

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE AGRONOMIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ZOOTECNIA**

**A METODOLOGÍA DE ANÁLISE DE DADOS ESPACIAIS
APLICADA A PLANTAS FORRAGEIRAS**

Mónica Graciela Cadenazzi Pascual
Eng. Agr. (UDELAR, ROU)

Tese apresentada como um dos requisitos ao
Grau de Doutor em Zootecnia,
Área de concentração Plantas Forrageiras

Porto Alegre (RS), Brasil
Agosto, 2000

AGRADECIMENTOS

Ao Prof. Dr. João Riboldi, por todos os ensinamentos, a ajuda e amizade demonstrada nestes anos.

Ao Prof. Ph.D. R.T. Clark, por sua enorme disponibilidade para ensinar e auxiliar em todo momento.

Ao Prof. Ph.D. Valério Pillar, pelo contínuo ensinamento de Ecologia Quantitativa.

À Profa. Dra Ilsi Boldrini, pela amizade destes anos.

A os Professores do Departamento de Forrageiras e de Estatística da UFRGS, e pelo convívio acadêmico de tantos anos.

A CNPq pelo fundamental auxílio recebido.

À `Universidad de la República, Facultad de Agronomía`, por permitir meu afastamento.

A meus colegas do Departamento de Plantas Forrageiras da UFRGS. Foram muitos os amigos destes anos, tantos que não posso mais que representá-los em dois colegas de muitos cursos e debates técnicos: a Lila e o Fernando.

A minha amiga Stella, quem me ensinou matemáticas sempre que foi preciso e a minha amiga Lara com sua grande hospitalidade nestes dias.

A minha amiga Dinara, quem sempre esteve presente com uma palavra de estímulo, e com sua grande hospitalidade de sempre.

A meus amigos da EEMAC, que fizeram fácil a minha partida e cálido o meu retorno.

A meus primeiros Professores de Estatística, Mónica Beltrami, Daniel Suazes e Orual Andina, amigos que influenciaram que esta disciplina se tornara tão importante na minha vida.

A meus sogros Romulo e Amelia, que entenderam a minha vocação.

A meus pais Carlos e Teresita, e a meus irmãos Laura e Carlitos, que desde longe sempre me alentaram.

A meu marido Pablo, quem sempre foi meu ponto de referência espacial, e quem sempre torce por meus logros.

Especialmente a meus filhos Pablo e Luciano, que sempre entenderam que a falta de tempo não é falta de amor.

A METODOLOGIA DE ANÁLISE DE DADOS ESPACIAIS APLICADA A PLANTAS FORRAGEIRAS

Autora: Mónica Graciela Cadenazzi Pascual
Orientador: Prof. Dr. João Riboldi

RESUMO

Dois experimentos e um levantamento por amostragem foram analisados no contexto de dados espaciais. Os experimentos foram delineados em blocos completos casualizados sendo que no experimento um (EXP 1) foram avaliados oito cultivares de trevo branco, sendo estudadas as variáveis Matéria Seca Total (MST) e Matéria Seca de Gramíneas (MSGRAM) e no experimento dois (EXP 2) 20 cultivares de espécies forrageiras, onde foi estudada a variável Percentagem de Implantação (%IMPL). As variáveis foram analisadas no contexto de modelos mistos, sendo modelada a variabilidade espacial através de semivariogramas exponenciais, esféricos e gaussianos. Verificou-se uma diminuição em média de 19% e 14% do Coeficiente de Variação (CV) das médias dos cultivares, e uma diminuição em média de 24,6% e 33,3% nos erros padrões dos contrastes ortogonais propostos em MST e MSGRAM. No levantamento por amostragem, estudou-se a associação espacial em *Aristida laevis* (Nees) Kunth, *Paspalum notatum* Fl e *Desmodium incanum* DC, amostrados em uma transecção fixa de quadros contíguos, a quatro tamanhos de unidades amostrais (0,1x0,1m; 0,1x0,3m; 0,1x0,5m; e 0,1x1,0m). Nas espécies *Aristida laevis* (Nees) Kunth e *Paspalum notatum* Fl, existiu um bom ajuste dos semivariogramas a tamanhos menores das unidades amostrais, diminuindo quando a unidade amostral foi maior. *Desmodium incanum* DC apresentou comportamento contrário, ajustando melhor os semivariogramas a tamanhos maiores das unidades amostrais.

A METHODOLOGY OF SPATIAL DATA ANALYSIS APPLIED ON FORAGE PLANT¹

Author: Mónica Graciela Cadenazzi Pascual
Adviser: Prof. Dr. João Riboldi

SUMMARY

Two experiments and one sample survey were analysed used methods for spatial statistical analysis. The experiments were in randomized blocks; in the first (EXP 1), measurement of total dry matter (MST) and dry weight of seeds (MSGRAM) were recorded for nine cultivars of white clover. In the second experiment (EXP 2), percentage of take (%IMPL) was recorded for 20 cultivars of forage plant species. Variables were analysed in the context of mixed models, with spatial variability modeled in terms of exponential, spherical and Gaussian semivariograms. A mean reduction of 19% and 13.7% was found for the Coefficient of Variation (CV) for the cultivar means, and a mean reduction of 24.6% and 33.3% was found for the standard errors of orthogonal contrast for MST and MSGRAM. In the sample survey, the spatial association was studied for *Aristida laevis*, *Paspalum notatum* and *Desmodium incanum*, sampled on a fixed transect for adjacent plots, for four sizes of sampling units (0.1 x 0.1m; 0.1 x 0.3 m; 0.1 x 0.5 m and 0.1 x 1 m). For the species *Aristida laevis* and *Paspalum notatum*, semivariograms fitted well for the smaller sampling units, but goodness of fit declined with increasing size of sampling unit. The opposite was found for *Desmodium incanum*, where semivariogram fit was better for larger sampling units.

¹ Doctoral thesis in Forrage Science, Facultad de Agronomía da Universidade Federal do Rio Grande do Sul, Porto Alegre, Brasil, (100p) – September, 2000.

SUMÁRIO

Página
1. INTRODUÇÃO.....	1
2. REVISÃO BIBLIOGRÁFICA.....	4
2.1. Um conceito intuitivo de dados espaciais.....	4
2.2 Modelos de amostragem: seu uso na área biológica.....	6
2.3 Os modelos de planejamento usados na área biológica.....	13
2.4. Análise de Dados Espaciais.....	17
2.4.1 A natureza dos dados espaciais.....	17
2.4.2 Semivariogramas.....	17
2.4.3. Isotropia e Anisotropia.....	21
2.4.4. Os modelos matemáticos.....	22
2.4.4.1. Modelo simples.....	22
2.4.4.2 Modelo com partição de variações.....	23
2.4.4.3.Modelo com refinamento pela mediana.....	24
2.4.4.4.Modelo de planejamento de experimentos.....	24
2.4.5 As estruturas de covariância.....	25
2.4.5.1 Autorregresiva.....	25
2.4.5.2 Espacial.....	26
2.4 Tendência atual.....	26
3. MATERIAL.....	28
3.1 Experimento 1.....	28
3.2 Experimento 2.....	30
3.3. Experimento 3.....	32

	Página
4 METODO.....	30
4.1 Variabilidade espacial.....	33
4.1.1 Modelagem probabilística e estacionaridade: a importância da análise descritiva.....	33
4.1.2 Estrutura de variabilidade: Semivariogramas.....	35
4.2 Análise de dados espaciais em planejamento de experimentos.....	36
4.2.1 Modelo Lineal Geral.....	36
4.2.2 Modelo Misto.....	38
4.2.3. Modelos mistos para dados espaciais.....	41
4.2.4. Estruturas de Covariância.....	43
4.2.5. Estimação e Seleção da Estrutura de Covariância....	44
4.3 Análise de dados espaciais em levantamentos por amostragem.....	45
4.3.1 Interpolação: Krigagem.....	46
4.3.2. Mapas.....	47
5 RESULTADOS E DISCUSSÃO.....	49
5.1 Experimento 1.....	49
5.1.1 Análise descritiva.....	49
5.1.2. Semivariogramas ajustados.....	55
5.1.3. Análise de Variância considerando a variabilidade espacial.....	56
5.2 Experimento 2.....	61
5.2.1 Análise descritiva.....	62
5.2.2. Semivariogramas ajustados.....	65

	Página
5.2.3. Análise de Variância considerando a variabilidade espacial.....	66
5.3 Experimento 3.....	69
6. CONCLUSÕES.....	74
7. REFÊRENCIAS BIBLIOGRÁFICAS.....	76
8. APÊNDICE.....	82

RELAÇÃO DE TABELAS

	Página
1 Definição das distâncias medidas para a construção de índices de homogeneidade, e sua simbologia neste texto.....	9
2 Estatísticas aplicadas no cálculo de vizinhos, suas distribuições assintóticas e referência principal (Modificada de Cressie, 1993).....	10
3 Caracterização das 20 espécies forrageiras utilizadas no experimento 2.....	30
4 Estatísticas descritivas dos resíduos de MST e MSGRAM no segundo ano.....	49
5 Coeficientes dos semivariogramas ajustados na variável MST no segundo ano, e valores dos criterios utilizados na seleção do modelo.....	56
6 Coeficientes dos semivariogramas ajustados na variável MSGRAM no segundo ano, e valores dos criterios utilizados na seleção do modelo.....	56
7 Contrastes ortogonais para cultivares de trevo branco.....	57
8 Resumo por cultivar da variável MST segundo o modelo de estimação empregado. Apresentam valores de médias, desvios estándar, valor da estatística t, nível mínimo de significância e coeficiente de variação.....	58
9 Contrastes da variável MST para os 4 modelos ajustados. Para cada modelo se apresentam as estimativas do contraste, os erros padrões do contraste, valor da estatística t e nível mínimo de significância.....	59
10 Resumo por cultivar da variável MSGRA segundo o modelo de estimação empregado. Apresentam-se médias, erro padrão, valor da estatística t, nível mínimo de significância e coeficiente de variação.....	60
12 Estatísticas descritivas dos resíduos de %IMPL no segundo corte.....	62

	Página
13 Coeficientes dos semivariogramas ajustados na variável %IMPL no segundo corte, e valores dos criterios utilizados na seleção do modelo.....	65
14 Contrastes ortogonais do experimento 2	66
15 Resumo por cultivar da variável %IMPL segundo o modelo de estimação empregado. Apresentam-se médias, erro padrão, valor da estatística t, nível mínimo de significância. e coeficiente de variação...	67
16 Contrastes da variável %IMPL para os 4 modelos ajustados. Para cada modelo apresentam-se as estimativas do contraste, erro padrão do contraste, valor da estatística t e nível mínimo de significância.....	68
17 Resultados do ajuste dos semivariogramas por espécie e por tamanho de unidade amostral. Se apresentam os valores calculados de r^2 e dos coeficientes a (range) e c . (sill).....	69

RELAÇÃO DE FIGURAS

	Página
1 Representação gráfica de um semivariograma teórico.....	18
2 Aleatorização do experimento 1.....	29
3 Aleatorização do experimento 2.....	31
4 Gráfico de probabilidade normal dos resíduos da variável MST no segundo ano.....	50
5 Gráfico de probabilidade normal dos resíduos da variável MSGRAM no segundo ano.....	50
6 Histograma dos resíduos da variável MST no segundo ano.....	51
7 Histograma dos resíduos da variável MSGRAM no segundo ano.....	51
8 “ Box-plot” dos resíduos de MST no segundo ano.....	52
9 “ Box-plot” dos resíduos de MSGRAM no segundo ano.....	52
10 Valores dos resíduos de MST por colunas da superfície amostrada.....	53
11 Valores dos resíduos de MST por linhas da superfície amostrada.....	53
12 Valores dos resíduos de MSGRAM por linhas da superfície amostrada.....	54
13 Valores dos resíduos de MSGRAM por colunas da superfície amostrada.....	54
14 Gráfico de probabilidade normal dos resíduos da variável %IMPL no segundo corte.....	63
15 Histograma dos resíduos da variável %IMPL no segundo ano.....	63
16 “ Box-plot” dos resíduos de %IMPL no segundo ano.....	64
17 Valores dos resíduos de %IMPL por colunas da superfície amostrada.....	64
18 Valores dos resíduos de %IMPL por linhas da superfície amostrada.....	65

	Página
19 Semivariograma exponencial da espécie <i>Paspalum notatum</i> Fl estudada em 400 unidades amostrais de 0,1 x 0,1 m.....	71
20 Semivariograma exponencial da espécie <i>Aristida laevis</i> (Nees) Kunth estudada em 400 unidades amostrais de 0,1 x 0,1 m.....	72
21 Semivariograma exponencial da espécie <i>Desmodium incanum</i> DC estudada em 40 unidades amostrais de 0,1 x 1,0 m.....	73

1. INTRODUÇÃO

Se denomina análise de dados espaciais a um conjunto de técnicas estatísticas que consideram na sua estrutura a referencia espacial dos pontos estudados, provenientes de amostragens ou de delineamentos experimentais clássicos.

A variabilidade espacial observada em muitos fenômenos naturais agronômicos e ecológicos é intensamente conhecida, observando-se muitas vezes no material experimental ou nas áreas sob estudo diversos padrões. O objetivo das estatísticas espaciais é modelar esse padrões. Nestes estudos, as unidades amostrais ou as parcelas experimentais ocupam localidades fixas distribuídas no espaço. Indivíduos, populações e comunidades vegetais, tipos de solos, quantidade de água e insolação, quase nunca tem uma distribuição uniforme ou aleatória, sendo mais comum a existência de agregados (patches), gradientes ou outras estruturas espaciais. Em geral, pode-se observar que unidades amostrais mais próximas no espaço, são mais semelhantes do que unidades mais distantes, determinando a existência de autocorrelação entre unidades amostrais. Esta autocorrelação espacial é importante e influi na estimativa de parâmetros, podendo ser sua caracterização fundamental para definir manejos e estratégias de utilização adequadas, visando obter altos níveis de produtividade na área agropecuária. O estudo da autocorrelação espacial também é considerado importante no contexto de diferentes teorias ecológicas, como as de sucessão, adaptação e estabilidade de comunidades.

Embora o conhecimento da existência dessa variabilidade espacial não seja recente, apenas nos últimos anos este fato é considerado no tratamento estatístico dos dados. Nestes casos, a suposição de independência entre unidades experimentais ou amostrais, comum a vários processos estocásticos não explica satisfatoriamente o comportamento da variável estudada, sendo necessário considerar as correlações espaciais no modelo.

Em experimentos, procura-se eliminar o efeito da autocorrelação das unidades amostrais através da utilização de aleatorização, repetição e controle local, nem sempre com sucesso. O uso efetivo de correções por covariáveis, pode diminuir o problema de autocorrelações entre unidades experimentais.

Em levantamentos por amostragem, no entanto, o controle da autocorrelação entre unidades amostrais é mais complexo.

O delineamento em blocos completos casualizados foi uma das primeiras técnicas de controle local, com o objetivo de eliminar o efeito da autocorrelação entre unidades experimentais. Mas experimentos que necessitam de blocos grandes, ou material experimental heterogêneo numa pequena escala, fazem com que a técnica não seja sempre efetiva. As alternativas utilizadas para remover o efeito da heterogeneidade espacial existente se baseiam em uso de blocos incompletos, análise de covariância, análise de vizinhos mais próximos ou análise espacial onde se modela a estrutura do erro experimental, incluindo as autocorrelações existentes.

Na maioria dos casos, existe um aumento na precisão das estimações de parâmetros quando a variação espacial é considerada no modelo. Assim, por exemplo, na experimentação com parcelas de maior magnitude, a inclusão de

animais em pastejo como fator de estudo, as modificações contínuas na estrutura vegetal (material experimental) pelo pastejo e as implicações que podem ter sobre algumas variáveis do estudo, associado a que em geral estes experimentos são planejados para longa duração no tempo, fazem com que o estudo e a determinação das estruturas de correlação espacial existente entre unidades experimentais seja importante e produza acréscimo na eficiência das estimativas de parâmetros.

No presente trabalho apresentar-se-á os fundamentos da análise de dados espaciais visando à detecção de autocorrelação espacial entre algumas espécies e populações vegetais de campo natural, bem como a avaliação de experimentos na área de melhoramento de plantas forrageiras com os seguintes objetivos:

- * Utilizar a metodologia de análise de dados espaciais para dados de levantamentos por amostragem e de experimentos de plantas forrageiras, verificando a existência de distribuição espacial em espécies e populações; ajustando semivariogramas teóricos à distribuições espaciais encontradas.

- * Avaliar o aumento da eficiência no uso de informação proveniente de experimentos em campo para estimar e testar hipótese sobre parâmetros e contrastes de interesse em experimentos clássicos de plantas forrageiras.

- * Implementar a metodologia da análise de dados espaciais através de aplicativos computacionais disponíveis.

2. REVISÃO BIBLIOGRÁFICA

2.1. Um conceito intuitivo de dados espaciais

Desde a pré-história o homem percebeu a natureza dos dados espaciais. As primeiras manifestações de estatísticas para dados espaciais aparecem na forma de mapas, onde se superpõem sobre o delineamento do terreno, informações de variáveis como ventos, quantidade de chuva, fertilidade, como forma de incrementar a informação. Os modelos espaciais aparecem depois no tempo. Cressie (1993) relata vários trabalhos, em resenha histórica, em que a heterogeneidade espacial foi percebida ou considerada. Casos como o de Student, que em 1907 estudou a distribuição das partículas através de um líquido, agregando os conteúdos por unidade de superfície, e encontrou uma distribuição de Poisson para o número de partículas por unidade de superfície. Fisher (1935) detectou a dependência espacial nos experimentos agrônômicos, e estudou formas de controlá-la, aplicando os princípios de casualização, bloqueamento e repetição. Segundo Yates (1938) o controle do material experimental não é efetivo numa escala diferente à estudada. Papadakis (1994) sugere uma correção pelos vizinhos como forma de incluir na análise a variabilidade espacial. Varias modificações de correção pelos vizinhos continuam sendo desenvolvidas (Besag e Kempton, 1986).

Na década de 50 os dados espaciais foram considerados uma extensão de séries temporais, mais várias características os separam. As características próprias dos dados espaciais são a presença de bordas, a direção não prédefinida do processo, a alta frequência de processos e estimações em longos espaços, a necessidade de definir estacionariedades e as freqüentes amostragens em malhas irregulares.

A presença de bordas que naturalmente existem em dados espaciais, faz que estas tenham um número diferente e inferior de vizinhos que os pontos

centrais. O efeito de borda geralmente domina as distribuições assintóticas que sejam calculadas a partir de medidas de distância ou similaridade. Enquanto a direção do processo é totalmente definida em séries temporais, é de difícil definição em dados espaciais coexistindo muitas vezes mais de uma direção. As relações de dependência em espaços amplos são comuns em dados espaciais, enquanto que as relações de curto prazo são salientes em series temporais. Por último, o estudo de dados espaciais baseia-se na teoria das variáveis regionalizadas pelo que existe a necessidade da suposição de estacionariedade para simular repetições, característica própria deste tipo de dados.

Atualmente a estatística está sendo incorporada em todos os campos de estudo, e muitos dos problemas que surgem são resultantes da heterogeneidade espacial. Assim, técnicas e idéias sobre dados espaciais que surgiram da experimentação agrícola, passam naturalmente muitas outras áreas variadas como geologia, ecologia, medicina e epimediología dentre outras.

2.2 Modelos de amostragem: seu uso na área biológica

A heterogeneidade do material é muito importante do ponto de vista biológico, e para efeitos da análise existem dois enfoques segundo os objetivos: remover a heterogeneidade ou quantificar e incorpora-la aos modelos.

Vários fatores são importantes na quantificação dessa heterogeneidade. A importância da escala da amostragem, do tamanho das unidades amostrais, da necessidade de uso de mais de uma escala em cada estudo e da detecção de padrões de distribuição de espécies e comunidades vegetais, e as medidas ou índices utilizados para avaliar dita heterogeneidade é apresentada por vários autores (Fortin e Gurevitch, 1993; Dutilleul, 1993; Levin, 1992; Legendre

e Fortin, 1989; Podani, 1984; Diggle 1981, 1979; Byth e Ripley 1980; e Perry e Mead, 1979).

Existem diferentes padrões espaciais definidos pela heterogeneidade. Dale e Zbigniewicz (1995) trabalhando com dados de vegetação, definem padrões espaciais em termos gerais como uma não aleatoriedade no arranjo espacial das espécies, que oferece um certo aumento na preditibilidade. O padrão espacial de uma espécie é geralmente estudado utilizando um ou mais métodos que examinam o efeito da distância ou do tamanho do bloco sobre uma variância, com baixos valores de variâncias indicando homogeneidade, e altos valores indicando heterogeneidade. Em vários estudos o objetivo não é o estudo do padrão espacial específico, sendo importante o padrão espacial de uma comunidade vegetal. O conceito de padrão espacial multidimensional de espécies esta baseado na percepção que, em duas dimensões a vegetação é como um mosaico onde se distinguem fases ou tipos de vegetação, definidos por diferentes combinações de densidades específicas ou diferentes combinações de ocorrência de espécies. Cada fase pode não ser totalmente homogênea mais existe alta similaridade dentro de fase e baixa entre fases. Este conceito esta relacionado com dependência espacial onde amostras de vegetação mais próximas tendem a ser mas semelhantes. Esta correlação espacial aparece como uma autocorrelacao nos dados. Os autores propõem uma análise de ordenação multiescalar e estudos da dimensão fractal da vegetação como forma de estudar o padrão espacial multiespecifico, e salientam a necessidade de utilizar várias técnicas de análise

simultaneamente, dado que cada uma delas mostra vantagens em pontos diferentes.

Os conceitos básicos de processos espaciais na análise de vegetação são também revisados por Podani (1984), que enfatiza as mudanças no espaço e no tempo, e a influência da escala de amostragem sobre os resultados obtidos. O autor define como distribuição espacial aquelas que ocorrem no espaço real representado pelas unidades amostrais mas também aquelas que ocorrem no espaço abstrato ou numérico, sendo estas últimas resultado do tipo de análise realizado.

A dependência na determinação de padrões e associações espaciais da amostragem realizada também é referida por Legendre (1993) e Legendre e Fortin (1989). Os autores apresentam várias metodologias para remover e quantificar a autocorrelação espacial e concluem que a consideração dos processos espaciais e temporais nos problemas ecológicos será um passo importante no desenvolvimento da pesquisa. Levin (1992) acrescenta que é fundamental conhecer os processos espaciais e temporais em padrões de vegetação do ponto de vista da dinâmica de populações e ecossistemas em várias escalas de estudo, enquanto Palmer (1988) e Kenkel e Walker (1993) propõem o estudo destes processos auxiliados pela geometria fractal. A análise fractal é também usada no estudo das características espaciais e temporais que afetam o rendimento de cereais, em experimentos de longa duração (Eghball e Varvel, 1997).

Dutilleul (1993); faz uma revisão dos tipos de experimentos na área ecológica, onde a heterogeneidade espacial está associada à informação,

ênfatizando a importância do conhecimento do material experimental e da adequação do delineamento utilizado em cada situação.

A heterogeneidade no material experimental ecológico pode ser resumida em três casos: por escala de estudo, por mosaicos ou por gradientes.

Freqüentemente existem combinações de heterogeneidade e homogeneidade segundo a escala de estudo, por exemplo uma heterogeneidade a pequena escala associada a uma homogeneidade a grande escala. As vezes, repetições ou amostras são localizadas em condições de microambiente e devem ser consideradas especialmente na interpretação dos resultados. Palmer (1988) mede este tipo de variações mediante semivariogramas e fractogramas.

Uma segunda forma de heterogeneidade espacial está representada por "patches" ou mosaicos. O caso mais simples de mosaico ocorre quando uma variável qualitativa apresenta duas classes (presença – ausência; água – terra). Naturalmente existem bem mais de duas classes, e assim se geram mosaicos diferentes a várias escalas. Uma forma de identificar o tamanho do mosaico, é combinando quadros contíguos e estudando a variação de alguma "função objetivo" em função do tamanho da gride (Greig-Smith, 1952). Também pode estudarse através de variogramas (Legendre e Fortin, 1989). Quando a escala experimental coincide com a escala do mosaico, a forma do experimento é facilmente determinada, se não houver coincidência entre ambas escalas o experimento é de difícil controle.

Outra forma de heterogeneidade espacial, são os gradientes ambientais. Estes podem ser em uma ou mais direções.

Quando se trabalha em amostragens por pontos e se quer avaliar a existência ou não de um padrão aleatório, utilizam-se estatísticas ou índices de homogeneidade. Existem várias estatísticas ou índices para medir os padrões espaciais aleatórios baseados em amostras aleatórias de n pontos de amostragem ou n eventos. Uma lista de algumas destas estatísticas está apresentada na Tabela 2, que é uma modificação daquela apresentada em Cressie (1993). Nestas estatísticas distinguem-se 5 tipos de distâncias que são definidas na Tabela 1 e que são indicadas por diferentes letras.

Tabela 1. Definição das distâncias medidas para a construção de índices de homogeneidade, e sua simbologia neste texto.

Distância	Pontos que une
W	Evento com evento mais próximo
X	Ponto de amostragem com evento mais próximo
X2	Ponto de amostragem com segundo evento mais próximo
Y	Ponto de amostragem –evento com evento mais próximo
Z	Ponto de amostragem –evento com evento mais próximo localizado num plano diferente ao do Ponto de amostragem-evento

Tabela 2. Estatísticas aplicadas no cálculo de vizinhos, suas distribuições assintóticas e referência principal (Modificada de Cressie, 1993).

Medidas Básicas	Estatística	Distribuição	Citada por:
W	$2(\lambda)^{1/2} \sum W_i / n$	$N(1, (4-\pi)/n\pi)$	Clark e Evans, 1954
	$2\pi\lambda \sum W_i^2$	χ_{2n}^2	Skellam, 1952
X	$\pi\lambda \sum X_i^2 / n$	$N(1, 1/n)$	Pielou, 1959
	$n \sum X_i^2 / (\sum X_i)^2$	Por simulação	Eberhardt (1967)
X, X₂	$\left\{ \sum (X_i^2 / X_{2,i}^2) \right\} / n$	$N(1/2, 1/12n)$	Holgate (1965 a)
	$(\sum X_i^2) / (\sum X_{2,i}^2)$	Beta (n, n)	Holgate (1965 a)
X, W	$\left\{ \sum (X_i^2 / (X_i^2 + W_i^2)) \right\} / n$	$N(1/2, 1/12n)$	Byth e Ripley (1980)
	$(\sum X_i^2) / (\sum W_i^2)$	$F_{2n, 2n}$	Hopkins (1954)
X, Z	$2n \sum (2X_i^2 + Z_i^2) / \left\{ \sum (\sqrt{2}X_i + Z_i) \right\}^2$	Por simulação	Hines e Hines (1979)
	$\sum \left\{ 2X_i^2 / (2X_i^2 + Z_i^2) \right\} / n$	$N(1/2, 1/12n)$	Besag e Gleaves (1973)
	$2 \sum \left\{ \min(2X_i^2, Z_i^2) / (2X_i^2 + Z_i^2) \right\} / n$	$N(1/2, 1/12n)$	Diggle et al. (1976)
	$2(\sum X_i^2) / (\sum Z_i^2)$	$F_{2n, 2n}$	Besag e Gleaves (1973)
	$-2 \sum \left\{ \log(2X_i^2 / (2X_i^2 + Z_i^2)) \right\}$	χ_{2n}^2	Cormack (1979)
X, Y	$\sum R_i / n'$	$N(1/2, 1/12n')$	Cox e Lewis (1976)
	$\bar{A}_X - \bar{A}_Y$	$N(0, (\bar{A}_X - \bar{A}_Y) / n')$	Satyamurthi (1979)

Os índices para detectar padrões espaciais estão sendo sempre avaliados quanto a sua performance, e outros são definidos.

Perry e Mead (1979), avaliam a potência do Índice de Dispersão como uma medida para detectar padrões espaciais. Este índice é freqüentemente

utilizado na área ecológica, expressado por $I = \left[\sum_{i=1}^n (x_i - \bar{x})^2 \right] \div \bar{x}$.

A avaliação foi realizada por simulação, assumindo um padrão aleatório de plantas, em mosaicos de plantas gerados a duas densidades. Os resultados

deste trabalho mostraram que este índice provê um poderoso teste de aleatoriedade, não sendo afetado pela densidade do mosaico nem pelo tamanho da unidade de amostragem.

Byth e Ripley (1980), estudam o teste de aleatoriedade de Hopkins, intensificando a amostragem pela introdução de uma amostragem sistemática.

A estatística de Hopkins é dada por

$$: H_F = \sum_{i=1}^m X_i^2 \div \sum_{i=1}^m W_i^2 ,$$

onde X_i é a distância entre i -ésimo ponto escolhido aleatoriamente na área amostral e seu vizinho mais próximo e W_i é a distancia entre i -ésimo evento escolhido aleatoriamente na área amostral e seu vizinho mais próximo. Eles definem um novo teste de aleatoriedade, e testam por simulação de Monte Carlo a potência de ambos os testes, sendo o novo superior. Cowling (1998) sugere o uso do teste de Hopkins em estudo de padrões espaciais sob transectos lineares, utilizando uma amostra da população e não um censo. O método mostrou-se poderoso para detectar grupos na vegetação.

Fortin e Gurevitch (1993); sugerem o teste de Mantel e o teste parcial de Mantel para avaliar a presença de correlação espacial nos dados. Estes métodos são baseados em matrizes de distância e em testes de aleatorização. Estes métodos diferem dos tradicionais que são testes de distribuição livre, baseados em poucas suposições. O teste de Mantel consiste em calcular a correlação linear entre duas matrizes de distância $n \times n$, onde uma matriz representa as distâncias espaciais entre unidades amostrais e a outra matriz representa as distâncias em outra ou outras variáveis de interesse no mesmo grupo de amostras. Esse coeficiente de correlação é testado por testes de

aleatoriedade (Manly, 1991), sendo sugerido um número mínimo de 20 unidades amostrais para poder detectar diferenças significativas.

A necessidade de medir a heterogeneidade aparece em muitos outros campos além do estudo de populações vegetais. Na área biológica, Diggle (1979), trabalhando com florestas discute os objetivos da análise de padrões espaciais por pontos, diferenciando os casos de amostragem e mapeamento. Ele discute a diferença entre assumir que os diferentes pontos de amostragem são independentes, ou considerar como um grupo de observações dependentes quando se tem o resultado de um mapeamento. Nos casos de amostragem são indicados como apropriados para o teste de hipóteses de distribuições aleatórias ou de estimações de intensidade ou número médio de eventos por área. No caso de mapeamento, hipóteses mais sofisticadas enquanto à distribuição espacial podem ser estudadas.

No contexto de ecologia de pastejo, Laca e Ortega (1995) consideram a heterogeneidade importante para explicar e entender as interrelações solo-planta-animal, lembrando a necessidade de dirigir o estudo em várias escalas. Propõem incluir no modelo teórico de pastejo uma quantificação de distribuição espacial de vegetação e de pastejo, o que aumentará a eficiência para predizer e manejar impactos de pastejo.

2.3 Os modelos de planejamento usados na área biológica

A análise de dados espaciais também tem sido utilizada com o objetivo de melhorar a estimacão de contrastes entre tratamentos em experimentos. Cullis e Gleeson (1989), trabalhando com experimentos de variedades em

programas de melhoramento, demonstraram um acréscimo na eficiência de estimação de contrastes entre variedades utilizando a análise do vizinho mais próximo quando comparada com os planejamentos em Blocos Completos Casualizados (BCC) e em Blocos Incompletos (BI). O uso do método de vizinho mais próximo em 1019 experimentos de variedades em cultivos de cereais conduzidos na Austrália de 1985 a 1987, resultou numa diminuição média de 42% nas variâncias das diferenças do rendimento, quando comparadas com o planejamento em BCC, e de 33% em relação aos delineamentos em BI. Os resultados mostram que as parcelas devem ser suficientemente largas para evitar a competição entre parcelas, e que o método do vizinho mais próximo é melhor no caso de experimentos com parcelas pequenas ou quando existem muitas parcelas no bloco.

Brownie et. al. (1993), trabalhando em programas de melhoramento de milho e soja, e utilizando diferentes métodos que incluem a variabilidade espacial, verificaram um aumento na eficiência de estimação de parâmetros. Esse aumento foi dependente do método utilizado (Papadakis, Análise da tendência ou Análise de covariância). Decisões intrínsecas de cada método, como tipo de covariância no método de Papadakis, grau do polinômio ajustado na tendência, ou estrutura de covariância utilizada, na análise de covariância modificaram o resultado final. Em comparação com o delineamento em BCC, a precisão melhorou com a inclusão na análise de referências espaciais. Os autores afirmam que o uso de técnicas de análise espacial melhorou a precisão, mas escolher a análise mais apropriada para um tipo de dados em particular pode ser trabalhoso.

Os mesmos resultados foram obtidos por Vollmann et al. (1996), que utilizaram o método do vizinho mais próximo e modificações deste em experimentos de melhoramento de espécies. Neste estudo, as correções melhoraram a eficiência das estimativas dos parâmetros em relação ao delineamento blocos casualizados, e também foram superiores a experimentos em látice, quando a heterogeneidade presente era de menor intensidade. Observam que em vários experimentos com parcelas compridas e estreitas, os ajustes longitudinais usando dois vizinhos, um de cada lado da parcela, foi superior ao ajuste por somente um vizinho.

Grondona et al (1996) analisam 35 experimentos de cereais utilizando procedimentos de análise espacial baseados em processos ARIMA, comparando com o delineamento em blocos incompletos. Foram comparados modelos com diferentes componentes de variação em larga escala e pequena escala, modelados uni ou bi-direcionalmente . Foram considerados 19 modelos espaciais e dois critérios foram utilizados para medir a adequabilidade do modelo espacial: o erro típico médio das diferenças e o erro quadrático médio de predição . Os modelos espaciais foram mais eficientes em reduzir a variabilidade residual que a análise em blocos incompletos, não sendo determinado um único modelo espacial que fora superior em todas as condições.

Cullis et al (1998) apresentam um estudo de um conjunto de experimentos varietais onde a análise incorpora uma estrutura de covariância para os erros em cada experimento. A análise produz preditores BLUE para os

genótipos e a interação genótipo por ambiente, e REML para os parâmetros espaciais e componentes de variância.

Casler (1999), compara delineamentos em blocos completos casualizados com delineamentos em latentes, análise de tendência e análise de vizinhos, em culturas forrageiras perenes de inverno. Os autores reportam um incremento na eficiência da precisão de 22 a 30% devido ao uso da análise de dados espaciais.

Zimmerman e Harville (1991), trabalhando com três diferentes ensaios de uniformidade, adotam um modelo que considera diretamente a heterogeneidade espacial, e estimam seus parâmetros por Máxima Verosimilhança ou variantes de Máxima Verosimilhança. Separam também a dependência espacial em dependência em “grande escala” ou tendências, que os autores modelam através de uma estrutura de médias, e a dependência em “pequena escala”, que modelam através da estrutura de correlação. Esta metodologia foi superior a blocos incompletos quando a relação entre parcelas vizinhas foi grande.

Ribeiro et al (1997), discutiram o método do variograma de mínimos quadrados generalizados que utiliza a variabilidade espacial na estimação dos contrastes entre tratamentos. Trabalharam com experimentos artificiais, nos quais sobrepos efeitos de tratamentos a ensaios de uniformidade. Essa metodologia tem a vantagem de que os “verdadeiros” valores de efeitos de tratamentos são conhecidos. Os autores concluíram que utilizando a informação de dependência espacial dos dados, através do método do variograma estimado por mínimos quadrados generalizados, foi possível obter

estimativas de parâmetros e contrastes mais próximas das reais e com menor variabilidade.

Grondona e Cressie (1991) usam a localização espacial de experimentos para aumentar a eficiência dos estimadores de contrastes num experimento industrial. Eles encontraram um incremento de 30% na eficiência de estimação de contrastes de tratamentos, sendo esses aumentos em eficiência maiores na presença de maiores dependências espaciais. Os autores trabalharam com seis experimentos simulados em blocos completos casualizados a partir de uma data original, cada um com 20 tratamentos e 4 repetições. Foram ajustados variogramas robustos na direção norte-sul. Realizou-se o ajuste de um semivariograma exponencial. A análise de variância foi estudada em quatro modelos: Delineamento em blocos casualizado, delineamento espacial incorporando uma matriz de covariâncias, modelo independente com variância constante e modelo espacial com incorporação de covariograma. O modelo que incluiu o covariograma foi o de melhor performance na estimação dos contrastes de tratamentos.

Cressie (1993) justifica a importância de uma abordagem que leve em consideração a dependência espacial. A detecção de estruturas de autocorrelação espacial em áreas experimentais pode permitir estimação mais eficiente dos contrastes entre tratamentos.

2.4. Análise de Dados Espaciais

Na análise de dados espaciais deve-se considerar a natureza dos dados e as formas de modelar a heterogeneidade espacial para incluir no modelo.

2.4.1 A natureza dos dados espaciais.

No caso de dados espaciais considera-se que em cada ponto da região tem-se uma variável aleatória diferente sendo que estes estudos baseiam-se na teoria de variáveis regionalizadas de Matheron (Cressie, 1993). No plano, cada dado é localizado por 2 coordenadas e os semivariogramas são uma das formas de modelar a distribuição espacial.

2.4.2 Semivariogramas

O semivariograma é a medida estatística padrão da variabilidade espacial, e é função da distância entre observações. É definida como a metade da variância da diferença entre duas observações separadas por uma distância dada “h”. O estimador clássico do semivariograma, também dito estimador de Matheron (Cressie, 1993), é expressado por:

$$\gamma_{(h)} = \frac{1}{2N(h)} \sum [z_{(s+h)} - z_{(s)}]^2$$

onde N(h) é o número de pares de pontos separados pela distância h, considerado um escalar.

A representação gráfica teórica de um semivariograma encontra-se na Figura 1. Os parâmetros do semivariograma são:

* o “**nugget**” ou **efeito pepita**, definido como o intercepto. É o valor do semivariograma a h=0. A notação padrão é **c₀** . e mede a variabilidade em escalas menores do que as estudadas.

* o “**sill**” ou **patamar**, que é o valor da semivariância máxima atingida. O patamar corresponde à variância de uma observação. A notação padrão é **C(s)**.

* o “**range**” ou **alcance**, que é a distância máxima a partir da qual não existe aumento na semivariância. Para distâncias menores que o patamar, existe correlação espacial, não existindo para distâncias maiores ou iguais ao patamar. A notação padrão é **a**.

Semivariogramas empíricos são ajustados aos dados, visando obter uma estimativa da forma da autocorrelação espacial.

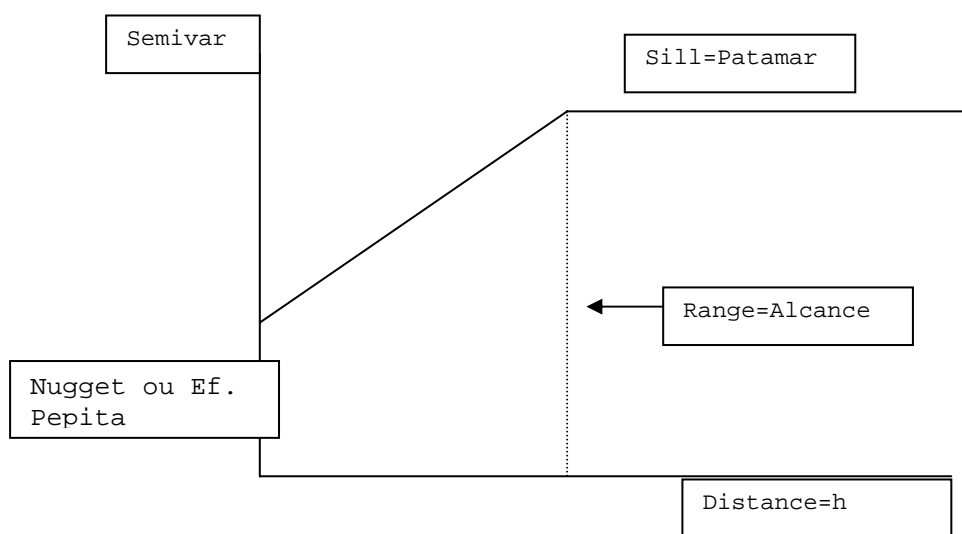


Figura 1. Representação gráfica de um semivariograma teórico.

O estimador de semivariograma de Matheron é um estimador obtido pelo método dos momentos, sendo não viesado sob a hipótese intrínseca de estacionariedade. Outros estimadores de semivariogramas foram propostos, como o estimador robusto de Cressie e Hawkins de expressão:

$$\hat{\gamma}_h = \frac{1}{2} \left[\frac{1}{N(h)} \sum_{N(h)} |z(s+h) - z(s)|^{1/2} \right]^4 \frac{0,457 + \frac{0,494}{N(h)}}{1}$$

Também podem-se utilizar para descrições os semivariogramas relativos. O semivariograma relativo local caracteriza-se pelo fato que no seu calculo participam os semivariogramas amostrais de cada local i divididos pelas médias n daquela região. O variograma relativo local corresponde à expressão:

$$\hat{\gamma}_{RL(h)} = \frac{\sum_{i=1}^n N_i(h) \frac{\hat{\gamma}_i(h)}{m_i^2}}{\sum_{i=1}^n N_i(h)}$$

O variograma relativo geral, é expresso por:

$$\hat{\gamma}_{RG(h)} = \frac{\hat{\gamma}(h)}{m^2(h)}$$

e o variograma relativo pareado por:

$$\hat{\gamma}_{RP(h)} = \frac{1}{2N(h)} \sum_{(i,j)/h_{ij} \approx h} \frac{[z(s+h) - z(s)]^2}{\left[\frac{z(s+h) + z(s)}{2} \right]^2}$$

Pela forma que estão construídos, todos os estimadores produzem resultados confiáveis em distâncias pequenas, perdendo esse poder nas distâncias maiores. Isso se deve ao fato que o número $N(h)$ de pares de pontos vai diminuindo a medida que h aumenta.

Em relação à forma do semivariograma empírico são ajustados semivariogramas teóricos.

Os modelos teóricos de semivariogramas mais utilizados, e mais disponíveis nos aplicativos computacionais, são o linear, o exponencial, o esférico e o gaussiano, caracterizados por:

Linear:
$$\gamma_{(h)} = c_0 + c_s h$$

Exponencial
$$\gamma_{(h)} = c_0 + c_s \left[1 - \exp\left(\frac{-h}{a}\right) \right]$$

Esférica
$$\gamma_{(h)} = c_0 + c_s \left[\left(\frac{3h}{2a}\right) - \frac{h^3}{2a^3} \right]$$

Gaussiano
$$\gamma_{(h)} = c_0 + c_s \left[1 - \exp\left(\frac{-h^2}{a^2}\right) \right]$$

Nestes modelos o vetor de parâmetros é $\underline{\theta} = (C_o, C_s, a)'$ cujos componentes são chamadas de efeito pepita, patamar e alcance; respectivamente. Todos os parâmetros devem ser não negativos.

Na escolha do modelo a ser ajustado para o semivariograma empírico, podem ser usados vários critérios, dentro dos quais, os critérios baseados em REML, o Critério de Informação de Akaike (AIC) e o Critério Bayesiano de Schwarz (SBC).

2.4.3. Isotropia e Anisotropia

Um fenômeno é dito isotrópico quando o padrão da estrutura de variabilidade espacial é o mesmo em todas as direções, e é anisotrópico quando este padrão muda dependente da direção do estudo, podendo existir fenômenos isotrópicos numa área do estudo e anisotrópicos em outra. A

metodologia aplicada no estudo de isotropia consiste em realizar um semivariograma geral em todas as direções, e compará-lo com semivariogramas dirigidos em diferentes direções. Quando o fenômeno é isotrópico, os semivariogramas geral e construídos para diferentes direções não apresentam diferenças relevantes, e um mesmo modelo teórico é adequado a todos eles (Ribeiro, 1995).

Dois tipos de anisotropia podem ocorrer: a geométrica e a zonal.

A anisotropia geométrica é a mais comum. No caso de modelos com patamar, variogramas de diferentes direções têm o mesmo patamar e alcances diferentes. Nos modelos lineares sem patamar, os coeficientes angulares são diferentes. Este tipo de anisotropia mostra o mesmo padrão realizado a diferentes escalas segundo a direção estudada. Este tipo de anisotropia pode ser padronizado utilizando como fator de correção os alcances.

A anisotropia zonal pode ser notada quando os variogramas direcionais possuem diferentes patamares. Neste caso o fenômeno apresenta diferentes formas de regionalização nas diferentes direções. Modelar a anisotropia significa considerar um maior número de parâmetros a serem estimados e é utilizado quando melhora efetivamente a descrição do fenômeno, sendo necessários nesse caso um número maior de dados.

2.4.4. Os modelos matemáticos

Vários modelos são apresentados na literatura para representar a variabilidade espacial: simples, com partição de variação, com refinamento pela

mediana e de planejamento de experimento. A caracterização dos diferentes modelos apresenta-se a seguir.

2.4.4.1. Modelo simples

Este modelo representa cada ponto como uma variável aleatória, caracterizada pela soma de um efeito médio da localização mais um erro. A autocorrelação espacial está representada pela inclusão da semivariância na variância do erro. Incluir a semivariância é uma forma de quantificar a autocorrelação espacial, enquanto que outra alternativa é incluir medidas de covariâncias ou de correlação (Cressie, 1993).

O modelo considerado é:

$$z_{(x)} = \mu_{(v)} + \varepsilon_{(x)}$$

onde:

$z_{(x)}$ é o valor de uma variável aleatória

$\mu_{(v)}$ é a média de z na localização v

$\varepsilon_{(x)}$ é um termo aleatório autocorrelacionado, com média zero e variância definida por:

$$\text{var}[\varepsilon_{(x)} - \varepsilon_{(x+h)}] = E\left\{\left[\varepsilon_{(x)} - \varepsilon_{(x+h)}\right]^2\right\}$$

onde h é o "lag", ou separação espacial entre os pontos considerados.

A média se assume constante: $E[z_{(x)} - z_{(x+h)}] = 0$

e a variância será:

$$V[z_{(x)} - z_{(x+h)}] = E[z_{(x)} - z_{(x+h)}]^2 = 2\gamma_{(h)}$$

A variância depende da separação entre os pontos (h) e não da posição de x , onde

γ é a semivariância, que em função de h é o semivariograma.

Modelos mais complexos, com inclusão de termos do delineamento, com consideração de efeitos fixos e aleatórios também serão considerados.

2.4.4.2 Modelo com partição de variações

O modelo é representado por:

$$z_{(s)} = \mu_{(s)} + \omega_{(s)} + \eta_{(s)} + \varepsilon_{(s)}$$

onde:

$\mu_{(s)} \equiv E\{z_{(s)}\}$ é a estrutura determinística, que pode ser chamada de variação de larga escala;

$\omega_{(s)}$ é um processo intrinsecamente estacionário de média zero, chamado de variação de pequena escala;

$\eta_{(s)}$ é um processo intrinsecamente estacionário de média zero, independente de ω , chamado de variação de micro escala;

$\varepsilon_{(s)}$ é um processo ruído branco de média zero, independente de ω e de η , chamado de erro de medida ou ruído. Denota-se $\text{var}[e_{(s)}] = C_{ME}$

2.4.4.3. Modelo com refinamento pela mediana

Outra forma de representar a variabilidade espacial foi proposta por Cressie (1993), que usa o refinamento pela mediana para estimar $\mu_{(s)}$. No caso

$\mu_{(.)} \equiv c^{te}$ o processo é dito intrinsecamente estacionário. A variabilidade pode ser descomposta da forma:

$$\gamma_z(h) = \gamma_\omega(h) + \gamma_\eta(h) + C_{ME}$$

Onde C_{ME} , a variância do erro de medida, não pode ser estimada dado o fato de que apenas uma observação é tomada em cada posição. Ela se soma à variação de microescala (aquela que ocorre em distâncias menores do que a dimensão da malha de amostragem), que também não pode ser detectada, compondo o efeito “nugget” ou efeito “pepita” Co. Desta forma a descomposição que pode ser estimada é:

$$\gamma_z(h) = \gamma_\omega(h) + C_o$$

2.4.4.4 Modelo de planejamento de experimentos

No contexto de experimentos, a adoção de metodologia de análise de dados espaciais, considera a estrutura clássica de modelos mistos, onde uma caracterização, de acordo com Zimmerman e Harville (1991), é dado por:

$$d = A\alpha + GT\tau + \varepsilon$$

onde

$d = Gy$ é um vetor de funções lineares da variável resposta y

α é um vetor de parâmetros desconhecidos

τ é um vetor de parâmetros desconhecidos

T é uma matriz de delineamento

A é uma matriz de constantes específicas ou de funções específicas de y

ε é um vetor de erros, sendo $E(\varepsilon) = 0$ e $Var(\varepsilon) = \sigma^2 \Sigma$, onde σ^2 é um parâmetro positivo e desconhecido e os elementos de Σ são constantes ou funções específicas de parâmetros desconhecidos. Na solução do problema adota-se mínimos quadrados generalizados, ou a metodologia de modelos mistos.

2.4.5 As estruturas de covariância

Os modelos propostos, apresentam uma matriz de variância-covariância entre pontos heterogênea, que permite modelar o padrão de variação espacial, sendo necessário estima-la. A escolha e o ajuste desta matriz, estará então determinada pela variação espacial presente em cada delineamento ou plano de amostragem.

Algumas das estruturas de covariância mais gerais utilizadas para modelar a variabilidade espacial são caracterizadas a seguir.

2.4.5.1 Autorregressiva

Nesta estrutura a covariância entre dois pontos está multiplicada pela correlação linear entre eles, sendo essa correlação decrescente exponencialmente em função da distância entre pontos.

$$\sigma^2 \begin{bmatrix} 1 & \rho & \rho^2 & \rho^3 \\ \rho & 1 & \rho & \rho^2 \\ \rho^2 & \rho & 1 & \rho \\ \rho^3 & \rho^2 & \rho & 1 \end{bmatrix}$$

2.4.5.2 Espacial

Esta estrutura em função da distância euclidiana, onde a covariância entre dois pontos está multiplicada por um coeficiente de correlação linear, relacionando exponencialmente com a distância euclidiana (d_{ij}) entre pontos, tendo σ^2 e ρ a mesma caracterização que em 2.4.5.1.

$$\sigma^2 \begin{bmatrix} 1 & \rho^{d_{12}} & \rho^{d_{13}} & \rho^{d_{14}} \\ \rho^{d_{12}} & 1 & \rho^{d_{23}} & \rho^{d_{24}} \\ \rho^{d_{13}} & \rho^{d_{23}} & 1 & \rho^{d_{34}} \\ \rho^{d_{14}} & \rho^{d_{24}} & \rho^{d_{34}} & 1 \end{bmatrix}$$

Ambas matrizes mostram covariâncias decrescentes em função da distância (d_{ij}) entre pontos. Em forma geral pode-se definir modelos de correlação espacial considerando-se:

$$Var(e_i) = \sigma_i^2 \quad \text{e} \quad Cov(e_i, e_j) = \sigma^2 [f(d_{ij})]$$

onde $f(d_{ij})$ são as funções de semivariograma teóricas, algumas das quais (linear, exponencial, esférica e gaussiana). foram descritas em 2.4.2.

2.5 Tendência atual

A aplicação da análise de dados espaciais no contexto de modelos mistos tem sido a forte tendência da última década, principalmente dada a disponibilidade de aplicativos computacionais específicos. Vários são os autores que desenvolveram ou aplicaram esta metodologia, alguns deles já comentado.

Também o uso de modelos Bayesianos no contexto de experimentos agrônômicos, é um dos tópicos que esta tendo maior atenção nos últimos anos.

Nos experimentos agronômicos, uma das principais particularidades dos experimentos em campo é a presença de padrões de fertilidade que ocorrem naturalmente ou através de práticas de manejo. A precisão das análises, da ordenação das médias e dos contrastes testados, dependerá da forma em que esses padrões não observáveis e estão representados na análise. Nesse contexto, é fundamental a representação espacial destes padrões. Alega-se que a adoção do paradigma Bayesiano simplifica a interpretação dos resultados, especialmente na ordenação e na seleção. A metodologia de cadeias de Markov por métodos de Monte Carlo, tem sido utilizada com frequência.

Segundo Besag e Higdon (1999) no paradigma Bayesiano é necessário especificar a distribuição “a priori” , a qual deve considerar as principais características da variabilidade espacial. No trabalho, eles adotaram um modelo de quatro vizinhos com uma autoregressão Gaussiana como o modelo base. Este estudo envolve medidas tanto discretas como contínuas, em experimentos de pequeno e grande porte, com número de observações que oscilam desde poucas dezenas até alguns milhares. Sempre o método Bayesiano foi superior aos métodos tradicionais, igualando-se em alguns casos aos métodos geoestatísticos.

3. MATERIAL

3.1 Experimento 1

Os dados referem-se a um experimento com oito cultivares de trevo-branco (*Trifolium repens* L.): 1.Regal, 2.Yi, 3.Bayucuí, 4.Huia, 5.Pitau, 6.Bagé, 7.Guaiba S1 e 8.Jacuí S2, sendo as duas últimas desenvolvidas no Rio Grande do Sul.

Os trabalhos de campo foram conduzidos durante os anos de 1983 a 1987 na Estação Experimental Agronômica da Universidade Federal do Rio Grande do Sul (EEA-UFRGS), em Eldorado do Sul, Rio Grande do Sul, sob responsabilidade do Prof. Dr. Nilton R. Paim. Paim e Riboldi (1994) descrevem com detalhes o experimento.

As nove cultivares foram testadas em associação com gramíneas, azevém anual (*Lolium multiflorum* Lam.) e pensacola (*Paspalum notatum* Fl. var. Saurae Parodi). Nos quatro anos de avaliação (83/84, 84/85, 85/85, 86/87), foram determinados rendimento de matéria seca de: trevo branco, azevém, pensacola, outras gramíneas e inços; e proteína bruta dentre outros. Os dados originais encontram-se no Apêndice 1. O delineamento experimental foi em blocos completos casualizados, com cinco repetições.

No presente trabalho, estudou-se a associação espacial de todas as variáveis listadas em cada ano do estudo.

O tamanho das parcelas utilizadas neste estudo foi de 2 m por 5 m, separadas entre si por 0,50 m dentro de cada bloco, e cada bloco separado entre si por 1,0 m.

Cada parcela foi georeferenciada pelas suas coordenadas, considerando o ponto médio das linhas e das colunas. A estrutura de aleatorização das cultivares nos diferentes blocos é apresentada na Figura 1, onde os números representam os cultivares.

Bloco 1

5	4	3	1	2	7	8	6
----------	----------	----------	----------	----------	----------	----------	----------

Bloco 2

1	5	6	2	7	3	4	8
----------	----------	----------	----------	----------	----------	----------	----------

Bloco 3

6	2	1	8	7	4	5	3
----------	----------	----------	----------	----------	----------	----------	----------

Bloco 4

8	3	5	7	2	4	6	1
----------	----------	----------	----------	----------	----------	----------	----------

Bloco 5

5	3	4	1	6	7	2	8
----------	----------	----------	----------	----------	----------	----------	----------

Figura 2. Aleatorização do experimento 1.

3.2 Experimento 2

Os dados referem-se a um experimento de comparação de 20 espécies forrageiras, caacterizadas na Tabela 3.

Tabela 3. Caraterização das 20 espécies forrageiras utilizadas no experimento 2.

Cultivar	<i>Espécie</i>	Cultivar	Origem
1	<i>Ornithopus compressus</i>	Encantada	Uruguai
2	<i>Trifolium repens</i>	Zapican	Uruguai
3	<i>Lotus pedunculatus</i>	Maku	Uruguai
4	<i>Trifolium alexandrinum</i>	-	Israel
5	<i>Trifolium pratense</i>	E116	Uruguai
6	<i>Trifolium repens</i>	Dusi	
7	<i>Trifolium vesiculosum</i>	Yucchi	Brasil
8	<i>Lotus pedunculatus</i>	Sunrise	Nova Zelanda
9	<i>Trifolium repens</i>	Prop	Nova Zelanda
10	<i>Lotus corniculatus</i>	San Gabriel	Uruguai
11	<i>Trifolium repens</i>	Bayucua	Uruguai
12	<i>Trifolium repens</i>	Sustain	Nova Zelanda
13	<i>Lotus pedunculatus</i>	LT2	Nova Zelanda
14	<i>Lotus tenuis</i>		
15	<i>Trifolium pratense</i>	Red quelí	
16	<i>Lotus subiflorus</i>	Rincon	Uruguai
17	<i>Festuca arundinacea</i>	Tacuabe	Uruguai
18	<i>Dactylis glomerata</i>	Overon	Uruguai
19	<i>Bromus auleticus</i>	Potrillo	Uruguai
20	<i>Lolium multiflorum</i>		

Os trabalhos de campo foram conduzidos durante o ano de 1998, sob responsabilidade do Prof. MSc. J.Bologna, da Estação Experimental “Mario Cassinoni” da “Universidad de la República”, Uruguai.

O experimento foi realizado em um solo basáltico de profundidades média. O delineamento experimental foi em blocos completos casualizados, com três repetições. As espécies foram introduzidas na pastagem nativa

(campo nativo) existente na área, sendo considerada uma testemunha sem introdução de espécies. Estudou-se a variável percentagem de implantação em três momentos de contagem.

O tamanho das parcelas utilizadas neste estudo foi de 2 m por 5 m. Os caminhos entre blocos foram de 4 m. Como coordenadas de cada parcela, foi considerado o ponto médio da linha e da coluna a que pertenciam. A estrutura de aleatorização dos tratamentos nos diferentes blocos é apresentada na Figura 3, onde os números correspondem as cultivares sendo T a testemunha de campo natural sem introdução de espécies.

Bloco 1:

4	9	11	13	5	7	14	16	3	12	8	2	1	6	10	15	T	18	17	20	19
---	---	----	----	---	---	----	----	---	----	---	---	---	---	----	----	---	----	----	----	----

Bloco 2:

7	11	14	8	4	12	1	15	9	16	3	13	6	5	2	10	T	19	17	18	20
---	----	----	---	---	----	---	----	---	----	---	----	---	---	---	----	---	----	----	----	----

Bloco 3:

16	5	9	12	8	15	3	6	1	2	13	14	4	7	10	11	T	17	19	20	18
----	---	---	----	---	----	---	---	---	---	----	----	---	---	----	----	---	----	----	----	----

Figura 3. Aleatorização do experimento 2.

3.3. Experimento 3

Os dados referem-se a um levantamento por amostragem em campo natural, realizado na EEA-UFRGS no ano 1994. Foi levantada a composição específica de 400 unidades amostrais contínuas de 0,1m por 0,1 m localizadas

sobre uma transecção linear numa pastagem natural. Em cada unidade amostral foi avaliada a percentagem de cada espécie presente, dentre outras variáveis. Cadenazzi (1996) descreve detalhadamente o estudo. Neste trabalho, dentre todas as espécies amostradas escolheram-se três de diferentes portes para descrever sua associação espacial: *Aristida laevis* (Nees) Kunth, *Desmodium incanum* DC. e *Paspalum notatum* Flügge.

4. MÉTODO

4.1 Variabilidade espacial

No caso de dados espaciais, ao contrário de outros procedimentos estatísticos em que entende-se os diferentes elementos de uma amostra como “repetições” ou realizações de uma variável aleatória, considera-se que em cada ponto da região tem-se uma variável aleatória diferente. Desta forma, para toda a região estudada tem-se um conjunto de variáveis aleatórias, denotado por:

$$\{z_{(s)}: s \in D \subset R^d\}$$

Onde: z é a variável aleatória em estudo;

s indica a posição no espaço;

D é a região em estudo e

R^d é o espaço d-dimensional no qual está contida a região

No plano $d=2$, e então os dados serão localizados por 2 coordenadas.

4.1.1 Modelagem probabilística e estacionaridade: a importância da análise descritiva

Contrariamente ao caso das variáveis aleatórias, as variáveis regionalizadas não permitem o estudo de sua distribuição ou estimação de momentos, ao obter-se uma realização de cada variável aleatória. Para contornar tal impedimento, adota-se na análise algum tipo de estacionaridade das variáveis, assumindo um comportamento das variáveis sem comprovação objetiva.

O tipo de estacionaridade mais forte e restritivo é a estacionaridade em sentido restrito que implica que as leis da função sejam invariantes para toda translação efetuada sobre os pontos de amostragem. Isto poderia ser representado como:

$$F(Z_1, Z_2, \dots, Z_n; s_1, s_2, \dots, s_n) = F(Z_1, Z_2, \dots, Z_n; s_1+h, s_2+h, \dots, s_n+h)$$

Neste tipo de estacionaridade assume-se que a função de distribuição depende apenas da geometria relativa dos pontos. As distribuições de primeira e segunda ordem sob este tipo de estacionaridade ficam:

Primeira ordem: $F_z(z, s) = F_z(z, s + h)$

Segunda ordem $F_{z_1, z_2}(z_1, z_2; s_1, s_2) = F_{z_1, z_2}(z_1, z_2; s_1 + h, s_2 + h)$
 $F_{z_1, z_2}(z_1, z_2; s_1, s_2) = F_{z_1, z_2}(z_1, z_2; s_1 - s_2) = F_{z_1, z_2}(z_1, z_2; h)$

A distribuição de primeira ordem independe de s, sendo em cada ponto da região a distribuição da variável aleatória igual.

Na distribuição de segunda ordem observa-se que a distribuição conjunta depende somente das distâncias entre dois pontos (h), portanto, diversos pontos separados pela mesma distância forneceriam as repetições necessárias para estimações.

A estacionaridade em sentido amplo é definida generalizando a de segunda ordem, sendo um processo estacionário em sentido amplo o que obedece às condições:

- i) $E \{ Z(s) \} = \mu, \forall s$
- ii) $C [Z_1(s_1), Z_2(s_1 + h)] = C(h) = E [Z_1(s_1) Z_2(s_1+h)]$

Outro tipo de estacionaridade fica definida a partir das diferenças entre as variáveis aleatórias em cada ponto, a estacionaridade dos incrementos, onde a variável utilizada é a diferença entre dois pontos: $[Z(s + h) - Z(s)]$. Um tipo de estacionaridade dos incrementos é a chamada estacionaridade intrínseca, definida pelas condições:

- i) $E \{ [Z(s_1) - Z(s_2)] \} = 0$

$$\text{ii) } E \{ [Z(s_1) - Z(s_2)]^2 \} = 2 \gamma(h)$$

Todo processo estacionário de segunda ordem é também intrinsecamente estacionário, mas a recíproca não é verdadeira.

A necessidade de assumir estacionariedade em algum sentido, faz com que seja importante um bom conhecimento do comportamento dos dados, que pode ser obtido com técnicas descritivas e exploratórias.

A inclusão na descrição de elementos espaciais como a distância de linha e coluna a que pertencem cada variável, auxilia na detecção de padrões. A localização de dados atípicos, quando ocorrem de forma concentrada, indicam sub-regiões com comportamento diferenciado do geral da área, e se ocorrem de maneira dispersa, podem indicar que o fenômeno tem alta variabilidade. A análise descritiva proporciona elementos que auxiliam na determinação de que tipo de estacionariedade podem ter os dados sob estudo, permitindo conhecer a estrutura de variabilidade do processo.

4.1.2 Estrutura de variabilidade: Semivariogramas

Os semivariogramas empíricos são ajustados aos dados, visando obter uma estimativa da forma da autocorrelação espacial. O semivariograma é uma função que descreve o valor da variância de z em função de h . Os parâmetros do semivariograma são o patamar (“sill”), que é o valor da semivariância máxima atingida; o alcance, que é a distância máxima a partir da qual não existe aumento na semivariância, e o efeito pepita (“nugget”) que mede a variabilidade em escalas menores do que as estudadas.

Em relação à forma do semivariograma empírico são ajustados semivariogramas teóricos, como o estimador clássico o de Matheron, ou o estimador robusto de Cressie e Hawkins que foram apresentados na revisão bibliográfica. Neste trabalho serão ajustados semivariogramas clássicos. Os modelos teóricos de semivariogramas mais utilizados, e mais disponíveis nos aplicativos computacionais, são o linear, exponencial, esférico e gaussiano. Nestes modelos o vetor de parâmetros é $\underline{\theta} = (C_o, C_s, a)'$ cujas componentes são chamadas de efeito pepita, patamar e alcance; respectivamente. Todos os parâmetros devem ser não negativos.

4.2 Análise de dados espaciais em planejamento de experimentos.

Em planejamento de experimentos, assume-se que cada parcela é atribuída a um tratamento, conforme um plano experimental que considera todas as possíveis combinações de parcelas - tratamentos. Mais recentemente, o componente espacial dos dados experimentais foi reconhecido e os modelos aleatórios foram adotados.

A metodologia de modelos mistos auxilia de forma importante a modelagem da variabilidade espacial nos modelos de planejamento de experimentos, incluindo na matriz de variâncias –covariâncias os valores de semivariogramas calculados.

4.2.1 Modelo Linear Geral

O modelo estatístico mais difundido é o modelo linear geral (GLM). Este pode ser representado da forma

$$y = X\beta + \varepsilon$$

onde y é o vetor de dados observado, X é uma matriz de planejamento conhecida que pode conter variáveis contínuas ou indicadoras, β é um vetor de efeitos fixos desconhecidos e ε é um vetor de erros aleatórios não observáveis, independentes e normalmente distribuídos.

Neste modelo, a estimação de β pode ser feita utilizando o método de mínimos quadrados, encontrando o valor de β que minimiza

$$(y - X\beta)'(y - X\beta) = \sum_{i=1}^n (y_i - X_i\beta)^2$$

onde y_i é o i -ésimo elemento de y , X_i é a i -ésima linha de X , e n é a dimensão de y , produzindo o sistema de equações normais:

$$X'X \hat{\beta} = X'y$$

cuja solução é dada por:

$$\hat{\beta} = (X'X)^{-1} X'y$$

onde $^{-1}$ denota a inversa generalizada.

Se os erros, ε são independentes e identicamente distribuídos seguindo uma distribuição normal com média zero e variância σ^2 , $\hat{\beta}$ é BLUE de β , ou melhor estimador linear imparcial (onde melhor significa de mínima variância). Se os erros cumprem com a condição de normalidade $(0, \sigma^2)$ e independência $\hat{\beta}$ é normalmente distribuído com média β e variância $(X'X)^{-1} \sigma^2$, sendo σ^2 estimado pelo quadrado médio do erro:

$$\hat{\sigma}^2 = \frac{(y - X\hat{\beta})'(y - X\hat{\beta})}{\lambda}$$

onde $\lambda = n - \text{rank}(X)$.

Geralmente as inferências no GLM consistem em calcular uma estimativa de uma combinação linear relevante de β e usar sua distribuição amostral para realizar testes de hipóteses ou construir intervalos de confiança.

Seja $L'\beta$ as combinações lineares relevantes, onde L é uma matriz de coeficientes estimáveis conhecida e de rank completo. Dada uma matriz L apropriada e sob suposições convenientes, o melhor estimador de $L'\beta$ é $L'\hat{\beta}$, e sua matriz de variâncias é $L'(X'X)^{-1}L\sigma^2$. Quando L tem rank 1, as quantidades acima se reduzem a escalares e pode-se construir a estatística t :

$$t = \frac{L'\hat{\beta}}{\sqrt{L'(X'X)^{-1}L\hat{\sigma}^2}}$$

Sob a suposição de normalidade para ε , t tem uma distribuição t de Student exata com λ graus de liberdade. O intervalo com coeficiente de $100(1-\alpha)\%$ de confiança associado é o seguinte:

$$L'\hat{\beta} \pm t_{\lambda, \alpha/2} \sqrt{L'(X'X)^{-1}L\hat{\sigma}^2}$$

Onde $t_{\lambda, \alpha/2}$ é o valor de t para λ graus de liberdade e probabilidade $\alpha/2$.

Quando o rank de L é maior que 1, constroi-se a estatística F :

$$F = \frac{\hat{\beta}'L[L'(X'X)^{-1}L]^{-1}L'\hat{\beta}}{\text{rank}(L)\hat{\sigma}^2}$$

Sob a suposição de normalidade para ε , a estatística F tem uma distribuição F exata com $\text{rank}(L)$ graus de liberdade no numerador e λ graus de liberdade no denominador.

4.2.2 Modelo Misto

O modelo misto permite uma especificação mais geral da matriz de covariância de ε que o modelo GLM. O modelo misto pode-se assumir que existe correlação entre erros, e que estes não sejam identicamente distribuídos embora assumam normalidade.

O modelo misto é escrito como:

$$y = X\beta + Zv + \varepsilon$$

onde cada termo é igual ao que foi definido no GLM, exceto a adição da matriz de incidência, Z, e o vetor de efeitos aleatórios desconhecidos, v. Assim como X, Z pode conter variáveis contínuas ou indicadoras. O nome "misto" vem do fato de que o modelo contém efeitos fixos, β , e efeitos aleatórios, v.

No modelo misto assume-se

$$E \begin{bmatrix} v \\ \varepsilon \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \text{e} \quad \text{Var} \begin{bmatrix} v \\ \varepsilon \end{bmatrix} = \begin{bmatrix} G & 0 \\ 0 & R \end{bmatrix}$$

A variância de y, portanto, é: $V = \text{Var}(y) = ZGZ' + R$

Pode-se modelar V através dos efeitos aleatórios agrupados pela matriz de incidência Z e pela especificação de estruturas de covariância para G e R. No modelo linear geral, a esperança de y é modelada através dos efeitos fixos β , e a extensão fornecida pelo modelo misto é que a variância de y pode ser

modelada através das matrizes Z , G e R . Modelos de efeitos aleatórios simples são um caso especial da especificação geral com Z contendo variáveis indicadoras, G contendo componentes de variância em uma estrutura diagonal e $R = \sigma^2 I$, onde I é uma matriz identidade de dimensão n . Este é o caso do modelo misto clássico, bastante útil especialmente em planejamento de experimentos com blocos casualizados e parcelas subdivididas. O GLM é, também, um caso especial do modelo misto onde $Z = \phi$ e $R = \sigma^2 I$.

A estimação dos efeitos fixos e aleatórios é mais complexa no modelo misto que no GLM. Neste caso, não se tem somente parâmetros desconhecidos em β , mas também em ν , G e R . O método de mínimos quadrados ordinários não é mais o melhor método de estimação. O método de mínimos quadrados generalizados (GLS) é muito mais apropriado. O método de GLS minimiza $(y - X\beta)' V^{-1} (y - X\beta)$. Entretanto, ele requer o conhecimento de V e, conseqüentemente, de G e R . Faltando esta informação, um procedimento usa o GLS aproximado, no qual toma-se alguma estimativa para V no problema de minimização. Obter uma estimativa de V através do método dos momentos é perigoso.

Em muitas situações, o melhor procedimento é usar os métodos de máxima verossimilhança (ML) e máxima verossimilhança restrita (REML), sob a suposição que ν e ε são normalmente distribuídos. Estes métodos constroem funções de verossimilhança e as maximizam em relação aos parâmetros desconhecidos. Utilizando poucos cálculos, é possível reduzir estes problemas de maximização para apenas os parâmetros em G e R . Isto pode ser escrito como:

$$ML: l(G, R) = -\frac{1}{2} \log|V| - \frac{n}{2} \log r V^{-1} r - \frac{n}{2} \left(1 + \log \frac{2\pi}{n} \right)$$

$$REML: l_R(G, R) = -\frac{1}{2} \log|V| - \frac{1}{2} \log|X V^{-1} X| - \frac{n-p}{2} \log r V^{-1} r - \frac{n-p}{2} \left(1 + \log \frac{2\pi}{n-p} \right)$$

onde $r = y - X(X'V^{-1}X)^{-1}X'V^{-1}y$ e p é o rank de X . Após a obtenção das estimativas, \hat{G} e \hat{R} , para G e R , soluciona-se as equações de modelos mistos (também conhecidas como equações de Henderson (Henderson, 1963)):

$$\begin{bmatrix} X'\hat{R}^{-1}X & X'\hat{R}^{-1}Z \\ Z'\hat{R}^{-1}X & Z'\hat{R}^{-1}Z + \hat{G}^{-1} \end{bmatrix} \begin{bmatrix} \hat{\beta} \\ \hat{v} \end{bmatrix} = \begin{bmatrix} X'\hat{R}^{-1}y \\ Z'\hat{R}^{-1}y \end{bmatrix}$$

para obter estimativas de β e v . Estas soluções também podem ser escritas

como:

$$\hat{\beta} = (X\hat{V}^{-1}X)^{-1}X\hat{V}^{-1}y$$

$$\hat{v} = \hat{G}Z\hat{V}^{-1}(y - X\hat{\beta})$$

que têm conexão com os estimadores empíricos de Bayes da distribuição relevante a posteriori (Laird e Ware, 1982). Caso G e R sejam conhecidas, $\hat{\beta}$ é comumente chamado de BLUE (melhor estimador linear imparcial) de β e \hat{v} é denominado BLUP (melhor preditor linear imparcial) de v (McLean et al., 1991 e Robinson, 1991). A matriz de covariância de $\hat{\beta}$ e \hat{v} seria:

$$C = \begin{bmatrix} X'R^{-1}X & X'R^{-1}Z \\ Z'R^{-1}X & Z'R^{-1}Z + G^{-1} \end{bmatrix}^{-1}$$

Entretanto, como se utilizam os estimadores \hat{R} e \hat{G} , a matriz de covariâncias obtida é:

$$\hat{C} = \begin{bmatrix} X' \hat{R}^{-1} X & X' \hat{R}^{-1} Z \\ Z' \hat{R}^{-1} X & Z' \hat{R}^{-1} Z + \hat{G}^{-1} \end{bmatrix}$$

Em geral, as estatísticas t e F usuais têm distribuição t de Student e F respectivamente. No entanto, pode-se ainda fazer bom uso dos testes aproximados. Para realizar inferências utilizando modelos mistos, consideram-se combinações lineares da forma:

$$L \begin{bmatrix} \beta \\ \nu \end{bmatrix}$$

Como qualquer combinação linear de ν é estimável, deve-se ter cuidado em buscar funções estimáveis apenas para os efeitos fixos do modelo.

Para matrizes L de rank 1, a estatística t geral é:

$$t = \frac{L \begin{bmatrix} \hat{\beta} \\ \hat{\nu} \end{bmatrix}}{\sqrt{L' \hat{C} L}}$$

Sob a suposição de normalidade de ν e ε , a estatística t segue uma distribuição t de Student, onde, especialmente em casos balanceados, seu grau de liberdade deve ser estimado. Assumindo que $\hat{\lambda}$ seja este grau de liberdade estimado, o intervalo com coeficiente de $100(1-\alpha)\%$ de confiança associado é o seguinte:

$$L \begin{bmatrix} \hat{\beta} \\ \hat{\nu} \end{bmatrix} \pm t_{\hat{\lambda}, \alpha/2} \sqrt{L' \hat{C} L}$$

Onde $t_{\hat{\lambda}, \alpha/2}$ é o valor da distribuição t para $\hat{\lambda}$ graus de liberdade e probabilidade $\alpha/2$.

Quando o rank de L é maior que 1, contrói-se a estatística F geral:

$$F = \frac{\begin{bmatrix} \hat{\beta} \\ \hat{v} \end{bmatrix} L(L'\hat{C}L)^{-1}L' \begin{bmatrix} \hat{\beta} \\ \hat{v} \end{bmatrix}}{\text{rank}(L)}$$

De forma análoga à estatística t, a estatística F segue uma distribuição F aproximada com rank(L) graus de liberdade no numerador e $\hat{\lambda}$ graus de liberdade no denominador.

O modelo misto geral simplifica e unifica diversas análises estatísticas comuns. Além da análise de dados espaciais e de modelos de efeitos aleatórios, o modelo misto é útil para análise de planejamento de experimentos de dados longitudinais, de planejamento em parcelas subdivididas, meta análises e modelos de variâncias heterogêneas. Além disso, extensões úteis do modelo misto estão disponíveis para modelos lineares generalizados e modelos não lineares.

4.2.3. Modelos mistos para dados espaciais

Em estatísticas espaciais, as observações podem estar correlacionadas em uma, duas ou três dimensões. Os exemplos mais freqüentes na área forrageira são dados em que se estuda a correlação em uma ou duas dimensões, na transecção ou no plano respetivamente. A inclusão de uma terceira dimensão, como a altura da forragem ou a profundidade dos cultivos é uma idéia interessante na hora de estudar estruturas de vegetação.

A maioria dos processos espaciais assume que os erros estão correlacionados tendo que ser estudada a forma dessa correlação expressa pela estrutura de covariância. As distâncias das observações das amostras, ou das parcelas experimentais, não são sempre eqüidistantes, nem com bordas

definidas, gerando as vezes grupos desbalanceados de solução mais complexa. O modelo misto tem características que fazem-no atrativo para este estudo de padrões espaciais. O modelo misto é baseado na identificação explícita de indivíduos, pelo que não exige dados balanceados; permite explicitar modelagem e análise de variação entre e dentro das unidades experimentais e freqüentemente os parâmetros individuais podem ter uma interpretação natural, relevante aos objetivos do estudo.

Para cada unidade experimental ij , o modelo misto tem a seguinte forma:

$$y_{ij} = X_{ij}\beta + Z_{ij}v_{ij} + \varepsilon_{ij}$$

Onde, β é considerado como fixo, v_{ij} e ε_{ij} são aleatórios, independentes entre si e um do outro, com $v_{ij} \sim N(0, G)$ e $\varepsilon_{ij} \sim N(0, R_{ij})$, sendo G uma matriz de covariância positiva definida ($p \times p$) e R_{ij} é uma matriz de covariância positiva definida ($n_{ij} \times n_{ij}$).

Assim, $y_{ij} \sim N(X_{ij}\beta, V_{ij})$, onde a matriz de variâncias e covariâncias de y_{ij} é: $\text{Var}(y_{ij}) = V_{ij} = Z_{ij}GZ'_{ij} + R_{ij}$

Neste trabalho, considerar-se-á a estrutura de covariância modelada através do R_{ij} .

Se definimos F como uma matriz $n \times n$ onde o elemento ij é $f(d_{ij})$, R tem duas formas:

Modelo sem efeito pepita: $R = \sigma^2 F$

Modelo com efeito pepita $R = I\sigma_1^2 + \sigma^2 F$

Sendo σ^2 a variância estimada chamada de “sill” ou patamar em estudo de dados espaciais, e σ_1^2 ou variância de microescala o “nugget”.

A estimação e inferência dos parâmetros do modelo misto normalmente é feita através do método de mínimo quadrados generalizados, dos métodos de máxima verossimilhança (ML) e máxima verossimilhança restrita (REML), ou através da estimação empírica de Bayes.

A estimação dos componentes de variância pode ser feita através dos métodos de ML ou de REML. Uma desvantagem no uso do método de ML particularmente no caso de experimentos espaciais, (onde p/n pode ser relativamente grande) é que as estimativas dos componentes de variância podem ser viesadas. Este viés, que é atribuído a perda de graus de liberdade da estimação de β , pode ser reduzido utilizando REML.

O método REML define as estimativas dos componentes de variância como a maximização da função de verossimilhança baseada nos $N - p$ contrastes de erros linearmente independentes, onde N é o total de observações de todas as unidades experimentais.

Neste trabalho, a estimação da variância e dos componentes de variância de R e G serão por REML, e as estimativas de β e υ pelas soluções das equações do modelo misto.

4.2.4. Estruturas de Covariância

No modelo misto, pode-se parametrizar diferentes estruturas de covariâncias para as matrizes G e R . Os modelos propostos adotam uma matriz de variância-covariância entre pontos heterogênea, que permite modelar o padrão de variação espacial, sendo necessário estimá-la. A escolha e o

ajuste desta matriz, estará então, determinada pela variação espacial presente em cada delineamento ou plano de amostragem.

A informação empírica que tenhamos, a estrutura dos dados e os programas computacionais disponíveis são importantes na definição da estrutura selecionada. Existem variadas estruturas de covariância que podem ser adequadas a diferentes condições, que são comentadas por Castro (1999). Em modelos espaciais, geralmente assume-se a covariância como uma função da distância entre as localizações s_i e s_j . Se a distância entre esses pontos for d_{ij} , os modelos resultantes seguem a forma geral:

$$\text{Cov}(e_i, e_j) = \sigma^2 [f(d_{ij})].$$

Os modelos para os quais $f(d_{ij})$ é a mesma para todas as localizações separadas pela mesma distância em qualquer direção, são chamados de modelos estacionários. Se adicionalmente, $f(d_{ij})$ não depende da direção, o modelo é chamado de isotrópico.

Neste trabalho, serão considerados modelos isotrópicos, e serão ajustados os semivariogramas linear, esférico, exponencial e gaussiano. Tradicionalmente nestes modelos não se considerava a ordenada na origem, dado que para distâncias iguais a zero, teoricamente corresponderiam semivariâncias iguais a zero. Em algumas situações, estes modelos não manifestam adequadamente as mudanças abruptas que acontecem em distâncias relativamente pequenas. Para modelar as modificações em pequena escala, geralmente se adiciona um outro parâmetro chamado de “nugget” ou efeito pepita.

Na escolha do modelo a ser ajustado para o semivariograma empírico, podem ser usados vários critérios.

4.2.5. Estimação e Seleção da Estrutura de Covariância

As variadas estruturas de covariâncias propostas, devem ser estimadas e um critério objetivo de seleção deve ser seguido.

A variância de y_{ij} foi expressada como $\text{Var}(y_{ij}) = V_{ij} = Z_{ij}GZ'_{ij} + R_{ij}$ e dada uma estrutura de covariância particular para G e R , pode-se considerar θ o vetor de parâmetros em V , com dimensão q . Geralmente θ é desconhecido, pode ser estimado por um procedimento de máxima verossimilhança (ML ou REML). Nestes casos, um algoritmo de otimização deve ser utilizado para estimar os componentes de θ . Um bom algoritmo deve realizar cálculos iterativos que não consumam muito tempo, deve ter rapidez e consistência de convergência e um bom método para avaliar dita convergência (Lindstrom e Bates, 1988). O algoritmo de Newton-Raphson possui boas características como algoritmo de otimização, assim como o algoritmo Scoring de Fisher, sendo muitas vezes superior ao algoritmo EM.

Para comparar varias estruturas de covariância o uso de técnicas estatísticas formais é necessário. Uma forma é maximizar critérios de informação REML tais como:

$$AIC_R = l_R(\hat{\theta}) - q$$

$$SBC_R = l_R(\hat{\theta}) - \frac{q}{2} \log(n - p)$$

$$HQC_R = l_R(\hat{\theta}) - q \log(n - p)$$

$$CAIC_R = l_R(\hat{\theta}) - \frac{q}{2}(\log(n-p) + 1)$$

onde q é o número de parâmetros de covariância, n é o número de observações e p é o rank da matriz de planejamento X . Esses critérios têm a forma de uma verossimilhança penalizada, onde a penalização é uma função do número de parâmetros de covariância. Neste trabalho, entre os testes que serão usados para a seleção do modelo, estão o teste da razão da verossimilhança o Critério de Informação de Akaike (AIC) e o Critério Bayesiano de Schwarz. (SBC). Os dois últimos critérios incluem uma penalização para o número de parâmetros de covariância a ser estimado, sendo a penalização do critério de Schwarz maior. O melhor modelo, será aquele que apresentar os maiores valores para os dois critérios.

Para avaliar a significância dos modelos de covariância frente a o modelo de erros independentes, será realizado o teste de $H_0 : \rho=0$. Calcula-se a diferença entre os critérios $-2REML_LL$ do modelo a testar e do modelo de erros independentes, e esta diferença tem uma distribuição χ^2 , sendo seus graus de liberdade a diferença entre os parâmetros estimados dos modelos a comparar.

4.3 Análise de dados espaciais em levantamentos por amostragem.

Na área de amostragem utiliza-se a teoria das variáveis regionalizadas com o objetivo de estimação de valores, inferências por médio de “kriggagem” e mapeamento das variáveis.

Os levantamentos por amostragem podem ser realizados sobre transeções lineares onde se assume variabilidade espacial em uma dimensão. Pode também ser sobre grades onde os pontos ou superfícies de amostragem estão localizados nas interseções de linhas e colunas, sendo as grades regulares ou não, e assumindo-se variabilidade espacial no plano ou duas dimensões. Também as amostragem podem ter um componente na terceira dimensão, quando sobre as grades anteriormente descritas se estuda a altura ou a profundidade de um processo.

A escala da amostragem, assim como a possibilidade de usar escalas múltiplas é fundamental para a descrição de processos e de padrões que podem ter diferentes manifestações segundo o padrão de estudo.

As descrições utilizando elementos de estatística descritiva e acrescentando componentes espaciais, como a localização de linhas e colunas, é também importante em estudos de amostragem.

O resumo da variabilidade espacial mediante o ajuste de semivariogramas, assim como o estudo da isotropia ou anisotropia quando possível, é relevante ao fenômeno estudado.

Todas as técnicas e comentários que foram realizados no caso de planejamento de experimentos com respeito a estes tópicos tem a mesma relevância em estudos de amostragem.

O ajuste de semivariogramas empíricos segue os mesmos lineamentos gerais, e o estudo da significância do modelo ajustado a diferentes distâncias ou "lag", podem ser realizado como no caso de modelos não lineares.

A diferença mais importante é que no caso de planejamento de experimentos o semivariograma integra uma matriz de variâncias-covariâncias, e no caso de levantamentos por amostragem o semivariograma é utilizado na predição e interpolação pelo método kriging, e na confecção de mapas.

4.3.1 Interpolação: Krigagem

Quando o interesse está na estimação de valores de pontos não amostrados usa-se um preditor que no caso de dados espaciais é o método de krigagem. Os pontos de amostragem utilizados para o krigagem podem ser tão pequenos como um ponto. Dado que as observações são dependentes a certas distâncias, utilizam-se os vizinhos de cada ponto para obter a estimativa, tendo cada vizinho um “peso” diferente segundo a distancia e a localização. Este preditor então é a média ponderada dos valores observados na vizinhança, ou seja

$$\hat{Z}_{(s_0)} = \sum_{i=1}^k \lambda_i z(s_i)$$

onde:

$\hat{Z}_{(s_0)}$ é o preditor para um ponto s_0 da região,

k é o número de vizinhos utilizados na predição,

λ_i é o peso atribuído a cada vizinho

$z(s_i)$ é o valor observado em cada vizinho

Existem diferentes tipos de krigagem: simples, ordinária, universal, indicadora e probabilística.

A característica própria da krigagem é o critério utilizado para obtenção dos pesos λ_i .

Se existe associação é fácil pensar que pontos mais próximos no espaço são mais semelhantes que pontos mais distantes, o que induz a pensar que qualquer critério de atribuição de pesos esteja relacionado com distâncias. No kriging, nem sempre a distância euclidiana é a distância mais adequada para definir o fenômeno. No caso de anisotropia, dois pontos separados por distâncias iguais ao ponto a ser estimado podem representar pesos diferentes na estimação em função da dependência da direção onde estão situados. É importante observar que a noção de distância deve incorporar de alguma forma a estrutura de variabilidade.

As distâncias entre os vizinhos, e as configurações de vizinhança onde pode existir equidistância dos vizinhos ou não, fazem com que as ponderações de cada um seja diferente.

O preditor de krigagem considera estas situações incluindo covariogramas ou semivariogramas na determinação dos pesos (Apêndice 7).

No método de Krigagem por blocos, estimam-se vários pontos de interesse, no somente um s_0 , para obter uma estimativa média e desse modo aumentar a precisão.

4.3.2. Mapas

A construção de mapas que facilitem uma idéia visual do comportamento da variável na área é um dos objetivos dos trabalhos de amostragem. Os mapas são construídos a partir de uma malha mais fina do que a malha

amostrada, obtida por aplicações sucessivas da krigagem. Os mapas podem ser tridimensionais ou bidimensionais como mapas de isolinhas. Existem varios aplicativos computacionais específicos desenvolvidos para a geração de mapas, sendo este um produto em constante modificação.

5. RESULTADOS E DISCUSSÃO

5.1 Experimento 1

Foi estudada a existência de associação espacial de todas as variáveis em todos os anos. O resumo dos resultados encontra-se no Apêndice 3. As variáveis Matéria Seca Total (MST) e Matéria Seca de Gramíneas (MSGRAM), no segundo ano apresentaram uma associação significativa a um padrão espacial. Faz-se uma descrição completa destas variáveis.

5.1.1 Análise descritiva

Uma análise descritiva geral foi realizada nas variáveis MST e MSGRAM no segundo ano, para resumir o comportamento dos dados. Na Tabelas 4 apresenta-se a análise descritiva dos resíduos do ajustamento.

Tabela 4. Estatísticas descritivas dos resíduos de MST e MSGRAM no segundo ano.

Estatística	Variável	
	MST	MSGRAM
No de dados	40	40
Média	0	0
Mínimo	-739,67	-504,92
Máximo	838,45	649,95
Mediana	38,28	-8,67
Quartil inferior	-215,08	-216,24
Quartil superior	211,46	181,36
Variância	95954,3	76281,0
Desvio padrão	309,77	276,19
Coeficiente de Variação	-	-
Coeficiente de Assimetria	0,12	0,22
Coeficiente de Curtose	0,67	-0,48

Nas Figuras 4 e 5 estão apresentados os gráficos de probabilidade normal dos resíduos. Não foi rejeitada a suposição de normalidade, sendo os

valores da estatística do teste de Shapiro-Wilks igual a 0,9783 ($Pr < W = 0,7312$) para MST e 0,9799 ($Pr < W = 0,7848$) para MSGRAM.

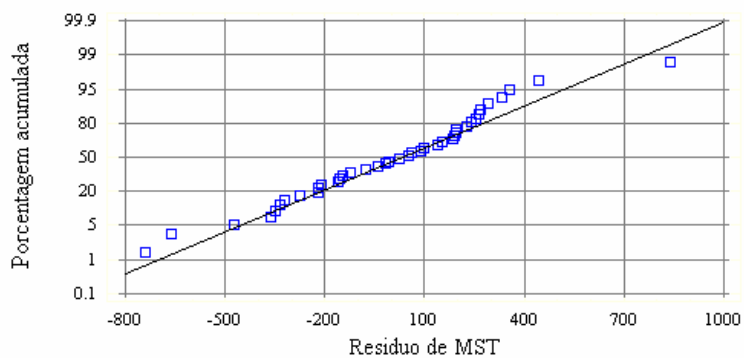


Figura 4. Gráfico de probabilidade normal dos resíduos da variável MST no segundo ano

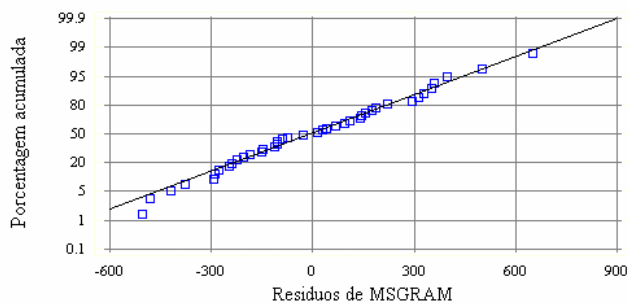


Figura 5. Gráfico de probabilidade normal dos resíduos de MSGRAM no segundo ano.

Nas figuras 6 a 9 encontram-se os histogramas e os “box-plot” dos resíduos de MST e MSGRAM

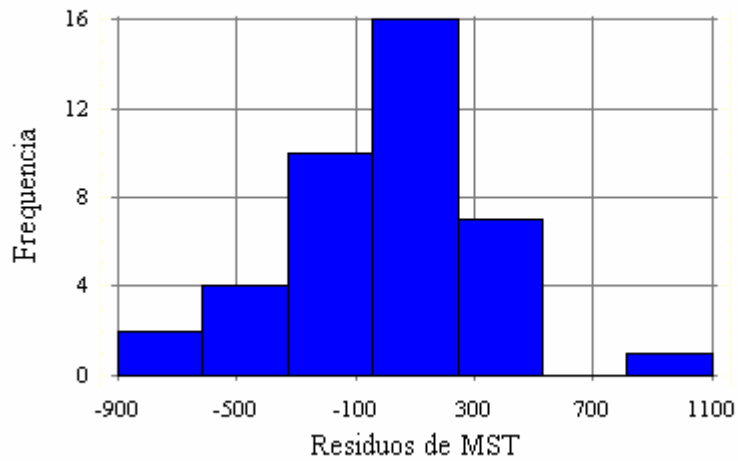


Figura 6. Histograma dos resíduos da variável MST no segundo ano.

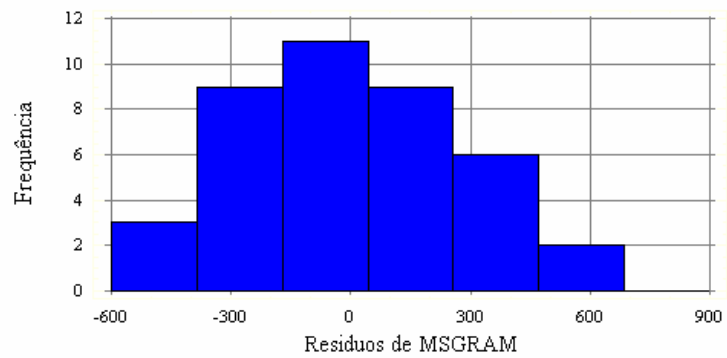


Figura 7. Histograma dos resíduos da variável MSGRAM no segundo ano.

Os histogramas de frequências e os “box-plot” são ferramentas descritivas univariadas que auxiliam na detecção de valores atípicos.

Verifica-se nas figuras 6 a 9 que nos resíduos de MST e MSGRAM não existem valores atípicos que poderiam causar perturbações na análise.

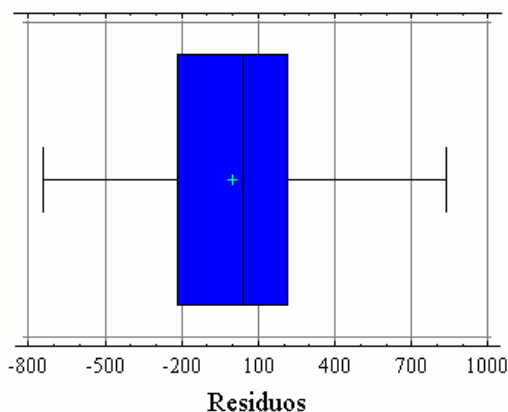


Figura 8. "Box-Plot" dos resíduos de MST no segundo ano.

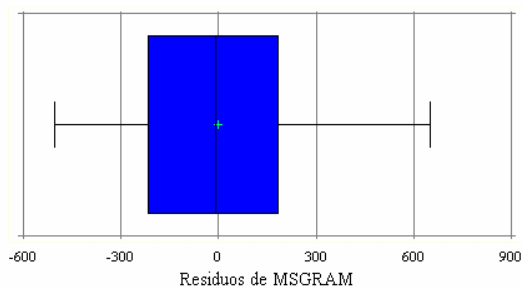


Figura 9. "Box-Plot" dos resíduos de MSGRAM no segundo ano.

Os gráficos apresentados nas Figuras 10 a 13 foram construídos com o objetivo de detetar o comportamento dos resíduos em relação às direções da amostragem realizada. Para cada linha e cada coluna foram atribuídos os resíduos encontrados. Tendência e heterogeneidade de variâncias que

coincidam com as coordenadas da amostragem podem ser identificadas desta forma. Neste caso, não se observam problemas de associação nas coordenadas X ou Y, uma vez que as linhas e colunas mostram valores médios próximos e dispersões semelhantes entre si, nas duas variáveis. Muito embora os gráficos de linhas estão construídos com 8 pontos para cada linha e os de colunas por 5 pontos para cada uma, essa diferença não é importante na descrição do fenômeno.

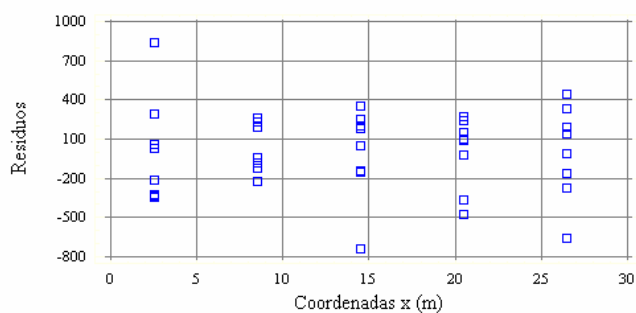


Figura 10 . Valores dos resíduos da MST por linhas da superfície amostrada.

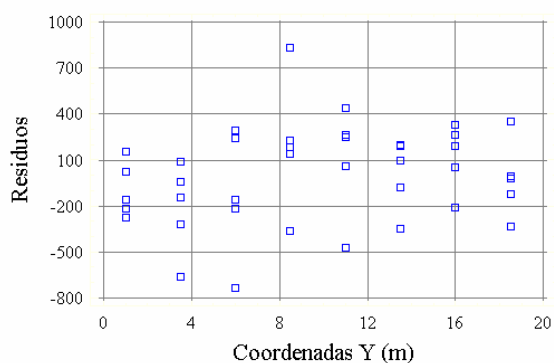


Figura 11. Valores dos resíduos da MST por colunas da superfície amostrada

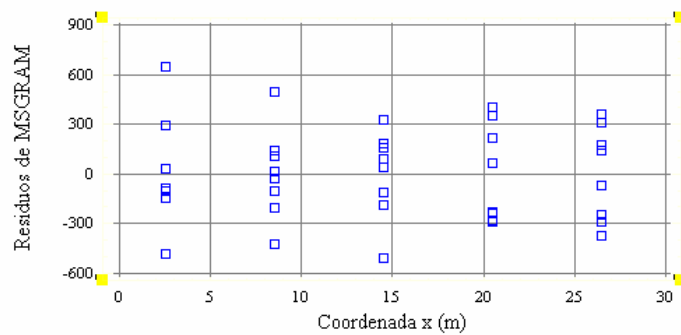


Figura 12. Valores dos resíduos de MSGRAM por linhas da superfície amostrada.

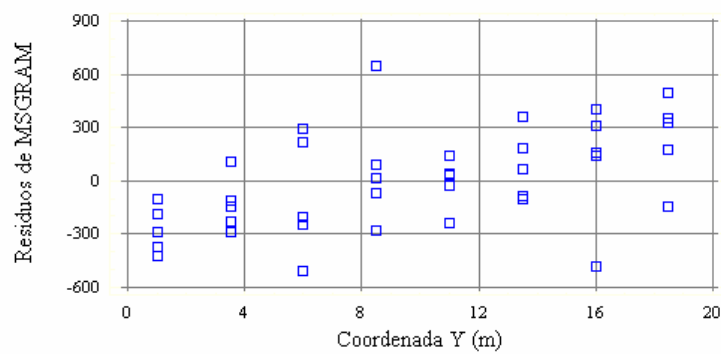


Figura 13. Valores dos resíduos de MSGRAM por coluna da superfície amostrada.

Estas avaliações descritivas não evidenciam perturbações no comportamento dos resíduos das variáveis estudadas, não indicando, portanto, a existência de problemas que invalidem suposições de estacionaridade.

5.1.2. Semivariogramas ajustados

No estudo da variação espacial de cada variável em cada ano, foram consideradas as posições dos resíduos no espaço para o cálculo das semivariâncias, utilizando-se o estimador:

$$\hat{\gamma}(h) = \sum_n \frac{(R_{ij} - R_{s,t})^2}{2N(h)}$$

onde:

R_{ij} é o resíduo situado nas coordenadas (i,j) da área em estudo;

R_{st} é o resíduo situado nas coordenadas (s,t) da área em estudo, separado de R_{ij} pela distância h ;

$N(h)$ é o número de pares de resíduos que estão separados pela distância h .

Uma vez calculadas as semivariâncias empíricas, quatro modelos de variação espacial: esférico, exponencial, gaussiano e linear foram ajustados considerando a presença ou ausência de nugget, e foram comparados com o modelo de independência na variação, pelos critérios de Akaike, Schwartz e $-2REML LL$, como salientado em 4.2.5.

Na variável MST, ajustaram significativamente todos os semivariogramas sem nugget, e o esférico e gaussiano com nugget, sendo não significativa a sua contribuição. Na variável MSGRAM, ajustaram significativamente os semivariogramas esférico, exponencial e gaussiano sem e com nugget.

Os modelos ajustados para MST e MSGRAM apresentam-se, respectivamente, nas Tabelas 5 e 6.

Tabela 5. Coeficientes dos semivariogramas ajustados na variável MST no segundo ano, e valores dos critérios utilizados na seleção do modelo.

Característica	Modelos sem nugget				Modelos com nugget				
	SPH	EXP	GAU	LIN	SPH	EXP (1)	GAU	LIN	IND.
Nugget					6423,6		6732	0	
Range	5,1*	1,9*	2,4*	2,4*	5,3*		2,5*	0	
Sill	96715*	96987*	97076*	97065*	89978*		90322*	95954	95954
Akaike	-281	-281	-281	-281	-281		-282	-284	-282
Schwarz's	-282	-283	-282	-282	-284		-284	-286	-283
-2REML LL	557	559	557	557	557		557	562	562
Valor de $\chi^2_{(1)}$	5	5	5	5	5		5	0	

(1) O modelo exponencial com nugget não convergiu.

(*) P < 0,05.

Tabela 6. Coeficientes dos semivariogramas ajustados na variável MSGRAM no segundo ano, e valores dos critérios utilizados na seleção do modelo.

Característica	Modelos sem nugget				Modelos com nugget				
	SPH	EXP	GAU	LIN	SPH	EXP	GAU	LIN	IND.
Nugget					28.3+	12	5.5*	0	
Range	9,6	4,2	2,5	0	30962*	23437	32424*	72814	
Sill	105746	89398*	75507*	76281	89942+	94076*	49688*	3467	76281
Akaike	-275	-274	-275	-278	-274	-274	-275	-279	-277
Schwarz's	-277	-275	-277	-280	-276	-277	-277	-282	-278
-2REML LL	546	543	547	553	543	543	544	553	553
Valor de $\chi^2_{(1)}$	7	10	6	0	10	10	9	0	

Em função dos resultados, os semivariogramas recomendados para correção seriam o esférico sem nugget para a inclusão no modelo misto de análise de variancia na variavel MST e o modelo exponencial sem nugget para a variavel MSGRAM. Embora estes resultados, também foram estudados os semivariogramas esfericos, exponencial, gaussiano e lineal para comparação.

5.1.3. Análise de Variância considerando a variabilidade espacial

Para complementação de resultados considerou-se a estrutura de contrastes ortogonais para cultivos de trevo branco apresentada na Tabela 7.

Na Tabela 8 apresentam-se os valores das médias, dos erros padrões e dos coeficientes de variação por cultivar da variável MST, considerando-se os diferentes modelos. Observa-se uma diminuição no erro padrão, mais não uma diferença na significancia, possivelmente pelo fato de serem todas sigificativamente diferente de zero já na situação sem correção.

Tabela 7. Contrastes ortogonais para cultivares de trevo branco.

Contraste 1	(Regal +Yi + Bayucué + Huia + Pitau + Bagé) vs (Guaíba S1 + Jacuí S2)
Contraste 2	Guaíba S1 Vs Jacuí S2
Contraste 3	(Regal) vs (Yi + Bayucué + Huia + Pitau + Bagé)
Contraste 4	(Yi + Bayucué) vs (Huia + Pitau + Bagé)
Contraste 5	Yi Vs Bayucué
Contraste 6	(Huia + Pitau) vs (Bagé)
Contraste 7	Huia Vs Pitau

Os valores estimados das médias das cultivares também estão corrigidos, uma vez que no cálculo utiliza-se a distância a que as cultivares encontram-se no experimento. Os coeficientes de variação são sempre menores com qualquer uma das correções espaciais, sendo em média um 19% menores que o coeficiente de variação médio da situação sem correção.

Tabela 8. Resumo por cultivar da variável MST segundo o modelo de estimação empregado. Apresentam valores de médias, erros padrão e coeficiente de variação.

Cultivar	1 (indep)	2 (sph)	3(exp)	4(gau)
Regal	2585,8 163,5 6,32	2508,5 129,7 5,17	2544,3 135,7 5,33	2507,6 129,3 5,16
Yi	3539,4 163,5 4,62	3573,3 131,9 3,68	3537,2 137,2 3,88	3571,1 131,7 3,69
Bayucuá	3086,2 163,5 5,29	3026,1 130,4 4,31	3043,2 137,1 4,51	3024,1 130,3 4,31
Huia	2465,8 163,5 6,63	2519,8 128,0 5,08	2473,6 136,4 5,51	2518,3 127,8 5,07
Pitau	2279,4 163,5 7,17	2341,8 131,2 5,60	2337,4 136,9 5,86	2345,4 131,0 5,59
Bagé	3109,2 163,5 5,26	3081,2 128,7 4,18	3100,2 135,9 4,38	3081,4 128,4 4,17
Guaíba S1	3407,0 163,5 4,80	3355,9 130,5 3,89	3359,5 137,5 4,09	3354,9 130,2 3,88
Jacuí S2	3705,2 163,5 4,41	3739,2 130,4 3,49	3729,7 136,1 3,65	3740,0 130,2 3,48

Todas as médias são significativamente diferentes de zero.

Na Tabela 9 apresentaram-se os resultados dos contrastes ortogonais de interesse para a variável MST. As comparações testadas mostram que nos contrastes 2, 3 e 5 existe uma modificação na significância, resultando significativas quando incorporada alguma correção espacial, não existindo diferenças importantes entre elas. Nos outros contrastes, as tendências se

mantiveram. Essa diferença resulta de uma diminuição nos erros padrões dos contrastes que foram em média de 24,6% menores quando comparadas ao modelo independente, sem correção espacial.

Tabela 9. Contrastes da variável MST para os 4 modelos ajustados. Para cada modelo se apresentam as estimativas do contraste, os erros padrões do contraste, e o nível mínimo de significância.

CONTRASTE	1(indep)	2(sph)	3(exp)	4(gau)
Contraste 1	-8541,6 1601,9 0,0001	-8469,4 1268,5 0,0001	-8463,5 1290,2 0,0001	-8473,6 1260,6 0,0001
Contraste 2	-298,2 231,2 0,2077	-383,2 168,4 0,0307	-370,22 177,7 0,0466	-385,0 167,0 0,0288
Contraste 3	-1551,0 895,5 0,0943	-2000,2 666,7 0,0056	-1769,8 693,8 0,0165	-2002,0 662,1 0,0053
Contraste 4	4168,0 895,5 0,0001	3912,3 636,1 0,0001	3918,8 675,5 0,0001	3895,1 630,7 0,0001
Contraste 5	453,2 231,2 0,0600	547,2 184,4 0,0061	494,1 187,0 0,0134	547,0 183,4 0,0059
Contraste 6	-457,0 400,5 0,2635	-383,4 290,3 0,1974	-490,5 306,5 0,1209	-390,2 288,0 0,1863
Contraste 7	186,4 231,2 0,4269	177,9 165,7 0,2921	136,1 175,5 0,4447	172,8 164,3 0,3019

Na Tabela 10 , apresenta-se o resumo por cultivar da variável MSGRA, que mantém a mesma tendência de resultados da variável MST. Neste caso, a diminuição do coeficiente de variaçãp foi em média de 13,7 %. Os contrastes estudados (Tabela 11) apresentam uma diminuição em média de 33,3% nos erros padrões, em relação ao modelo independente. Neste caso, embora exista uma diminuição do desvio padrão, o nível de significancia dos contrastes não variou com respeito ao caso de erros independentes.

Tabela 10. Resumo por cultivar da variável MSGRA segundo o modelo de estimação empregado. Apresentam-se médias, erro padrão e coeficiente de variação.

<i>Cultivar</i>	<i>1 (indep)</i>	<i>2 (sph)</i>	<i>3(exp)</i>	<i>4(gau)</i>
Regal	1891,2 145,7 7,70	1794,2 123,9 6,91	1831,0 134,1 7,32	1837,6 112,3 6,11
Yi	2300,0 145,7 6,33	2254,9 125,8 5,58	2268,2 138,7 6,11	2304,3 114,9 4,99
Bayucua	2007,0 145,7 7,26	1998,6 130,2 6,51	1991,6 137,7 6,91	2002,2 113,2 5,65
Huia	1830,6 145,7 7,96	1803,5 128,2 7,11	1789,4 137,3 7,67	1818,8 110,7 6,09
Pitau	1654,0 145,7 8,81	1812,9 129,9 7,17	1769,4 137,4 7,77	1725,0 114,0 6,61
Bagé	2147,4 145,7 6,78	2119,6 124,6 5,88	2126,4 134,9 6,34	2108,1 111,2 5,27
Guaíba S1	2218,4 145,7 6,57	2218,2 126,2 5,69	2197,3 138,9 6,32	2207,4 113,3 5,13
Jacuí S2	2218,6 145,7 6,57	2232,7 126,7 5,67	2229,5 134,4 6,03	2251,5 113,1 5,02

Todas as médias são significativamente diferentes de zero.

Tabela 11. Contrastes da variável MSGRAM para os 4 modelos ajustados. Para cada modelo se apresentam os valores do contraste, do desvio estándar do contraste e nível mínimo de significância.

<i>CONTRASTE</i>	<i>1 (indep)</i>	<i>2 (sph)</i>	<i>3(exp)</i>	<i>4(gau)</i>
Contraste 1	-2961,6 1428,2 0,0474	-3138,2 982,8 0,0035	-3009,1 1005,4 0,0057	-3161,4 1092,1 0,0073
Contraste 2	-0,2 206,1 0,9992	-14,5 128,0 0,9105	-32,1 134,1 0,8121	-44,1 143,1 0,7302
Contraste 3	-483,0 798,4 0,5501	-1018,6 508,5 0,0549	-790,1 523,7 0,1426	-770,5 568,5 0,1862
Contraste 4	1657,0 798,4 0,0473	1288,5 468,8 0,0104	1409,1 500,0 0,0088	1615,5 538,8 0,0056
Contraste 5	293,0 206,1 0,1663	256,2 140,2 0,0785	276,5 146,5 0,0695	302,0 159,0 0,0679
Contraste 6	-140,2 357,0 0,6976	-325,5 214,4 0,1402	-316,9 227,6 0,1748	-195,5 245,9 0,4333
Contraste 7	176,6 206,1 0,3989	-9,4 122,0 0,9391	20,0 129,8 0,8785	93,8 140,1 0,5088

As reduções observadas tanto no coeficiente de variação (19 % para MST e 13,7 % MSGRAM) como para os erros padrões dos contrastes (24,6% para MST é 33,3% para MSGRAM) são menores daquelas verificadas por Cullis e Gleeson (1989) que observaram uma redução de 42% das variâncias das medias dos rendimentos das cultivares. Estudos de simulação seriam necessarios para auxiliar na identificação dessas diferenças.

5.2 Experimento 2.

Foi estudada a existencia de associação espacial da variável porcentagem de implantação (%IMPL) em todos os momentos de corte. O resumo dos resultados apresentam-se no Apêndice 4. A variável %IMPL, no segundo corte apresenta uma associação significativa a um padrão espacial.

5.2.1 Análise descritiva

Como forma de resumo do comportamento dos dados foi realizada uma análise descritiva geral para os resíduos, sendo os resultados apresentados na Tabela 12.

Tabela 12. Estatísticas descritivas dos resíduos de %IMPL no segundo corte.

Estatística	%IMPL
No de dados	60
Média	0
Mínimo	-0,27
Máximo	0,43
Mediana	-0,004

Quartil inferior	-0,07
Quartil superior	0,04
Variância	0,019
Desvio padrão	0,14
Coefficiente de Assimetria	0,74
Coefficiente de Curtose	1,53

Na Figura 14 apresenta-se o gráfico de probabilidade normal dos resíduos, onde percebe-se alguma discrepância em relação a distribuição normal (Test $W=0,94$; Prob < 0,02).

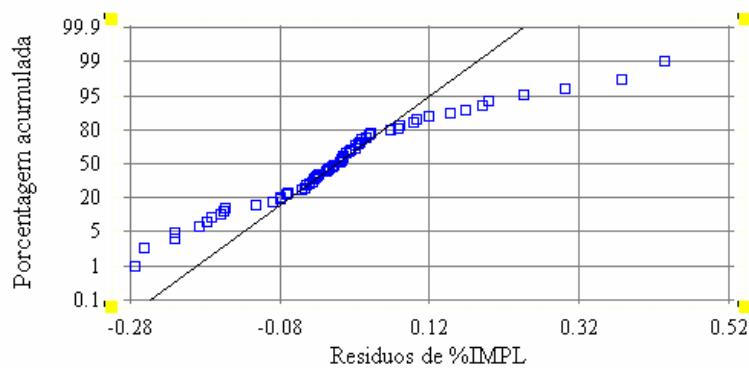


Figura 14. Gráfico de probabilidade normal dos resíduos de %IMPL no segundo corte.

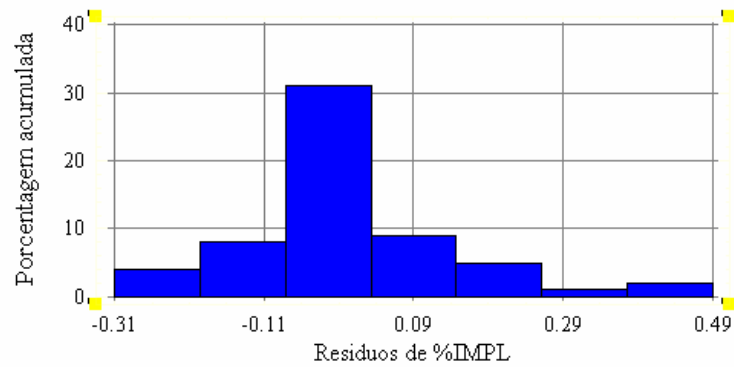


Figura 15. Histograma dos resíduos da variável %IMPL no segundo corte.

Nas Figuras 15 e 16 encontram-se os histogramas e os “box-plot” dos resíduos da variável %IMPL.

A Figura 16 indica alguns valores atípicos. Seis pontos são considerados fora da amplitude interquartilica. Esses pontos não foram retirados da análise.

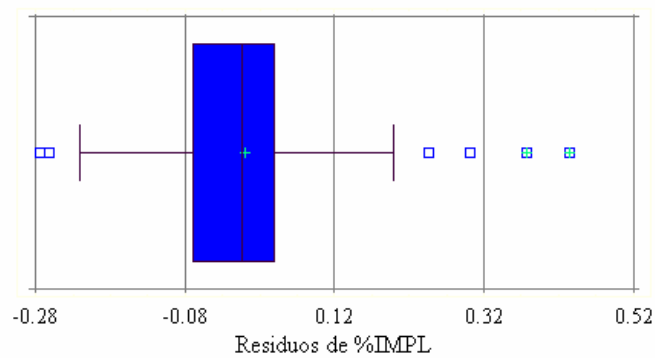


Figura 16. “Box-plot” dos resíduos da variável %IMPL.

Nas Figuras 17 e 18 não são observadas tendências associadas às posições espaciais de linha e coluna.

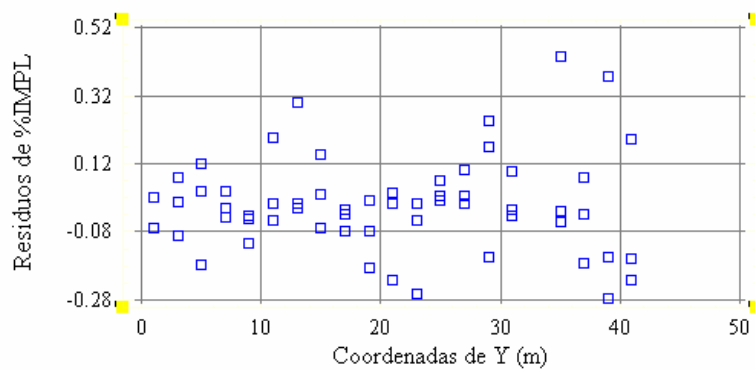


Figura 17. Valores dos resíduos da variável %IMPL por colunas da superfície amostrada.

Na Figura 17, pode-se observar uma maior concentração de pontos com valores altos de resíduos à direita do gráfico, na coordenada Y = 40.

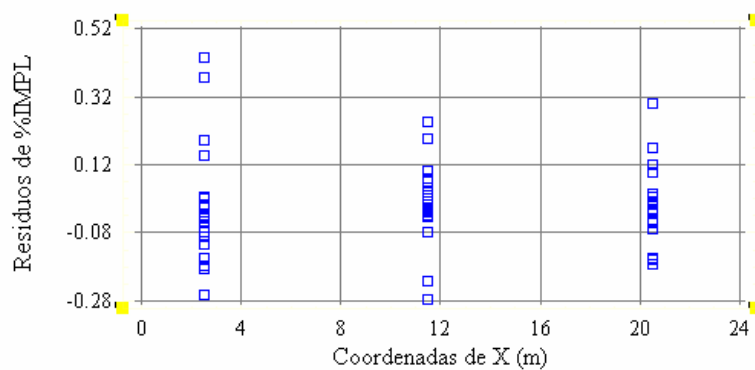


Figura 18. Valores dos resíduos da variável %IMPL por linhas da superfície amostrada.

5.2.2. Semivariogramas ajustados

Como no caso do Experimento 1, foram ajustados os semivariogramas exponencial, esférico, gaussiano e linear, com e sem nugget, aos resíduos do ajustamento. Os modelos ajustados para %IMPL apresentam-se na Tabela 13.

Tabela 13. Coeficientes dos semivariogramas ajustados na variável %IMPL no segundo corte, e valores dos criterios utilizados na seleção do modelo.

Caraterística	<i>Modelos sem Nugget</i>				<i>Modelos com Nugget</i>				
	<i>SPH</i>	<i>EXP</i>	<i>GAU</i>	<i>LIN</i>	<i>SPH</i>	<i>EXP</i>	<i>GAU</i>	<i>LIN</i>	<i>INDEP</i>
Nugget	-	-	-	-	0.01	0.01	0.01	0.02	-
Range	3.17	1.44	1.53	0	12.7	4.17	7.17	0	-
Sill	0.02	0.02	0.02	0.02	0.008	0.009	0.008	0	0.02
Akaike	29.9	30.3	29.9	28.7	30.5	29.9	30.5	27.7	29.7
Schwarz's	27.8	28.3	27.8	26.7	27.4	26.8	27.3	24.6	28.7
-2REML LL	-63.8	-64.7	-63.8	-61.5	-67.1	-65.8	-66.9	-61.5	-61.5
Valor de $\chi^2_{(1)}$	2.3	3.2	2.3	0	5.6*	4.3	5.4	0	

O semivariograma esférico com nugget foi significativamente diferente quando comparado com o modelo de erros independentes. A avaliação do ajuste do modelo, se realizou por test χ^2 neste caso com 1 g.l.

5.2.3. Análise de Variância considerando a variabilidade espacial

Foram planejados 6 contrastes ortogonais orientdos a responder questionamientos biológicos, os quais são apresentados na Tabela 14.

Tabela 14. Contrastes ortogonais do experimento 2.

Contraste 1 (Leguminosas vr Gramíneas)	$(C1+C2+C3+C4+C5+C6+C7+C8+C9+C10+C11+C12+C13+C14+C15+C16)$ vr $(C17+C18+C19+C20)$
Contraste 2 (Lotus vr T.alexandrinum)	$(C3+C8+C10+C13+C14+C16)$ vr C4
Contraste 3 (Ornithopus vr T. Vesiculosum)	C1 vr C7
Contraste 4 (Gramíneas perennes vr Gramínea anual)	$(C17+C18+C19)$ vr C20
Contraste 5 (Entre gramíneas perennes)	$(C17+C18)$ vr C19
Contraste 6 (T. branco vr Resto de Leguminosas)	$(C2+C6+C9+C11+C12)$ vr $(C1+C3+C4+C5+C7+C8+C10+C13+C14+C15+C16)$

Na Tabela 15 apresentam-se os resultados dos contrastes, segundo o modelo de independência de erros e os modelos que incorporam os semivariogramas ajustados.

Existe modificação na significância no caso de quatro cultivares, sendo também neste caso o modelo gaussiano de semivariograma que registra o maior número de diferenças.

Tabela 15. Resumo por cultivar da variável %IMPL segundo o modelo de estimação empregado. Apresentam-se médias, erros padrões, nível mínimo de significância. e coeficiente de variação.

<i>Cultivar</i>	<i>1 (indep)</i>	<i>2 (sph)</i>	<i>3(exp)</i>	<i>4(gau)</i>
C1	0.0982 0.0998 0.3314	0.0724 0.0387 0.0694	0.0653 0.0476 0.1778	1.7498 0.0280 0.0001
C2	0.8065 0.0998 0.0001	0.7626 0.0382 0.0001	0.7555 0.0469 0.0001	2.4096 0.0280 0.0001
C3	0.5156 0.0998 0.0001	0.5653 0.0374 0.0001	0.5524 0.0467 0.0001	2.1349 0.0279 0.0001
C4	0.3367 0.0998 0.0017	0.2889 0.0380 0.0001	0.29069 0.0467 0.0001	2.0727 0.0280 0.0001
C5	0.5613 0.0998 0.0001	0.5029 0.0376 0.0001	0.4936 0.0462 0.0001	2.0962 0.0280 0.0001
C6	0.2336 0.0998 0.0246	0.1622 0.0382 0.0001	0.1597 0.0471 0.0016	1.6347 0.0279 0.0001
C7	0.1810 0.0998 0.0776	0.0650 0.0378 0.0938	0.0857 0.0460 0.0706	1.7597 0.0279 0.0001
C8	0.1812 0.0998 0.0773	0.2410 0.0378 0.0001	0.2269 0.0469 0.0001	2.1532 0.0280 0.0001
C9	0.2925 0.0998 0.0057	0.3103 0.0369 0.0001	0.2986 0.0459 0.0001	1.7551 0.0279 0.0001
C10	0.4654 0.0998 0.0001	0.3257 0.0371 0.0001	0.3391 0.0457 0.0001	1.8032 0.0279 0.0001
C11	0.3605 0.0998 0.0009	0.2921 0.0373 0.0001	0.2934 0.0458 0.0001	1.6159 0.0279 0.0001
C12	0.5305 0.0998 0.0001	0.5555 0.0377 0.0001	0.5460 0.0470 0.0001	2.2700 0.0280 0.0001
C13	0.0520 0.0998 0.6053	0.0932 0.0372 0.0168	0.0848 0.0464 0.0758	1.5588 0.0279 0.0001
C14	0.4428 0.0998 0.0001	0.3949 0.0369 0.0001	0.3894 0.0460 0.0001	1.9636 0.0279 0.0001
C15	0.2741 0.099 0.0092	0.2264 0.0369 0.0001	0.2202 0.0458 0.0001	1.9241 0.0280 0.0001
C16	0.4785 0.0998 0.0001	0.5523 0.0384 0.0001	0.5373 0.04713 0.0001	2.1220 0.0279 0.0001

Tabela 15.

<i>Cultivar</i>	<i>1 (indep)</i>	<i>2 (sph)</i>	<i>3(exp)</i>	<i>4(gau)</i>
C17	0.4161 0.0998 0.0002	0.4137 0.0472 0.0001	0.3808 0.0530 0.0001	1.8466 0.0278 0.0001
C18	0.6742 0.0998 0.0001	0.7964 0.0474 0.0001	0.7613 0.0532 0.0001	2.3137 0.0281 0.0001

C19	0.4150 0.0998 0.0002	0.3927 0.0470 0.0001	0.3671 0.0532 0.0001	1.9148 0.0279 0.0001
C20	0.5895 0.0998 0.0001	0.7513 0.0482 0.0001	0.7118 0.0541 0.0001	2.7386 0.0280 0.0001

Na Tabela 16 apresentam-se os resultados dos 6 contrastes ortogonais descritos. As comparações testadas mostram que nos contraste 2, 4 e 5 existe uma modificação na significância. Estes contrastes resultaram significativos aõ incorporar alguma correção espacial. Nos outros contrastes, as tendencias se mantiveram.

Tabela 16. Contrastes da variável %IMPL para os 4 modelos ajustados. Para cada modelo apresentam-se as estimativas do contraste, erros padrões dos contrastes e nível mínimo de significância.

CONTRASTE	1 (indep)	2 (sph)	3 (exp)	4 (gau)
Contraste 1	-10.2728 3.5721 0.0066	-16.0243 2.9080 0.0001	-14.1834 3.0282 0.0001	-16.9254 0.2303 0.0001
Contraste 2	0.1153 0.6470 0.8595	0.4387 0.1823 0.0211	0.3859 0.1921 0.0517	-0.7006 0.0044 0.0001
Contraste 3	-0.0828 0.1412 0.5608	0.0073 0.0419 0.8614	-0.0203 0.0440 0.6460	-0.0098 0.0005 0.0001
Contraste 4	-0.2631 0.3458 0.4515	-0.6510 0.0824 0.0001	-0.6262 0.0876 0.0001	-2.14074 0.0090 0.0001
Contraste 5	0.2603 0.2445 0.2938	0.4246 0.0668 0.0001	0.4078 0.0701 0.0001	0.3308 0.0087 0.0001
Contraste 6	-6.5262 2.9619 0.0337	-6.2698 0.7191 0.0001	-6.1582 0.7599 0.0001	0.1522 0.0072 0.0001

5.3 Experimento 3

Foi estudada a existencia de associação espacial das variáveis porcentagem de cobertura (%COB) de três espécies provenientes de uma amostragem de campo nativo. Para estudar o efeito da escala de amostragem na determinação da variabilidade espacial, foram ajustados semivariogramas para cada espécie para quatro tamanhos de unidade amostral (0,1 x0,1 m ; 0,1 x 0,3 m; 0,1 x 0,5 m; 0,1 x 1,0 m). Estes tamanhos foram escolhidos por serem de uso frequente em estudos de vegetação.As frequencias específicas foram

calculadas a partir da transecção contínua inicial. Assim tem-se 400 unidades amostrais de 0,1 x 0,1 m; 133 unidades de 0,1 x 0,3 m; 80 unidades de 0,1 x 0,5 m e 40 unidades de 0,1 x 1,0 m.

Para cada espécie foram ajustados semivariogramas exponencial, esférico, gaussiano e linear, sendo os resultados apresentados na Tabela 17 e nas Figuras 19 a 21.

Tabela 17. Resultados do ajuste dos semivariogramas por espécie e por tamanho de unidade amostral. Se apresentam os valores calculados de R^2 e dos coeficientes a (range) e $c_{(s)}$ (sill)

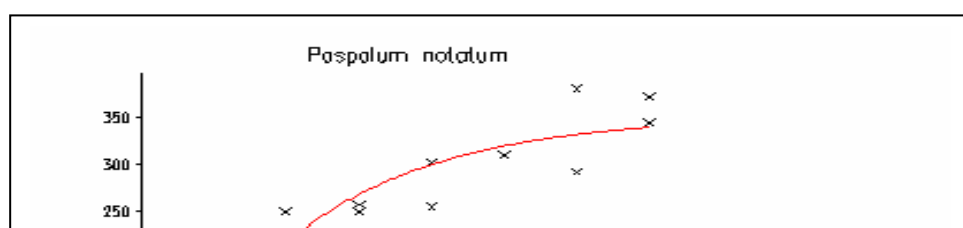
<i>Aristida laevis</i>		0,1 x 0,1 m	0,1 x 0,3 m	0,1 x 0,5 m	0,1 x 1,0 m
XP	R^2	91,3 %	6,0%	2,0%	Não ajusta
	a	0,2217	0,1912	0,236	
	$c_{(s)}$	445	2638	194	
ESF	R^2	95,0 %	13,6%	Não ajusta	Não ajusta
	a	0,4341	0,58		
	$c_{(s)}$	403,5	2638		
GAUS	R^2	95,1 %	14,5%	3,3 %	Não ajusta
	a	0,1869	0,285	0,368	
	$c_{(s)}$	400,2	2643	194,1	
LIN	R^2	20,0 %	Não ajusta	Não ajusta	Não ajusta
	h	757,8			

Continuação Tabela 17.

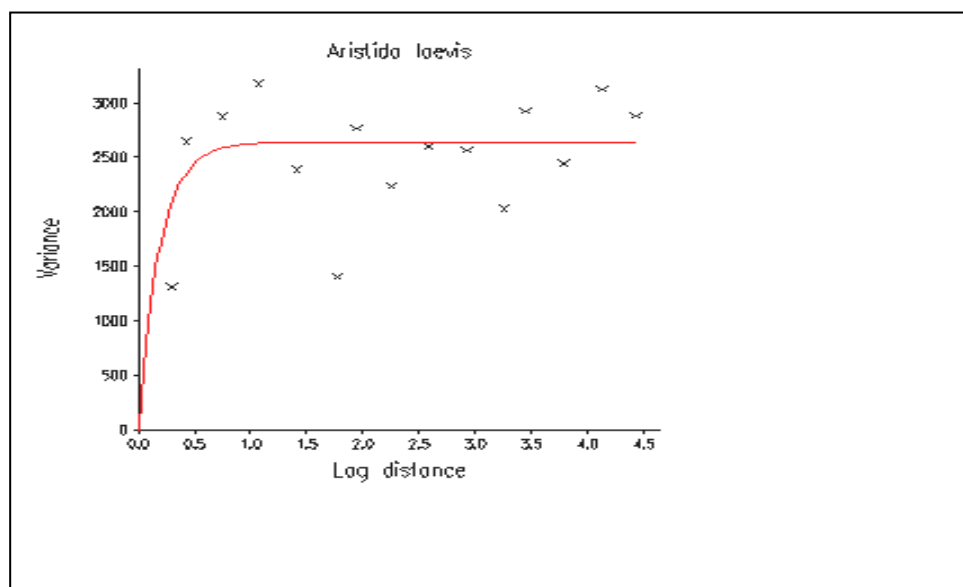
<i>Paspalum notatum</i>		0,1 x 0,1 m	0,1 x 0,3 m	0,1 x 0,5 m	0,1 x 1,0m
XP	R^2	84,1 %	11,5 %	1,8 %	Não ajusta
	a	0,2093	0,2672	0,258	
	$c_{(s)}$	352	1960	165,4	
ESF	R^2	68,6 %	Não ajusta	Não ajusta	Não ajusta
	a	0,4266			
	$c_{(s)}$	0,1519			
GAUS	R^2	58,7 %	18,2 %	4,0 %	No ajusta
	a	0,1519	0,3696	0,3949	
	$c_{(s)}$	307,9	1963	165,6	
LIN	R^2	Não ajusta	Não ajusta	Não ajusta	Não ajusta
	h				
<i>Desmodium incanum</i>		0,1 x 0,1 m	0,1 x 0,3 m	0,1 x 0,5 m	0,1 x 1,0m

XP	R^2	Não ajusta	36,1 %	58,0 %	52,6
	a		0,561	1,118	1,759
	$C_{(s)} \cdot$		1217	122	93,11
ESF	R^2	6,0 %	30,5 %	52,5 %	40,2
	a	0,1695	1,184	2,448	2,942
	$C_{(s)} \cdot$	162,4	1178	119,6	88,19
GAUS	R^2	Não ajusta	30,5 %	46,6 %	38,9
	a		0,498	1,005	1,258
	$C_{(s)} \cdot$		1172	118,9	87,85
LIN	R^2	Não ajusta	Não ajusta	Não ajusta	Não ajusta

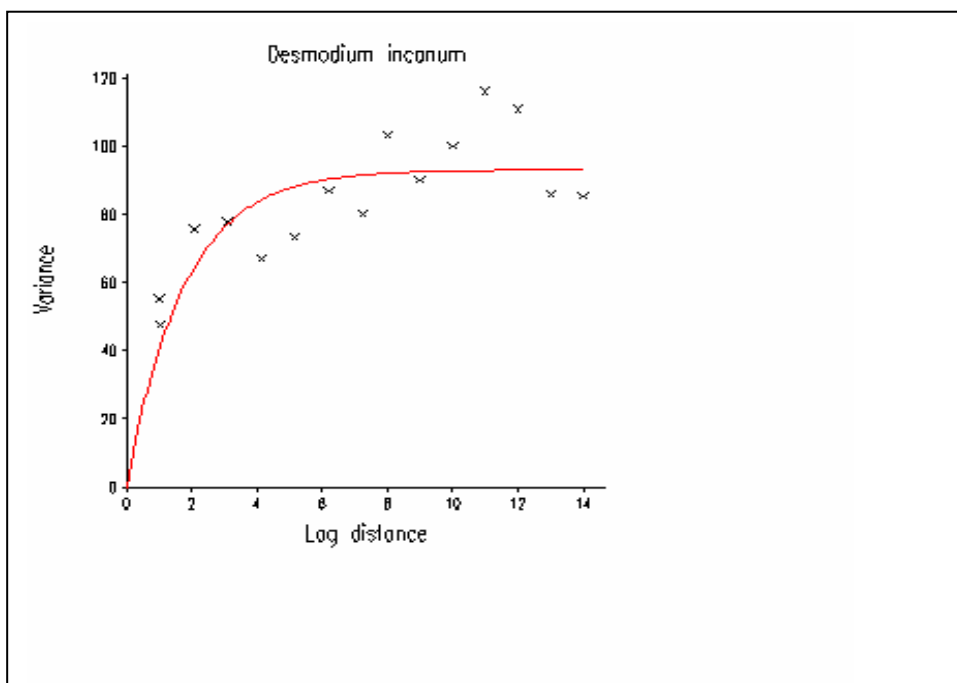
Na espécie *Paspalum notatum* Fl., é uma espécie de gramínea presente de contribuição importante em termos de frequência. Neste estudo observa-se um bom ajuste a modelos espaciais quando se trata de unidades amostrais pequenas (Tabela 17), diminuindo esse ajuste na medida que aumenta o tamanho amostral. Isso pode dever-se ao fato que as ausências desta espécie não são aleatórias, sendo que ocorrem em “rachas” que podem ser captadas com unidades amostrais pequenas. Esse comportamento segundo Palmer (1988) seria o de heterogeneidade a pequena escala e homogeneidade a escala maior. O mesmo comportamento observa-se na espécie *Aristida laevis* (Nees) Kunth, que apresenta associação espacial a tamanhos de unidade amostral pequena, sendo homogeneizada sua distribuição na medida que aumenta o tamanho da unidade amostral. Esta espécie é uma gramínea cespitosa, de porte ereto, que se apresenta em forma de maciegas no campo. Neste trabalho, a espécie apresentava no campo um aspecto de manchas ou “patches” a escala visual, que se verifica na análise quando podem-se manifestar as discontinuidades como zeros. Quando a unidade aumenta de tamanho, a espécie aparece em mais unidades amostrais, diminuindo as discontinuidades e homogeneizando seu padrão.



O caso contrario observa-se na leguminosa *Desmodium incanum* DC. Esta é também uma espécie muito frequente, que aparece em quase todas as unidades amostrais neste estudo, não apresentando manchas na estimacão visual. Não se observa associacão espacial quando se trabalha com unidades amostrais pequenas, embora seja detectada uma associacão na medida que o tamanho da unidade vai aumentando. Neste caso, ao ser esta uma especie de grande frequência e contribuicão nesta vegetacão, a associacão espacial observada pode ser devida à presenca de um gradiente vegetacional, onde existem aumentos ou diminuicões na densidade da espécie.



Quando o objetivo é detectar padrões uniespecíficos de variabilidade espacial, estas estruturas de semivariogram indicam que tamanhos menores de unidades amostrais tendem a detectar associações espaciais em *Aristida laevis* (Nees) Kunth e *Paspalum notatum* Fl, e tamanhos maiores em *Desmodium incanum* DC. Esta constatação é dependente de varios aspectos importantes, dentre os quais a vegetação acompanhante, o tamanho e desenvolvimento das plantas, o tipo e estrutura da amostragem realizada.



É interessante observar a existencia de associação espacial em caso em que não é detectado visualmente tal configuração, mas que se manifestam dependentes da escala do estudo.

6. CONCLUSÕES

A análise realizada modelando a variabilidade espacial, foi efetiva na redução dos erros padrões das média dos tratamentos e dos contrastes, melhorando sensivelmente suas estimações.

O ajuste de semivariogramas é fundamental e determinante para o sucesso na análise. Diante disso sua modelagem deve ser cuidadosa sendo importante a incorporação dos conhecimentos de modelagem sob o ponto de vista estatístico e o conhecimento do fenômeno a modelar.

Existe dependência da escala de amostragem nos ajustes dos semivariogramas, sendo necessário trabalhar com várias escalas simultâneas, tanto em levantamentos por amostragem como em delineamento de experimentos.

A baixa qualidade do ajuste dos semivariogramas, que as vezes se observa no caso de delineamento de experimentos, pode ser devido à homogeneidade gerada pela escala adotada, que se caracteriza por parcelas próximas mais que a falta de variabilidade espacial. A disponibilidade de subamostras fixas georeferenciadas no experimento, capazes de detectar variações a escalas menores, poderia auxiliar nesses casos.

Em levantamentos por amostragem, o conhecimento de associações espaciais pode auxiliar na determinação de estratégias de amostragem para uma espécie.

Os aplicativos computacionais que incluem a variabilidade espacial, assim como os que modelam estruturas de associação, estão cada vez mais disponíveis, pelo que esta metodologia pode ter maior difusão. Diante disso, sugere-se a incorporação da metodologia de análise de dados espaciais, como

um procedimento rotineiro na avaliação de experimentos e de levantamentos por amostragem.

7. REFÊRENCIAS BIBLIOGRÁFICAS

- BAIRD, D. ; MEAD, R. The empirical efficiency and validity of two neighbour models. **Biometrics**, Raleigh, v.47, p.1473-1487, 1991.
- BEALS, E.W. Vegetational change along altitudinal gradients. **Science**, v.165, p. 981-985.
- BESAG, J.; KEMPTON, R.A. Statistical analysis of field experiments using neighboring plots. **Biometrics**, Raleigh, v.42, p.231-251, 1986.
- BESAG, J.E.; GLEAVES, J.T. On the detection of spatial pattern in plant communities. **Bulletin of the International Statistical Institute**, v. 45, Book 1, p. 153-158, 1973.
- BESAG, J; HIGDON, D. Bayesian analysis of agricultural field experiments. **Journal of the Royal Statistic Society B**. v.61, p.690-717, 1999.
- BOUXIN, G. Permutation and Monte Carlo tests applied to the measurement of horizontal patterns in vegetation. **Biometrie-Praximetrie**, v.32, p.115-140, 1992.
- BROWNIE, C.; BOWMAN, D.T.; BURTON, J.W.; Estimating spatial variation in analysis of data from yield trials: a comparison of methods. **Agronomy Journal**, Madison, v.85, p.1244-1253, 1993.
- BYTH, K., RIPLEY, D. On sampling spatial patterns by distance methods. **Biometrics**, Raleigh, v.36, p. 279-284, 1980.
- BYTH, K.; RIPLEY, B.D. On sampling spatial patterns by distance methods. **Biometrics**, Raleigh, v. 36, p. 279-284, 1980.
- CADENAZZI, M. **Otimização de amostragem para estudo de padrões em vegetação campestre**. . Porto Alegre, 1996, 179p. Dissertação de Mestrado em Zootecnia,- Faculdade de Agronomia, Universidade Federal do Rio Grande do Sul, Porto Alegre, 1999.

CASLER, M.D. Spatial variation affects precision of perennial cool-season forage grass trials. **Agronomy Journal**, Madison, v.91, p. 75-81, 1999.

CASTRO, S.M.J. **A metodologia de análise de dados longitudinais em plantas forrageiras no contexto de estruturas de covariância**. Porto Alegre, 1999, 171p. Dissertação de Mestrado em Zootecnia,- Faculdade de Agronomia, Universidade Federal do Rio Grande do Sul, Porto Alegre, 1999.

CLARK, P.J.; EVANS, F.C. Distance to nearest neighbor as a measure of spatial relationships in populations. **Ecology**, Washington, v. 35, p. 445-453, 1954.

CORMACK, R.M. Spatial aspects of competition between individuals. **In: Spatial and Temporal Analysis in Ecology**. Ed. R.M. Cormack e J.K. Ord. International Co-operative Publishing House, Fairland, M.D. p. 151-152.

COWLING, A. Spatial methods for line transect surveys. **Biometrics**, Raleigh, v.54, p.828-839, 1998.

COX, T.F.; LEWIS, T. A conditional distance ratio method for analyzing spatial patterns. **Biometrika**, London, v. 63, p. 483-491, 1976.

CRESSIE, N. **Statistics for spatial data**. New York, John Willey, 900 p, 1993.

CULLIS, B.; GOGEL, B.; VERBYLA, A.; THOMPSON, R. Spatial analysis of multi-environment early generation variety trials. **Biometrics**, Raleigh, v.54, p.1-18, 1998.

CULLIS, B.R.; GLEESON, AC. Efficiency of neighbour analysis for replicated variety trials in Australia. **Journal of Agricultural Science**, Cambridge, v.113, p.233-239, 1989.

CULLIS, B.R.; GLEESON, AC.; Spatial analysis of field experiments- an extension to two dimensions. **Biometrics**, Raleigh, v.47, p.1449-1460, 1991.

DALE, M.R.T.; MAH, M. The use of wavelets for spatial pattern analysis in ecology. **Journal of Vegetation Science**, Uppsala, v. 9, p. 805-814, 1998.

DALE, M.R.T.; ZBIGNIEWICS, M.W. The evaluation of multi-species pattern. **Journal of Vegetation Science**, Uppsala, v. 6, p. 391-398, 1995.

DIGGLE, P.J. Binary mosaics and the spatial pattern of heather. **Biometrics**, Raleigh, v.37, p. 531-539, 1981.

DIGGLE, P.J. On parameter estimation and goodness-of-fit testing for spatial point patterns. **Biometrics**, Raleigh, v.35, p. 87-101, 1979.

DIGGLE, P.J.; BESAG, J.E.; GLEAVES, J.T. Statistical analysis of spatial point patterns by means of distance methods. **Biometrics**, Raleigh, v. 32, p. 659-667, 1976.

DUTILLEUL, P. Spatial heterogeneity and the desing of ecological field experiments. **Ecology**, Washington, v.76, p.1646-1658, 1993.

EBERHARDT, L.L. Some developments in 'distance sampling'. **Biometrics**, Raleigh, v. 23, p.207-216, 1967.

EGHBALL, B.; VARVEL, G.E. Fractal analysis of temporal yield variability of crop sequences: Implications for site-specific management. **Agronomy Journal**, Madison, v.89, p. 851-855, 1997.

FISHER, R.A. **The Desing of Experiments**. Oliver e Boyd, Edinburgo, 1935.

FORTIN, M.J. Effects of quadrant size and data measurement on the detection of boundaries. **Journal of Vegetation Science**, Uppsala, v. 10, p. 43-50, 1999.

FORTIN, M.J.; GUREVITCH, J. Mantel test, Spatial structure in field experiments. In: **Desing and analysis of ecological experiments**. Ed. Scheiner, S.M. e Gurevitch, J. Chapman & Hall, New York, p.342-359, 1993.

GENSTAT. **Reference manual supplement**, Genstat 5 Committee of the Statistics Department, IACR, Rothamsted, 147 p.

GILMOUR, AR.; CULLIS, B.R.; VERBYLA, AP. Accounting for natural extraneous variation in the analysis of field experiments. **Journal of Agricultural, Biological and Environmental Statistics**, v.2, p.269-293, 1997.

GREIG-SMITH, P. The use of random and contiguous quadrants in the study of the structure of plant communities. **Annals of Botany**, London, v.16, p. 293-316, 1952.

GRONDONA, M.O , CRESSIE, N. Using spatial consideration in the analysis of experiments. **Technometrics**, Washington, v. 33, p. 381-392, 1991.

GRONDONA, M.O; CROSSA, J.; FOX, P.N.; PFEIFFER, W.H. Analysis of variety yield trials using two-dimensional separable ARIMA processes. **Biometrics**, Raleigh, v.52, p.763-770, 1996.

HENDERSON, C.R. Selection index and expected genetic advance. In: HANSON, W.D.; ROBINSON, H.F. **Statistical genetics and plant breeding**. Washington, National Academy of Science. National Research Council, p.141-163, Publication 982, 1963.

HINES, W.G.S.; HINES, R.J.O. The Eberhardt statistic and the detection of nonrandomness of spatial point distributions. **Biometrika**, London, v. 66, p. 73-79, 1979.

HOLGATE, P. Some new test for randomness. **Journal of Ecology**, v.53, p.261-266, 1965.

HOPKINS, B. A new method for determining the type of distribution of plant individuals. **Annals of Botany**, London, v. 18, p.213-227, 1954.

HURLBERT, S.H. Pseudoreplicação and the desing of ecological field experiments. **Ecological Monographs**, Durhan, v. 54, p. 187-211, 1984.

- KENKEL, N.C.; WALKER, D.J. Fractals and ecology. **Abstracta Botanica**. Budapest, v. 17 p. 53-70, 1993.
- LACA, E.A.; ORTEGA, I.M. Integrating foraging mechanisms across spatial and temporal scales. In: **Fifth International Rangeland Congress**, Vol. II, p. 129-132, 1995.
- LAIRD, N.M.; WARE, J.H. Random-effects models for longitudinal data, **Biometrics**, Raleigh, v. 38, p. 963-974, 1982.
- LEGENDRE, P. Spatial autocorrelation: Trouble or new paradigm?. **Ecology**, New York, v.74, p. 1659-1673, 1993.
- LEGENDRE, P.; FORTIN, M.J. Spatial pattern and ecological analysis. **Vegetatio**, The Hague, v.80, p. 107-138, 1989.
- LEVIN, S. A. The problem of pattern and scale in ecology. **Ecology**, Washington, v.73, p. 1943-1967, 1992.
- LINDSTROM, M.J.; BATES, D.M. Newton-Raphson and EM algorithms for linear mixed-effects models for repeated-measures data, **Journal of the American Statistical Association**, Washington, v.83, p.1014-1022, 1988.
- MANLY, B.F.J. **Randomization and Monte Carlo methods in Biology**. Chapman and Hall, London, 281 p.
- McLEAN, R.A.; SANDERS, W.L.; STROUP, W.W. A unified approach to mixed linear models, **The American Statistician**, Alexandria, v.45, p. 54-64, 1991.
- PAIM, N.; RIBOLDI, J. Duas novas cultivares de travo-branco comparadas com outras disponíveis no Rio Grande do Sul, em associação com gramíneas. **Pesquisa Agropecuária Brasileira**, Brasília, v.29, p.43-53, 1994.

- PALMER, M.W.. Fractal geometry: a tool for describing spatial patterns of plant communities. **Vegetatio**, The Hague, v.75, p. 91-102, 1988.
- PAPADAKIS, J. **Advances in the nearest neighbour method of field experiments**. Weighted adjustment at treatment level that gives also the error separately for each treatment and terminates with vain experiments. Editor: Papadakis, J. Athens, 15p. 1994.
- PEARCE, S.C. Field experimentation on rough land: the method of Papadakis reconsidered. **Journal of Agricultural Science**, Cambridge, v.131, p.1-11, 1998.
- PEREIRA, S.M. **Métodos geoestatísticos e procedimentos de lógica nebulosa: uma aplicação para a obtenção de um mapa de fertilidade de solo**. Rio de Janeiro, 1994, 139p. Dissertação de Mestrado em Ciências Estatística,-Instituto de Matemática, Universidade Federal de Rio de Janeiro, Rio de Janeiro, 1994.
- PERRY, J.N. ; MEAD, R.. On the power of the index of dispersion test to detect spatial pattern. **Biometrics**, Raleigh, v.35, p. 613-622, 1979.
- PIELOU, E.C. The use of point-to-plant distances in the pattern of plant population. **Journal of Ecology**, Oxford, v. 47, p. 607-613, 1959.
- PODANI, J. Spatial processes in the analysis of vegetation: Theory and review. **Acta Botanica Hungarica**, v. 30, p.75-118, 1984.
- RIBEIRO JUNIOR, P.J. **Métodos geoestatísticos no estudo da variabilidade espacial de parâmetros do solo**. Piracicaba:USP, 1995. 96f. Dissertação de Mestrado em Agronomia - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, 1995.
- RIBEIRO, P.J.; SIMM, E.B.; XAVIER, L.; PINTO, P.C.; VIDAL, E. **Análise espacial de experimentos**.Universidade Federal do Paraná. Relatório Técnico, 21p. 1997.
- RIPLEY, B.D. **Statistical inference for spatial processes**. Cambridge University Press, Cambridge, 148 p. 1988.

ROBINSON, G.K. That BLUP is a good thing: the estimation of random effects. **Statistical Science**, v. 6, p. 15-51, 1991.

SAS Institute. **Getting Started with Proc Mixed**. Cary, SAS, 121p., 1994.

SATYAMURTHI, K.R. Density, derived from measured distances, for studying the spatial patterns. **Sankhya B**, v.40, p.197-203.

SKELLAM, J.G. Studies in statistical ecology, I. Spatial pattern. **Biometrika**, London, v.39, p.346-362, 1952.

STROUP, W.W.; BAENZIGER, P.S.; MULITZE, D.K. Removing spatial variation from wheat yield trials: a comparison of methods. **Crop Science**. Madison, V.86, p.62-66, 1994.

VER HOEF, J.M.; CRESSIE, N. Spatial statistics: analysis of field experiments. In: **Desing and analysis of ecological experiments**. Ed. Scheiner, S.M. e Gurevitch, J. Chapman & Hall, New York, p.319-339, 1993.

VOLLMANN, J.; BUERSTMAYR, H.; RUCKENBAUER, P. Efficient control of spatial variation in yield trials using neighbour plot residuals. **Experimental Agriculture**, Cambridge, v. 32, p. 185-197, 1996.

WU, T.; DUTILLEUL, P. Validity and efficiency of neighbor analyses in comparison with classical complete and incomplete block analyses of field experiments. **Agronomy Journal**, Madison, v. 91, p. 721-731, 1999.

YATES, F. The comparative advantages of systematic and randomized arrangements in the desing of agricultural and biological experiments. **Biometrika**, London, v. 30, p.444-466, 1938.

ZIMMERMAN, D.L.; HARVILLE, D.A A random field approach to the analysis of Field-plot experiments and other spatial experiments. **Biometrics**, Raleigh, v.47, p.223-239, 1991.

Apêndice 1. Dados originais do experimento 1.

ANO	BLOC	CULTIVARES	MSTBR	MSGRAM	MSINCOS	MSTO	MSTMI	PARCE	FILA	COL
O						T		LA		
1	1	Regal	2361	1972	520	4853	4333	1	2.5	7.5
1	1	Yi	1994	1816	673	4483	3810	2	2.5	10
1	1	Bayucua	2252	1970	290	4512	4222	3	2.5	5
1	1	Huia	1340	1488	170	2998	2828	4	2.5	2.5
1	1	Pitau	167	1854	552	2573	2021	5	2.5	1
1	1	Bagé	2547	1417	113	4077	3964	6	2.5	17.5
1	1	GuaíbaS1	2591	1458	619	4668	4049	7	2.5	12.5
1	1	JacuiS2	2622	1507	223	4352	4129	8	2.5	15
1	2	Regal	2105	1657	771	4533	3762	9	8.5	1
1	2	Yi	3671	1420	478	5569	5091	10	8.5	7.5
1	2	Bayucua	2260	1814	407	4481	4074	11	8.5	12.5
1	2	Huia	1476	1618	125	3219	3094	12	8.5	15
1	2	Pitau	64	1829	909	2802	1893	13	8.5	2.5
1	2	Bagé	2262	1973	527	4762	4235	14	8.5	5
1	2	GuaíbaS1	3644	2080	917	6641	5724	15	8.5	10
1	2	JacuiS2	2777	1327	180	4284	4104	16	8.5	17.5
1	3	Regal	2281	1325	404	4010	3606	17	14.5	5
1	3	Yi	2297	1346	357	4000	3643	18	14.5	2.5
1	3	Bayucua	1833	1229	479	3541	3062	19	14.5	17.5
1	3	Huia	836	1734	326	2896	2570	20	14.5	12.5
1	3	Pitau	66	1555	656	2277	1621	21	14.5	15
1	3	Bagé	1728	1529	832	4089	3257	22	14.5	1
1	3	GuaíbaS1	2352	1643	356	4351	3995	23	14.5	10
1	3	JacuiS2	2957	2496	763	6216	5453	24	14.5	7.5
1	4	Regal	1497	1659	438	3594	3156	25	20.5	17.5
1	4	Yi	2148	1611	185	3944	3759	26	20.5	10
1	4	Bayucua	1356	1641	380	3377	2997	27	20.5	2.5
1	4	Huia	1198	1777	163	3138	2975	28	20.5	12.5
1	4	Pitau	167	1671	690	2528	1838	29	20.5	5
1	4	Bagé	2631	1617	400	4648	4248	30	20.5	15
1	4	GuaíbaS1	2229	2819	706	5754	5048	31	20.5	7.5
1	4	JacuiS2	2248	1680	493	4421	3928	32	20.5	1
1	5	Regal	2211	2097	315	4623	4308	33	26.5	7.5
1	5	Yi	1913	1929	369	4211	3842	34	26.5	15
1	5	Bayucua	711	1863	596	3170	2574	35	26.5	2.5
1	5	Huia	620	1573	333	2526	2193	36	26.5	5
1	5	Pitau	70	1931	990	2991	2001	37	26.5	1
1	5	Bagé	1776	2584	177	4537	4360	38	26.5	10
1	5	GuaíbaS1	1777	1499	322	3598	3276	39	26.5	12.5
1	5	JacuiS2	1971	1837	504	4312	3808	40	26.5	17.5
2	1	Regal	325	2471	757	3553	2796	1	2.5	7.5
2	1	Yi	945	2259	525	3729	3204	2	2.5	10
2	1	Bayucua	836	2230	442	3508	3066	3	2.5	5
2	1	Huia	63	1611	597	2271	1674	4	2.5	2.5
2	1	Pitau	456	1477	500	2433	1933	5	2.5	1
2	1	Bagé	587	1931	386	2904	2518	6	2.5	17.5
2	1	GuaíbaS1	757	2059	370	3186	2816	7	2.5	12.5

Continuação Apêndice 1

2	1 JacuíS2	1426	1668	530	3624	3094	8	2.5	15
2	2 Regal	171	1279	691	2141	1450	9	8.5	1
2	2 Yi	943	2122	477	3542	3065	10	8.5	7.5
2	2 Bayucua	560	1711	514	2785	2271	11	8.5	12.5
2	2 Huia	137	1783	512	2432	1920	12	8.5	15
2	2 Pitau	86	1572	356	2014	1658	13	8.5	2.5
2	2 Bagé	618	1750	297	2665	2368	14	8.5	5
2	2 GuaíbaS1	413	1998	1036	3447	2411	15	8.5	10
2	2 JacuíS2	565	2528	267	3360	3093	16	8.5	17.5
2	3 Regal	43	1300	333	1676	1343	17	14.5	5
2	3 Yi	748	2099	377	3224	2847	18	14.5	2.5
2	3 Bayucua	707	2251	313	3271	2958	19	14.5	17.5
2	3 Huia	22	1931	539	2492	1953	20	14.5	12.5
2	3 Pitau	110	1722	329	2161	1832	21	14.5	15
2	3 Bagé	474	1875	434	2783	2349	22	14.5	1
2	3 GuaíbaS1	543	2173	774	3490	2716	23	14.5	10
2	3 JacuíS2	554	2226	940	3720	2780	24	14.5	7.5
2	4 Regal	38	2376	502	2916	2414	25	20.5	17.5
2	4 Yi	705	2194	514	3413	2899	26	20.5	10
2	4 Bayucua	1122	1916	486	3524	3038	27	20.5	2.5
2	4 Huia	43	2031	838	2912	2074	28	20.5	12.5
2	4 Pitau	464	2009	396	2869	2473	29	20.5	5
2	4 Bagé	656	2681	387	3724	3337	30	20.5	15
2	4 GuaíbaS1	830	2073	489	3392	2903	31	20.5	7.5
2	4 JacuíS2	1609	2064	533	4206	3673	32	20.5	1
2	5 Regal	69	2030	544	2643	2099	33	26.5	7.5
2	5 Yi	368	2826	595	3789	3194	34	26.5	15
2	5 Bayucua	52	1927	364	2343	1979	35	26.5	2.5
2	5 Huia	7	1797	418	2222	1804	36	26.5	5
2	5 Pitau	13	1490	417	1920	1503	37	26.5	1
2	5 Bagé	365	2500	605	3470	2865	38	26.5	10
2	5 GuaíbaS1	447	2789	284	3520	3236	39	26.5	12.5
2	5 JacuíS2	677	2607	332	3616	3284	40	26.5	17.5
3	1 Regal	766	3807	321	4894	4573	1	2.5	7.5
3	1 Yi	1234	4145	115	5494	5379	2	2.5	10
3	1 Bayucua	364	4001	114	4479	4365	3	2.5	5
3	1 Huia	363	3351	408	4122	3714	4	2.5	2.5
3	1 Pitau	873	3649	261	4783	4522	5	2.5	1
3	1 Bagé	648	3019	214	3881	3667	6	2.5	17.5
3	1 GuaíbaS1	805	3582	86	4473	4387	7	2.5	12.5
3	1 JacuíS2	1086	3137	309	4532	4223	8	2.5	15
3	2 Regal	648	2647	491	3786	3295	9	8.5	1
3	2 Yi	1213	3072	206	4491	4285	10	8.5	7.5
3	2 Bayucua	877	3376	210	4463	4253	11	8.5	12.5
3	2 Huia	755	2793	736	4284	3548	12	8.5	15
3	2 Pitau	321	2862	122	3305	3183	13	8.5	2.5
3	2 Bagé	1084	3094	351	4529	4178	14	8.5	5
3	2 GuaíbaS1	1657	3431	430	5518	5088	15	8.5	10
3	2 JacuíS2	358	2808	127	3293	3166	16	8.5	17.5
3	3 Regal	191	2767	111	3069	2958	17	14.5	5

Continuação Apêndice 1

3	3 Yi	421	2841	225	3487	3262	18	14.5	2.5
3	3 Bayucua	199	2613	71	2883	2812	19	14.5	17.5
3	3 Huia	29	2883	142	3054	2912	20	14.5	12.5
3	3 Pitau	126	2796	134	3056	2922	21	14.5	15
3	3 Bagé	413	3165	143	3721	3578	22	14.5	1
3	3 GuaíbaS1	232	3522	260	4014	3754	23	14.5	10
3	3 JacuíS2	543	4273	199	5015	4816	24	14.5	7.5
3	4 Regal	220	3073	161	3454	3293	25	20.5	17.5
3	4 Yi	687	3947	283	4917	4634	26	20.5	10
3	4 Bayucua	580	3722	142	4444	4302	27	20.5	2.5
3	4 Huia	238	3117	237	3592	3355	28	20.5	12.5
3	4 Pitau	664	3709	57	4430	4373	29	20.5	5
3	4 Bagé	766	3560	218	4544	4326	30	20.5	15
3	4 GuaíbaS1	1476	4397	161	6034	5873	31	20.5	7.5
3	4 JacuíS2	1077	3761	170	5008	4838	32	20.5	1
3	5 Regal	280	3063	251	3594	3343	33	26.5	7.5
3	5 Yi	645	3706	149	4500	4351	34	26.5	15
3	5 Bayucua	307	3219	76	3602	3526	35	26.5	2.5
3	5 Huia	270	4206	103	4579	4476	36	26.5	5
3	5 Pitau	170	3051	102	3323	3221	37	26.5	1
3	5 Bagé	794	4618	270	5682	5412	38	26.5	10
3	5 GuaíbaS1	557	3953	122	4632	4510	39	26.5	12.5
3	5 JacuíS2	691	3435	64	4190	4126	40	26.5	17.5
4	1 Regal	763	6016	132	6911	6779	1	2.5	7.5
4	1 Yi	747	5780	258	6785	6527	2	2.5	10
4	1 Bayucua	492	4844	201	5537	5336	3	2.5	5
4	1 Huia	665	4380	161	5206	5045	4	2.5	2.5
4	1 Pitau	1104	4966	309	6379	6070	5	2.5	1
4	1 Bagé	1265	4837	165	6267	6102	6	2.5	17.5
4	1 GuaíbaS1	644	5043	156	5843	5687	7	2.5	12.5
4	1 JacuíS2	799	5412	283	6494	6211	8	2.5	15
4	2 Regal	1192	4680	225	6097	5872	9	8.5	1
4	2 Yi	1022	6104	323	7449	7126	10	8.5	7.5
4	2 Bayucua	759	4506	181	5446	5265	11	8.5	12.5
4	2 Huia	713	4718	298	5729	5431	12	8.5	15
4	2 Pitau	909	4855	154	5918	5764	13	8.5	2.5
4	2 Bagé	747	5730	492	6969	6477	14	8.5	5
4	2 GuaíbaS1	447	6937	589	7973	7384	15	8.5	10
4	2 JacuíS2	442	4164	208	4814	4606	16	8.5	17.5
4	3 Regal	50	3888	72	4010	3938	17	14.5	5
4	3 Yi	822	4599	88	5509	5421	18	14.5	2.5
4	3 Bayucua	82	3993	184	4259	4075	19	14.5	17.5
4	3 Huia	33	4291	139	4463	4324	20	14.5	12.5
4	3 Pitau	116	3923	149	4188	4039	21	14.5	15
4	3 Bagé	488	4444	26	4958	4932	22	14.5	1
4	3 GuaíbaS1	54	4917	143	5114	4971	23	14.5	10
4	3 JacuíS2	74	6454	265	6793	6528	24	14.5	7.5
4	4 Regal	446	4678	156	5280	5124	25	20.5	17.5
4	4 Yi	376	5256	71	5703	5632	26	20.5	10
4	4 Bayucua	786	5040	143	5969	5826	27	20.5	2.5

Continuação Apêndice 1

4	4 Huia	225	4661	90	4976	4886	28	20.5	12.5
4	4 Pitau	303	4872	42	5217	5175	29	20.5	5
4	4 Bagé	240	5683	442	6365	5923	30	20.5	15
4	4 GuaíbaS1	985	7228	182	8395	8213	31	20.5	7.5
4	4 JacuiS2	765	5599	105	6469	6364	32	20.5	1
4	5 Regal	634	5412	122	6168	6046	33	26.5	7.5
4	5 Yi	1014	5972	79	7065	6986	34	26.5	15
4	5 Bayucua	894	4631	79	5604	5525	35	26.5	2.5
4	5 Huia	485	4516	71	5072	5001	36	26.5	5
4	5 Pitau	312	4616	199	5127	4928	37	26.5	1
4	5 Bagé	623	6198	125	6946	6821	38	26.5	10
4	5 GuaíbaS1	919	5619	156	6694	6538	39	26.5	12.5
4	5 JacuiS2	1208	4769	145	6122	5977	40	26.5	17.5

Apêndice 2. Dados originais do experimento 2. As variáveis listadas são data, solo, bloco, cultivar, parcela, número de plantas por m2, posição na fila (m) e posição na coluna (m).

data	solo	bloco	cult	parcela	nplm2	fila	col
2	2	1	1	61	69	2.5	25
2	2	1	2	62	404	2.5	23
2	2	1	3	63	288.5	2.5	17
2	2	1	4	64	327	2.5	1
2	2	1	5	65	242.5	2.5	9
2	2	1	6	66	204	2.5	27
2	2	1	7	67	134.5	2.5	11
2	2	1	8	68	254	2.5	21
2	2	1	9	69	181	2.5	3
2	2	1	10	70	227	2.5	29
2	2	1	11	71	165	2.5	5
2	2	1	12	72	215.5	2.5	19
2	2	1	13	73	4	2.5	7
2	2	1	14	74	311.5	2.5	13
2	2	1	15	75	223	2.5	31
2	2	1	16	76	338.5	2.5	15
2	2	1	17	77	523	2.5	37
2	2	1	18	78	484.5	2.5	35
2	2	1	19	79	484.5	2.5	41
2	2	1	20	80	311.5	2.5	39
2	2	2	1	81	38.5	11.5	13
2	2	2	2	82	769	11.5	29
2	2	2	3	83	181	11.5	21
2	2	2	4	84	257.5	11.5	9
2	2	2	5	85	350	11.5	27
2	2	2	6	86	246	11.5	25
2	2	2	7	87	184.5	11.5	1
2	2	2	8	88	277	11.5	7
2	2	2	9	89	219	11.5	17
2	2	2	10	90	315	11.5	31
2	2	2	11	91	384.5	11.5	3
2	2	2	12	92	450	11.5	11
2	2	2	13	93	11.5	11.5	23
2	2	2	14	94	323	11.5	5
2	2	2	15	95	261.5	11.5	15
2	2	2	16	96	204	11.5	19
2	2	2	17	97	642.5	11.5	37
2	2	2	18	98	165	11.5	39
2	2	2	19	99	300	11.5	35
2	2	2	20	100	111.5	11.5	41
2	2	3	1	101	65.5	20.5	17
2	2	3	2	102	634.5	20.5	19
2	2	3	3	103	565.5	20.5	13
2	2	3	4	104	350	20.5	25
2	2	3	5	105	327	20.5	3
2	2	3	6	106	169	20.5	15
2	2	3	7	107	246	20.5	27
2	2	3	8	108	238.5	20.5	9
2	2	3	9	109	404	20.5	5
2	2	3	10	110	488.5	20.5	29
2	2	3	11	111	450	20.5	31
2	2	3	12	112	350	20.5	7
2	2	3	13	113	46	20.5	21
2	2	3	14	114	299.5	20.5	23
2	2	3	15	115	288.5	20.5	11

Continuação Apêndice 2.

2	2	3	16	116	238.5	20.5	1
2	2	3	17	117	538.5	20.5	35
2	2	3	18	118	238.5	20.5	41
2	2	3	19	119	219	20.5	37
2	2	3	20	120	150	20.5	39

Fatores utilizados para converter de número de plantas por m2 a %IMPL.

CULTIVAR	FATOR
1	587
2	747
3	669
4	925
5	546
6	883
7	1040
8	1415
9	916
10	738
11	924
12	638
13	394
14	703
15	940
16	544
17	1365
18	439
19	806
20	324

Apêndice 3. Dados originais do experimento 3. São apresentados o número da unidade amostral (UA), a percentagem de cobertura das espécies *Aristida laevis*, *Desmodium incanum* e *Paspalum notatum* e a distância (m) desde o início da transecta.

UA	Arist_lae	Desm_inc	Pano	dist	UA	Arist_lae	Desm_inc	Pano	dist
1	0	25	45	0.05	201	0	20	30	20.05
2	0	20	25	0.15	202	1	20	20	20.15
3	5	5	60	0.25	203	1	25	20	20.25
4	35	30	30	0.35	204	1	20	3	20.35
5	100	15	10	0.45	205	15	30	25	20.45
6	50	0	20	0.55	206	1	25	25	20.55
7	15	20	30	0.65	207	25	20	0	20.65
8	0	25	30	0.75	208	100	20	0	20.75
9	0	35	25	0.85	209	100	10	25	20.85
10	1	15	40	0.95	210	50	15	25	20.95
11	5	20	30	1.05	211	20	20	25	21.05
12	10	10	50	1.15	212	1	35	40	21.15
13	20	25	30	1.25	213	0	30	35	21.25
14	10	25	40	1.35	214	0	0	90	21.35
15	0	30	40	1.45	215	0	0	90	21.45
16	0	25	30	1.55	216	0	25	25	21.55
17	0	15	50	1.65	217	0	25	25	21.65
18	0	40	40	1.75	218	0	25	30	21.75
19	0	40	40	1.85	219	0	20	60	21.85
20	0	20	70	1.95	220	0	20	25	21.95
21	0	20	20	2.05	221	1	30	25	22.05
22	0	40	30	2.15	222	20	30	0	22.15
23	0	20	30	2.25	223	55	0	0	22.25
24	0	10	25	2.35	224	100	10	0	22.35
25	0	35	20	2.45	225	70	0	0	22.45
26	0	30	0	2.55	226	30	15	0	22.55
27	1	40	0	2.65	227	10	0	0	22.65
28	10	20	25	2.75	228	5	25	0	22.75
29	20	0	30	2.85	229	0	20	20	22.85
30	50	40	25	2.95	230	0	0	25	22.95
31	50	20	20	3.05	231	0	25	30	23.05
32	50	40	20	3.15	232	0	25	35	23.15
33	5	15	25	3.25	233	0	20	15	23.25
34	1	30	35	3.35	234	0	25	15	23.35
35	0	70	50	3.45	235	0	20	0	23.45
36	0	40	20	3.55	236	0	20	15	23.55
37	0	25	30	3.65	237	0	20	30	23.65
38	0	25	20	3.75	238	0	25	40	23.75
39	0	25	30	3.85	239	1	10	30	23.85
40	1	20	30	3.95	240	1	30	30	23.95
41	0	30	50	4.05	241	20	15	0	24.05
42	0	50	30	4.15	242	1	25	20	24.15
43	0	30	40	4.25	243	1	0	0	24.25
44	0	35	50	4.35	244	15	0	0	24.35

Continuação Apêndice 3.

45	0	30	70	4.45	245	50	15	0	24.45
46	0	20	25	4.55	246	100	10	0	24.55
47	0	20	35	4.65	247	100	0	0	24.65
48	0	15	40	4.75	248	45	5	20	24.75
49	0	20	40	4.85	249	1	20	25	24.85
50	0	25	35	4.95	250	1	20	25	24.95
51	0	20	35	5.05	251	1	25	30	25.05
52	1	30	30	5.15	252	0	20	25	25.15
53	5	0	30	5.25	253	0	0	40	25.25
54	15	15	0	5.35	254	0	0	50	25.35
55	15	15	0	5.45	255	0	0	50	25.45
56	0	15	0	5.55	256	0	0	0	25.55
57	15	25	0	5.65	257	0	0	0	25.65
58	0	30	0	5.75	258	0	0	0	25.75
59	0	30	0	5.85	259	0	0	0	25.85
60	0	20	0	5.95	260	1	0	0	25.95
61	0	30	0	6.05	261	10	0	0	26.05
62	0	20	0	6.15	262	0	10	20	26.15
63	0	10	25	6.25	263	10	10	20	26.25
64	0	15	30	6.35	264	15	1	0	26.35
65	0	0	20	6.45	265	1	5	20	26.45
66	0	20	20	6.55	266	0	15	25	26.55
67	0	15	20	6.65	267	10	0	50	26.65
68	0	15	20	6.75	268	1	20	20	26.75
69	0	5	30	6.85	269	0	0	25	26.85
70	0	20	50	6.95	270	0	0	25	26.95
71	0	15	30	7.05	271	0	20	0	27.05
72	0	0	30	7.15	272	0	20	0	27.15
73	0	0	20	7.25	273	0	20	0	27.25
74	50	0	35	7.35	274	0	25	0	27.35
75	50	0	30	7.45	275	0	60	0	27.45
76	35	0	35	7.55	276	0	20	0	27.55
77	25	0	0	7.65	277	0	15	15	27.65
78	40	10	0	7.75	278	0	25	20	27.75
79	60	25	0	7.85	279	5	25	25	27.85
80	0	0	0	7.95	280	10	25	35	27.95
81	1	15	0	8.05	281	15	15	90	28.05
82	1	30	0	8.15	282	20	15	70	28.15
83	1	30	0	8.25	283	20	20	20	28.25
84	1	20	0	8.35	284	40	0	0	28.35
85	0	30	0	8.45	285	10	20	25	28.45
86	15	0	0	8.55	286	1	0	25	28.55
87	30	0	0	8.65	287	1	15	20	28.65
88	15	20	0	8.75	288	0	15	20	28.75
89	20	10	0	8.85	289	0	0	0	28.85
90	0	30	30	8.95	290	0	20	0	28.95
91	5	30	25	9.05	291	0	20	0	29.05
92	0	20	20	9.15	292	0	25	0	29.15
93	5	20	35	9.25	293	0	50	0	29.25
94	10	25	0	9.35	294	0	25	15	29.35

Continuação Apêndice 3.

95	15	10	0	9.45	295	0	15	20	29.45
96	30	35	0	9.55	296	5	20	30	29.55
97	25	25	0	9.65	297	5	25	0	29.65
98	20	20	0	9.75	298	0	20	0	29.75
99	50	25	0	9.85	299	0	30	20	29.85
100	100	15	0	9.95	300	0	30	0	29.95
101	100	0	0	10.05	301	0	30	0	30.05
102	35	20	0	10.15	302	0	25	10	30.15
103	15	0	0	10.25	303	0	0	40	30.25
104	0	20	20	10.35	304	0	15	35	30.35
105	20	20	10	10.45	305	0	40	0	30.45
106	40	15	20	10.55	306	0	20	25	30.55
107	5	25	20	10.65	307	0	20	15	30.65
108	0	20	30	10.75	308	0	25	25	30.75
109	0	15	20	10.85	309	0	20	20	30.85
110	0	30	25	10.95	310	0	15	30	30.95
111	0	15	0	11.05	311	0	20	25	31.05
112	0	15	25	11.15	312	0	0	20	31.15
113	0	20	25	11.25	313	0	10	0	31.25
114	0	20	25	11.35	314	0	15	0	31.35
115	1	25	30	11.45	315	0	30	0	31.45
116	0	20	25	11.55	316	0	25	0	31.55
117	15	25	0	11.65	317	0	25	0	31.65
118	0	20	30	11.75	318	5	20	0	31.75
119	0	25	20	11.85	319	1	20	0	31.85
120	0	0	15	11.95	320	0	20	0	31.95
121	0	30	20	12.05	321	0	15	0	32.05
122	0	20	20	12.15	322	0	15	0	32.15
123	10	20	15	12.25	323	15	25	20	32.25
124	5	10	25	12.35	324	0	25	25	32.35
125	15	20	20	12.45	325	0	25	20	32.45
126	40	30	20	12.55	326	25	15	10	32.55
127	15	25	20	12.65	327	25	20	0	32.65
128	20	25	25	12.75	328	20	25	25	32.75
129	20	0	0	12.85	329	1	20	20	32.85
130	10	20	5	12.95	330	0	20	0	32.95
131	15	20	25	13.05	331	0	20	30	33.05
132	10	0	30	13.15	332	0	60	20	33.15
133	1	30	25	13.25	333	0	20	50	33.25
134	5	0	5	13.35	334	0	25	80	33.35
135	1	25	30	13.45	335	0	20	80	33.45
136	0	25	25	13.55	336	0	30	30	33.55
137	0	15	0	13.65	337	0	20	30	33.65
138	0	25	25	13.75	338	0	25	30	33.75
139	0	30	15	13.85	339	0	20	30	33.85
140	0	30	25	13.95	340	0	15	5	33.95
141	0	40	0	14.05	341	0	0	15	34.05
142	0	30	0	14.15	342	0	35	20	34.15
143	5	30	20	14.25	343	0	50	25	34.25
144	0	20	70	14.35	344	0	25	25	34.35

Continuação Apêndice 3.

145	0	50	30	14.45	345	0	25	30	34.45
146	0	25	20	14.55	346	5	25	35	34.55
147	0	90	15	14.65	347	15	20	0	34.65
148	0	30	25	14.75	348	50	35	0	34.75
149	0	30	25	14.85	349	100	25	0	34.85
150	0	30	20	14.95	350	100	0	0	34.95
151	0	40	25	15.05	351	30	0	20	35.05
152	0	40	0	15.15	352	0	20	20	35.15
153	0	30	0	15.25	353	0	20	20	35.25
154	0	40	0	15.35	354	0	20	30	35.35
155	0	25	20	15.45	355	0	30	40	35.45
156	0	15	20	15.55	356	0	25	50	35.55
157	0	15	30	15.65	357	0	30	25	35.65
158	0	20	20	15.75	358	0	60	25	35.75
159	0	20	0	15.85	359	0	40	80	35.85
160	0	25	0	15.95	360	0	60	30	35.95
161	0	30	0	16.05	361	0	65	25	36.05
162	0	30	0	16.15	362	20	80	25	36.15
163	0	20	0	16.25	363	5	70	0	36.25
164	0	30	30	16.35	364	0	90	0	36.35
165	5	25	60	16.45	365	0	60	0	36.45
166	10	15	30	16.55	366	0	65	0	36.55
167	5	30	25	16.65	367	0	25	30	36.65
168	10	0	30	16.75	368	0	60	25	36.75
169	1	20	30	16.85	369	0	45	0	36.85
170	1	30	15	16.95	370	25	40	0	36.95
171	10	25	30	17.05	371	25	30	0	37.05
172	10	15	20	17.15	372	5	20	10	37.15
173	1	25	25	17.25	373	10	20	15	37.25
174	0	25	25	17.35	374	20	90	25	37.35
175	25	30	20	17.45	375	15	50	40	37.45
176	60	25	20	17.55	376	0	30	40	37.55
177	20	30	25	17.65	377	1	90	55	37.65
178	1	20	30	17.75	378	5	90	25	37.75
179	0	20	30	17.85	379	30	25	25	37.85
180	5	20	30	17.95	380	60	20	25	37.95
181	25	15	15	18.05	381	50	15	30	38.05
182	0	25	15	18.15	382	20	15	0	38.15
183	15	25	15	18.25	383	0	25	25	38.25
184	40	25	15	18.35	384	0	20	15	38.35
185	10	15	35	18.45	385	0	30	0	38.45
186	5	60	25	18.55	386	0	20	0	38.55
187	10	20	30	18.65	387	0	25	0	38.65
188	15	30	15	18.75	388	0	25	0	38.75
189	15	30	25	18.85	389	0	0	25	38.85
190	15	20	30	18.95	390	0	5	10	38.95
191	10	25	25	19.05	391	0	25	35	39.05
192	10	30	20	19.15	392	0	30	25	39.15
193	10	25	0	19.25	393	0	20	90	39.25
194	20	20	25	19.35	394	0	40	80	39.35

Continuação Apêndice 3.

195	1	55	20	19.45	395	0	20	0	39.45
196	0	25	30	19.55	396	5	0	85	39.55
197	0	25	25	19.65	397	20	25	60	39.65
198	0	15	35	19.75	398	70	0	20	39.75
199	0	25	60	19.85	399	15	0	20	39.85
200	0	30	35	19.95	400	0	20	15	39.95

Apêndice 4. Obtenção de um estimador pontual do valor s_0 , com s_i vizinhos.

Considera-se que:

$$E \{ \hat{Z}_{(s_0)} - Z(s_0) \} = 0$$

$$E \left\{ \sum_i \lambda_i Z(s_i) - Z(s_0) \right\} = 0$$

$$\sum_i \lambda_i E \{ Z(s_i) \} - E \{ Z(s_0) \} = 0 \quad E \left\{ \sum_i \lambda_i Z(s_i) - Z(s_0) \right\} = 0$$

Pela hipótese intrínseca pode-se dizer:

$$\begin{aligned} \mu \sum_i \lambda_i - \mu &= 0 \\ \sum_i \lambda_i &= 1 \end{aligned}$$

A soma de pesos igual a um, garante a não tendenciosidade do preditor.

A variância de estimação pode ser descomposta na forma:

$$\begin{aligned} \text{Var} \{ \hat{Z}(s_0) - Z(s_0) \} &= E \{ [\hat{Z}(s_0) - Z(s_0)]^2 \} = \\ &= E \{ [\hat{Z}(s_0)]^2 \} - 2E \{ \hat{Z}(s_0) Z(s_0) \} + E \{ [Z(s_0)]^2 \} = \\ &= E \left\{ \left[\sum_i \lambda_i Z(s_i) \right] \left[\sum_j \lambda_j Z(s_j) \right] \right\} - 2E \left[\sum_i \lambda_i Z(s_i) Z(s_0) \right] + E \{ [Z(s_0)]^2 \} = \\ &= \sum_i \sum_j \lambda_i \lambda_j E [Z(s_i) Z(s_j)] - 2 \sum_i \lambda_i E [Z(s_i) Z(s_0)] + E \{ [Z(s_0)]^2 \} = \end{aligned}$$

Pode-se utilizar conceito do semivariograma para substituir nesta equação.

Lembremos que:

$$2\gamma_{ij} = 2\gamma(s_i, s_j) = 2\gamma(h) = E \{ [Z(s_i) - Z(s_j)]^2 \}$$

onde os pontos s_i e s_j estão separados uma distância h .

Também tem-se:

$$E \{ [Z(s_i) - Z(s_j)]^2 \} = E \{ [Z(s_i)]^2 \} + E \{ [Z(s_j)]^2 \} - 2E [Z(s_i) Z(s_j)]$$

$$E[Z(s_i)Z(s_j)] = \frac{1}{2} \left\{ E\{[Z(s_i)]^2\} + E\{[Z(s_j)]^2\} \right\} - \gamma(s_i, s_j)$$

Substituindo e continuando a expansão da variancia de estimação temos:

$$\begin{aligned} &= \sum_i \sum_j \lambda_i \lambda_j \left\{ \frac{1}{2} \left\{ E\{[Z(s_i)]^2\} + E\{[Z(s_j)]^2\} \right\} - \gamma(s_i, s_j) \right\} - \\ &- 2 \sum_i \lambda_i \left\{ \frac{1}{2} \left\{ E\{[Z(s_i)]^2\} + E\{[Z(s_0)]^2\} \right\} - \gamma(s_i, s_0) \right\} + E\{[Z(s_0)]^2\} = \\ &= \frac{1}{2} \sum_i \lambda_i E\{[Z(s_i)]^2\} \left(\sum_j \lambda_j \right) + \frac{1}{2} \sum_j \lambda_j E\{[Z(s_j)]^2\} \left(\sum_i \lambda_i \right) - \\ &- \sum_i \sum_j \lambda_i \lambda_j \gamma(s_i, s_j) - \sum_i \lambda_i E\{[Z(s_i)]^2\} - E\{[Z(s_0)]^2\} \left(\sum_j \lambda_j \right) + 2 \sum_i \lambda_i \gamma(s_i, s_0) + E\{[Z(s_0)]^2\} \end{aligned}$$

lembrando que : $\left(\sum_i \lambda_i \right) = \left(\sum_j \lambda_j \right) = 1$, podemos escrever:

$$\text{Var}\{\hat{Z}(s_0) - Z(s_0)\} = - \sum_i \sum_j \lambda_i \lambda_j \gamma(s_i, s_j) + 2 \sum_i \lambda_i \gamma(s_i, s_0) - 2 \nu \left(\sum_i \lambda_i \right)$$

Devemos encontrar o mínimo da variancia, com a restrição de que a soma dos pesos seja 1. Realizando as derivadas parciais destas equações com respeito a λ_i e η , se forma um sistema de equações do tipo:

$$\begin{cases} \sum_i \hat{\lambda}_i \gamma(s_i, s_j) + \hat{\eta} = \gamma(s_i, s_j) \\ \sum_i \hat{\lambda}_i = 1 \end{cases}$$

Pode ser expresado como:

$$\begin{bmatrix} \gamma(s_1, s_1) & \gamma(s_1, s_2) & \Lambda & \gamma(s_1, s_k) & 1 \\ \gamma(s_2, s_1) & \gamma(s_2, s_2) & \Lambda & \gamma(s_2, s_k) & 1 \\ \text{M} & \text{M} & \text{M} & \text{M} & 1 \\ \gamma(s_k, s_1) & \gamma(s_k, s_2) & \Lambda & \gamma(s_k, s_k) & 1 \\ 1 & 1 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} \hat{\lambda}_1 \\ \hat{\lambda}_2 \\ \text{M} \\ \hat{\lambda}_k \\ \hat{\eta} \end{bmatrix} = \begin{bmatrix} \gamma(s_1, s_0) \\ \gamma(s_2, s_0) \\ \text{M} \\ \gamma(s_k, s_0) \\ 1 \end{bmatrix}$$

$$A \begin{bmatrix} \hat{\lambda} \\ \hat{\eta} \end{bmatrix} = b \qquad \begin{bmatrix} \hat{\lambda} \\ \hat{\eta} \end{bmatrix} = A^{-1}b$$

Os coeficientes $\gamma(s_i, s_j)$ da matriz A e $\gamma(s_i, s_0)$ do vetor b são obtidos substituindo a distância h_{ij} entre os vizinhos s_i e s_j e s_i e s_0 respectivamente e substituindo esse valor no modelo de semivariograma ajustado.