesse mesmo período, Fisher enfatizou a importância de se fazerem repetições e a atribuição aleatória dos tratamentos às unidades experimentais.

A casualização é um procedimento inerente à locação dos tratamentos nas unidades experimentais, de tal modo que todas as unidades experimentais tenham a mesma probabilidade de receber um dado tratamento, por meio de sorteio. Quando a casualização foi proposta por Fisher, teve por finalidade tanto contornar o problema da variabilidade ambiental (continuidade espacial) entre as unidades experimentais, quanto criar uma distribuição de amostragem que suportasse uma modelagem da incerteza presente nas conclusões dos experimentos.

Segundo Yates (1964), Fisher percebeu, após a análise de vários experimentos arranjados sistematicamente, que muitas suposições eram questionáveis. Com frequência, elas se demonstravam falsas com relação à independência dos erros provenientes das unidades experimentais. Isto poderia ser evitado com a casualização dos tratamentos. Assim a casualização, pelo menos aproximadamente, proporcionaria erros experimentais independentes e identicamente distribuídos, com média nula e variância fixa.

Kempthorne (1977) também concorda que a justificativa inferencial para a prática da casualização na pesquisa experimental, concentra-se na criação de uma distribuição de amostragem objetiva para os erros experimentais. Assim, o comportamento assintótico dessa distribuição aproxima-se de normais não-correlacionadas.

Por causa da casualização, condição julgada imprescindível e defendida até hoje, experimentos e levantamentos amostrais não-casualizados têm sido considerados como inadequados para inferências estatísticas válidas (Pontes, 2002). Conseqüentemente, ainda hoje as principais técnicas utilizadas no planejamento de experimentos agrícolas e na análise de dados continuam alicerçados nos princípios fisherianos.

De acordo com Pearce¹ citado por Teixeira (2001), apesar da ênfase que têm tido os delineamentos aleatorizados, há casos em que a aleatorização é impraticável ou não adequada. Segundo o autor, os ensaios com espaçamentos, são difíceis de serem casualizados, pois tratamentos extremos não devem ocorrer juntos. A aleatorização de espaçamentos implica na utilização de grandes áreas tanto de parcelas como de bordaduras, além de ser praticamente impossível aleatorizar, em se tratando de plantio mecanizado. Como solução de tais dificuldades, o autor aponta para o uso dos delineamentos sistemáticos propostos por Nelder (1962).

Os delineamentos sistemáticos são usados nas mais diversas áreas, como: florestal, horticultura, solos, etc.

Na área florestal, os delineamentos sistemáticos são freqüentemente usados para estudos preliminares e têm o objetivo de testar o maior número de espaçamentos possíveis. No entanto, Bar-Hen (2002) cita duas limitações na utilização desse delineamento. A primeira é que o arranjo sistemático (não casualizado) das plantas, não permite o uso das análises convencionais. A segunda é a alta sensibilidade para valores perdidos. Quando uma planta é perdida, o espaçamento das plantas vizinhas é alterado e, com isso, esses valores não podem ser incluídos no conjunto de dados e informações consideráveis são excluídas das análises.

O objetivo deste trabalho é aplicar a metodologia geoestatística para identificação de dependência espacial em um experimento em delineamento sistemático tipo “leque”, levando-se em consideração:

- a) a eliminação dos dados das plantas vizinhas aos valores perdidos;
- b) as informações de ocorrência de parcelas perdidas e o tempo que ocorrem.

¹PEARCE, S.C. **The agriculture field experiment**. New York: John Wiley, 1983. 335p.

2 REVISÃO DE LITERATURA

2.1 Delineamentos sistemáticos

Com a introdução dos conceitos de repetição, aleatorização dos tratamentos, controle local e, em alguns casos, a bordadura em estatística e experimentação, por Fisher, houve uma reformulação das bases teóricas envolvidas no planejamento experimental, sendo que, o uso da aleatorização em experimentos de campo foi o principal deles. Entretanto, a instalação dos experimentos fisherianos, atualmente, pode na prática, restringir o número ou tamanho de ensaios, dadas as dificuldades relativas à disponibilidade de área e de recursos para instalação, manutenções e avaliações.

Segundo Minami (1977), experimentos que ocupam grandes áreas são estatisticamente indesejáveis e, muitas vezes, tornam-se impraticáveis devido à falta de recursos como mão-de-obra, área experimental, e pela dificuldade de serem controlados. Em decorrência, os pesquisadores tendem a diminuir o número de tratamentos, prejudicando a execução do projeto todo.

Além disso, mesmo em se querendo testar um grande número de tratamentos, a área experimental total requerida seria tão grande que certamente não se conseguiria garantir a homogeneidade dentro dos blocos, inflacionando o erro experimental e, conseqüentemente, diminuindo a sensibilidade na comparação dos efeitos dos tratamentos. Visando superar, em parte, tais restrições, pesquisadores (Nelder, 1962; Bleasdale, 1966) propuseram no início da década de 60, delineamentos que não levam em consideração o princípio da aleatorização, em que as bordaduras seriam diminuídas, ou eliminadas, e ao se disporem os tratamentos em gradientes crescentes, poderiam dispensar as plantas de bordaduras, pois os efeitos das parcelas anterior e posterior em relação a uma determinada parcela, seriam compensatórios (Stape, 1995).

Um delineamento sistemático consiste numa malha de pontos, cada um representando a posição de uma planta, sendo que cada área por planta e/ou retangularidade, ou

seja, a razão entre as distâncias inter e intra linhas, é constante e muda conforme um modelo consistente nas diferentes partes da malha. Para que a planta seja facilmente colocada na posição correta no campo, Nelder (1962) impõe a condição de que a malha de pontos seja fixada pela intersecção de duas linhas, as paralelas, ou linhas retas, e os arcos de círculos concêntricos. A segunda condição é que os contornos da área e da forma sejam simples para que o experimento possa ser colhido em parcelas limitadas pelos arcos destes contornos. Uma terceira é que a densidade e as retangularidades em estudo variem dentro da malha para que se possa estudar um número maior de diferentes combinações. Esta condição limita as formas de delineamento possíveis e somente se conhecem cinco formas. Destas cinco, quatro se baseiam em raios e arcos circulares concêntricos, enquanto que a quinta se baseia em retângulos.

Pela não casualização, os delineamentos sistemáticos ferem o princípio básico na alocação dos tratamentos, não garantindo a independência dos erros associados a cada unidade experimental, o que impossibilita a sua análise estatística de forma tradicional. Stape (1995) ressalta que a falta de aleatorização vem sendo, sem dúvida alguma, a principal preocupação na recomendação ou não, desse delineamento no estudo de espaçamentos, por impossibilitar uma correta estimativa e comparação entre os efeitos de tratamentos, devido à impossibilidade de se estimar o erro experimental. O autor sugere o uso de modelos de regressão nos delineamentos sistemáticos, tendo como variável dependente alguma medida de crescimento ou produção, e como variável independente a densidade de plantas. No entanto, essas regressões, embora úteis na comparação relativa entre tratamentos dentro do próprio ensaio, não possibilitam a determinação do desvio das estimativas de produção por área no delineamento e aquelas que seriam auferidas em plantios comerciais sob aquela densidade. Essa dificuldade tem sido contornada elegendo-se tratamentos promissores no delineamento sistemático e repetindo-os, agora num delineamento tradicional casualizado, para definição correta dos espaçamentos e suas estimativas.

Uma das vantagens do delineamento sistemático é o seu uso para ensaios preliminares (Leeuwen, 1994). Ocupam muito menos espaço, permitem incluir tratamentos extremos e são fáceis de observar com grande valor demonstrativo. Willey & Rao (1981) ressaltam ainda que são especialmente úteis no estudo inicial de respostas básicas, têm potencial considerável, mas são relativamente inexplorados.

Numa revisão sobre delineamentos sistemáticos, dois tipos básicos foram identificados por Freeman (1988). O primeiro em que a sistematização é necessária em função

de uma particularidade inerente ao material amostral disponível, como experimentos em indústrias e na área médica; e o segundo grupo em que a sistematização é devida à existência de níveis quantitativos contínuos nos tratamentos, os quais são dispostos numa ordem crescente ou decrescente, nas unidades experimentais, objetivando reduzir ou eliminar as bordaduras, diminuir a área experimental e aumentar o número de tratamentos em teste. Em seu estudo Chalita (1991) concluiu que o uso da sistematização quando não há fator quantitativo, não se justifica e não traz ganho algum à experimentação.

Porém, umas das maiores críticas aos delineamentos sistemáticos, com fator contínuo, para Mead (1988), é que faltam métodos de análise que possibilitem utilizar apropriadamente os dados experimentais obtidos nestes delineamentos, e também se devem a uma visão restrita da experimentação na sua realidade prática. Mas a recomendação dos delineamentos sistemáticos para Rao et al. (1990) é dependente do balanço entre a precisão estatística e a sua praticidade agrônômica.

Para a instalação de um delineamento sistemático, Huxley (1985), Pearce (1989) e Rao et al. (1990) apontam os seguintes critérios: a área deverá ser uniforme, plana e com tratos culturais adequadamente realizados.

A escolha de uma área uniforme para instalação de um experimento sistemático é baseada apenas na opinião do pesquisador, segundo Cox (1958), e quando um resultado surpreendente ocorre, o pesquisador imediatamente questiona a validade do arranjo sistemático. O autor ainda classifica a determinação de arranjo dos tratamentos dentro de um determinado bloco em aleatório, sistemático e subjetivo. No aleatório, cada tratamento é associado a uma parcela experimental, aleatoriamente, e sem reposição. No sistemático, a ordem dos tratamentos está sequencialmente definida em função de se ter um bom conhecimento das variações não controladas e de haver uma forte justificativa para esse arranjo. No subjetivo, o pesquisador associa, por critérios pessoais, os tratamentos às unidades experimentais. O autor condena o arranjo subjetivo, que diferentemente dos anteriores, possibilita a existência de tendenciosidade do pesquisador, e a eliminação do “viés” na estimativa do erro do sistemático pode, provavelmente, ser obtida por uma modificação do método de análise, mas que os diferentes métodos não deixam claro qual deva ser utilizada. Finalmente, pondera que os delineamentos sistemáticos devem ser tolerados dentro de limites bem definidos de conhecimento do pesquisador e de sua praticidade. Muitos experimentos sistemáticos são sugeridos e colocados em campo. Como exemplo têm-se os trabalhos de Lutwick et al. (1980), Huxley

& Mangui (1978), Bauder et al. (1975), Freyman & Dolman (1971), Cleaver et al. (1970), Fox (1973), Freeman (1964) entre outros. Entretanto, inexistem trabalhos efetuando análises estatísticas, para se porem em prova seus aspectos positivos e negativos.

Apesar das preocupações relativas à estimativa do erro experimental e, conseqüentemente, à análise estatística desse tipo de experimento, Wright (1976) considera que as grandes vantagens práticas dos delineamentos sistemáticos com fator quantitativo contínuo, aliadas ao forte apelo intuitivo de sua concepção mais compacta e abrangente, levaram também a área florestal, espelhando-se nas áreas hortícola e agrícola, a utilizar-se dos mesmos em seus estudos de espaçamentos.

Ainda que os delineamentos de espaçamento sistemático não sejam muito usados, há muitas razões pelas quais eles deveriam tornar-se mais populares no futuro. O espaço de terra para ensaios é, muitas vezes, limitado e, ao comparar os delineamentos sistemáticos com os aleatórios, temos as seguintes vantagens: delineamentos mais compactos, redução de heterogeneidade local e inexistência de bordadura. Entre as desvantagens relaciona o risco de gradientes sistemáticos causando “viés” nas estimativas dos efeitos e no erro, e no risco de interação entre parcelas vizinhas (Nelder, 1962; Rogers, 1972). Segundo Stevens (1949), um experimento sistemático não atinge o seu próprio objetivo - a eliminação do efeito da heterogeneidade do campo - a não ser que seja submetido a uma análise estatística rigorosa. Essa análise, geralmente dá mais trabalho que a dos delineamentos modernos que contêm o princípio da aleatorização. Mas atualmente está afirmação não se justifica, pois os recursos computacionais estão cada vez mais sofisticados e simples de serem implementados.

Em junho de 1961, foi realizada uma conferência em Oxford, com o objetivo de evitar a utilização desordenada de experimentos sistemáticos, como os propostos por Dawkins (1960). Segundo Scott (1962), reforçou-se a objeção aos delineamentos sistemáticos, pela impossibilidade de sua análise estatística e sua utilização prática. Uma forte objeção apontada por Fisher (1935) aos delineamentos sistemáticos, é a dificuldade em se estimar o erro experimental, ressaltando ainda que o arranjo dos tratamentos pode coincidir com algum padrão de variação não controlada, viciando as estimativas dos efeitos dos tratamentos e a estimativa da variância das estimativas desses efeitos. Mas Yates (1938) ressalta uma importante observação sobre a discussão entre aleatorização e sistematização. O autor acha que não se deve discutir mais o fim da aleatorização, mas sim avaliar a sua real necessidade em certos casos.

2.1.1 Delineamento sistemático tipo “leque”

Dos cinco delineamentos sistemáticos propostos por Nelder (1962), destaca-se o “leque”, que se baseia num sistema de raios e arcos de círculos concêntricos, onde a retangularidade é constante. Nesse delineamento, a área por planta aumenta com o aumento da distância à origem, sendo que os raios separados por um ângulo constante e os arcos separados por uma progressão geométrica da distância radial. Os três principais valores que definem esse delineamento são r_0 (primeira distância radial), α (fator da progressão geométrica radial) e θ (ângulo entre raios), também destacando-se a retangularidade τ e a área associada por planta A_n (Figura 1).

Para se usar o delineamento proposto por Nelder (1962) e reforçado por Bleasdale (1966), é necessário primeiro calcular o fator de espaçamento α através de:

$$\log \alpha = \frac{\log A_m - \log A}{2N - 2}. \quad (1)$$

em que,

- i) α = fator de espaçamento;
- ii) A_m = a área máxima por planta;
- iii) A = área mínima por planta e;
- iv) $n = 1, 2, \dots, N, N + 1$ = número de arcos a serem plantados em que $(N+1)$ será o último arco representando a bordadura.

Nelder (1962) chama a atenção para o fato de que α não deve exceder 1,111. Se isso acontecer, será necessário aumentar o número de arcos ou diminuir a diferença entre o maior espaçamento e o menor.

O ângulo entre os raios (θ) é calculado da seguinte forma:

$$\theta = \tau(\alpha - 1)\sqrt{\alpha}. \quad (2)$$

em que τ é a retangularidade da disposição das plantas (constante nesse delineamento).

A distância da primeira planta, medida a partir do centro será:

$$r_0 = \sqrt{\frac{2A}{\theta(\alpha^3 - \alpha)}}. \quad (3)$$

A distância das plantas subsequentes ao longo dos raios a partir do centro, r_1 , r_2, \dots, r_n é dada por:

$$r_n = r_0 \alpha^n. \quad (4)$$

A área por planta, associada a cada arco, no delineamento é dada por:

$$A_n = \frac{r_n^2 \theta (\alpha^2 - 1)}{2\alpha}. \quad (5)$$

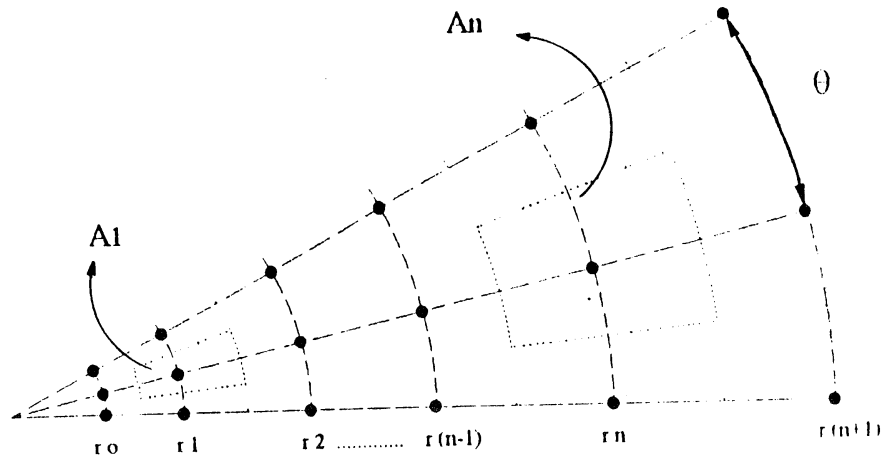


Figura 1 - Valores do raio inicial (r_0), raios dos tratamentos (r_1 a r_n), ângulo entre os raios (θ) e área associada a cada planta (A_1 a A_n) no delineamento sistemático em leque

2.2 Geoestatística

Compreender a distribuição espacial de dados oriundos de fenômenos ocorridos no espaço constitui hoje um grande desafio para a elucidação de questões centrais em diversas

áreas do conhecimento, seja em saúde, em ambiente, em geologia, em agronomia, entre tantas outras. A ênfase da análise espacial é estudar propriedades e relações, levando em conta a localização espacial do fenômeno em observação, ou seja, a idéia central é incorporar o espaço à análise que se deseja fazer (Câmara et al., 2001).

A geoestatística é um tópico especial da estatística aplicada que trata de problemas referentes às variáveis regionalizadas, aquelas que têm comportamento espacial mostrando características intermediárias entre as variáveis verdadeiramente aleatórias e as totalmente determinísticas (Landim, 1998).

Para Ribeiro Júnior (1995) o fato da teoria das variáveis regionalizadas ter sido apresentada com uma notação quase própria e com um novo nome de geoestatística, acabou por gerar algumas confusões. Muitos autores se referem a estatística clássica e a geoestatística como se fosse possível fazer uma distinção. A geoestatística é construída a partir de conceitos básicos da estatística, em particular processos estocásticos (funções aleatórias). Assim, as variáveis têm em comum uma dupla característica: são aleatórias já que os valores numéricos observados podem variar consideravelmente de um ponto a outro no espaço e são espaciais porque apesar de muito variáveis dentro do espaço, os valores numéricos observados não são inteiramente independentes (Guerra, 1988).

Assim, Ribeiro Júnior (1995) ressalta que a geoestatística não se refere a um tipo especial, diferente ou alternativo de estatística. O fato é que cada observação é descrita não apenas pelo seu valor, mas também por informação de sua posição, expressa por um sistema de coordenadas. Considerando que observações mais próximas geograficamente tendem a ter valores mais similares e que tal fato pode ser avaliado por medidas de associação, a análise geoestatística determina o grau de associação (correlação) entre amostras, com base na direção e na distância entre elas. A dependência espacial (autocorrelação espacial) faz com que os pontos de amostragem mais próximos entre si sejam mais semelhantes do que aqueles mais distantes (Montagna, 2001).

Na geoestatística não se pode trabalhar com a suposição de que as observações são independentes e igualmente distribuídas, usualmente empregadas em procedimentos da estatística clássica. A independência entre as observações, que é derivada do conceito de aleatoriedade, pode ser refutada por meio de medidas de autocorrelação, que assinalam a dependência espacial ou temporal das amostras.

Portanto, o conceito fundamental da abordagem geoestatística remete à de-

terminação do modo e do grau de variabilidade espacial (regionalizada) expressos por um semivariograma. Nele são conduzidos ajustes, com o objetivo de determinar um modelo de variabilidade dentro de uma região, fornecendo parâmetros adequados ao entendimento da estrutura de dependência espacial (Vieira, 2000). Assim, deixar de considerar a dependência espacial não apenas viola pressupostos mas também significa deixar de observar importantes aspectos que determinam a ocorrência do fenômeno estudado. Considerar a estrutura de variabilidade espacial, contribui para ampliar a fração explicada da variabilidade dos dados.

2.2.1 Geoestatística aplicada em experimentos de campo

O maior desafio quando se conduz um experimento é o de se compararem os tratamentos com a maior precisão possível, para se ter segurança nas inferências a serem feitas a partir dos resultados alcançados. A precisão de um experimento está totalmente ligada a pequenas variações nas unidades experimentais, antes de serem aplicados os tratamentos, ou induzidas involuntariamente, durante a execução do experimento, que causam heterogeneidade entre as parcelas, também conhecida como variação casual, variação ambiental ou, simplesmente, erro experimental (Pilon, 2004).

Segundo Cressie e Hartfield (1996), a dependência espacial em dados de ensaios de campo tem sido reconhecida desde meados de 1930. De acordo com esses autores, nos ensaios de campo podem ser obtidos estimadores precisos para os efeitos dos tratamentos, reduzindo ou controlando a variação residual não devida aos tratamentos. A primeira abordagem, e mais comum, é usar delineamentos de complexidade crescente, tais como delineamentos em blocos; a segunda, que não é incompatível com a primeira, é usar modelos alternativos e análises que procuram modelar diretamente a heterogeneidade espacial.

Brus (2000) denomina a primeira abordagem como sendo baseada no delineamento, a qual contrastaria com a segunda abordagem baseada no modelo. Segundo o autor, a primeira é não-paramétrica, pois tem validade, qualquer que seja o modelo dos dados. Também salienta que a casualização desenvolvida pelo pesquisador nos delineamentos amostrais ou experimentais é de um tipo diferente daquela pressuposta pela adoção de um modelo, a qual não é gerada pela intervenção do pesquisador (como é o caso de pressupor $\varepsilon_{ij} \sim N(0, \sigma^2)$ não correlacionados).

Apesar de a casualização desenvolvida num experimento ou observação equivaler unicamente à abordagem baseada no delineamento, a existência dos dois tipos de abordagem

deve ser assumida tacitamente na experimentação de campo, se esta deseja ser realmente útil. Considere-se a validade do TCL (Teorema do limite) na abordagem baseada no delineamento apenas o tamanho das amostras, que é de fundamental importância para a aproximação da normal como distribuição dos dados de médias e totais. Porém, em muitas situações práticas, as parcelas experimentais são pequenas e, mesmo unitárias (como exemplo temos experimentação com animais), além de ser pequeno o número usual de repetições. Nesse caso, é necessário adotar uma abordagem baseada em modelo para tornar possível as inferências estatísticas (Pontes, 2002).

Segundo Brus & Gruijter (1993), na abordagem baseada no delineamento, ensaios fisherianos, por exemplo, os pontos a serem amostrados são selecionados aleatoriamente (com a suposição de independência e identicamente distribuídos), utilizando-se um delineamento de amostragem bem definido. A qualidade das medidas, tais como tendências e variância, é definida em termos da esperança sobre as realizações do delineamento de amostragem. Em contrapartida, os mesmos autores descrevem que, na abordagem baseada no modelo, a qual tem a geoestatística como representante, os valores são considerados como realizações de um campo aleatório (processo estocástico). Nele, geralmente, as variáveis aleatórias estão espacialmente dependentes, ou, equivalentemente, o processo estocástico que pode ter gerado os dados é descrito por um modelo, tendo a vantagem, por exemplo, de estimar com eficiência as médias espaciais de pequenas sub-regiões de um atributo particular em estudo.

2.2.2 Semivariograma

Alguns métodos geoestatísticos de estimadores da autocorrelação espacial são usados como ferramentas de continuidade espacial, como: o variograma ou semivariograma, o covariograma e o correlograma.

A função semivariograma deve o seu nome a Matheron (entre 1957 e 1962), bem como o seu tratamento e interpretações teórica e prática, embora fosse uma função conhecida anteriormente de acordo com Valente (1989).

O semivariograma é uma função matemática definida para representar o nível de dependência entre duas unidades experimentais e é usado para modelar dois valores correlacionados no espaço ou no tempo. Assim a construção do semivariograma para identificar a dependência espacial origina-se nas diferenças entre duas unidades experimentais georeferenciadas no espaço, separadas por uma distância u . Quando o semivariograma depende

apenas da distância $\|u\|$, o semivariograma é dito isotrópico (ou seja, depende apenas da norma $\|x_i - x_j\|$). Caso contrário, o semivariograma é dito anisotrópico (depende não só da distância, mas também da direção da reta que passa pelos pontos x_i e x_j) (Pontes, 2002).

Num estudo em que se deseja averiguar a dependência espacial de amostras georeferenciadas podem ocorrer três tipos de semivariogramas: observado ou experimental (obtido a partir das amostras colhidas no campo), verdadeiro (real, mas desconhecido) e teórico (de referência, utilizado para o ajuste do modelo) (Guerra, 1988). O semivariograma observado ou experimental é obtido sobre os dados, ou seja, proveniente do conjunto de dados da pesquisa realizada. É dado pela expressão (6) (Ribeiro Júnior et al., 2003):

$$\hat{\gamma}(u) = \frac{1}{2N_u} \sum_{(i,j) \in M_u} (Y(x_i) - Y(x_j))^2, \quad (6)$$

em que,

- i. $u = \|x_i - x_j\|$ é a distância euclidiana entre x_i e x_j .
- ii. $Y(x_i)$ e $Y(x_j)$ são as medidas de interesse nas posições x_i e x_j .
- iii. $M_u \equiv \{[x_i = (l, c), x_j = (l', c')] : \sqrt{(l - l')^2 + (c - c')^2} = u\}$, é o conjunto de todos os pares de posições que estão separadas por uma distância euclidiana u , freqüentemente denominada *lag* e
- iv. N_u é o número total de pares de posições em M_u .

É comum serem considerados pertencentes a N_u os pares de parcelas $[x_i, x_j]$ tal que, $\|x_i - x_j\| \approx u$.

O semivariograma verdadeiro é aquele que representa a situação real dos dados (sempre desconhecido). Assim, na análise estrutural o objetivo é estudar qual é o semivariograma teórico que melhor se ajusta ao semivariograma experimental, de forma que a partir do semivariograma teórico, realizam-se inferências sobre o semivariograma verdadeiro (Guerra, 1988).

O método de ajuste de semivariogramas teóricos a semivariogramas experimentais é denominado de método das aproximações sucessivas, pois o processo de ajuste finaliza quando as discrepâncias entre os valores experimentais e teóricos forem mínimas (Oliveira, 2003).

Na presença de dependência espacial, em geral, duas parcelas próximas são mais semelhantes do que duas mais afastadas permitindo afirmar que $\gamma(u)$ aumenta à medida que u aumenta até estabilizar. O ponto, em que os dados deixam de apresentar dependência espacial, é chamado de patamar ou *sill* e, no semivariograma, a distância, a partir da origem até o patamar é chamada de alcance ou *range* denotado por ϕ . Pela definição $\gamma(u)$ para $u = 0$ deverá ser zero, mas na prática observa-se que há casos em que $u = 0$ e $\gamma(u)$ aproxima-se de um valor positivo, denominado efeito pepita, ou efeito *nugget* simbolizado por τ^2 . Pode ocorrer também que $\gamma(u)$ cresça indefinidamente, indicando a ausência de dependência espacial ou, o que é comum ocorrer, seus valores variarem sempre aleatoriamente em torno do patamar, caracterizando a situação de independência espacial ou efeito pepita puro (Teixeira, 2001).

Na Figura 2, podem-se identificar os parâmetros do semivariograma, com efeito pepita, da seguinte forma: τ^2 é o efeito pepita ou *nugget*, σ^2 é a variância ou *sill*, $\sigma^2 + \tau^2$ a variância total e patamar também conhecida como *sill* total e ϕ que é o alcance ou *range*. Alguns autores afirmam que o fato de ocorrer efeito pepita, pode estar associado a algum erro de medição (Vieira, 1996; Guerra, 1988).

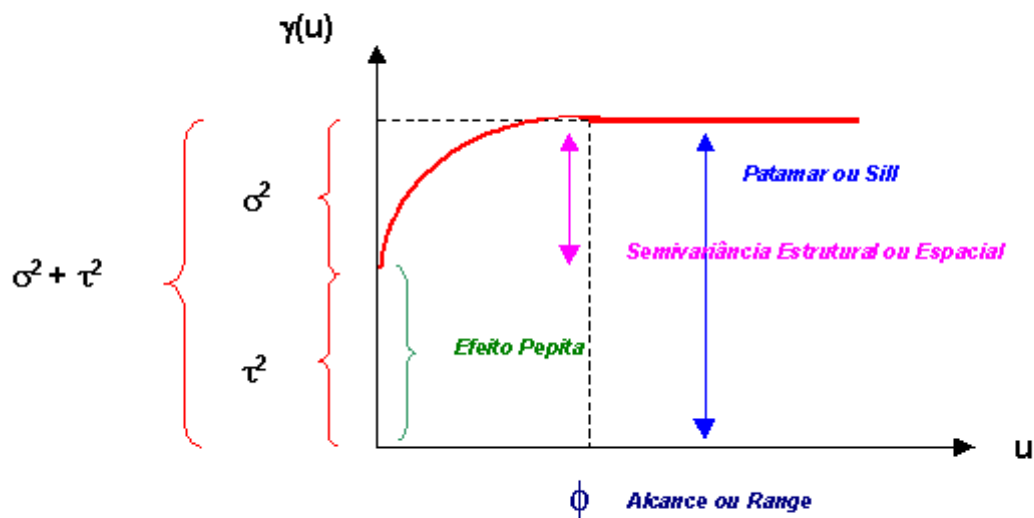


Figura 2 - Semivariograma com efeito pepita

2.3 Modelos geoestatísticos

O formato básico para dados geoestatísticos univariados é

$$(x_i, y_i), \quad i = 1, \dots, n$$

sendo que x_i identifica a posição espacial tipicamente em um plano bidimensional e y_i é uma medição escalar tomada na posição x_i . Uma característica da geoestatística é que a variável de medição y_i pode, em princípio, estar localizada em qualquer lugar numa região de estudo A .

No contexto de delineamento experimental, x_i representa a posição central de uma unidade experimental. Usualmente, é assumido que o delineamento de amostragem para as posições x_i é estocasticamente independente do processo que gera as medições y_i e que a região A é um subconjunto fixo de pontos em \mathfrak{R}^2 . A forma básica de um modelo geoestatístico é um processo estocástico de valor real $\{Y(x) : x \in A\}$ o qual é tipicamente considerado como uma realização parcial de um processo estocástico $\{Y(x) : x \in \mathfrak{R}^2\}$.

Um processo estocástico pode ser definido como uma família de variáveis aleatórias que descrevem a evolução através do tempo (ou de outra condição de avaliação) de algum sistema (Mota, 2002).

Freqüentemente, o processo de medição Y_i pode ser registrado como uma versão ruidosa de uma variável aleatória subjacente $S(x_i)$, o valor na posição x_i do processo de interesse $\{S(x) : x \in A \subset \mathfrak{R}^2\}$, sendo $S(x)$ chamado de sinal. Esse modelo básico pode ser estendido para um modelo com dois componentes: um processo estocástico $S(x)$, e um modelo estatístico para medições, $y = (y_1, \dots, y_n)^T$ condicional a $\{S(x) : x \in A \subset \mathfrak{R}^2\}$.

Um processo estocástico é considerado Gaussiano, se a distribuição conjunta de $S(x_1), \dots, S(x_n)$ é multivariada Gaussiana para qualquer inteiro n , e um conjunto de posições x_i . O processo é estacionário se, a esperança e a variância de $S(x)$ são as mesmas para todo x , e a correlação entre $S(x_i)$ e $S(x_j)$ depende unicamente da distância entre as posições x_i e x_j . O processo é estacionário e isotrópico se, adicionalmente, esta correlação depende unicamente de $u = \|x_i - x_j\|$, a distância euclidiana entre x_i e x_j , para qualquer par de inteiros $0 < i, j \leq n$.

O modelo Gaussiano, estacionário e isotrópico para um conjunto de dados (x_i, x_j) , $i, j = 1, \dots, n$, é definido pelas seguintes suposições (Diggle & Ribeiro Júnior, 2000):

1) $\{S(x) : x \in A \subset \mathfrak{R}^2\}$ é um processo Gaussiano estacionário com média μ , variância σ^2 ,

função de covariância $cov\{S(x_i), S(x_j)\} = \gamma(x_i, x_j) = \sigma^2 \rho(u)$, e função de correlação $\rho(u) = corr\{S(x_i), S(x_j)\}$, em que $u = \|x_i - x_j\|$;

- 2) Condicionalmente a $\{S(x) : x \in \mathfrak{R}^2\}$, os y_i são realizações mutuamente independentes dos Y_i , normalmente distribuídos com média condicional $E[Y_i|S(\cdot)] = S(x_i)$ e variância condicional τ^2 .

Este modelo pode ser expresso pela seguinte forma:

$$Y_i = S(x_i) + Z_i, \quad i = 1, \dots, n,$$

sendo que Z_i são erros aleatórios mutuamente independentes distribuídos normalmente com média zero e variância τ^2 .

Para especificar esse modelo, são necessários somente os momentos de primeira e segunda ordem, ou seja, a função de média $E[Y(x)] = \mu(x)$ e a função de covariância $Cov(Y(x_i), Y(x_j)) = \gamma(x_i - x_j) = E[(Y(x_i) - E(Y(x_i)))(Y(x_j) - E(Y(x_j)))]$. A função de variância pode ser descrita como $\sigma^2 = VarY(x)$. Portanto, resumidamente um processo gaussiano $Y(x)$ é estacionário se a função da média é constante, $E[Y(x)] = \mu(x)$ para todo x , e a função de covariância depende somente da diferença entre os dois pontos amostrados, $Cov(Y(x_i), Y(x_j)) = \gamma(x_i - x_j)$. O processo é isotrópico se adicionalmente a covariância depende só da distância, de forma que $\gamma(x_i - x_j) = \gamma(\|x_i - x_j\|)$, sendo que $\|\cdot\|$ é a distância euclidiana. Para um processo estacionário, a variância de $Y(x)$ é constante e a função de covariância pode ser escrita como $\gamma(u) = \sigma^2 \rho(u)$, sendo σ^2 a variância e $\rho(\cdot)$, a função de correlação.

Uma outra classe de modelos muito útil na prática são os processos gaussianos nos quais a média é a variável, podendo depender de covariáveis e erros aleatórios, porém, com estrutura de covariância estacionária. Para tais situações, $Y(x) - \mu(x)$ é um processo estacionário gaussiano com média zero.

A matriz de covariâncias especificada através da função de correlação $\rho(u)$ deve ser positiva definida. Essa restrição assegura que para qualquer inteiro m , qualquer conjunto de posições x_i e qualquer constante real a_i , a combinação linear $\sum_{i=1}^m a_i S(x_i)$ terá uma variância não-negativa.

2.3.1 Funções de correlação

É, usualmente, exigido que o modelo de função de correlação $\rho(u)$ incorpore as características dadas por:

- 1) $\rho(u)$ é monótona não crescente em u (as correlações entre duas medidas decrescem ou permanecem constantes com o aumento da distância entre as duas locações amostrais correspondentes),
- 2) $\rho(u) \rightarrow 0$ com $u \rightarrow \infty$ e
- 3) pelo menos um parâmetro no modelo controla a taxa com que $\rho(u)$ tende a zero.

Alguns exemplos de famílias paramétricas que reúnem essas características e que podem ser assumidas como uma função de correlação, apresentadas por Diggle & Ribeiro Júnior (2000):

a) A Família Esférica, definida por

$$\rho(u; \phi) = \begin{cases} 1 - \frac{3}{2} \left(\frac{u}{\phi} \right) + \frac{1}{2} \left(\frac{u}{\phi} \right)^3 & \text{se } 0 \leq u \leq \phi \\ 0 & \text{se } u > \phi \end{cases} \quad (7)$$

em que $\phi > 0$ determina a taxa para qual a correlação tende a zero ao aumentar u . O nome desta família recorre ao fato que $\rho(u; \phi)$ tem uma interpretação geométrica como o volume de intersecção de duas esferas tridimensionais, cujos centros são separados por uma distância u . Essa família não é muito recomendada, pois não proporciona flexibilidade ao modelo. Além disso, é de relevância duvidosa em planos bidimensionais.

b) A Família Exponencial Potência, definida por

$$\rho(\mu) = \exp \left\{ - \left(\frac{u}{\phi} \right)^\kappa \right\}, \quad (8)$$

com $\phi > 0$ e $0 < \kappa \leq 2$.

O parâmetro ϕ determina a taxa para qual a correlação tende a zero ao se aumentar u . O parâmetro, κ , é chamado de ordem do modelo e determina a suavidade analítica do sinal $S(x)$. A função de correlação exponencial é correspondente ao caso em que $\kappa = 1$. O caso em que $\kappa = 2$ é denominado algumas vezes, função de correlação Gaussiana.

Essa família freqüentemente ajusta com razoável qualidade, estruturas de correlações empíricas de dados espaciais.

c) A Família Matérn, cuja definição é dada por

$$\rho(u; \phi, \kappa) = \{2^{\kappa-1} \Gamma(\kappa)\}^{-1} \left(\frac{u}{\phi}\right)^{\kappa} K_{\kappa}\left(\frac{u}{\phi}\right), \quad (9)$$

em que $K_{\kappa}(\cdot)$ denota a função Bessel modificada de terceiro tipo de ordem κ . Os parâmetros ϕ e κ são como descritos nas famílias anteriores. Essa família é a mais recomendada devido a sua flexibilidade.

2.3.2 Modelos teóricos de semivariogramas

A sensibilidade dos semivariogramas, para detectar a variabilidade espacial das amostras, está diretamente ligada ao melhor ajuste dos dados experimentais ao modelo teórico do semivariograma. Modelos teóricos de semivariogramas são superpostos à seqüência de pontos obtidos no semivariograma experimental, de modo que a curva que melhor se ajusta aos pontos obtidos represente a magnitude, o alcance e a intensidade da variabilidade espacial da variável estudada (Zimback, 2003). Alguns modelos teóricos de semivariograma são:

a) Modelo Esférico ou Modelo de Matheron

Esse modelo apresenta um crescimento rápido na origem, alcança um patamar a uma distância u finita e é representado por:

$$\gamma(u) = \begin{cases} \tau^2 + \sigma^2 \left(\frac{3u}{2\phi} - \frac{1}{2} \left(\frac{u}{\phi} \right)^3 \right) & \text{se } 0 < u \leq \phi, \\ \tau^2 + \sigma^2 & \text{se } u > \phi, \\ 0 & \text{se } u = 0, \end{cases} \quad (10)$$

com $\tau^2 \geq 0$, $\sigma^2 \geq 0$, $\phi \geq 0$ e $u \geq 0$.

b) Modelo Exponencial ou de Formery

Esse modelo apresenta um comportamento linear na origem, alcança seu patamar somente assintoticamente e é representado por:

$$\gamma(u) = \begin{cases} \tau^2 + \sigma^2 \left[1 - \exp\left(-\frac{u}{\phi}\right) \right] & \text{se } 0 < u \leq \phi, \\ \tau^2 + \sigma^2 & \text{se } u > \phi, \\ 0 & \text{se } u = 0. \end{cases} \quad (11)$$

c) Modelo Matérn

Diggle & Ribeiro Júnior (2000) comentam sobre as diferentes estruturas de covariância utilizadas no estudo espacial e sugerem uma nova estrutura da função de correlação denominada família Matérn, descrita na equação (9). Essa função apresenta um parâmetro κ chamado de ordem do modelo Matérn o qual determina a suavização analítica do sinal $S(x)$ num modelo Gaussiano.

2.3.3 Métodos de estimação dos parâmetros do semivariograma

Existem diversos métodos de ajuste de modelos semivariográficos. McBrantney & Webster² citados por Ortiz (2002) comentam o ajuste de semivariogramas abordando os métodos de mínimos quadrados (ordinários, ponderados e generalizados). Segundo os autores, o método de mínimos quadrados ponderados tem melhor realismo e eficiência computacional.

O método de ajuste de modelo de semivariograma mais comum é o ajuste a “sentimento”. Por meio desse escolhe-se um modelo e seus parâmetros em que é 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). Como sugerido pelo autor, o processo deve ser repetido até que se obtenha um modelo adequado. Recomenda que seja utilizado o ajuste empírico como valor inicial para estimar os parâmetros do modelo de interesse por outro método em que as estimativas dos parâmetros sejam mais precisas. Menciona também, que este ajuste subjetivo é questionado e será cada vez menos utilizado.

O método de mínimos quadrados ordinários é citado por esse mesmo autor como inadequado, porque os pontos da nuvem de dados correspondem às estimativas provenientes de diferentes pares de pontos e estas estimativas não possuem a mesma precisão e não são independentes. Para o método de quadrados mínimos ordinário tem-se a seguinte expressão:

$$L(\tau^2, \sigma^2, \phi) = \sum_{N(u)} \{\hat{\gamma}(u) - \gamma(u; (\tau^2, \sigma^2, \phi))\}^2, \quad (12)$$

- i) $u = \|x_i - x_j\|$ é a distância euclidiana entre x_i e x_j ;
- ii) $N(u)$ é o número total de pares de posições em $M(u)$;
- iii) $\hat{\gamma}(u)$ é o valor da semivariância obtida a partir do semivariograma empírico para as parcelas que estão separadas à distância u e

²McBRATNEY, A.B.; WEBSTER, R. Choosing functions for semi-variograms of soil properties and fitting them to sampling estimates. **Journal of Soil Science**, v.77 p.617-639, 1986.

iv) $\gamma(u; (\tau^2, \sigma^2, \phi))$ corresponde ao valor das semivariâncias esperadas a partir do semivariograma teórico para as parcelas que estão separadas pela distância u .

Para contornar o problema do método de mínimos quadrados ordinários McBratney & Webster³ citados por Pilon (2004) utilizaram o método de mínimos quadrados generalizados. Com esse método, obtiveram estimativas eficientes assintoticamente, mas com a exigência de inversão de uma matriz de grandes dimensões. A expressão utilizada para o ajuste dos parâmetros é dada por:

$$G((\tau^2, \sigma^2, \phi); \hat{\gamma}) = \{\hat{\gamma} - \gamma(\tau^2, \sigma^2, \phi)\}' \sum^{-1} \{\hat{\gamma} - \gamma(\tau^2, \sigma^2, \phi)\} \quad (13)$$

- i) $\hat{\gamma}$ representa o vetor de estimativas de semivariâncias $n_u \times 1$, em que N_u é o número de distâncias u ;
- ii) $\gamma(\tau^2, \sigma^2, \phi)$ é o vetor, $n_u \times 1$, das semivariâncias estimadas que depende do vetor das estimativas dos parâmetros;
- iii) \sum é a matriz de variâncias e covariâncias de dimensões $n_u \times n_u$ entre as estimativas de $\gamma(u; \tau^2, \sigma^2, \phi)$.

Ainda McBratney & Webster (1986) consideraram o método de mínimos quadrados ponderados mais real e computacionalmente eficiente e o modelo é ajustado minimizando a seguinte equação:

$$W(\tau^2, \sigma^2, \phi) = \sum_{N(u)} \frac{\{\hat{\gamma}(u) - \gamma(u; (\tau^2, \sigma^2, \phi))\}^2}{\text{var}[\hat{\gamma}(u)]}, \quad (14)$$

em que, $\text{var}[\hat{\gamma}(u)]$ é a variância da estimativa da semivariância.

Barry et al.⁴ citados por Diggle & Ribeiro Júnior (2000) propuseram o método de mínimos quadrados n-ponderados, que atribui pesos aos resíduos somente de acordo com os números de parcelas que estão separadas pela distância u . Tal método consiste na minimização da expressão:

$$W_p(\tau^2, \sigma^2, \phi) = \sum_{M(u)} N(u) \{\hat{\gamma}(u) - \gamma(u; (\tau^2, \sigma^2, \phi))\}^2, \quad (15)$$

³McBRATNEY, A.B.; WEBSTER, R. Choosing functions for semi-variograms of soil properties and fitting them to sampling estimates. **Journal of Soil Science**, v.77, p.617-639. 1986.

⁴BARRY, J.T.; CROWDER, M.J.; DIGGLE, P.J. **Parametric estimation of the variogram**. Lancaster: University Technical Report, 1997.

em que, u , $N(u)$, $\hat{\gamma}(u)$ e $\gamma(u; (\tau^2, \sigma^2, \phi))$ já foram definidos na equação (12).

O método de máxima verossimilhança consiste em maximizar a função de probabilidade, em relação aos efeitos fixos e aos componentes de variância dos efeitos aleatórios do modelo (Oliveira, 2003).

A estimação dos parâmetros de modelos baseados em verossimilhança produz de maneira geral estimadores não viesados e eficientes quando aplicados para grandes amostras. Diggle e Ribeiro Júnior (2000), afirmaram ainda que ao utilizar estes métodos deve-se considerar a pressuposição Gaussiano e ilustram este estudo com um modelo linear Gaussiano em que a média é definida por um modelo de regressão linear.

2.3.4 Critério de Akaike na seleção de modelos

Em geral, na seleção de modelos, alguns autores têm utilizado o Critério de Akaike (AIC). O AIC é baseado na teoria de decisão e pode ser definido como (Oliveira, 2003):

$$AIC = -2l + 2p \quad (16)$$

em que:

l é o \ln da função de verossimilhança;

p é o número de parâmetros do modelo considerado.

De acordo com esse critério o melhor modelo é aquele que apresenta o menor valor de AIC.

3 MATERIAL E MÉTODOS

3.1 Material

3.1.1 Localização do experimento

Os dados de volume sólido por planta a serem utilizados neste trabalho são provenientes de um experimento de espaçamento de *Eucalyptus dunnii* em delineamento sistemático tipo “leque”, instalado em março de 1988 na fazenda de Experimental de Ibiti, de propriedade da Ripasa S/A Celulose e Papel, situada no município de Itararé, Estado de São Paulo. Estes mesmos dados foram utilizados em sua Dissertação de Mestrado do Prof^o Dr^o José Luiz Stape em 1995. As informações como caracterização do material genético, solo, bioclimática, práticas silviculturais, entre outras, poderão ser encontradas na Dissertação citada.

3.1.2 Avaliações dendrométricas

O experimento foi avaliado do segundo ao sexto ano. Em cada uma das avaliações coletaram-se dados individuais referentes a CAP (circunferência à altura do peito, 1,30m em relação ao solo) e altura total das plantas úteis. Neste trabalho serão utilizados os dados referentes ao sexto ano, idade comercial de corte da espécie.

3.1.3 Variável de interesse

Para se obter a equação do volume sólido por planta, com casca, foram abatidas, aos 70 meses, árvores de *Eucalyptus dunnii*, do povoamento comercial com espaçamento de 3,0m x 3,0m, que circundam o experimento.

Estas árvores foram cubicadas de metro em metro, da base de corte até o diâmetro mínimo de 6cm, calculando-se o volume de cada seção pela fórmula de Smalian, e o volume de ponteiro como se fosse um cone. Assim ajustaram-se equações volumétricas tendo como variável dependente o volume individual, como variáveis independentes o DAP

(diâmetro à altura do peito) e altura total, conforme descreve Stape (1995). Então para o cálculo da variável de interesse, utilizou-se um modelo de equação volumétrica dado por:

$$V = \exp[-10,257052 + 0,9964961 \times \ln(DAP^2 \times H)] \quad (17)$$

em que,

- i) V é volume individual da árvore (m^3);
- ii) DAP (Diâmetro à altura do peito) = $\frac{CAP}{\pi}$ (cm) e
- iii) H é a altura (m).

3.2 Métodos

3.2.1 Instalação e caracterização do experimento sistemático tipo “leque”

O experimento foi instalado em delineamento sistemático tipo “leque” conforme o modelo 1a proposto por Nelder (1962). Nesse delineamento, a retangularidade é constante por toda a malha de pontos e é medida pela relação do comprimento da parcela para a largura até o ponto onde a planta estará. r_n é o raio dos n arcos e θ é o ângulo constante entre os sucessivos raios. Os raios r_1, r_2, \dots formam uma progressão geométrica (α) que pode ser escrita como $r_n = r_0 \alpha^n$. A área por planta aumenta para fora do centro e os n arcos são proporcionais a r_n^2 . Nesse delineamento, o maior raio (N), é definido por r_0, α e θ , que também definirão o maior e menor espaçamento e a retangularidade.

O número de combinações possíveis entre os valores de r_0, α e θ para os delineamentos, foi definido dentro das faixas comerciais de espaçamento para o *Eucalyptus dunnii*, optando-se pelos seguintes valores $r_0 = 5,30$, $\alpha = 1,21$ e $\theta = 10^\circ$. A partir desses valores, obteve-se o delineamento caracterizado na Tabela 1.

O experimento com *Eucalyptus dunnii* foi instalado com 36 repetições (raios) contando com bordadura interna com raio 5,30m e com uma bordadura externa com raio de 43,60m (Figura 3).

Tabela 1. Valores da distância radial, área por planta e densidade de plantas para 10 tratamentos (espaçamentos) no delineamento sistemático tipo “leque”

Tratamento	Distância Radial (m)	Área (m^2 /planta)	Densidade(plantas/ha)
1	6,42	1,40	7143
2	7,78	2,05	4878
3	9,42	3,01	3322
4	11,41	4,41	2268
5	13,82	6,47	1546
6	16,74	9,50	1053
7	20,27	13,93	718
8	24,55	20,43	489
9	29,73	29,97	334
10	36,00	43,95	228

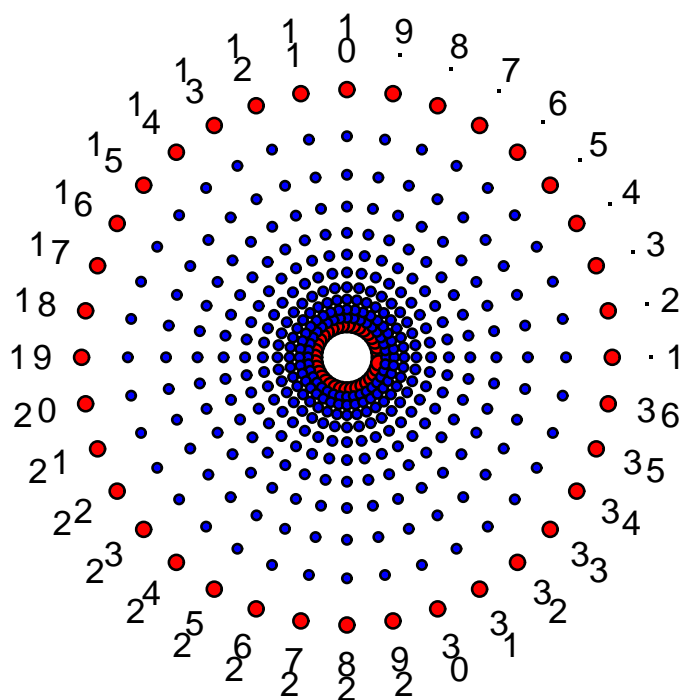


Figura 3 - Croqui de instalação do experimento sistemático “leque” com 10 tratamentos e 36 repetições com *Eucalyptus dunnii*

3.2.2 Falhas e bifurcações

As falhas, mortes e bifurcações foram anotadas nas fichas de avaliação de campo, considerando-se como falha de plantio a inexistência da árvore na primeira avaliação, e como mortalidade as futuras falhas ocorridas ao longo do crescimento, ocorrendo nesse experimento somente a mortalidade das plantas.

Devido à inexistência de bordadura entre plantas e entre linhas, a ocorrência de falhas e mortalidade de algumas árvores faz com que as plantas vizinhas, tenham uma área disponível de crescimento diferente daquela inicialmente estabelecida para os tratamentos (Nelder,1962). Mas se as ocorrências forem poucas e aleatórias dentro do experimento, Gomez & De Datta⁵ citado por Stape (1995) propõem eliminar, da análise experimental, todas as plantas vizinhas a estas falhas ou mortes.

Neste trabalho os dados da ocorrência de mortalidade serão analisados, considerando-se:

Modelo I: eliminação dos dados das plantas vizinhas a estas mortas, conforme Gomez & De Datta (1972) propõem;

Modelo II: as informações de mortes das plantas serão consideradas como uma covariável no modelo;

Modelo III: além da covariável morte das plantas, também será levado em consideração o tempo da ocorrência da morte.

Tabela 2 mostra os números dos raios das plantas mortas e a Tabela 3, os números dos raios das plantas que serão eliminadas no modelo I.

⁵GOMEZ, K.A.; DE DATTA, S.K. Missing hills in rice experimentation plots. **Agronomy Journal**, n.64, p.163-164, 1972.

Tabela 2. Número do raio das plantas devido à morte no delineamento sistemático tipo “leque”, aos 6 anos

Tratamento	N ^o do raio
1	6
3	11
4	3
6	22
9	36
Total	5

Tabela 3. Números dos raios das plantas que serão eliminadas, e suas vizinhas devido a morte no delineamento sistemático tipo “leque”, aos 6 anos

Tratamento	N ^o do raio
1	5-6-7
2	6-11
3	3-10-11-12
4	2-3-4-11
5	3-22
6	21-22-23
7	22
8	36
9	1-35-36
10	36
Total	24

3.2.3 Análise Clássica

Para a realização da análise de variância (ANOVA) de acordo com o modelo de Gauss-Markov Normal, será utilizado o modelo linear de um delineamento inteiramente casualizado dado pela expressão:

$$y_{ij} = \mu + t_i + \varepsilon_{ij}, \quad i = 1, \dots, I \quad \text{e} \quad k = 1, \dots, J, \quad (18)$$

em que, y_{ij} é o valor observado na i -ésima parcela que recebeu o j -ésimo tratamento, μ é a média geral, t_i é o efeito do i -ésimo tratamento e ε_{ij} é o erro experimental associado à observação y_{ij} .

A pressuposição feita em relação ao erro experimental foi a que definiu esse modelo. Considerou-se que os erros são variáveis aleatórias independentes e têm distribuição normal com média zero e variância constante σ_a^2 . O modelo pode ser expresso da seguinte forma matricial:

$$Y = X\alpha + \varepsilon$$

em que, $Y = (y(x_1), \dots, y(x_n))$ é o vetor dos valores da variável resposta; X é a matriz delimitação; α é o vetor das médias desconhecidas dos tratamentos; ε é o vetor dos erros independentes e normalmente distribuídos. Dessa forma, no modelo Gauss-Markov Normal, os erros são apresentados, matricialmente, por

$$\varepsilon \sim N(0, I_n \sigma_a^2),$$

em que I_n é a matriz identidade $n \times n$.

3.2.3.1 Modelo considerando a eliminação das plantas vizinhas

Neste modelo será considerada a eliminação das plantas vizinhas (Modelo I). Assim será considerado que o delineamento sistemático segue o modelo Gauss-Markov, e assume as pressuposições de i.i.d, e será dado por:

$$Y_{ij} = \mu + a_i + \varepsilon_{ij}, \quad i = 1, \dots, I \quad \text{e} \quad j = 1, \dots, n_i, \quad (19)$$

em que, Y_{ij} é o valor observado que recebeu o i -ésimo tratamento na j -ésima repetição, μ é a média geral, a_i é o efeito do i -ésimo tratamento e ε_{ij} é o erro experimental associado à observação Y_{ij} .

3.2.4 Análise de Covariância

Em alguns experimentos, pode ser muito difícil e até impossível obter unidades experimentais semelhantes. Quando a condição inicial da unidade experimental for conhecida e puder ser medida, e ainda, se for conhecido que esta condição inicial (uma variável) tenha

influência sobre a variável resposta, pode-se utilizar essa informação para corrigir a variável resposta. Segundo Anjos (2004), esse procedimento pode ser feito utilizando-se a técnica de análise de covariância. Para Pimentel Gomes (2000), a análise de covariância é uma técnica muito interessante, embora seu uso não seja muito comum. Tem ela por fim utilizar uma ou mais variáveis auxiliares na interpretação dos dados referentes a uma variável na qual se tem interesse, e também complementa o controle local e pode até substituí-lo em alguns casos.

A covariável que será inserida no modelo necessita estar correlacionada com a variável resposta para que se possa fazer uso de tal análise. A ANCOVA (Análise de Covariância) permite que se faça um ajuste do efeito de uma variável resposta que sofreu influência de uma variável ou uma causa de variação não controlada. Permite, portanto, um controle do erro experimental, aumentando a precisão do experimento. É possível também fazer o ajuste das médias dos tratamentos em função da(s) covariável(eis) e, em alguns casos, estimar observações perdidas durante o experimento.

Considerando um experimento com um fator e uma covariável, o modelo estatístico pode ser escrito da seguinte maneira:

$$Y_{ij} = \mu + \alpha_i + \beta(X_{ij} - \bar{X}) + \varepsilon_{ij}, \quad i = 1, \dots, I \quad \text{e} \quad j = 1, \dots, J, \quad (20)$$

em que, Y_{ij} é o valor observado que recebeu o i -ésimo tratamento na j -ésima repetição, μ é a média geral, α_i é o efeito do i -ésimo tratamento, X_{ij} é o valor observado da covariável, \bar{X} é a média da covariável, β é o coeficiente de regressão linear entre a covariável (X) e a variável resposta (Y) com $\beta \neq 0$. Neste caso, assume-se que a variável resposta e a covariável estão relacionadas linearmente e ε_{ij} é o erro experimental associado à observação Y_{ij} .

3.2.4.1 Modelo com a covariável morte da planta

Neste modelo foi incluído o efeito de morte da planta como covariável (Modelo II). O modelo (21) considerando apenas as informações de mortes é dado por:

$$Y_{ij} = \mu + a_i + \beta(P_{ij} - \bar{P}) + \varepsilon_{ij}, \quad i = 1, \dots, I \quad \text{e} \quad j = 1, \dots, n_i, \quad (21)$$

em que, Y_{ij} é o valor observado que recebeu o i -ésimo tratamento na j -ésima repetição, μ é a média geral, a_i é o efeito do i -ésimo tratamento, β é o coeficiente de regressão entre a

covariável (mortes) e a variável resposta, P_{ij} é o valor observado da covariável (mortes), \bar{P} é a média da covariável (mortes) e ε_{ij} é o erro experimental associado à observação Y_{ij} .

3.2.4.2 Modelo com as covariáveis morte da planta e tempo

Neste modelo além da covariável morte das plantas, também será levado em consideração o tempo da ocorrência, da morte (Modelo III), assim o modelo é dado por:

$$Y_{ijk} = \mu + a_i + \beta(P_{ij} - \bar{P}) + \beta_1(T_{ik} - \bar{T}) + \varepsilon_{ijk}, \quad i = 1, \dots, I \quad j = 1, \dots, n_i \quad e \quad k = 1, \dots, K, (22)$$

em que, Y_{ijk} é o valor observado que recebeu o i -ésimo tratamento na j -ésimo repetição, no k -ésimo tempo, μ é a média geral, a_i é o efeito do i -ésimo tratamento, β é o coeficiente de regressão entre a covariável (mortes) e a variável resposta, β_1 é o coeficiente de regressão entre a covariável (tempo) e a variável resposta, P_{ij} é o valor observado da covariável (mortes), \bar{P} é a média da covariável (mortes), T_{ik} é o valor observado da covariável (tempo), \bar{T} é a média da covariável (tempo) e ε_{ijk} é o erro experimental associado à observação Y_{ijk} .

3.2.5 Construção das covariáveis

A covariável considerada no modelo II, foi construída de acordo com o número de plantas mortas, onde zero foi atribuído às plantas que não perderam vizinhas e um às que perderam uma vizinha. Já na covariável tempo foram levados em consideração os anos de avaliação, sendo que a planta perdida no segundo ano ao final do sexto ano receberá o valor cinco da covariável, se a planta perdida ocorrer somente no sexto ano o valor da covariável será um.

3.2.6 Transformação de dados

A aplicação válida dos testes de significância na análise de variância requer que os erros experimentais sejam independentes e normalmente distribuídos com média zero e variância σ^2 . Na prática, porém, estas condições ideais nem sempre são satisfeitas. Assim o estudo de transformações adequadas aos diversos tipos de dados, para que sejam satisfeitas as exigências da análise de variância, vem sendo feito há muito tempo e por diversos autores (Demétrio, 1978).

Box & Cox⁶ citados por Demétrio & Zocchi (2003) fizeram um estudo sobre as transformações aplicadas aos dados e sugeriram um método baseado em uma função de verossimilhança e sua distribuição. Obteve-se uma família de transformações dada por

$$Y(\lambda) = \begin{cases} \frac{Y^\lambda - 1}{\lambda} & \text{se } \lambda \neq 0 \\ \log(Y) & \text{se } \lambda = 0 \end{cases} \quad (23)$$

sendo λ o parâmetro e Y a variável resposta. Na ausência de uma transformação, $\lambda = 1$.

Nesse trabalho, tais estimativas serão obtidas por um procedimento numérico, que se encontra disponível no programa estatístico R.

Após determinar a necessidade da transformação, serão aplicados os testes de Shapiro-Wilk, Bartlett, Hartley e uma análise de resíduos (Barbin, 2003) para verificar a normalidade e homogeneidade.

3.2.7 Análise exploratória espacial

Na análise dos dados a pesquisa geoestatística, para certificar-se que existe realmente uma estrutura de dependência espacial, deve ser precedida por uma análise exploratória.

O objetivo dessa análise também é tão somente ter-se um indicativo do comportamento dos dados associado às suas posições na região amostrada e permite: identificar “outliers”, realizar mapas de contorno, gráficos tridimensionais, aplicar estatísticas baseadas em janelas móveis, avaliar normalidade.

Outra ferramenta para a investigação de dados espacialmente atípicos é a análise do efeito proporcional, que é o aumento da variância proporcionalmente ao aumento da média. Na natureza este efeito proporcional é observado em fenômenos com distribuições assimétricas, lognormal por exemplo. A análise do efeito proporcional consiste em construir gráficos de dispersões da média e da variância no sentido das linhas e colunas. Em uma situação que não apresenta tendência, tem-se uma expectativa dos pontos em torno da média. Para o caso de dispersão que varia no espaço, os pontos afastam-se da média gradualmente.

⁶BOX, G.E.P.; COX, D.R. An Analysis of Transformations. **Journal of the Royal Statistical Society**, v.26, n.2, p.211-252, 1964.

3.2.8 Modelo geoestatístico

O modelo geoestatístico modela a dependência espacial através de um componente de erro aleatório, assumindo que os dados, $y = (y_1, \dots, y_n)$, seguem o modelo linear Gaussiano dado por

$$Y_i = \mu(x_i) + S(x_i) + Z_i, \quad i = 1, \dots, n, \quad (24)$$

em que,

- i) x_i é a posição espacial da i -ésima parcela;
- ii) Y_i é o valor observado na posição x_i ;
- iii) $\mu(x_i)$ é determinada a partir de um modelo de regressão linear dado por

$$\mu(x_i) = \sum_{k=1}^p f_k(x_i) \alpha_k, \quad (25)$$

em que, $f_k(\cdot)$ são observações espacialmente relacionadas às variáveis explanatórias;

- iv) $S(\cdot)$ é um processo estacionário Gaussiano, com média zero, variância σ^2 e função de correlação $\rho(u; \phi)$, ou seja, $S \sim NMV(0, \sigma^2 R(\phi))$, em que $R(\phi)$ é a matriz de covariância $n \times n$ com (i, j) -ésimo elemento, $\rho(u_{ij})$, sendo $u_{ij} = \|x_i - x_j\|$ a distância euclidiana entre x_i e x_j ;
- v) Z_i são variáveis aleatórias, independentes e normalmente distribuídas com média zero e variância τ^2 .

Esse modelo pode ser escrito na forma matricial e expresso por

$$Y = X\mu + S(x) + \varepsilon \quad (26)$$

em que, $Y = (y(x_1), \dots, y(x_n))$ é o vetor dos valores da variável resposta nas posições (x_1, \dots, x_n) ; μ é o vetor das médias desconhecidas dos tratamentos; $S(x)$ é um processo estocástico estacionário e, ε é o vetor dos erros independentes e normalmente distribuídos.

Assim, para Y é assumida a distribuição normal multivariada, isto é, $Y \sim NMV(X\mu, \sigma^2 R(\phi) + \tau^2 I_n)$, em que I_n é uma matriz identidade $n \times n$, $R(\phi)$ é a matriz de covariância de dimensão $n \times n$, cujos elementos são fornecidos pela função de correlação $\rho(u, \phi)$.

O modelo geoestatístico decompõe o erro, em dois componentes, um de variáveis dependentes, $S \sim NMV(0, \sigma^2 R(\phi))$, que indicam índices ambientais não medidos e outro de variáveis independentes, $Z \sim N(0, I_n \tau^2)$, que ocorrem devido a erros de medidas.

Existem várias escolhas para a função de correlação $\rho(u)$. Nesse trabalho serão avaliadas três funções de correlação: a função exponencial e as funções Matérn, com $\kappa = 1$ e $\kappa = 2$ descritas no item 2.3.1.

3.2.9 Construção do semivariograma empírico

Para a construção do semivariograma, o primeiro passo consiste na realização de uma análise de variância clássica ANOVA ou ANCOVA. Obtidas as estimativas dos efeitos dos tratamentos \hat{a}_i , os resíduos r_{ij} serão calculados da seguinte forma:

1. Modelo I:

$$r_{ij} = Y_{ij} - \hat{\mu} - \hat{a}_i, \quad (27)$$

2. Modelo II:

$$r_{ij} = Y_{ij} - \hat{\mu} - \hat{a}_i - \hat{\beta}(P_{ij} - \bar{P}), \quad (28)$$

3. Modelo III:

$$r_{ijk} = Y_{ijk} - \hat{\mu} - \hat{a}_i - \hat{\beta}(P_{ij} - \bar{P}) - \hat{\beta}_1(T_{ik} - \bar{T}), \quad (29)$$

em que, r_{ij} é o resíduo, Y_{ij} é o valor observado que recebeu o i -ésimo tratamento da j -ésima parcela, $\hat{\mu}$ é a média geral, \hat{a}_i é a média do i -ésimo tratamento e $\hat{\beta}$, $\hat{\beta}_1$, P_{ij} , \bar{P} , T_{ik} e \bar{T} já foram descritos conforme equação (21) e (22).

Considerando as posições dos resíduos no espaço, as semivariâncias entre os resíduos serão calculadas utilizando o seguinte estimador:

$$\hat{\gamma}(u) = \frac{1}{2N_u} \sum_{(i,j) \in M_u} (R(x_i) - R(x_j))^2, \quad (30)$$

em que,

- i) $u = \|x_i - x_j\|$ é a distância euclidiana entre x_i e x_j ;
- ii) $R(x_i)$ e $R(x_j)$ são os resíduos da análise de variância das i -ésima e da j -ésima parcelas;
- iii) $M_u \equiv \{[x_i = (l, c), x_j = (l', c')] : \sqrt{(l - l')^2 + (c - c')^2} = u\}$, é o conjunto de todos os pares de posições que estão separadas a uma distância euclidiana u , freqüentemente denominada *lag* e;
- iv) N_u é o número total de pares de posições em M_u .

3.2.10 Ajuste de um modelo de semivariograma teórico para as semivariâncias

Uma vez calculadas as semivariâncias empíricas, será ajustado um modelo de semivariograma teórico a partir de um gráfico contendo as semivariâncias estimadas por (30), ou seja, um semivariograma teórico será ajustado aos valores do semivariograma empírico. Neste trabalho serão utilizadas os modelos teóricos exponencial e Matérn, descritos no item 2.3.2.

3.2.11 Estimação de máxima verossimilhança

Os parâmetros σ^2 , ϕ e τ^2 do semivariograma serão estimados pelo método de máxima verossimilhança.

Considerando o modelo apresentado em (24) e as estruturas de covariância exponencial e família Matérn e assumindo que $y = (y_1, \dots, y_n)$ segue uma distribuição multivariada Gaussiana tem-se que a função de verossimilhança é expressa por:

$$L(\alpha, \theta; y) = (2\pi)^{-\frac{n}{2}} |G(\theta)|^{-\frac{1}{2}} \exp\left[-\frac{1}{2}(y - X\alpha)' \{G(\theta)\}^{-1} (y - X\alpha)\right], \quad (31)$$

em que y é o vetor de observações experimentais; X é a matriz do delineamento; α é o vetor de parâmetros desconhecidos associados aos efeitos fixos; θ são os parâmetros da matriz de covariância, $G(\theta) = \sigma^2 R(\phi) + \tau^2 I_n$, com $\theta = (\tau^2, \sigma^2, \phi)$.

A função de verossimilhança é uma função dos parâmetros desconhecidos, portanto a estimação de máxima verossimilhança consiste em encontrar os valores $(\hat{\alpha}, \hat{\theta})$ de (α, θ) mais prováveis de terem produzido os valores observados y .

Desta forma, procura-se o vetor (α, θ) que satisfaça a seguinte expressão:

$$L[(\hat{\alpha}, \hat{\theta}); y] = f[(y; \hat{\alpha}, \hat{\theta})] = \text{máx}\{f(y; \alpha, \theta)\} \quad \alpha, \theta \in \Theta\}.$$

Então, o logaritmo da função de verossimilhança para (α, θ) é dado por

$$\ell(\alpha, \theta) = -\frac{1}{2}\{n\log(2\pi) + \log|G(\theta)| + (y - X\alpha)' \{G(\theta)\}^{-1}(y - X\alpha)\}, \quad (32)$$

Por conveniência computacional, é possível a extração de um fator escalar a partir de $G(\theta)$, como por exemplo:

$$G(\theta) = \sigma^2\{\nu^2 + R(\phi)\} = \sigma^2V, \quad (33)$$

em que $\nu^2 = \frac{\tau^2}{\sigma^2}$. Obtém-se o logaritmo da função de verossimilhança $l(\alpha, \theta', \sigma^2; y)$, em que $\theta' = (\nu^2, \phi)$, dado por:

$$l(\alpha, \theta', \sigma^2; y) = -\frac{1}{2}\{n\log(2\pi) + n\log(\sigma) + \log|V| + (y - X\alpha)' \{V\}^{-1}(y - X\alpha)\} \quad (34)$$

Diferenciando-se parcialmente $l(\alpha, \theta', \sigma^2; y)$ em relação a α e σ^2 e igualando-se as derivadas a zero, obtém-se

$$\hat{\alpha}(\theta) = [X'(V)^{-1}X]^{-1}X'(V)^{-1}y, \quad (35)$$

e matriz de covariância $\hat{v}ar(\hat{\alpha}) = \hat{\sigma}^2(X'V^{-1}X)^{-1}$, em que $\hat{\sigma}^2 = n^{-1}(y - X\hat{\alpha})'V^{-1}(y - X\hat{\alpha})$.

As estimativas de θ' são obtidas substituindo-se $\hat{\alpha}$ e $\hat{\sigma}^2$ no logaritmo da função de verossimilhança apresentada na equação (34). Dessa forma, a maximização do logaritmo da função de verossimilhança só depende de dois parâmetros desconhecidos, ν^2 e ϕ , simplificando muito as rotinas de maximizações. Nesse trabalho, tais estimativas foram obtidas por um procedimento numérico por meio da subrotina geoR (Ribeiro Junior & Diggle, 2004), que se encontra disponível no programa estatístico R.

3.2.12 Critério de Akaike na seleção de modelos

Foi utilizado, para seleção de modelos, o Critério de Akaike (AIC) conforme apresentado no item 2.3.4, o qual é baseado na teoria de decisão e pode ser definido pela eq. (16). O melhor modelo é aquele que apresentou menor valor de AIC.

3.2.13 Conjunto de dados e programas

O conjunto de dados e os programas utilizados neste trabalho encontram-se nos Anexos A, B, C e D.

4 RESULTADOS E DISCUSSÃO

4.1 Transformação de dados

As estimativas dos parâmetros de transformação ($\hat{\lambda}$) para os modelos I, II e III, foram respectivamente 0,4363, 0,4393 e 0,4454, indicando a necessidade de transformação. Como as estimativas foram próximas de 0,5, que estão dentro do intervalo de confiança usaremos a transformação $\sqrt{Y_{ij}}$.

Observando-se a tabela 5, verifica-se a falta de normalidade pelo teste de Shapiro-Wilk, rejeitando a hipótese de nulidade em favor da hipótese alternativa. Pelo teste de Bartlett rejeita-se a hipótese de homogeneidade de variâncias ao nível de 1% de probabilidade, porém sabe-se que esse é sensível a falta de normalidade e não se rejeita a homogeneidade de variâncias pelo teste de Hartley (Tabela 6).

Tabela 4. Resultados do teste de Shapiro-Wilk para os resíduos da variável volume (m^3) nos modelos I, II e III

Modelo	Shapiro-Wilk	
	Valor do Teste	Probabilidade
I	$W = 0,9755$	$p < 0,0001$
II	$W = 0,9756$	$p < 0,0001$
III	$W = 0,9732$	$p < 0,0001$

Tabela 5. Resultados dos testes de Bartlett e Hartley para a variável volume (m^3) dos modelos I, II e III

Testes	Bartlett		Hartley	
Modelo	Valor do Teste	Probabilidade	Valor do Teste	F_{tab}
I	$k = 21,6458$	$p < 0,05$	$F_{max} = 3,0262^{NS}$	$F_{tab} = 3.6$
II e III	$k = 24,1141$	$p < 0,01$	$F_{max} = 3,2649^{NS}$	$F_{tab} = 3.76$

(1)^{NS} Não significativo ao nível de 1% de probabilidade.

Para os modelos analisados os gráficos de resíduos padronizados vs. valores estimados, “box-plot”, gráfico de probabilidade normal e distribuição de frequências encontram-se no Anexo E.

Os gráficos da probabilidade normal mostram que existem ainda alguns pontos afastados, indicando valores periféricos, candidatos a “outliers” e os “box-plots” confirmam essas presenças. Assim confirma-se que a transformação realizada não produziu a distribuição normal esperada.

Mas Cressie⁷ citado por Mello (2004) discute que um fundamento comum da modelagem estocástica é que os dados tenham distribuição normal, sob a qual toda inferência geoestatística linear tem-se desenvolvido. Observa-se, porém, que a geoestatística não exige normalidade dos dados para ser aplicada. Se houver normalidade, as inferências realizadas poderão ganhar algumas propriedades estatísticas ótimas.

Observando-se os dados e considerando-se que ocorram 5 possíveis “outliers” em 355 e 336 dados, desconsideraremos a leve assimetria apresentanda nos gráficos de distribuição de frequências (Anexo E) e assumiremos que os dados seguem aproximadamente uma distribuição normal.

4.2 Análise exploratória espacial

A análise exploratória espacial fornece uma idéia do comportamento da dispersão dos dados levando em conta a posição geográfica de cada observação (Câmara et al., 2001).

Devido à análise de resíduos apontar um comportamento assimétrico dos modelos I, II e III, a existência do efeito proporcional entre média e variância é investigada por meio da figura 4, em que as medidas de posição maiores correspondem as medidas de dispersão

⁷CRESSIE, N. **Statistics for spatial data**. New York: Wiley, 1993. 900p.

mais elevadas. A figura 4 apresenta o efeito proporcional nítido, não podendo assumir um comportamento estabilizado em torno de uma constante.

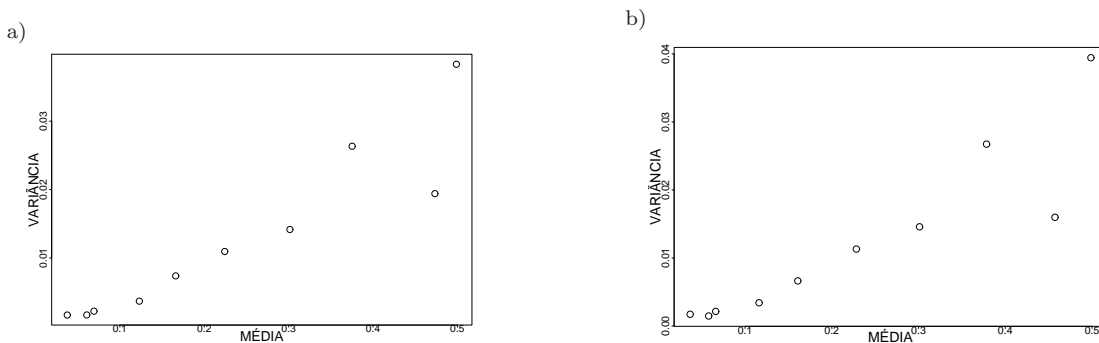


Figura 4 - Gráfico da média x variância para o modelo I (a), e modelos II e III (b)

O semivariograma empírico pode dar uma idéia errada se for calculado com dados com variação espacial na média. Ajustar uma média não constante para dados espaciais tende a inflacionar a estimativa resultante da variação do processo estocástico no processo adjacente. Uma solução proposta por Diggle & Ribeiro Júnior (2000) é assumir que $\mu(x)$ é determinada a partir de um modelo de regressão linear dado por

$$\mu(x) = \sum_{k=1}^p f_k(x)\alpha_k, \quad (36)$$

em que, $f_k(\cdot)$ são observações espacialmente relacionadas às variáveis explanatórias;

Diante do exposto, apresentamos as análises descritivas espaciais da variável transformada, considerando as variáveis explanatórias. Para o modelo I as variáveis explanatórias inseridas serão os tratamentos, para o II serão os tratamentos e a covariável de morte da planta e para o modelo III os tratamentos e as covariáveis de morte da planta e tempo de ocorrência.

A figura 5 mostra a localização geográfica da variável volume quando as variáveis explanatórias são inseridas, sendo que uma visão dos dados é que aquela em que cores mais fortes são associadas aos maiores valores da variável estudada.

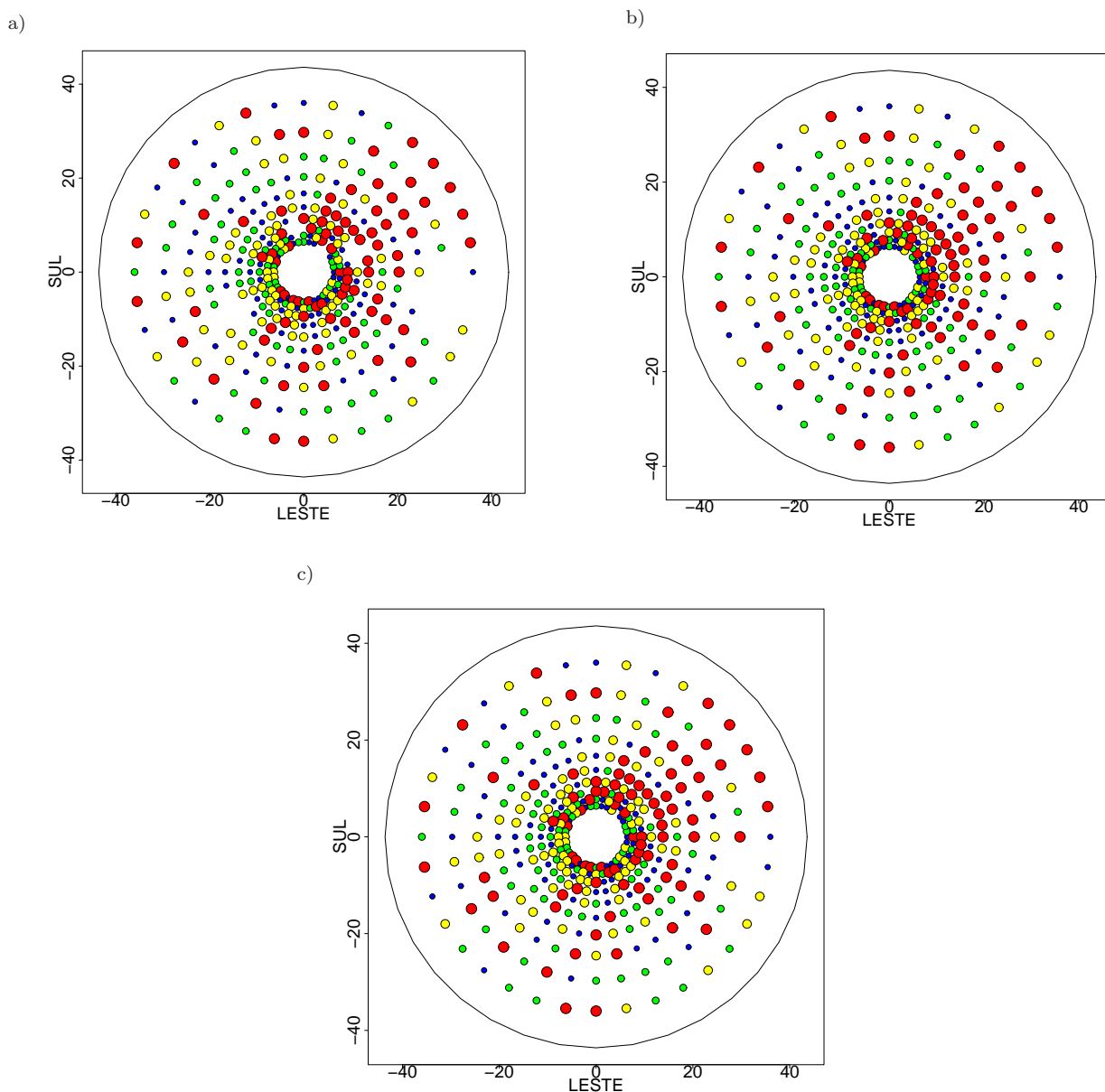


Figura 5 - Localização geográfica dos resíduos da variável volume considerando as variáveis explanatórias para os modelos I (a), II (b) e III (c)

4.3 Resultados para os modelos estudados

As figuras 6, 7 e 8 ilustram os três semivariogramas teóricos ajustados aos valores do semivariograma empírico construído a partir dos resíduos da análise clássica para os modelos I, II e III.

Os gráficos ajustados encontram-se no limite do intervalo do envelope simulado, apresentando alguns pontos fora do envelope, ocorrendo a dependência espacial.

No entanto, pode-se afirmar que a dependência espacial é considerada fraca, o que não justificaria considerá-la com a aplicação de um modelo geostatístico.

Assim, não se levando em consideração a dependência espacial, ou seja, não ocorrendo a autocorrelação espacial entre as parcelas, os dados poderão ser submetidos a uma análise clássica.

Conforme a tabela 6, os valores das estimativas dos parâmetros sill e nugget ($\hat{\sigma}^2$, $\hat{\tau}^2$), foram próximos para todos modelos e estruturas de covariâncias testadas. Para o parâmetro estimado range ($\hat{\phi}$), os valores foram sempre menores nos modelos com estrutura de covariância Matérn com $\kappa = 2$. Estes resultados foram confirmados com o valor máximo do logaritmo da função de máxima verossimilhança (Log). Assim, a função de correlação que apresentou melhor desempenho foi a Matérn com $\kappa = 2$ para os três modelos considerados.

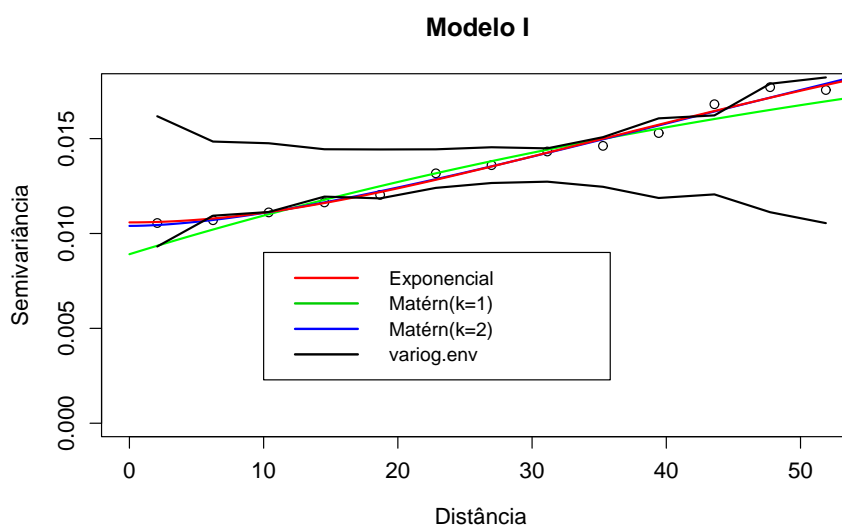


Figura 6 - Ajuste dos semivariogramas teóricos aos valores do semivariograma empírico e envelope simulado para o Modelo I. Parâmetros ajustados por máximo verossimilhança considerando-se as estruturas de covariância exponencial, Matérn com $\kappa = 1$ e Matérn com $\kappa = 2$

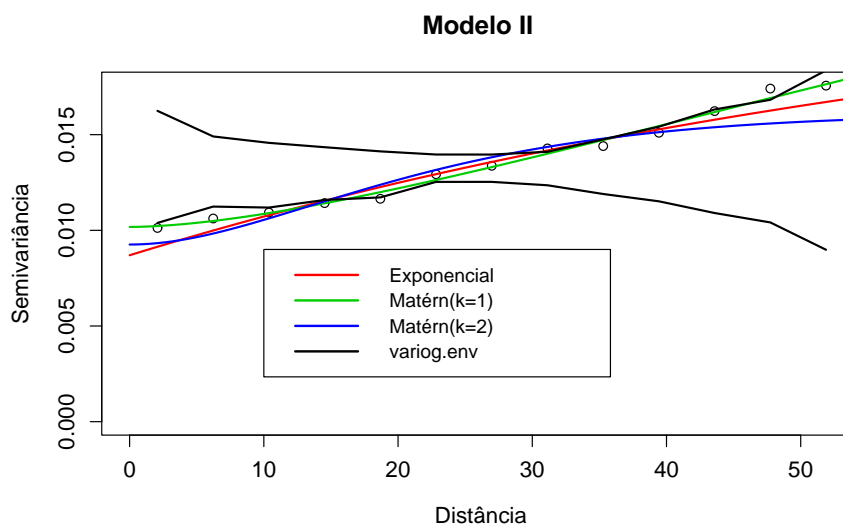


Figura 7 - Ajuste dos semivariogramas teóricos aos valores do semivariograma empírico e envelope simulado para o Modelo II. Parâmetros ajustados por máximo verossimilhança considerando-se as estruturas de covariância exponencial, Matérn com $\kappa = 1$ e Matérn com $\kappa = 2$

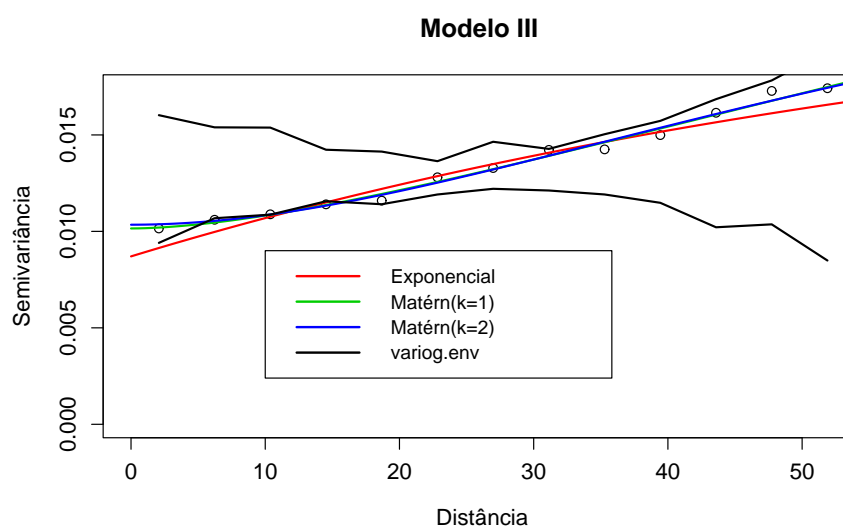


Figura 8 - Ajuste dos semivariogramas teóricos aos valores do semivariograma empírico e envelope simulado para o Modelo III. Parâmetros ajustados por máximo verossimilhança considerando-se as estruturas de covariância exponencial, Matérn com $\kappa = 1$ e Matérn com $\kappa = 2$

Tabela 6. Resultados das estimativas dos parâmetros, $(\hat{\sigma}^2, \hat{\tau}^2, \hat{\phi})$, logaritmo da função de verossimilhança (Log), número de parâmetros do modelo considerado (Npars) e critério de Akaike (AIC) para os modelos I, II e III, obtidas pelo método de estimação de parâmetros do modelo geoestatístico utilizando diferentes modelos de função de correlação

Modelo	Função de Correlação	$\hat{\sigma}^2$	$\hat{\tau}^2$	$\hat{\phi}$	Log	Npars	AIC
I	Exponencial	0,0007	0,0128	14,7680	248,8784	13	-472,7569
	Matérn ($\kappa = 1$)	0,0007	0,0128	10,8015	249,1508	13	-472,3016
	Matérn ($\kappa = 2$)	0,0008	0,0128	7,4183	249,4095	13	-472,8189
II	Exponencial	0,0009	0,0124	15,7573	267,1691	14	-506,3381
	Matérn ($\kappa = 1$)	0,0009	0,0125	11,4108	267,4671	14	-506,9342
	Matérn ($\kappa = 2$)	0,0009	0,0125	7,8339	267,7342	14	-507,4685
III	Exponencial	0,0008	0,0124	14,7581	267,7417	15	-505,4835
	Matérn ($\kappa = 1$)	0,0008	0,0125	10,8284	268,0239	15	-506,0477
	Matérn ($\kappa = 2$)	0,0008	0,0125	7,5082	268,2830	15	-506,5661

Os valores das estimativas dos parâmetros range ($\hat{\phi}$) foram próximos para os modelos com a função de correlação Matérn com $\kappa = 2$, comparando-se esses modelos e seguindo o critério de Akaike, o modelo II foi o mais adequado, pois apresentou menor valor de AIC (Tabela 6).

5 CONCLUSÕES

Face as metodologias empregadas nesse trabalho e os resultados obtidos, pode-se concluir que:

- a) os métodos geoestatísticos foram adequados para verificar se há de dependência espacial na análise do experimento em delineamento sistemático tipo “leque”;
- b) a análise indicou uma pequena (fraca) contribuição do efeito espacial, não se identificando dependência espacial nos modelos testados;
- c) o modelo mais adequado foi o modelo II que envolve a variável morte de planta, de acordo com o critério de Akaike;

Essas conclusões só podem ser consideradas para esse conjunto de dados. Deve-se avaliar dois pontos nesse trabalho. Os dados não apresentam dependência espacial ou está mascarado por algum fator não investigado? Por esses motivos, como sugestões para outros experimentos, a dependência espacial deve ser verificada de acordo com as condições e variáveis que possam estar influenciando, e os estudos de simulação devem ser considerados.

ANEXOS

ANEXO A - Conjunto de dados.

Tabela 7. Conjunto de dados para o tratamento 1, aos 6 anos. Ap - Área por Planta, N° Raio - Número do raio das plantas, X1 e X2 (m) - Coordenadas, Trat - Tratamento, H (m) - Altura, Dap (cm) - Diâmetro altura do peito, Vol (m^3) - Volume, Mp - Covariável morte da planta e T - Covariável tempo

Ap	N° Raio	Coordenadas		Trat	H (m)	Dap (cm)	Vol (m^3)	Covariáveis	
		X1 (m)	X2 (m)					Mp	T
1,40	1	6,42000000	0,00000000	1	7,2	3,5	0,00304839	0	0
1,40	2	6,32246577	1,11482130	1	13,2	7,3	0,02413581	0	0
1,40	3	6,03282662	2,19576932	1	15,2	7,3	0,02777902	0	0
1,40	4	5,55988309	3,21000000	1	8,9	3,8	0,00443595	0	0
1,40	5	4,91800532	4,12669645	1	17,5	7,3	0,03196663	1	5
1,40	6	4,12669645	4,91800532	1	NA	NA	NA	NA	NA
1,40	7	3,21000000	5,55988309	1	18,9	10,8	0,075337859	1	5
1,40	8	2,19576932	6,03282662	1	7,7	3,2	0,002726235	0	0
1,40	9	1,11482130	6,32246577	1	9,2	3,8	0,00458494	0	0
1,40	10	0,00000000	6,42000000	1	12,4	7,6	0,024573312	0	0
1,40	11	-1,11482130	6,32246577	1	11,9	4,8	0,009438555	0	0
1,40	12	-2,19576932	6,03282662	1	8,9	3,8	0,004435947	0	0
1,40	13	-3,21000000	5,55988309	1	19,2	11,8	0,09130104	0	0
1,40	14	-4,12669645	4,91800532	1	15,2	8,3	0,035878693	0	0
1,40	15	-4,91800532	4,12669645	1	12,2	5,4	0,01223565	0	0
1,40	16	-5,55988309	3,21000000	1	10,9	4,6	0,00794477	0	0
1,40	17	-6,03282662	2,19576932	1	19,4	13,4	0,118855432	0	0
1,40	18	-6,32246577	1,11482130	1	15,7	7,3	0,028689547	0	0
1,40	19	-6,42000000	0,00000000	1	16,4	8,3	0,038700915	0	0
1,40	20	-6,32246577	-1,11482130	1	18,8	8,9	0,050961101	0	0
1,40	21	-6,03282662	-2,19576932	1	12,9	5,9	0,015431886	0	0
1,40	22	-5,55988309	-3,21000000	1	13,9	7,6	0,02753487	0	0
1,40	23	-4,91800532	-4,12669645	1	18,9	12,4	0,099217589	0	0
1,40	24	-4,12669645	-4,91800532	1	19,9	12,6	0,107832706	0	0
1,40	25	-3,21000000	-5,55988309	1	8,7	4,5	0,00607426	0	0
1,40	26	-2,19576932	-6,03282662	1	18,9	12,1	0,094491029	0	0
1,40	27	-1,11482130	-6,32246577	1	17,9	10,5	0,067468966	0	0
1,40	28	0,00000000	-6,42000000	1	10,4	5,3	0,010054595	0	0
1,40	29	1,11482130	-6,32246577	1	21,2	15,3	0,169117224	0	0
1,40	30	2,19576932	-6,03282662	1	5,4	2,2	0,000907177	0	0
1,40	31	3,21000000	-5,55988309	1	7,7	3,5	0,003259317	0	0
1,40	32	4,12669645	-4,91800532	1	14,2	7,0	0,023874994	0	0
1,40	33	4,91800532	-4,12669645	1	16,7	8,3	0,039406356	0	0
1,40	34	5,55988309	-3,21000000	1	10,5	5,7	0,011734984	0	0
1,40	35	6,03282662	-2,19576932	1	16,3	8,6	0,041286394	0	0
1,40	36	6,32246577	-1,11482130	1	8,2	3,7	0,003876611	0	0

NA - dado perdido (planta morta)

ANEXO A - Conjunto de dados.

Tabela 8. Conjunto de dados para o tratamento 2, aos 6 anos. Ap - Área por Planta, N° Raio - Número do raio das plantas, X1 e X2 (m) - Coordenadas, TRAT - Tratamento, H (m) - Altura, Dap (cm) - Diâmetro altura do peito, Vol (m^3) - Volume, Mp - Covariável morte da planta e T - Covariável tempo

Ap	N° Raio	Coordenadas		Trat	H (m)	Dap (cm)	VOL (m^3)	Covariáveis	
		X1 (m)	X2 (m)					Mp	T
2,05	1	7,78000000	0,00000000	2	21,8	15,0	0,167157524	0	0
2,05	2	7,66180432	1,35098282	2	15,7	8,3	0,03705471	0	0
2,05	3	7,31080858	2,66091671	2	19,7	11,1	0,082921981	0	0
2,05	4	6,73767764	3,89000000	2	14,9	7,0	0,025047707	0	0
2,05	5	5,95982577	5,00088760	2	19,9	13,1	0,116528878	0	0
2,05	6	5,00088760	5,95982577	2	19,3	9,1	0,054680623	1	5
2,05	7	3,89000000	6,73767764	2	19,7	12,9	0,11187827	0	0
2,05	8	2,66091671	7,31080858	2	20,7	10,8	0,082486595	0	0
2,05	9	1,35098282	7,66180432	2	4,9	2,5	0,001062399	0	0
2,05	10	0,00000000	7,78000000	2	17,5	9,5	0,054037678	0	0
2,05	11	-1,35098282	7,66180432	2	22,3	14,5	0,159807242	1	2
2,05	12	-2,66091671	7,31080858	2	13,7	7,0	0,023037219	0	0
2,05	13	-3,89000000	6,73767764	2	14,1	7,2	0,025076561	0	0
2,05	14	-5,00088760	5,95982577	2	19,1	10,3	0,069269177	0	0
2,05	15	-5,95982577	5,00088760	2	17,6	9,4	0,053211234	0	0
2,05	16	-6,73767764	3,89000000	2	20,0	11,5	0,090334244	0	0
2,05	17	-7,31080858	2,66091671	2	20,4	9,9	0,068352297	0	0
2,05	18	-7,66180432	1,35098282	2	15,3	7,8	0,03190778	0	0
2,05	19	-7,78000000	0,00000000	2	17,7	10,7	0,069274423	0	0
2,05	20	-7,66180432	-1,35098282	2	18,3	9,9	0,061339376	0	0
2,05	21	-7,31080858	-2,66091671	2	17,7	9,2	0,051267332	0	0
2,05	22	-6,73767764	-3,89000000	2	18,8	9,5	0,058037332	0	0
2,05	23	-5,95982577	-5,00088760	2	19,4	10,2	0,068998597	0	0
2,05	24	-5,00088760	-5,95982577	2	12,0	4,8	0,009517591	0	0
2,05	25	-3,89000000	-6,73767764	2	17,2	9,2	0,049824104	0	0
2,05	26	-2,66091671	-7,31080858	2	17,7	10,0	0,06053566	0	0
2,05	27	-1,35098282	-7,66180432	2	14,1	6,4	0,01982994	0	0
2,05	28	0,00000000	-7,78000000	2	19,4	10,7	0,075903504	0	0
2,05	29	1,35098282	-7,66180432	2	17,5	9,5	0,054037678	0	0
2,05	30	2,66091671	-7,31080858	2	19,3	13,1	0,113027569	0	0
2,05	31	3,89000000	-6,73767764	2	19,5	13,1	0,114194714	0	0
2,05	32	5,00088760	-5,95982577	2	10,0	4,8	0,007936395	0	0
2,05	33	5,95982577	-5,00088760	2	18,3	11,1	0,077048948	0	0
2,05	34	6,73767764	-3,89000000	2	7,1	3,3	0,002673552	0	0
2,05	35	7,31080858	-2,66091671	2	15,5	8,6	0,039266992	0	0
2,05	36	7,66180432	-1,35098282	2	7,1	3,2	0,002514516	0	0

ANEXO A - Conjunto de dados.

Tabela 9. Conjunto de dados para o tratamento 3, aos 6 anos. Ap - Área por Planta, N° Raio - Número do raio das plantas, X1 e X2 (m) - Coordenadas, Trat - Tratamento, H (m) - Altura, Dap (cm) - Diâmetro altura do peito, Vol (m^3) - Volume, Mp - Covariável morte da planta e T - Covariável tempo

Ap	N° Raio	Coordenadas		Trat	H (m)	Dap (cm)	Vol (m^3)	Covariáveis	
		X1 (m)	X2 (m)					Mp	T
3,01	1	9,42000000	0,00000000	3	19,5	12,1	0,097480069	0	0
3,01	2	9,27688903	1,63576583	3	12,0	4,9	0,009916855	0	0
3,01	3	8,85190448	3,22182975	3	20,8	14,3	0,145023785	1	5
3,01	4	8,15795930	4,71000000	3	15,0	7,0	0,025215221	0	0
3,01	5	7,21613865	6,05505928	3	18,9	10,5	0,071224614	0	0
3,01	6	6,05505928	7,21613865	3	11,8	5,1	0,010561528	0	0
3,01	7	4,71000000	8,15795930	3	20,0	11,8	0,095091647	0	0
3,01	8	3,22182975	8,85190448	3	15,9	7,6	0,031481889	0	0
3,01	9	1,63576583	9,27688903	3	21,5	14,3	0,149887008	0	0
3,01	10	0,00000000	9,42000000	3	22,0	12,7	0,121062526	1	2
3,01	11	-1,63576583	9,27688903	3	NA	NA	NA	NA	NA
3,01	12	-3,22182975	8,85190448	3	16,5	8,8	0,043750519	1	2
3,01	13	-4,71000000	8,15795930	3	19,7	11,8	0,093670232	0	0
3,01	14	-6,05505928	7,21613865	3	18,7	10,2	0,066517522	0	0
3,01	15	-7,21613865	6,05505928	3	10,0	6,0	0,01238124	0	0
3,01	16	-8,15795930	4,71000000	3	15,3	9,2	0,044338461	0	0
3,01	17	-8,85190448	3,22182975	3	21,0	15,0	0,161044399	0	0
3,01	18	-9,27688903	1,63576583	3	5,7	2,1	0,000872621	0	0
3,01	19	-9,42000000	0,00000000	3	16,6	9,2	0,048092037	0	0
3,01	20	-9,27688903	-1,63576583	3	7,7	2,9	0,002240572	0	0
3,01	21	-8,85190448	-3,22182975	3	17,0	11,1	0,071594008	0	0
3,01	22	-8,15795930	-4,71000000	3	18,0	9,2	0,052133199	0	0
3,01	23	-7,21613865	-6,05505928	3	19,3	11,8	0,091774895	0	0
3,01	24	-6,05505928	-7,21613865	3	18,4	11,3	0,080275262	0	0
3,01	25	-4,71000000	-8,15795930	3	4,0	2,2	0,00067269	0	0
3,01	26	-3,22182975	-8,85190448	3	18,1	10,8	0,07215989	0	0
3,01	27	-1,63576583	-9,27688903	3	16,7	10,5	0,062961214	0	0
3,01	28	0,00000000	-9,42000000	3	19,3	13,5	0,120010098	0	0
3,01	29	1,63576583	-9,27688903	3	11,4	5,4	0,01143603	0	0
3,01	30	3,22182975	-8,85190448	3	10,8	5,4	0,010836186	0	0
3,01	31	4,71000000	-8,15795930	3	19,0	11,8	0,090353302	0	0
3,01	32	6,05505928	-7,21613865	3	18,8	11,5	0,084932601	0	0
3,01	33	7,21613865	-6,05505928	3	19,1	11,3	0,083318311	0	0
3,01	34	8,15795930	-4,71000000	3	19,7	12,4	0,103402253	0	0
3,01	35	8,85190448	-3,22182975	3	20,4	12,3	0,105349458	0	0
3,01	36	9,27688903	-1,63576583	3	21,5	14,6	0,15621921	0	0

NA - dado perdido (planta morta)

ANEXO A - Conjunto de dados.

Tabela 10. Conjunto de dados para o tratamento 4, aos 6 anos. Ap - Área por Planta, N° Raio - Número do raio das plantas, X1 e X2 (m) - Coordenadas, Trat - Tratamento, H (m) - Altura, Dap (cm) - Diâmetro altura do peito, Vol (m^3) - Volume, Mp - Covariável morte da planta e T - Covariável tempo

Ap	N° Raio	Coordenadas		Trat	H (m)	Dap (cm)	Vol (m^3)	Covariáveis	
		X1 (m)	X2 (m)					Mp	T
4,41	1	11,41000000	0,00000000	4	21,2	14,0	0,141687423	0	0
4,41	2	11,23665646	1,98132570	4	21,9	15,6	0,181574097	1	5
4,41	3	10,72189279	3,90244983	4	NA	NA	NA	NA	NA
4,41	4	9,88134985	5,70500000	4	24,2	16,7	0,229746835	1	5
4,41	5	8,74056710	7,33420662	4	22,2	14,0	0,148346832	0	0
4,41	6	7,33420662	8,74056710	4	22,4	14,6	0,162735234	0	0
4,41	7	5,70500000	9,88134985	4	21,9	13,1	0,128197297	0	0
4,41	8	3,90244983	10,72189279	4	22,5	15,4	0,181795313	0	0
4,41	9	1,98132570	11,23665646	4	21,2	13,5	0,131781199	0	0
4,41	10	0,00000000	11,41000000	4	22,5	18,3	0,25640017	0	0
4,41	11	-1,98132570	11,23665646	4	21,9	15,6	0,181574097	1	2
4,41	12	-3,90244983	10,72189279	4	17,4	8,9	0,047178917	0	0
4,41	13	-5,70500000	9,88134985	4	21,4	12,7	0,117772232	0	0
4,41	14	-7,33420662	8,74056710	4	20,3	13,4	0,124349583	0	0
4,41	15	-8,74056710	7,33420662	4	13,9	7,3	0,02541114	0	0
4,41	16	-9,88134985	5,70500000	4	19,9	13,7	0,127407739	0	0
4,41	17	-10,72189279	3,90244983	4	20,0	13,4	0,122518296	0	0
4,41	18	-11,23665646	1,98132570	4	18,7	11,5	0,084482411	0	0
4,41	19	-11,41000000	0,00000000	4	16,6	8,6	0,042043579	0	0
4,41	20	-11,23665646	-1,98132570	4	18,7	12,1	0,093494609	0	0
4,41	21	-10,72189279	-3,90244983	4	20,2	15,0	0,154930458	0	0
4,41	22	-9,88134985	-5,70500000	4	17,0	9,5	0,052499077	0	0
4,41	23	-8,74056710	-7,33420662	4	19,7	12,7	0,108447943	0	0
4,41	24	-7,33420662	-8,74056710	4	19,2	12,6	0,104052651	0	0
4,41	25	-5,70500000	-9,88134985	4	20,5	14,0	0,137025185	0	0
4,41	26	-3,90244983	-10,72189279	4	22,0	16,9	0,213946963	0	0
4,41	27	-1,98132570	-11,23665646	4	19,7	13,4	0,120686913	0	0
4,41	28	0,00000000	-11,41000000	4	14,2	7,3	0,025957639	0	0
4,41	29	1,98132570	-11,23665646	4	17,1	8,9	0,046368313	0	0
4,41	30	3,90244983	-10,72189279	4	20,4	14,3	0,142244545	0	0
4,41	31	5,70500000	-9,88134985	4	21,9	16,2	0,195758155	0	0
4,41	32	7,33420662	-8,74056710	4	18,9	11,1	0,079566148	0	0
4,41	33	8,74056710	-7,33420662	4	19,8	12,4	0,103925294	0	0
4,41	34	9,88134985	-5,70500000	4	19,2	11,5	0,086733279	0	0
4,41	35	10,72189279	-3,90244983	4	22,5	16,2	0,201102345	0	0
4,41	36	11,23665646	-1,98132570	4	9,2	4,6	0,006709663	0	0

NA - dado perdido (planta morta)

ANEXO A - Conjunto de dados.

Tabela 11. Conjunto de dados para o tratamento 5, aos 6 anos Ap - Área por Planta, N° Raio - Número do raio das plantas, X1 e X2 (m) - Coordenadas, Trat - Tratamento, H (m) - Altura, Dap (cm) - Diâmetro altura do peito, Vol (m^3) - Volume, Mp - Covariável morte da planta e T - Covariável tempo

Ap	N° Raio	Coordenadas		Trat	H (m)	Dap (cm)	Vol (m^3)	Covariáveis	
		X1 (m)	X2 (m)					Mp	T
6,47	1	13,82000000	0,00000000	5	22,2	17,5	0,231429734	0	0
6,47	2	13,61004315	2,39981781	5	22,4	16,7	0,21271583	0	0
6,47	3	12,98655201	4,72671838	5	24,2	21,0	0,362708614	1	5
6,47	4	11,96847107	6,91000000	5	23,1	16,7	0,219339548	0	0
6,47	5	10,58673420	8,88332476	5	24,0	20,4	0,339528585	0	0
6,47	6	8,88332476	10,58673420	5	24,5	20,1	0,3364935	0	0
6,47	7	6,91000000	11,96847107	5	22,7	18,1	0,253067629	0	0
6,47	8	4,72671838	12,98655201	5	23,0	17,2	0,231618672	0	0
6,47	9	2,39981781	13,61004315	5	21,8	14,6	0,158391322	0	0
6,47	10	0,00000000	13,82000000	5	8,2	3,5	0,003470196	0	0
6,47	11	-2,39981781	13,61004315	5	20,2	15,3	0,161167281	0	0
6,47	12	-4,72671838	12,98655201	5	22,6	18,5	0,263175633	0	0
6,47	13	-6,91000000	11,96847107	5	21,0	15,8	0,178615498	0	0
6,47	14	-8,88332476	10,58673420	5	15,7	8,9	0,042584821	0	0
6,47	15	-10,58673420	8,88332476	5	19,0	13,1	0,111276772	0	0
6,47	16	-11,96847107	6,91000000	5	19,3	12,4	0,101309997	0	0
6,47	17	-12,98655201	4,72671838	5	18,1	14,2	0,124506503	0	0
6,47	18	-13,61004315	2,39981781	5	15,4	8,3	0,036349116	0	0
6,47	19	-13,82000000	0,00000000	5	20,0	14,2	0,137528135	0	0
6,47	20	-13,61004315	-2,39981781	5	20,1	14,8	0,150096585	0	0
6,47	21	-12,98655201	-4,72671838	5	20,3	16,2	0,181504433	0	0
6,47	22	-11,96847107	-6,91000000	5	20,5	14,8	0,153073012	1	1
6,47	23	-10,58673420	-8,88332476	5	19,2	12,1	0,09598559	0	0
6,47	24	-8,88332476	-10,58673420	5	21,3	15,3	0,169912144	0	0
6,47	25	-6,91000000	-11,96847107	5	22,3	19,3	0,282556141	0	0
6,47	26	-4,72671838	-12,98655201	5	20,4	11,8	0,09698675	0	0
6,47	27	-2,39981781	-13,61004315	5	20,9	16,1	0,184558361	0	0
6,47	28	0,00000000	-13,82000000	5	20,7	15,0	0,158751768	0	0
6,47	29	2,39981781	-13,61004315	5	18,5	11,1	0,077888046	0	0
6,47	30	4,72671838	-12,98655201	5	20,5	16,1	0,181038402	0	0
6,47	31	6,91000000	-11,96847107	5	19,0	11,1	0,079985654	0	0
6,47	32	8,88332476	-10,58673420	5	20,2	17,7	0,21547509	0	0
6,47	33	10,58673420	-8,88332476	5	18,5	12,3	0,095570232	0	0
6,47	34	11,96847107	-6,91000000	5	19,9	10,8	0,07930966	0	0
6,47	35	12,98655201	-4,72671838	5	18,9	11,5	0,085382783	0	0
6,47	36	13,61004315	-2,39981781	5	22,8	15,6	0,189009372	0	0

ANEXO A - Conjunto de dados.

Tabela 12. Conjunto de dados para o tratamento 6, aos 6 anos. Ap - Área por Planta, N° Raio - Número do raio das plantas, X1 e X2 (m) - Coordenadas, Trat - Tratamento, H (m) - Altura, Dap (cm) - Diâmetro altura do peito, Vol (m^3) - Volume, Mp - Covariável morte da planta e T - Covariável tempo

Ap	N° Raio	Coordenadas		Trat	H (m)	Dap (cm)	Vol (m^3)	Covariáveis	
		X1 (m)	X2 (m)					Mp	T
9,50	1	16,74000000	0,00000000	6	23,5	16,2	0,210008225	0	0
9,50	2	16,48568179	2,90687048	6	23,3	18,1	0,259732897	0	0
9,50	3	15,73045446	5,72541719	6	24,9	23,2	0,45512701	0	0
9,50	4	14,49726525	8,37000000	6	24,5	19,7	0,323279522	0	0
9,50	5	12,82358398	10,76026457	6	21,0	13,7	0,134425033	0	0
9,50	6	10,76026457	12,82358398	6	9,7	7,6	0,019239219	0	0
9,50	7	8,37000000	14,49726525	6	23,2	18,5	0,270137793	0	0
9,50	8	5,72541719	15,73045446	6	22,9	19,7	0,302238905	0	0
9,50	9	2,90687048	16,48568179	6	22,0	17,2	0,221582806	0	0
9,50	10	0,00000000	16,74000000	6	18,4	13,7	0,117836494	0	0
9,50	11	-2,90687048	16,48568179	6	20,9	18,5	0,243445935	0	0
9,50	12	-5,72541719	15,73045446	6	19,2	14,3	0,133905661	0	0
9,50	13	-8,37000000	14,49726525	6	16,8	10,8	0,06699463	0	0
9,50	14	-10,76026457	12,82358398	6	19,4	14,0	0,129697676	0	0
9,50	15	-12,82358398	10,76026457	6	20,8	23,2	0,380426172	0	0
9,50	16	-14,49726525	8,37000000	6	22,0	18,1	0,245290702	0	0
9,50	17	-15,73045446	5,72541719	6	22,5	19,1	0,279223932	0	0
9,50	18	-16,48568179	2,90687048	6	22,0	19,1	0,273040456	0	0
9,50	19	-16,74000000	0,00000000	6	19,3	15,0	0,148051258	0	0
9,50	20	-16,48568179	-2,90687048	6	19,4	14,3	0,135295599	0	0
9,50	21	-15,73045446	-5,72541719	6	19,0	14,6	0,138113996	1	1
9,50	22	-14,49726525	-8,37000000	6	NA	NA	NA	NA	NA
9,50	23	-12,82358398	-10,76026457	6	19,8	16,2	0,177049351	1	1
9,50	24	-10,76026457	-12,82358398	6	20,0	15,3	0,159577129	0	0
9,50	25	-8,37000000	-14,49726525	6	22,6	22,6	0,392202037	0	0
9,50	26	-5,72541719	-15,73045446	6	21,1	16,6	0,198028065	0	0
9,50	27	-2,90687048	-16,48568179	6	20,5	17,2	0,206525985	0	0
9,50	28	0,00000000	-16,74000000	6	14,2	9,2	0,041161488	0	0
9,50	29	2,90687048	-16,48568179	6	23,5	21,8	0,379503373	0	0
9,50	30	5,72541719	-15,73045446	6	21,4	16,6	0,200833695	0	0
9,50	31	8,37000000	-14,49726525	6	20,5	15,3	0,163552405	0	0
9,50	32	10,76026457	-12,82358398	6	23,5	20,4	0,3324796	0	0
9,50	33	12,82358398	-10,76026457	6	19,8	15,0	0,151873171	0	0
9,50	34	14,49726525	-8,37000000	6	22,5	21,0	0,337315162	0	0
9,50	35	15,73045446	-5,72541719	6	23,7	20,7	0,345198194	0	0
9,50	36	16,48568179	-2,90687048	6	22,7	18,9	0,275849005	0	0

NA - dado perdido (planta morta)

ANEXO A - Conjunto de dados.

Tabela 13. Conjunto de dados para o tratamento 7, aos 6 anos. Ap - Área por Planta, N° Raio - Número do raio das plantas, X1 e X2 (m) - Coordenadas, Trat - Tratamento, H (m) - Altura, Dap (cm) - Diâmetro altura do peito, Vol (m^3) - Volume, Mp - Covariável morte da planta e T - Covariável tempo

Ap	N° Raio	Coordenadas		Trat	H (m)	DAP (cm)	Vol (m^3)	Covariáveis	
		X1 (m)	X2 (m)					Mp	T
13,93	1	20,27000000	0,00000000	7	24,5	25,3	0,532261846	0	0
13,93	2	19,96205315	3,51984855	7	24,6	24,5	0,501276193	0	0
13,93	3	19,04756941	6,93274830	7	18,1	12,4	0,095032303	0	0
13,93	4	17,55433492	10,13500000	7	25,2	23,4	0,46853841	0	0
13,93	5	15,52772086	13,02930483	7	26,5	25,3	0,575553514	0	0
13,93	6	13,02930483	15,52772086	7	24,2	20,4	0,342348035	0	0
13,93	7	10,13500000	17,55433492	7	22,7	24,7	0,470248203	0	0
13,93	8	6,93274830	19,04756941	7	20,2	15,3	0,161167281	0	0
13,93	9	3,51984855	19,96205315	7	23,5	20,1	0,322806204	0	0
13,93	10	0,00000000	20,27000000	7	20,2	18,5	0,235320331	0	0
13,93	11	-3,51984855	19,96205315	7	20,2	15,9	0,174008792	0	0
13,93	12	-6,93274830	19,04756941	7	22,1	19,1	0,27427719	0	0
13,93	13	-10,13500000	17,55433492	7	21,5	19,7	0,283824149	0	0
13,93	14	-13,02930483	15,52772086	7	16,7	13,1	0,097850656	0	0
13,93	15	-15,52772086	13,02930483	7	21,0	17,2	0,211545341	0	0
13,93	16	-17,55433492	10,13500000	7	21,5	17,8	0,231881218	0	0
13,93	17	-19,04756941	6,93274830	7	21,6	17,8	0,232955947	0	0
13,93	18	-19,96205315	3,51984855	7	22,0	22,0	0,361889024	0	0
13,93	19	-20,27000000	0,00000000	7	20,5	16,9	0,199409007	0	0
13,93	20	-19,96205315	-3,51984855	7	22,0	20,1	0,302271405	0	0
13,93	21	-19,04756941	-6,93274830	7	20,5	18,5	0,238802852	0	0
13,93	22	-17,55433492	-10,13500000	7	21,5	20,7	0,313261398	1	1
13,93	23	-15,52772086	-13,02930483	7	23,1	20,7	0,336489231	0	0
13,93	24	-13,02930483	-15,52772086	7	22,9	21,6	0,363115832	0	0
13,93	25	-10,13500000	-17,55433492	7	20,7	15,8	0,176072725	0	0
13,93	26	-6,93274830	-19,04756941	7	22,6	21,3	0,348523833	0	0
13,93	27	-3,51984855	-19,96205315	7	20,0	15,9	0,17229194	0	0
13,93	28	0,00000000	-20,27000000	7	23,3	23,6	0,440743849	0	0
13,93	29	3,51984855	-19,96205315	7	23,5	20,1	0,322806204	0	0
13,93	30	6,93274830	-19,04756941	7	21,3	18,8	0,256171185	0	0
13,93	31	10,13500000	-17,55433492	7	22,0	21,6	0,348893917	0	0
13,93	32	13,02930483	-15,52772086	7	21,0	19,7	0,277246446	0	0
13,93	33	15,52772086	-13,02930483	7	22,6	19,4	0,289308441	0	0
13,93	34	17,55433492	-10,13500000	7	23,9	23,6	0,452053201	0	0
13,93	35	19,04756941	-6,93274830	7	19,9	15,3	0,158782032	0	0
13,93	36	19,96205315	-3,51984855	7	22,8	19,1	0,282933786	0	0

ANEXO A - Conjunto de dados.

Tabela 14. Conjunto de dados para o tratamento 8, aos 6 anos. Ap - Área por Planta, N° Raio - Número do raio das plantas, X1 e X2 (m) - Coordenadas, Trat - Tratamento, H (m) - Altura, Dap (cm) - Diâmetro altura do peito, Vol (m^3) - Volume, Mp - Covariável morte da planta e T - Covariável tempo

Ap	N° Raio	Coordenadas		Trat	H (m)	Dap (cm)	Vol (m^3)	Covariáveis	
		X1 (m)	X2 (m)					Mp	T
20,43	1	24,55000000	0,00000000	8	24,5	21,6	0,388394448	0	0
20,43	2	24,17703034	4,26306275	8	19,5	19,4	0,24975362	0	0
20,43	3	23,06945382	8,39659451	8	24,9	25,1	0,532432442	0	0
20,43	4	21,26092364	12,27500000	8	25,0	26,3	0,586706622	0	0
20,43	5	18,80639108	15,78043580	8	27,2	27,4	0,692446876	0	0
20,43	6	15,78043580	18,80639108	8	25,7	27,9	0,678405321	0	0
20,43	7	12,27500000	21,26092364	8	21,4	22,0	0,352053432	0	0
20,43	8	8,39659451	23,06945382	8	24,5	21,5	0,384819041	0	0
20,43	9	4,26306275	24,17703034	8	22,2	20,1	0,305009654	0	0
20,43	10	0,00000000	24,55000000	8	22,7	21,6	0,359955579	0	0
20,43	11	-4,26306275	24,17703034	8	21,5	22,9	0,383115523	0	0
20,43	12	-8,39659451	23,06945382	8	20,7	24,5	0,422060777	0	0
20,43	13	-12,27500000	21,26092364	8	21,2	21,5	0,333155116	0	0
20,43	14	-15,78043580	18,80639108	8	19,7	21,5	0,309662434	0	0
20,43	15	-18,80639108	15,78043580	8	20,7	20,8	0,304556437	0	0
20,43	16	-21,26092364	12,27500000	8	24,7	25,0	0,523985311	0	0
20,43	17	-23,06945382	8,39659451	8	20,8	14,3	0,145023785	0	0
20,43	18	-24,17703034	4,26306275	8	20,7	16,9	0,201347611	0	0
20,43	19	-24,55000000	0,00000000	8	21,5	22,9	0,383115523	0	0
20,43	20	-24,17703034	-4,26306275	8	21,6	22,6	0,374907413	0	0
20,43	21	-23,06945382	-8,39659451	8	25,0	27,4	0,636628259	0	0
20,43	22	-21,26092364	-12,27500000	8	20,3	25,5	0,448287692	0	0
20,43	23	-18,80639108	-15,78043580	8	18,2	12,4	0,095555499	0	0
20,43	24	-15,78043580	-18,80639108	8	22,5	22,9	0,400870986	0	0
20,43	25	-12,27500000	-21,26092364	8	18,2	12,1	0,091003395	0	0
20,43	26	-8,39659451	-23,06945382	8	22,7	20,4	0,321200125	0	0
20,43	27	-4,26306275	-24,17703034	8	25,7	24,5	0,523610716	0	0
20,43	28	0,00000000	-24,55000000	8	23,5	23,6	0,444513745	0	0
20,43	29	4,26306275	-24,17703034	8	22,2	26,4	0,525169533	0	0
20,43	30	8,39659451	-23,06945382	8	22,0	18,8	0,264559979	0	0
20,43	31	12,27500000	-21,26092364	8	21,0	18,8	0,252575694	0	0
20,43	32	15,78043580	-18,80639108	8	25,7	25,9	0,584933853	0	0
20,43	33	18,80639108	-15,78043580	8	18,2	18,8	0,219008724	0	0
20,43	34	21,26092364	-12,27500000	8	22,7	25,1	0,485547573	0	0
20,43	35	23,06945382	-8,39659451	8	9,2	11,0	0,0381344	0	0
20,43	36	24,17703034	-4,26306275	8	20,2	19,6	0,264029435	1	5

ANEXO A - Conjunto de dados.

Tabela 15. Conjunto de dados para o tratamento 9, aos 6 anos. Ap - Área por Planta, N° Raio - Número do raio das plantas, X1 e X2 (m) - Coordenadas, Trat - Tratamento, H (m) - Altura, Dap (cm) - Diâmetro altura do peito, Vol (m^3) - Volume, Mp - Covariável morte da planta e T - Covariável tempo

Ap	N° Raio	Coordenadas		Trat	H (m)	Dap (cm)	Vol (m^3)	Covariáveis	
		X1 (m)	X2 (m)					Mp	T
29,97	1	29,73000000	0,00000000	9	24,0	30,2	0,742054656	1	5
29,97	2	29,27833450	5,16256030	9	23,0	22,3	0,388629768	0	0
29,97	3	27,93706159	10,16825885	9	24,0	24,5	0,48909226	0	0
29,97	4	25,74693523	14,86500000	9	25,0	27,1	0,622811862	0	0
29,97	5	22,77450129	19,11007562	9	25,5	27,4	0,649315768	0	0
29,97	6	19,11007562	22,77450129	9	22,7	24,5	0,46269004	0	0
29,97	7	14,86500000	25,74693523	9	22,7	27,7	0,590940695	0	0
29,97	8	10,16825885	27,93706159	9	21,2	22,9	0,377788326	0	0
29,97	9	5,16256030	29,27833450	9	22,0	27,5	0,564568043	0	0
29,97	10	0,00000000	29,73000000	9	23,5	27,5	0,602921957	0	0
29,97	11	-5,16256030	29,27833450	9	23,0	27,1	0,573154347	0	0
29,97	12	-10,16825885	27,93706159	9	22,0	25,8	0,497146663	0	0
29,97	13	-14,86500000	25,74693523	9	21,0	22,9	0,374236715	0	0
29,97	14	-19,11007562	22,77450129	9	18,0	21,2	0,275213483	0	0
29,97	15	-22,77450129	19,11007562	9	21,5	24,8	0,4490757	0	0
29,97	16	-25,74693523	14,86500000	9	19,8	21,6	0,314120472	0	0
29,97	17	-27,93706159	10,16825885	9	21,8	22,9	0,38844246	0	0
29,97	18	-29,27833450	5,16256030	9	23,2	22,3	0,391997265	0	0
29,97	19	-29,73000000	0,00000000	9	20,5	18,3	0,233685257	0	0
29,97	20	-29,27833450	-5,16256030	9	23,0	24,8	0,480293049	0	0
29,97	21	-27,93706159	-10,16825885	9	22,6	21,3	0,348523833	0	0
29,97	22	-25,74693523	-14,86500000	9	24,2	28,3	0,657331046	0	0
29,97	23	-22,77450129	-19,11007562	9	23,7	24,2	0,471284436	0	0
29,97	24	-19,11007562	-22,77450129	9	22,5	27,4	0,573177001	0	0
29,97	25	-14,86500000	-25,74693523	9	22,1	23,2	0,404116952	0	0
29,97	26	-10,16825885	-27,93706159	9	23,8	26,7	0,575702187	0	0
29,97	27	-5,16256030	-29,27833450	9	21,3	21,6	0,337831022	0	0
29,97	28	0,00000000	-29,73000000	9	22,5	23,7	0,429265358	0	0
29,97	29	5,16256030	-29,27833450	9	21,4	23,6	0,404924037	0	0
29,97	30	10,16825885	-27,93706159	9	22,2	23,2	0,405939114	0	0
29,97	31	14,86500000	-25,74693523	9	20,6	24,2	0,409840903	0	0
29,97	32	19,11007562	-22,77450129	9	18,0	18,1	0,200833559	0	0
29,97	33	22,77450129	-19,11007562	9	24,2	29,9	0,733476604	0	0
29,97	34	25,74693523	-14,86500000	9	19,3	25	0,409783895	0	0
29,97	35	27,93706159	-10,16825885	9	24,2	29,9	0,733476604	1	5
29,97	36	29,27833450	-5,16256030	9	NA	NA	NA	NA	NA

NA - dado perdido (planta morta)

ANEXO A - Conjunto de dados.

Tabela 16. Conjunto de dados para o tratamento 10, aos 6 anos. Ap - Área por Planta, N° Raio - Número do raio das plantas, X1 e X2 (m) - Coordenadas, Trat - Tratamento, H (m) - Altura, Dap (cm) - Diâmetro altura do peito, Vol (m^3) - Volume, Mp - Covariável morte da planta e T - Covariável tempo

Ap	N° Raio	Coordenadas		Trat	H (m)	Dap (cm)	Vol (m^3)	Covariáveis	
		X1 (m)	X2 (m)					Mp	T
36,00	1	36,00000000	0,00000000	10	8,7	11,8	0,041485694	0	0
36,00	2	35,45307911	6,25133437	10	24,3	30,6	0,771260238	0	0
36,00	3	33,82893432	12,31272515	10	25,0	29,3	0,727639011	0	0
36,00	4	31,17691451	18,00000000	10	24,2	29,3	0,704434837	0	0
36,00	5	27,57759995	23,14035392	10	24,3	32,0	0,843183095	0	0
36,00	6	23,14035392	27,57759995	10	22,4	27,9	0,59157972	0	0
36,00	7	18,00000000	31,17691451	10	24,2	24,4	0,48915018	0	0
36,00	8	12,31272515	33,82893432	10	10,2	15,3	0,08157658	0	0
36,00	9	6,25133437	35,45307911	10	20,5	28,5	0,565028766	0	0
36,00	10	0,00000000	36,00000000	10	18,2	18,5	0,212098761	0	0
36,00	11	-6,25133437	35,45307911	10	20,0	20,1	0,274883973	0	0
36,00	12	-12,31272515	33,82893432	10	21,3	30,4	0,667573466	0	0
36,00	13	-18,00000000	31,17691451	10	22,7	25,5	0,501090981	0	0
36,00	14	-23,14035392	27,57759995	10	20,7	22,9	0,368909075	0	0
36,00	15	-27,57759995	23,14035392	10	24,4	27,1	0,607916122	0	0
36,00	16	-31,17691451	18,00000000	10	16,4	18,1	0,183041383	0	0
36,00	17	-33,82893432	12,31272515	10	21,4	26,9	0,525600507	0	0
36,00	18	-35,45307911	6,25133437	10	26,0	30,1	0,798372179	0	0
36,00	19	-36,00000000	0,00000000	10	21,2	25,0	0,449977242	0	0
36,00	20	-35,45307911	-6,25133437	10	23,7	27,7	0,616880133	0	0
36,00	21	-33,82893432	-12,31272515	10	20,0	22,1	0,332087974	0	0
36,00	22	-31,17691451	-18,00000000	10	22,0	27,1	0,548319992	0	0
36,00	23	-27,57759995	-23,14035392	10	21,0	25,5	0,463690802	0	0
36,00	24	-23,14035392	-27,57759995	10	18,0	21,6	0,285659451	0	0
36,00	25	-18,00000000	-31,17691451	10	22,6	23,2	0,413227474	0	0
36,00	26	-12,31272515	-33,82893432	10	19,4	26,1	0,448808823	0	0
36,00	27	-6,25133437	-35,45307911	10	24,1	28,5	0,66387687	0	0
36,00	28	0,00000000	-36,00000000	10	25,3	31,7	0,86143399	0	0
36,00	29	6,25133437	-35,45307911	10	21,9	27,9	0,578420567	0	0
36,00	30	12,31272515	-33,82893432	10	20,9	24,4	0,422664949	0	0
36,00	31	18,00000000	-31,17691451	10	20,1	25,5	0,443886464	0	0
36,00	32	23,14035392	-27,57759995	10	20,7	26,4	0,489805158	0	0
36,00	33	27,57759995	-23,14035392	10	22,4	25,1	0,479152981	0	0
36,00	34	31,17691451	-18,00000000	10	22,2	25,9	0,505532863	0	0
36,00	35	33,82893432	-12,31272515	10	21,4	26,1	0,494907593	0	0
36,00	36	35,45307911	-6,25133437	10	19,2	27,7	0,500119855	1	5

ANEXO B - Programa R para Modelo I.

```
#####
#                               PROGRAMA - MODELO I                               #
#####

## LENDO O ARQUIVO DE DADOS:
  a <- read.table ("Anova6M1.txt", header=T)

## ESPECIFICANDO TRATAMENTO
  trat <- as.factor (a$Trat)

## VALOR DO PARÂMETRO ESTIMADO PARA A TRANSFORMAÇÃO BOX-COX:
  require(MASS)
  box <- boxcox (a$Volume~trat)
  box <- boxcox(a$Volume~trat, lam=seq (0,0.9, l=50))
  lambda <- box$x[which(box$y==max(box$y))]
  lambda
  Volumet <- sqrt(a$Volume)
  d <- aov(Volumet ~ trat)

## GEOESTATÍSTICA
  require(geoR)
  res <- residuals(d)

## ESCRIVENDO OS RESÍDUOS COMO DADOS GEOESTATÍSTICOS
  res.coord <- data.frame(a$X1,a$X2,res)
  resigeo <- as.geodata(res.coord,coords.col=1:2,data.col=3)
  X<-model.matrix(d)
```

ANEXO B - Programa R para Modelo I.

```
## SEMIVARIOGRAMA EMPÍRICO:
  dist.eu <- max(dist(resigeo$coords))
  varemp <- variog(resigeo,coords=resigeo$coords,
  resigeo$data,max.dist=(dist.eu)*0.75,
  messages=FALSE)
## SEMIVARIOGRAMAS TEÓRICOS:
  varteo <- variofit(varemp,ini.cov.pars=c(0.06,10),
  nug=0.01,cov.model="exp",
  limits=pars.limits(phi=c(0, dist.eu)),
  messages=FALSE)
  varteo1 <- variofit(varemp,ini.cov.pars=c(0.06,10),
  nug=0.01,cov.model="mat",kappa=1,
  limits=pars.limits(phi=c(0, dist.eu)),
  messages=FALSE)
  varteo2 <- variofit(varemp,ini.cov.pars=c(0.06,10),
  nug=0.01,cov.model="mat",kappa=2,
  limits=pars.limits(phi=c(0, dist.eu)),
  messages=FALSE)
## ENVELOPE SIMULADO:
  ano.v.mc <- variog.mc.env(resigeo, obj=varemp)
```

ANEXO B - Programa R para Modelo I.

```

## GRÁFICO DE AJUSTE DOS SEMIVARIOGRAMAS TEÓRICOS AO
SEMIVARIOGRAMA EMPÍRICO E ENVELOPE SIMULADO:

  plot(varemp, main="", ylab=" ", xlab=" ", cex.axis=1.2,cex=1)
  lines(varteo,col="green",lwd=1.8)
  lines(varteo1,col="blue",lwd=1.8)
  lines(varteo2, col="red",lwd=1.8)
  lines(ano.v.mc,col=25,lwd=1.8)
  legend(10,0.04, legend=c("Exponencial", "Matérn(k=1)",
    "Matérn(k=2)", "variog.env"),col=c(2,3,4,25),lty=c(1,1,1,1),
    lwd=c(2,2,2,2))
  title(main="Modelo I", xlab= " Distância",
    ylab= "Semivariância",cex.lab=1.2,cex.main=1.3)

## ESTIMAÇÃO DOS PARÂMETROS ESPACIAIS POR MÁXIMA VEROSSIMILHANÇA:

volume.coords <- data.frame(resigeo$coords,Volumet,trat)
volume.geo <- as.geodata(volume.coords,coords.col=1:2,
  data.col=3,covar.col=4)

mv.volume <- likfit(volume.geo,trend=~trat,
  ini.cov=c(0.01,10),nug=0.01,cov.model="exp",
  message=FALSE)

mv.volume1 <- likfit(volume.geo,trend=~trat,
  ini.cov=c(0.01,10),nug=0.01,cov.model="mat",
  kappa=1, message=FALSE)

mv.volume2 <- likfit(volume.geo,trend=~trat,
  ini.cov=c(0.01,10),nug=0.01,cov.model="mat",
  kappa=2, message=FALSE)

```

ANEXO B - Programa R para Modelo I.

```
## CRITÉRIO DE AKAIKE PARA SELEÇÃO DE MODELOS:
```

```
  compara <- cbind
    (c(mv.volume$log, mv.volume1$log, mv.volume2$log),
     c(mv.volume$npars, mv.volume1$npars, mv.volume2$npars),
     c(mv.volume$AIC, mv.volume1$AIC, mv.volume2$AIC))
  dimnames(compara) <- list(c("ML", "ML1", "ML2"),
    c("logL", "npars", "AIC"))

  compara
```


ANEXO C - Programa R para Modelo II.

```
#####
#                               PROGRAMA - MODELO II                               #
#####

## LENDO O ARQUIVO DE DADOS:
      c <- read.table ("Cov6M2.txt",header=T)

## ESPECIFICANDO TRATAMENTO E COVARIÁVEL MORTE DA PLANTA:
      Trat <- as.factor(c$Trat)
      Perdida <- c$Perdida

## VALOR DO PARÂMETRO ESTIMADO PARA A TRANSFORMAÇÃO BOX-COX:
      require(MASS)
      box <- boxcox (c$Volume~Trat+Perdida)
      box <- boxcox(c$Volume~Trat+Perdida, lam=seq(0.3,0.9,l=21))
      lambda <- box$x[which(box$y==max(box$y))]
      lambda
      Volumet <- sqrt(c$Volume)
      ct <- aov(Volumet ~ Trat + Perdida)

## GEOESTATÍSTICA:
      require(geoR)
      res <- residuals(ct)

## ESCRIVENDO OS RESÍDUOS COMO DADOS GEOESTATÍSTICOS:
      res.coord <- data.frame(c$X1,c$X2,res,Trat,Perdida)
      resigeo <- as.geodata(res.coord,coords.col=1:2,
                           data.col=3, covar.col=4:5)
      X <- model.matrix(ct)
```

ANEXO C - Programa R para Modelo II.

SEMIVARIOGRAMA EMPÍRICO:

```
dist.eu <- max(dist(resigeo$coords))
varemp <- variog(resigeo,coords=resigeo$coords,
                resigeo$data, max.dist=(dist.eu)*0.75,
                messages=FALSE)
```

SEMIVARIOGRAMAS TEÓRICOS:

```
varteo <- variofit(varemp,ini.cov.pars=c(0.06,10),
                  nug=0.01,cov.model="exp",
                  limits=pars.limits(phi=c(0,dist.eu)),
                  messages=FALSE)
varteo1 <- variofit(varemp,ini.cov.pars=c(0.06,10),
                   nug=0.01,cov.model="mat",kappa=1,
                   limits=pars.limits(phi=c(0,dist.eu)),
                   messages=FALSE)
varteo2 <- variofit(varemp,ini.cov.pars=c(0.06,10),
                   nug=0.01,cov.model="mat",kappa=2,
                   limits=pars.limits(phi=c(0,dist.eu)),
                   messages=FALSE)
```

ENVELOPE SIMULADO:

```
ano.v.mc <- variog.mc.env(resigeo, obj=varemp)
```

ANEXO C - Programa R para Modelo II.

```
## GRÁFICO DE AJUSTE DOS SEMIVARIOGRAMAS TEÓRICOS AO
```

```
SEMIVARIOGRAMA EMPÍRICO E ENVELOPE SIMULADO:
```

```
plot(varemp, main="", ylab=" ", xlab=" ", cex.axis=1.2,cex=1)
lines(varteo,col=2,lwd=1.8)
lines(varteo1,col=3,lwd=1.8)
lines(varteo2, col=4,lwd=1.8)
lines(ano.v.mc,col=25,lwd=1.8)
legend(10,0.04, legend=c("Exponencial", "Matérn(k=1)",
"Matérn(k=2)", "variog.env"), col=c(2,3,4,25), lty=c(1,1,1,1),
lwd=c(2,2,2,2))
title(main="Modelo II", xlab= " Distância",
ylab="Semivariância", cex.lab=1.2,cex.main=1.3)
```

```
## ESTIMAÇÃO DOS PARÂMETROS ESPACIAIS POR MÁXIMA VEROSSIMILHANÇA:
```

```
volume.coords <- data.frame(resigeo$coords,Volumet,
Perdida,Trat)
volume.geo <- as.geodata(volume.coords,coords.col=1:2,
data.col=3, covar.col=4:5)
mv.volume <- likfit(volume.geo,trend=~Trat+Perdida,
ini.cov=c(0.01,10), nug=0.01, cov.model="exp",
message=FALSE)
mv.volume1 <- likfit(volume.geo,trend=~Trat+Perdida,
ini.cov=c(0.01,10),nug=0.01,cov.model="mat",
kappa=1, message=FALSE)
mv.volume2 <- likfit(volume.geo,trend=~Trat+Perdida,
ini.cov=c(0.01,10),nug=0.01,cov.model="mat",
kappa=2, message=FALSE)
```


ANEXO D - Programa R para Modelo III.

```
#####
#                               PROGRAMA - MODELO III                               #
#####

## LENDO O ARQUIVO DE DADOS:
      c <- read.table ("Cov6M3.txt", header=T)
## ESPECIFICANDO TRATAMENTO E COVARIÁVEIS MORTE DA
PLANTA E TEMPO:
      Trat <- as.factor(c$Trat)
      Perdida<-c$Perdida
      Tempo <- c$Tempo
## VALOR DO PARÂMETRO ESTIMADO PARA TRANSFORMAÇÃO BOX-COX:
      require(MASS)
      box <- boxcox (c$Volume~Trat + Perdida + Tempo )
      box <- boxcox(c$Volume~Trat+Perdida + Tempo,
                    lam=seq (0.3,0.9, l=21))
      lambda <- box$x[which(box$y==max(box$y))]
      lambda
      Volumet <- sqrt(c$Volume)
      ct <- aov(Volumet ~ Trat + Perdida + Tempo)
## GEOESTATÍSTICA
      require(geoR)
      res <- residuals(ct)
## ESCREVENDO OS RESÍDUOS COMO DADOS GEOESTATÍSTICOS:
      res.coord <- data.frame(c$X1,c$X2,res,Trat,Perdida,Tempo)
      resigeo <- as.geodata(res.coord,coords.col=1:2,
                            data.col=3,covar.col=4:6)
      X <- model.matrix(ct)
```

ANEXO D - Programa R para Modelo III.

```
## SEMIVARIOGRAMA EMPÍRICO:
```

```
dist.eu <- max(dist(resigeo$coords))
varemp <- variog(resigeo,coords=resigeo$coords,
                resigeo$data,max.dist=(dist.eu)*0.75,
                messages=FALSE)
## SEMIVARIOGRAMAS TEÓRICOS:
varteo <- variofit(varemp,ini.cov.pars=c(0.06,10),
                  nug=0.01,cov.model="exp",
                  limits=pars.limits(phi=c(0, dist.eu)),
                  messages=FALSE)
varteo1 <- variofit(varemp,ini.cov.pars=c(0.06,10),
                   nug=0.01,cov.model="mat",kappa=1,
                   limits=pars.limits(phi=c(0, dist.eu)),
                   messages=FALSE)
varteo2 <- variofit(varemp,ini.cov.pars=c(0.06,10),
                   nug=0.01,cov.model="mat",kappa=2,
                   limits=pars.limits(phi=c(0, dist.eu)),
                   messages=FALSE)
## ENVELOPE SIMULADO:
ano.v.mc <- variog.mc.env(resigeo, obj=varemp)
```

ANEXO D - Programa R para Modelo III.

```
## GRÁFICO DE AJUSTE DOS SEMIVARIOGRAMAS TEÓRICOS AO
SEMIVARIOGRAMA EMPÍRICO E ENVELOPE SIMULADO:
```

```
plot(varemp, main="", ylab=" ", xlab=" ", cex.axis=1.2,cex=1)
lines(varteo,col=2,lwd=1.8)
lines(varteo1,col=3,lwd=1.8)
lines(varteo2, col=4,lwd=1.8)
lines(ano.v.mc,col=25,lwd=1.8)
legend(10,0.04, legend=c("Exponencial", "Matérn(k=1)",
"Matérn(k=2)", "variog.env"),col=c(2,3,4,25),lty=c(1,1,1,1),
lwd=c(2,2,2,2))
title(main="Modelo III", xlab= " Distância",
ylab= "Semivariância",cex.lab=1.2,cex.main=1.3)
```

```
## ESTIMAÇÃO DOS PARÂMETROS POR MÁXIMA VEROSSIMILHANÇA:
```

```
volume.coords <- data.frame(resigeo$coords,Volumet,
Perdida,Tempo,Trat)
volume.geo <- as.geodata(volume.coords,coords.col=1:2,
data.col=3,covar.col=4:5)
mv.volume <- likfit(volume.geo,trend=~Trat-1+Perdida + Tempo,
ini.cov=c(0.01,10),nug=0.01,cov.model="exp",
message=FALSE)
mv.volume1 <- likfit(volume.geo,trend=~Trat-1+Perdida + Tempo,
ini.cov=c(0.01,10),nug=0.01,cov.model="mat",kappa=1,
message=FALSE)
mv.volume2 <- likfit(volume.geo,trend=~Trat-1+Perdida + Tempo,
ini.cov=c(0.01,10),nug=0.01,cov.model="mat",kappa=2,
message=FALSE)
```

ANEXO D - Programa R para Modelo III.

```
## CRITÉRIO DE AKAIKE PARA SELEÇÃO DE MODELOS:
```

```
  compara <- cbind
            (c(mv.volume$log, mv.volume1$log, mv.volume2$log),
             c(mv.volume$npars, mv.volume1$npars, mv.volume2$npars),
             c(mv.volume$AIC, mv.volume1$AIC, mv.volume2$AIC))
  dimnames(compara) <- list
            (c("ML", "ML1", "ML2"),
             c("logL", "npars", "AIC"))

  compara
```


ANEXO E - Análise de Resíduos para os modelos I, II e III.

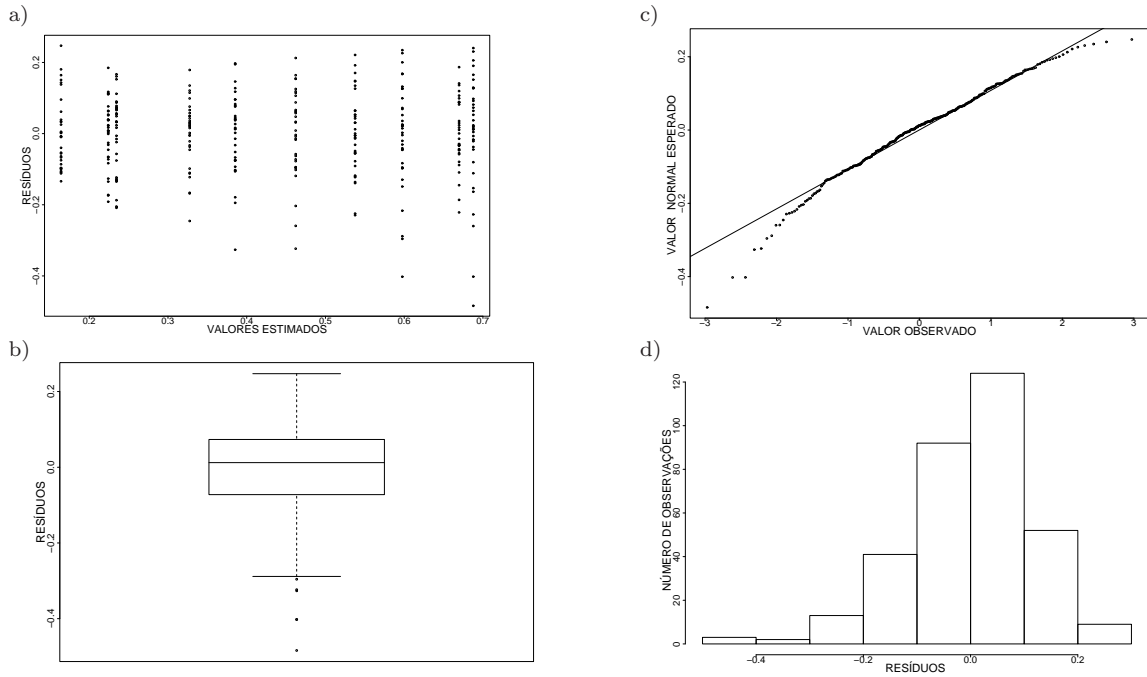


Figura 9 - Resíduos vs valores observados (a), “box-plot” (b), gráfico de probabilidade normal (c) e distribuição de freqüências (d) para o modelo I

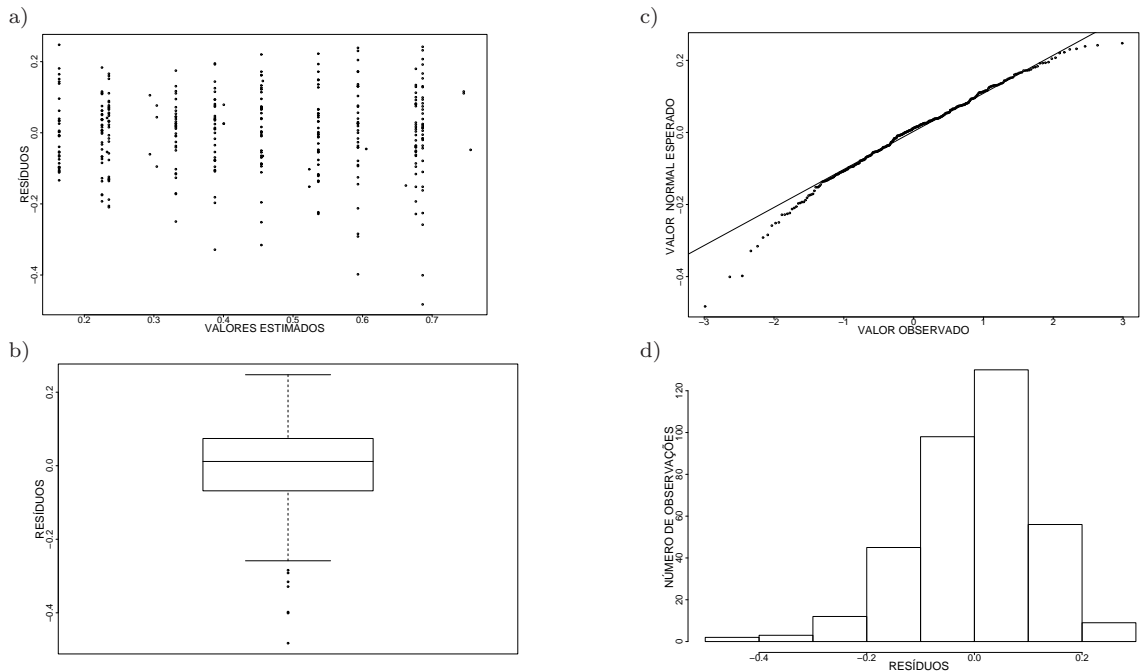


Figura 10 - Resíduos vs valores observados (a), “box-plot” (b), gráfico de probabilidade normal (c) e distribuição de freqüências (d) para o modelo II

ANEXO E - Análise de Resíduos para os modelos I, II e III.

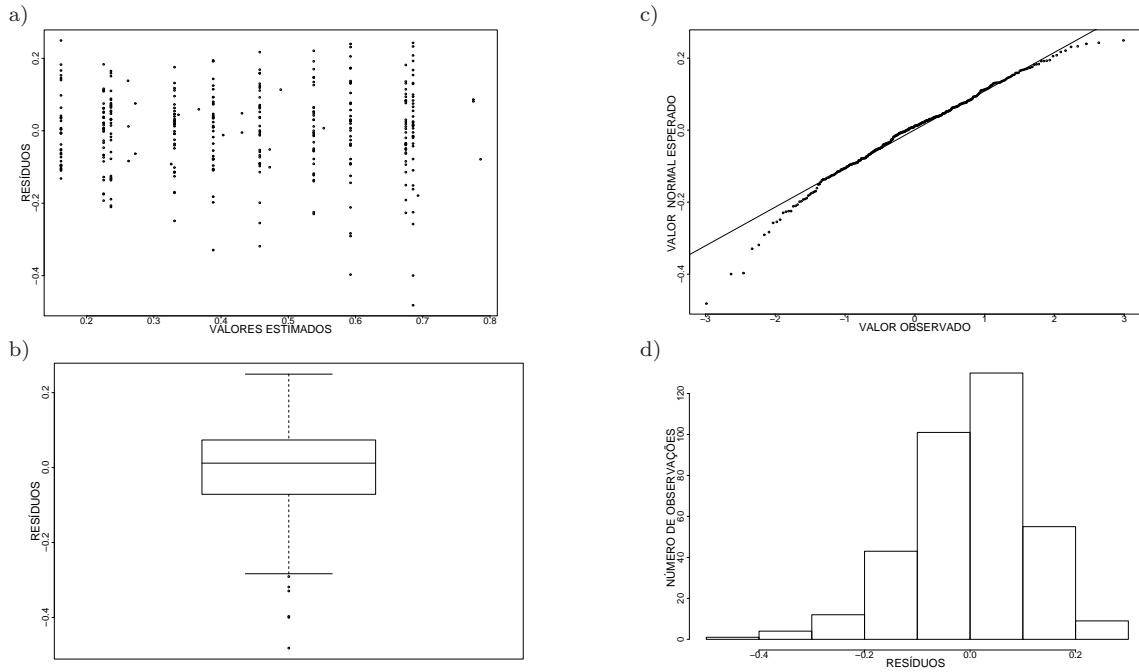


Figura 11 - Resíduos vs valores observados (a), “box-plot” (b), gráfico de probabilidade normal (c) e distribuição de freqüências (d) para o modelo III

REFERÊNCIAS BIBLIOGRÁFICAS

- ANJOS, A. **Notas de aula - Planejamento de experimentos II**. Curitiba, 2004. 91p.
- BARBIN, D. **Planejamento e Análise Estatística de Experimentos Agronômicos**. Arapongas: Midas, 2003, 194p.
- BAR-HEN A. Influence of missing data on compact designs for spacing experiments **Journal of Applied Statistics**, v.29, n.8, p.1229-1240, 2002.
- BAUDER, J.W.; HANKS, R.J.; JAMES, D.W. Crop production function determinations as influenced by irrigation and nitrogen fertilization using a continuous variable design. **Soil Science Society of America Proceedings**, n.39: p.1187-92, 1975.
- BLEASDALE, J.K.A. Systematic designs for spacing experiments. **Experimental Agriculture**, n.3: p.73-85, 1966.
- BRUS, D.J.; De GRUIJTER, J.J. Design-based versus model-based estimates of mean spacial - theory and application enviromental soil science. **Envirometrics**, v.4, n.2, p.123-152, 1993.
- BRUS, D.J. Using regression models in desig-based estimation of spacial means of soil properties. **European Journal of Soil Science**, v.51, n.1, p.159-172, 2000.
- CÂMARA, G.; MONTEIRO, A.M.; FUKS, S.D.; FELGUEIRAS, C. **Análise espacial de dados geográficos**. 2.ed. São José dos Campos: INPE, 2001.
- CHALITA, M.A.C. Delineamentos sistemáticos. Piracicaba, 1991. 72p. Dissertação (Mestrado) - Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo.
- CLEAVER, T. J; GREENWOOD; WOOD, J. T. Systematically arranged fertilizer experiments. **Journal of Horticultural Science**, n.45: p.457-69, 1970.

- COX, D.R. **Planning of experiments**. New York, John Wiley: 1958. 308p.
- CRESSIE, N.; HARTFIELD; M.N. Conditionally specified gaussian model for spacial statistical analysis of field trials. **Journal of Agricultural, Biological, and Environmental Statistics**, v.1, n.1, p.60-77, 1996.
- DAWKINS, H.C. The pudden clinal plot: thinning experiments without surrounds. **Empire Forestry Review**, n.39, p.168-171, 1960.
- DEMÉTRIO, C.G.B. Transformação de dados. Efeitos sobre a análise de variância. Piracicaba, 1978. 113p. Dissertação (Mestrado) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo.
- DEMÉTRIO, C.G.B.; ZOCCHI, S.S. Modelos de regressão e covariância. **Notas de aula: Regressão e Covariância**, Piracicaba, 2003. 229p.
- DIGGLE, P.J.; RIBEIRO JÚNIOR, P.J. **Model based geostatistics**. Caxambu, ABE, 2000. 129p.
- FISHER, R. **The design of experiments**. Edinburg: Oliver and Boyd, 1935. 252p.
- FOX, R.L. Agronomic investigations using continuous function experimental designs - nitrogen fertilization of sweet corn. **Agronomy Journal**, n.65: p.454-456, 1973.
- FREEMAN, G.H. The use of systematic design for a spacing trial with a tropical tree crop. **Biometrics**, n.20, p.200-203, 1964.
- FREEMAN, G.H. Systematic design. In: KOTS, S. JOHNSON, N.L.; READ, C.B. (Ed.). **Encyclopedia of Statistical Sciences**. New York: John Wiley, v.9, p.143-147, 1988.
- FREYMAN, S.; DOLMAN, D. A simple systematic design for planting density experiments with set row widths. **Canadian Journal of Plant Science**, n.51, p.340-342, 1971.
- GREENBERG, B.G. "Why Randomize?". **Biometrics**, v.7, n.3, p.309-322, 1951.
- GUERRA, P. A. G. **Geoestatística operacional**. Brasília: MME/DNPM, 1988. 145p.
- HUXLEY, P.A. Systematic designs for field experimentation with multipurpose trees. **Agroforestry Systems**, n.3, p.197-207, 1985.

- HUXLEY, P.A.; MAINGU, Z. Use of a systematic spacing design as an aid to the study of inter-cropping: some general considerations. **Experimental Agricultural**, n.14, p.49-56, 1978.
- KEMPTHORNE, O. "Why Randomize?". **Journal of Statistical Planning and Inference**, v.1, p.1-25, 1977.
- LANDIM, P. M. B. **Análise estatística de dados geológicos**. São Paulo, Ed. UNESP, 1998. 226p.
- LEEUEWEN, J.V. Planejamento de ensaios com sistemas agroflorestais. In: Anais, In Congresso Brasileiro sobre Sistemas Agroflorestais; Encontro sobre Sistemais Agroflorestais nos Países do Mercosul, Porto Velho, 1994. Colombo: EMBRAPA, PR, v.1, p.463-473.
- LUTWICK, L.E.;KOZUB, G.C.; SMITH, A.D. A two-factor systematic design for fertilizer studies. **Canadian Journal of Soil Science** , n.60, p.657-663, 1980.
- McBRATNEY, A.B.; WEBSTER, R. Choosing functions for semi-variograms of soil properties and fitting them to sampling estimates. **Journal of Soil Science**, v.77, p.617-639, 1986.
- MEAD, R. **The design of experiment**: statistical principles for practical application studies. Cambridge, University Press, 1988, 620p.
- MELLO, J.M. Geoestatística aplicada ao inventário florestal. Piracicaba, 2004. 111p. Tese (Doutorado) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo.
- MINAMI, K. Análise de crescimento e densidade de população de *Solanum melonngera* L. - Berinjela, cultivada em delineamento sistemático e convencional. Piracicaba, 1977. 81p. Dissertação (Mestrado) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo.
- MONTAGNA, M.A. Distribuição espacial e amostragem seqüencial da mosca branca *Bemisia tabaci* Raça B (HOMOPTERA: ALEYRODIDAE) no agroecossistema do melão. Ribeirão Preto, 2001. 83p. Dissertação (Mestrado) - Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo.

- MOTA, J.M.A. Cadeias de Markov. **Inferência Bayesiana**, Piracicaba, 2002, 35p. (Notas de aula)
- NELDER, J.A. New kinds of systematic designs for spacing experiments. **Biometrics**, n.18: p.283-307, 1962.
- OLIVEIRA, M.C.N. Métodos de estimação de parâmetros em modelos geoestatísticos com diferentes estruturas de covariância: uma aplicação ao teor de cálcio no solo. Piracicaba, 2003. 140p. Tese (Doutorado) - Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo.
- ORTIZ, G.C. Aplicação de métodos geoestatístico para identificar a magnitude e a estrutura da variabilidade espacial de variáveis físicas do solo. Piracicaba, 2002. 75p. Dissertação (Mestrado) - Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo.
- PEARCE, S.C. **The agriculture field experiments**: a statistical examination of theory and practices. Chichester, John Willey, 1989, 335p.
- PILON, A. A. Métodos para incorporação da dependência espacial em análise de dados. Piracicaba, 2004. 137p. Dissertação (Mestrado) - Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo.
- PIMENTEL GOMES, F.P. **Curso de Estatística Experimental**. Piracicaba, 14Ed. 2000, 477p.
- PONTES, J.M. A geoestatística: aplicações em experimentos de campo. Lavras, 2002, 82p. Dissertação (Mestrado) - Universidade Federal de Lavras.
- RAO, M.R.; SHARMA, M.M.; ONG, C.K. A study of the potencial of hedgerow intercropping in semi-arid India using a two-way systematic design. **Agroforestry Systems**, n.11, p.243-258, 1990.
- RIBEIRO JÚNIOR, P.J. Métodos geoestatísticos no estudo da variabilidade espacial de parâmetros do solo. Piracicaba, 1995. 99p. Dissertação (Mestrado) - Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo.
- RIBEIRO JÚNIOR, P.J.; DIGGLE, P.J. **The geoR package functions for geostatistical data analysis**: . 2004, 119p.

- RIBEIRO JÚNIOR, P.J.; CHRISTENSEN, O.F.; DIGGLE P.J. Geoestatistical software - geoR and geoRglm. <http://www.ci.tuwien.ac.at/Conferences/DSC-2003/>,15p.
- ROGERS, I.S. Practical considerations in the use of systematic spacing designs. **Australian Journal of Experimental Agriculture and Animal Husbandry**, n.12, p.306-9, 1972.
- SCOOT, J.F. Statistical objections to clinal plots. **Empire Forestry Review**, Oxford, n.41, p.17-18, 1962.
- STAPE, J.L. Utilização de delineamento sistemático tipo “leque” no estudo de espaçamentos florestais. Piracicaba, 1995. 86p. Dissertação (Mestrado) - Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo.
- STEVENS, W.L. Análise estatística do ensaio de variedades de café. *Bragantia*, n.9: p.103-123, 1949.
- TEIXEIRA, F. das C. Análise estatística espacial aplicada à um ensaio florestal. Lavras, 2001. 77p. Dissertação (Mestrado) - Universidade Federal de Lavras.
- VALENTE, J. M. G. P. **Geomatemática – Lições de geoestatística**. Ouro Preto: Ed. da Fundação Gorceix, v.3, p.511-713, 1989.
- VIEIRA, S.R. **Geoestatística**. Curso de extensão universitária. Botucatu: UNESP, 1996, 71p.
- VIEIRA, S.R. **Geoestatística em estudos de variabilidade espacial do solo**. Tópicos em ciência do solo. Viçosa: Sociedade Brasileira de Ciência do Solo, 2000.
- WILLEY, R. W.; RAO, M.R. A systematic design to examine effects of plant population and spatial arrangement in intercropping, illustrated by an experiment on chickpea/safflower. **Experimental Agriculture**, n.17: p.68-73, 1981.
- WRIGHT, H. L. Experiments as a source of data for growth models. In: IUFRO WORLD CONGRESS, 16, Oslo, **Proceedings**, Oslo: IUFRO, 1976 v.1, p.60-73.
- YATES, F. - The comparative advantages of systematic and randomized arrangements in the design of agricultural and biological experiments. **Biometrika**, n.30, p.440-466, 1938.
- YATES, F. - Sir Ronald Fisher and the design of experiments. **Biometrics**, n.2, p.307-321, 1964.

ZIMBACK, C.R.L. **Geoestatística**. Curso de extensão universitária. Botucatu: UNESP, 2003. 25p.