ANDRÉIA LUIZA GONZAGA FERES

ANÁLISE ESTATÍSTICA ESPACIAL NA AVALIAÇÃO DE PRODUTIVIDADE NO MELHORAMENTO GENÉTICO DO FEIJOEIRO

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Estatística Aplicada e Biometria, para obtenção do título de *Magister Scientiae*.

VIÇOSA
MINAS GERAIS - BRASIL
2009

ANDRÉIA LUIZA GONZAGA FERES

ANÁLISE ESTATÍSTICA ESPACIAL NA AVALIAÇÃO DE PRODUTIVIDADE NO MELHORAMENTO GENÉTICO DO FEIJOEIRO

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Estatística Aplicada e Biometria, para obtenção do título de *Magister Scientiae*.

APROVADA: 18 de fevereiro de 2009.

_____	_____
José Eustáquio Souza Carneiro (Coorientador)	Nerilson Terra Santos (Coorientador)
_____	_____
Paulo Roberto Cecon	Sebastião Martins Filho

_____

Antônio Policarpo Souza Carneiro
(Orientador)

AGRADECIMENTOS

Primeiro a Deus, pela oportunidade de mais uma vez colocar uma tarefa em minha vida me proporcionando um novo recomeço.

A minha mãe que mesmo ausente se faz presente em todas as minhas buscas.

A minha família e amigos que compreenderam minha constante ausência em vários momentos.

A todos os professores do departamento de Estatística que contribuíram para que eu me tornasse uma pessoa melhor em minha profissão. E principalmente àqueles que vi cumprindo o seu papel com muita dedicação e sabedoria, ao perceber que o conhecimento é adquirido todos os dias e não uma coisa pronta e acabada.

Ao Paulo Martins (Paulinho-DPI), por todos os bons-dias seguidos de um sorriso, que me fazia repensar principalmente nos dias em que as coisas não iam bem.

Ao professor José Eustáquio Souza Carneiro e a todas as pessoas com quem convivi na agronomia. Em especial a Vanessa, José Ângelo e Lelisângela que sempre estavam dispostos a responder minhas dúvidas.

Especialmente ao Lucas e ao Gilmar que me ajudaram incondicionalmente em todas as etapas desta dissertação.

E ao meu orientador, Antônio Policarpo Souza Carneiro, pela colaboração, boa vontade e maneira como me orientou.

BIOGRAFIA

ANDRÉIA LUIZA GONZAGA FERES**,** filha de Clarice Feres Gonzaga e Luiz Gonzaga, nascida em 9 de agosto de 1970, em Ponte Nova-MG.

Em janeiro de 1994 graduou-se em Licenciatura Plena em matemática pela Universidade Federal de Viçosa (UFV).

Entre 1995 e 2001, trabalhou no COLÉGIO CIP- CON/OBJETIVO - Governador Valadares-MG, como professora do ensino médio e pré-vestibular.

Em 1998, concluiu sua especialização (“Lato-Sensu”) Fundamentos da Matemática, pela Universidade Pontifícia Católica (PUC) – Belo Horizonte-Minas Gerais.

Em 2002, foi nomeada em caráter efetivo como professora de ensino fundamental e médio junto a Secretária de Educação de Minas Gerais, com a qual mantém vínculo empregatício até a presente data.

Em 2007, iniciou seu curso de Mestrado na área de Estatística Aplicada, junto a Universidade Federal de Viçosa (UFV) – Viçosa-MG, sob orientação do professor Dr. Antônio Policarpo Souza Carneiro, tendo defendido em fevereiro de 2009 a dissertação intitulada “Análise espacial na avaliação de produtividade no melhoramento genético do feijoeiro”.

SUMÁRIO

RESUMO..... vi

ABSTRACT..... viii

1. INTRODUÇÃO..... 01

2. OBJETIVO GERAL..... 03

2.1. Objetivo Específico..... 03

3. REFERENCIAL TÉORICO..... 04

3.1. Programas de melhoramento genético do feijoeiro..... 04

3.2. Delineamentos experimentais tradicionais..... 06

3.3. Dependência espacial..... 06

3.4. Métodos de Análise Espacial..... 10

3.4.1. Método de Papadakis..... 10

3.4.2. Método de médias móveis..... 10

3.4.3. Método para erros dependentes (ED)..... 11

4. REFERÊNCIAS BIBLIOGRÁFICAS..... 12

CAPÍTULO 1

CARACTERIZAÇÃO DA DEPENDÊNCIA ESPACIAL EM EXPERIMENTOS PARA AVALIAÇÃO DA PRODUTIVIDADE NO FEIJOEIRO (*Phaseolus vulgaris L.*)..... 14

RESUMO..... 14

ABSTRACT..... 15

1. INTRODUÇÃO..... 16

2. MATERIAL E MÉTODO..... 17

3. RESULTADOS E DISCUSSÃO..... 23

4. CONCLUSÃO..... 41

5. REFERÊNCIAS BIBLIOGRÁFICAS..... 42

CAPÍTULO 2
EFICIÊNCIA DE MÉTODOS DE ANÁLISE ESTATÍSTICA ESPACIAL PARA AVALIAÇÃO DA PRODUTIVIDADE DE FAMÍLIAS DO FEIJOEIRO..... 43

RESUMO..... 43

ABSTRACT..... 45

1. INTRODUÇÃO. 46

2. MATERIAL E MÉTODO..... 48

3. RESULTADOS E DISCUSSÃO..... 55

4. CONCLUSÕES..... 73

5. REFERÊNCIAS BIBLIOGRÁFICAS..... 74

ANEXOS..... 75

RESUMO

FERES, Andréia Luiza Gonzaga, M.Sc., Universidade Federal de Viçosa, fevereiro de 2009. **Análise estatística espacial na avaliação de produtividade no melhoramento genético do feijoeiro**. Orientador: Antônio Policarpo Souza Carneiro. Coorientadores: José Eustáquio Souza Carneiro e Nerilson Terra Santos.

A dependência espacial é a propensão pela qual o valor de uma variável observada em certa posição tende a assemelhar-se mais aos valores vizinhos do que ao restante das observações do conjunto amostral. O objetivo deste trabalho foi a avaliação da eficiência de alguns métodos de análise estatística espacial para melhorar a precisão experimental em ensaios de seleção de famílias em programa de melhoramento genético do feijoeiro. Foram utilizados dados para produtividade de oito experimentos montados em látice, conduzidos na estação experimental de Coimbra, pertencente ao Departamento de Fitotecnia da Universidade Federal de Viçosa (UFV). Para cada análise, foram analisados os métodos de análise de vizinhança: método das médias móveis, método de Papadakis, reaplicação das médias móveis, reaplicação do método Papadakis e o método com erros dependentes em comparação com a análise tradicional em látice, que desconsidera a dependência espacial entre parcelas quanto ao ajuste dos modelos e à classificação das famílias avaliadas. Foi verificada a existência da dependência espacial nos experimentos pelo teste de Durbin-Watson, semivariograma empírico e teórico com o ajustamento de um modelo geoestatístico para resíduos, verificando a coerência dos dados e percebendo a ocorrência da

dependência espacial em seis dos oito experimentos avaliados (capítulo 1). Dos seis experimentos, cinco apresentaram dependência espacial para análise em blocos e um para análise em látice e em blocos, com alcance variando de 2,6 a 33,3 m. No capítulo 2, foram utilizados o teste de Durbin-Watson para verificar a dependência espacial entre resíduos para as diferentes metodologias, além do critério de informação de Akaike (AIC) e do teste da razão da verossimilhança (LRT) para comparar o ajuste dos modelos. Utilizaram-se o coeficiente de variação e a eficiência relativa para comparar a eficiência das metodologias testadas em relação à análise em látice. Também foram utilizados o coeficiente de correlação de Spearman e a eficiência de seleção para comparar as classificações de famílias do feijoeiro obtidas na análise em látice e pelos métodos de análise espacial. Verificou-se que os métodos de análise espacial ajustaram-se melhor aos dados em relação aos métodos que assumem erros independentes. Quanto à precisão experimental, a eficiência da análise em látice foi maior em relação à análise em blocos quando há maior dependência espacial. Os métodos de análise espacial, médias móveis, Papadakis e o método com erros dependentes apresentaram razoável eficiência, enquanto as metodologias propostas, reaplicações das médias móveis e Papadakis foram altamente eficientes em relação à análise em látice, sendo mais eficiente o método reaplicação de Papadakis cuja vizinhança é composta por uma parcela superior e uma inferior em relação à parcela referência. Os métodos testados apresentaram razoável nível de concordância em relação às famílias selecionadas na análise em látice. As metodologias de reaplicação de médias móveis e reaplicação de Papadakis mostram-se altamente eficientes para melhorar a precisão experimental, além de conseguir garantir a independência entre resíduos, sendo uma excelente alternativa de análise em relação à análise em látice.

ABSTRACT

FERES, Andréia Luiza Gonzaga, M.Sc., Universidade Federal de Viçosa, February 2009. **Geostatistic analyses in the assessment the beans families productivity**. Adviser: Antônio Policarpo Souza Carneiro. Co-advisers: José Eustáquio Souza Carneiro and Nerilson Terra Santos.

Spatial dependence is a tendency that a variable´s value, observed in certain position, leads to approximate to its neighbors' value than the value of the rest of the elements observed in the sample. The aim of this work was to evaluate the efficiency of some methods of geostatistic analyses to improve the experimental accuracy of tests that selects families for genetic improvement program of beans plantation. Date set were used to evaluate the productivity of eight experiments set in lattice design, located in Coimbra's Experimental Station controlled by the Plant Science Department at Federal University of Viçosa – UFV. For each study was analyzed the neighborhood methods: Moving Means Method, Papadakis Method, Moving Means Replication, Papadakis Method Replication and the method that contains errors dependent in the comparison between traditional analyses in lattice which do not consider the spatial dependence among plots according to its fitting model, and the family classification in study. The existence of experiment's spatial dependence was observed through the Durbin-Watson's test, empirical and theoretical semivariogram with a geostatistic fitting model for residues checking the coherence of data and noticing the occurrence of spatial dependence in 6 out of 8 experiments in study (Chapter1). Five out of the six experiments showed

spatial dependence for analyses in blocks and one for analyses in lattice and blocks, reaching 2.6 to 33.3m. In chapter 2, it was used the Durbin-Watson's test to observe the spatial dependence among residues for different methodologies, besides the Akaike Information Criteria (AIC) and the Maximum Likelihood Method to compare the fitting model. The variation coefficient and the relative efficiency were used to compare the efficiency of tested methodologies related to the analyses in lattice. The Spearman's Correlation Coefficient and the efficiency of selection were also used to compare the classification of beans plantation families acquired through the analyses in lattice and through the spatial analyses method. It was noticed that the spatial analyses method had a better fitting to the data than the methods that assume independent errors. The experiments' accuracy, the analyses in lattice efficiency was higher compared to the analyses in blocks when there is higher spatial dependence. The spatial analyses method, Moving Means Method, Papadakis Method and the method with independent errors showed reasonable efficiency, while the proposed methodologies, Moving Means Replication and Papadakis were highly efficient than the analyses in lattice. Papadakis replication Method showed to be more efficient once its neighborhood is constituted by a superior and inferior plot compared to the reference plot. The tested methods showed reasonable concordance level compared to the selected families for analyses in lattice. The Moving Means Replication and Papadakis Replication methodologies showed highly efficient to improve the experiments accuracy, besides to guarantee the independence among residues, being an excellent alternative for analyses compared to the analyses in lattice.

1. INTRODUÇÃO

O feijão é um dos alimentos básicos do povo brasileiro, constituindo-se em uma das principais fontes de proteína, especialmente para as classes menos favorecidas. A média atual de consumo de feijão é de 17 kg /habitante/ano, bem inferior ao do início da década de 70, de 25 kg/ano (BORÉM e CARNEIRO, 2006). A preferência do consumidor é regionalizada e diferenciada principalmente quanto à cor e ao tipo de grão.

O feijoeiro comum (*Phaseolus vulgaris L.*) é cultivado ao longo do ano, na maioria dos estados brasileiros, proporcionando constante oferta do produto no mercado. É cultivado tanto para subsistência quanto em cultivos altamente tecnificados. Vários fatores são importantes para a obtenção de produto de qualidade, envolvendo cuidados que vão desde a fase de pré-produção, como a seleção da época mais adequada ao plantio, até a fase de comercialização, envolvendo questões relacionadas ao armazenamento (EMBRAPA, 2008).

O estado de Minas Gerais além de ser um grande produtor de feijão destaca-se pela variedade de tipos comerciais utilizados pelos produtores. Nas regiões Norte, Sul e Noroeste do estado, predominam feijões do tipo carioca. Na Zona da Mata, além do feijão preto, cultiva-se também o feijão vermelho, com maior aceitação e valor no mercado regional, chegando a superar até duas vezes o preço dos feijões dos grupos carioca e preto (SILVA, 2005).

Normalmente os cultivares de feijão são avaliados em diferentes ambientes, visando a dar subsídios à tomada de decisão na recomendação. Atualmente, o desenvolvimento, avaliação e recomendação de cultivares de feijão para diferentes regiões do estado de Minas Gerais são responsabilidades dos programas de melhoramento do feijoeiro de três instituições de pesquisa mineiras: Empresa de Pesquisa Agropecuária de Minas Gerais (Epamig), Universidade Federal de Lavras (Ufla) e Universidade Federal de Viçosa (UFV) em parceria com a Embrapa Arroz e Feijão. A UFV e a Epamig são responsáveis pelos ensaios de famílias, progênies e cultivares nas regiões da Zona da Mata e Metalúrgica (SILVA, 2005).

Nos programas de melhoramento de feijão, a fase inicial de seleção envolve a avaliação de um grande número de progênies. A avaliação destas famílias em experimentos com repetições torna-se difícil devido à exigência de

grandes áreas experimentais. Com poucas repetições e a necessidade de grandes áreas, tais experimentos dependem de formas mais sofisticadas de planejamento e de análises que garantam boa precisão experimental.

Na maioria das vezes, o delineamento em blocos ao acaso torna-se inviável devido à existência de grande heterogeneidade dentro dos blocos em avaliações que envolvem grande número de genótipos. Duarte (2000) cita que a análise de variância tradicional confia à casualização a tarefa de neutralizar os efeitos prejudiciais desse tipo de correlação, mas com frequência não o faz de forma adequada. Caso ocorra a dependência espacial entre parcelas e sua existência seja desconsiderada, tal comportamento pode impedir que a análise estatística constitua uma ferramenta eficaz do melhorista para a seleção de genótipos realmente superiores.

O delineamento em látice é bastante utilizado em programas de melhoramento de feijão, visando a aumentar a precisão experimental. Existem vários tipos de látice, porém um dos mais utilizados é o látice quadrado que apresenta algumas desvantagens em sua utilização. Para satisfazer a exigência de ser quadrado, o pesquisador pode ser levado a reduzir o número de famílias a serem testadas, podendo ocasionar o descarte de uma família promissora ou ter que completar com testemunhas que não têm nenhuma ligação com o experimento, ocasionando nesta situação a necessidade de uma maior área experimental, além de aumento nos custos.

Na tentativa de manter número adequado de repetições sem sacrificar a precisão experimental, pode ser utilizado pelos melhoristas como alternativa o método de análise espacial que leva em conta a posição das parcelas na área experimental (COSTA, 2003).

Algumas formas de modelagens espaciais têm sido sugeridas atualmente por serem mais apropriadas para experimentos que apresentam dependência espacial entre os erros experimentais. Entre elas, a análise de vizinhança (métodos médias móveis e Papadakis) e a análise considerando o modelo para erros dependentes.

2. OBJETIVO GERAL

Aplicação e avaliação da eficiência de alguns métodos de análise estatística espacial para melhorar a precisão experimental em ensaios de seleção de famílias em programa de melhoramento genético do feijoeiro na Zona da Mata de Minas Gerais.

2.1. Objetivo Específico

- Diagnosticar a existência da dependência espacial;
- Avaliar a eficiência da análise estatística espacial na classificação de famílias em ensaios de produtividade do feijoeiro em experimentos montados em látice;
- Comparar a eficiência de algumas metodologias de análise estatística espacial: método das médias móveis (M), método de Papadakis (P), reaplicação de médias móveis (RM) e Papadakis (RP) e erros dependentes (ED) em relação à análise padrão que desconsidera a dependência espacial;
- Desenvolver e avaliar uma nova metodologia de cálculo do índice ambiental, utilizado nas metodologias M e P, que leve em consideração todas as parcelas e não apenas algumas parcelas vizinhas;
- Propor uma metodologia de análise espacial que seja tão eficiente ou superior à análise em látice.

3. REFERENCIAL TÉORICO

3.1. Programas de melhoramento genético do feijoeiro

Os primeiros trabalhos de pesquisa com feijão no Brasil tiveram início na década de 30.

A década de 70 foi marcada pela criação da Embrapa Arroz e Feijão, evento importante no melhoramento genético do feijoeiro (RAMALHO et al. 2004).

De 1981 a 1997, durante a vigência do Sistema Brasileiro de Avaliação e Recomendação de Cultivares, instituído pelo Ministério da Agricultura, Pecuária e Abastecimento - Mapa, a rede de avaliação de linhagens de feijão foi conduzida pelas instituições do Sistema Nacional de Pesquisa Agropecuária - SNPA, universidades, empresas de pesquisa, setor cooperativo e iniciativa privada. Estas instituições empreendiam um programa integrado e dinâmico de avaliação de linhagens e cultivares, dentro das Comissões Técnicas Regionais de Feijão - CTs Feijão, para as regiões Sul, Sudeste/Centro-Oeste e Norte/Nordeste. Nestes 16 anos, as CTs Feijão, por meio de reuniões anuais, discutiram os resultados da rede de avaliação de linhagens e lançaram 34 novas cultivares de feijão, homologadas pelas Comissões Regionais de Avaliação e Recomendação de Cultivares CRCs/MA.

Com a lei n° 9.456 de 25/4/97, que trata da Proteção de Cultivares e o Decreto n° 2.366 de 5/11/97 que a regulamentou, o cenário acima descrito sofreu algumas mudanças, em que as relações institucionais, que tinham caráter cooperativo, se deslocaram para um ambiente de "competição", agora embasadas em contratos e convênios de cooperação técnica acompanhados de Planos Anuais de Trabalho. Foram extintas as CRCs e instituído o Registro Nacional de Cultivares em que a indicação de uma nova cultivar é de exclusiva responsabilidade do obtentor. Mesmo com a Lei de Proteção de Cultivares e com o Registro Nacional de Cultivares e sua publicação pelo Mapa, a rede oficial pública de avaliação de cultivares de feijão continua com o objetivo de informar, de maneira transparente, o desempenho das novas cultivares nas diversas regiões brasileiras, esperando, com isso, contribuir para que a

assistência técnica e o agricultor possam escolher a melhor cultivar para suas condições (EMBRAPA, 2003).

O Brasil é o principal produtor e consumidor mundial de feijão, e o melhoramento genético desta cultura já vem tendo uma importante influência na distribuição de renda e na qualidade da alimentação da população.

O melhoramento de plantas visa à obtenção de cultivares superiores aos existentes em cultivo, seja na produção de grãos, de fibras, na resistência às pragas e moléstias, seja, ainda, na maior riqueza de proteínas, óleo ou outras características de interesse. Os estudos de melhoramento geralmente abrangem vários ciclos de seleção na fase inicial, seguidos, na fase final, dos testes regionais das melhores variedades e híbridos. No estágio inicial, o uso de técnicas experimentais que possibilitem a obtenção de estimativas confiáveis são de fundamental importância (CONAGIN, 1997).

O objetivo específico de um programa de melhoramento é promover a adaptação de cultivares às diferentes regiões produtoras, com maior estabilidade de produção para cultivo nas diferentes épocas de semeadura, resistência a doenças e pragas e qualidade do produto final.

A Embrapa Arroz e Feijão coordena o programa nacional de melhoramento genético do feijoeiro, gerando populações e linhagens com características de interesse agronômico. As variações observadas na preferência dos consumidores orientam a pesquisa tecnológica e direcionam a produção e comercialização do produto, pois as regiões brasileiras são bem definidas quanto à preferência do grão de feijão consumido. Algumas características como a cor, o tamanho e o brilho do grão podem determinar seu consumo (EMBRAPA, 2003).

Atualmente, o desenvolvimento, avaliação e recomendação de cultivares de feijão para diferentes regiões do estado de Minas Gerais são responsabilidades das instituições de pesquisa que compõem o convênio de cooperação técnica, envolvendo a UFV, Ufla, Epamig e Embrapa Arroz e Feijão.

3.2. Delineamentos experimentais tradicionais

O delineamento experimental é parte do planejamento utilizado na experimentação agrícola e indica a forma como os tratamentos serão designados nas unidades experimentais. Eles são muito utilizados em programas de melhoramento vegetal em que é normal a avaliação de um grande número de tratamentos (linhagens, progênies, híbridos). Os melhoristas utilizam os delineamentos com o objetivo de controlar a heterogeneidade ambiental. Entre os mais utilizados estão o delineamento em blocos ao acaso e o delineamento em látice.

O delineamento em blocos é recomendado sempre que existir heterogeneidade entre as parcelas de um experimento ou uma variação no material, sendo utilizado como controle local. A área ou o material é subdivido em blocos que sejam homogêneos, e cada bloco deve conter todos os tratamentos. Porém, o delineamento em blocos às vezes pode perder sua eficiência quando violada a pressuposição de homogeneidade dentro dos blocos, exigindo que o melhorista opte por outro tipo de delineamento que possua mais controle local, como os delineamentos em blocos incompletos, entre eles o látice quadrado que vem sido muito utilizado na área de experimentação agrícola.

O látice quadrado foi introduzido por Yates (1936), e é um delineamento que vai subdividir a repetição em blocos menores, permitindo dispor de um número $V=K^2$ de cultivares em blocos de k parcelas, decorrendo daí que o número de tratamentos deve ser um número quadrado perfeito. Nos látices quadrados, os tratamentos de um bloco em uma repetição se distribuem por todos os blocos de qualquer das outras repetições (PIMENTEL GOMES, 1990)

3.3. Dependência espacial

A dependência espacial é a tendência de que o valor de uma variável observada em certa posição tem de assemelhar-se mais aos valores vizinhos do que ao restante das observações do conjunto amostral. A preocupação com

a dependência espacial é bastante antiga, segundo trabalhos publicados no início do século XX, conforme mostra Vieira (1995).

A metodologia de análise de dados proposta por Fisher utilizada na agricultura é baseada nas seguintes pressuposições: normalidade da variável, homogeneidade da variância e independência entre os resíduos. A normalidade e a homogeneidade podem ser verificadas por meio de testes específicos, e caso não sejam satisfeitas tais pressuposições, pode-se fazer a transformação dos dados para que a hipótese seja atendida. Já a suposição de independência entre resíduos deve ser garantida pela casualização. Duarte (2000) cita que a análise de variância tradicional confia à casualização a tarefa de neutralizar os efeitos prejudiciais desse tipo de correlação, mas, com frequência, não o faz de forma adequada, caso ocorra dependência espacial entre parcelas e sua existência seja desconsiderada. Este autor alerta que tal comportamento pode impedir que a análise estatística constitua uma ferramenta eficaz do melhorista para a seleção de genótipos realmente superiores.

As funções de semivariâncias são as mais utilizadas na avaliação da existência da dependência espacial. A semivariância é definida como

$$S(h) = 1/2\,\{\text{Var}\,[Z(t)-Z(t+h)]\},$$

em que Var [Z(t)-Z(t +h)]} é a variância da diferença entre os valores assumidos pela variável Z nas posições t e t +h . O estimador da semivariância é definido como estimador de Matheron:

$$\hat{S}(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(t+h) - Z(t)]^2 ,$$

em que N(h) é o número de pares separados pela distância h, citado por Duarte (2000). Esta função permite gerar o semivariograma experimental, um gráfico que representa uma função de semivariâncias em relação às suas respectivas distâncias e tem o papel de descrever a estrutura da variabilidade espacial. Um semivariograma permite verificar e modelar a dependência espacial de uma variável.

No semivariograma expresso pela Figura 1, observa-se que as semivariâncias aumentam com o aumento da distância(h), indicando

observações menos similares com maiores distâncias, e se estabilizam a uma certa distância chamada de alcance (a), e todos os valores superiores ao alcance possuem distribuição aleatória, portanto, são independentes entre si. O valor da S(h) constante é chamado de patamar, representado pela letra C no semivariograma.

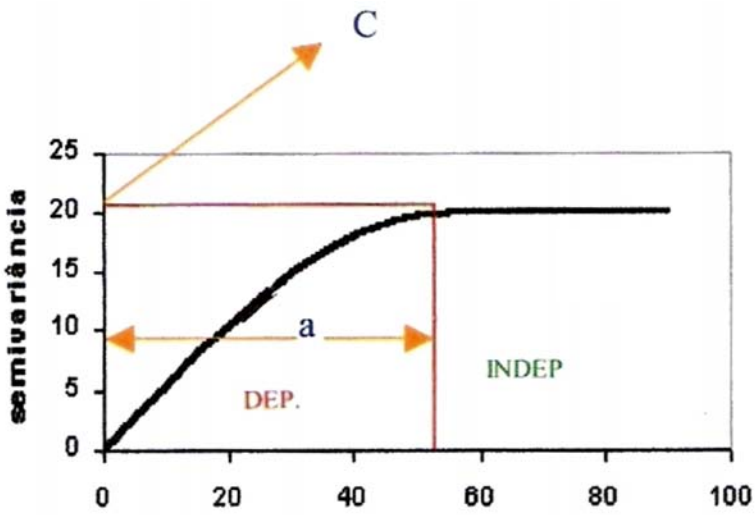

Figura 1 - Semivariograma sem o efeito pepita: C = patamar, a = alcance e h = distância entre observações

A utilização de dados amostrais na estimativa da semivariância e na construção do semivariograma revela que, frequentemente, para h=0, a semivariância S(0) difere de zero. A impossibilidade de fazer reamostragem exatamente sobre o mesmo ponto já amostrado e erros como erros de amostragem, erros de análise de laboratório etc. são justificativas desta descontinuidade na origem. Quando $S(0) \neq 0$, surge um novo termo no semivariograma chamado efeito pepita (C_0) e , neste caso, o patamar é dado por: $C+C_0$, sendo C a contribuição que é a diferença entre o patamar e o efeito pepita (Figura 2) (GUIMARÃES, 2004).

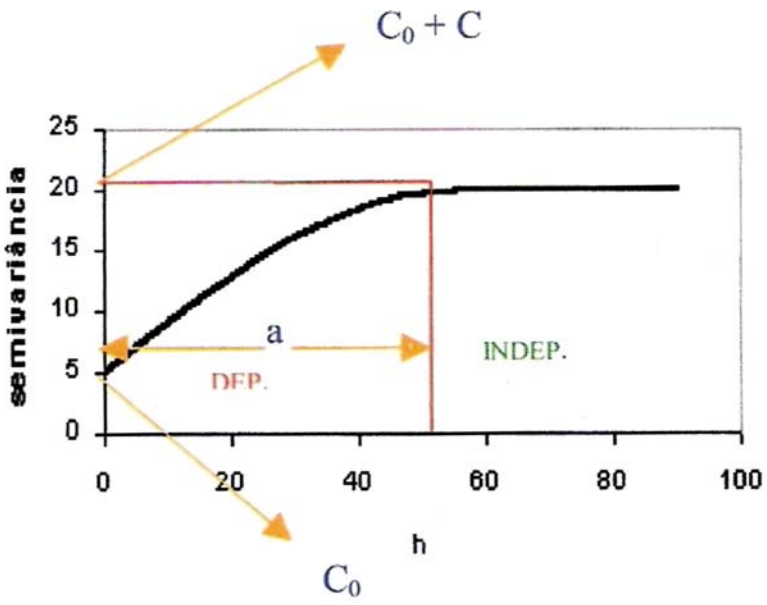

Figura 2 - Semivariograma experimental com efeito pepita: C_0 = efeito pepita, C = contribuição, $C+C_0$ = patamar, h = distância entre observações a = alcance.

Ao construir o semivariograma, deve-se verificar se ele apresenta comportamento anisotrópico, em que os parâmetros C, C_0, a e o modelo dependem da direção h, ou isotrópico, em que um semivariograma é suficiente para descrever o comportamento da variável em qualquer direção (GUIMARÃES, 2004).

Para a construção do semivariograma, gráfico que mostra o comportamento das semivariâncias (S(h)) em função da distância h, deve-se ajustar uma função contínua ao semivariograma experimental, obtida das amostras colhidas no campo, definindo os parâmetros contribuição (C), alcance (a) e efeito pepita (C_0). As funções de semivariogramas mais utilizadas são os modelos geoestatísticos: linear, exponencial, gaussiano e esférico (GUIMARÃES, 2004).

A verificação da existência da dependência espacial entre os resíduos nos permite utilizar as diferentes formas de análise espacial, que são metodologias de análises de dados indicadas quando os resíduos estão autocorrelacionados.

3.4. Métodos de Análise Espacial

Entre os métodos mais citados em análise espacial, destacam-se os métodos de análise de vizinhança: Papadakis (PAPADAKIS, 1937) e médias móveis (RICKEY, 1924) que consideram a dependência espacial na forma de análise de covariância. Outro método de ajuste para dependência espacial é o uso de modelo com erros dependentes que considera a dependência espacial por meio da matriz não diagonal de covariâncias residuais, levando em consideração a distância entre parcelas no campo experimental.

DUARTE (2000) comenta que este tipo de análise pode ser utilizado alternativamente ou como complemento às análises usuais de blocos.

3.4.1. Método de Papadakis

O método Papadakis (P) é utilizado com o objetivo de reduzir e corrigir a dependência entre erros. Para isso, busca eliminar tendências existentes nos solos, por meio do desempenho das parcelas vizinhas.

A metodologia utilizada por Papadakis corresponde a uma análise de covariância em que um índice ambiental é utilizado como covariável, ajustando a produção de cada parcela pela produção das parcelas vizinhas. Neste método, o índice ambiental é calculado como a média dos resíduos das parcelas vizinhas.

A análise de diferentes formas de cálculo do índice ambiental mostrou que o número e a localização das parcelas vizinhas interferem na eficiência do método de Papadakis (CARGNELUTTI FILHO, 2003).

3.4.2. Método de médias móveis

O método de médias móveis (M) ajusta o valor observado em cada parcela pelos valores das parcelas vizinhas, mediante análise de covariância (BARTLET, 1938). Na metodologia de médias móveis, o índice ambiental é utilizado como covariável, sendo calculado como a média simples dos valores observados nas parcelas vizinhas e não pela média dos resíduos das parcelas vizinhas, como no método Papadakis.

Costa (2003) afirma que os métodos de vizinhança médias móveis e Papadakis melhoraram o delineamento com controle local.

3.4.3. Método para erros dependentes (ED)

No método ED, assume-se que há dependência espacial entre os erros aleatórios e que eles têm distribuição normal com média zero e matriz de variâncias e covariâncias residuais (R), ε_{ij} ~N (0,R). A matriz de variâncias e covariâncias R é definida por:

i) R= I σ^2 (modelo com erros independentes)

ii) R= I C_0 + F σ^2 (modelo com erros dependentes),

sendo I a matriz identidade; σ^2, a variância residual no modelo com erros independentes e contribuição no modelo com erros dependentes; C_0 , o efeito pepita; e F, a matriz formada pelos elementos da função de distância, f(h), que corresponde aos modelos geoestatísticos usados para representar a variabilidade espacial.

4. REFERÊNCIAS BIBLIOGRÁFICAS

BORÉM, A. CARNEIRO, J. E. S. A cultura. In: VIEIRA, C.; PAULAJR.; T. J. de. **Feijão.** 2. Ed. Viçosa; UFV, 2006. P.13-18.

CONAGIN, A.; AMBROSANO, G. M. B; NAGAI, V. **Poder discriminativo da posição de classificação e dos testes estatísticos na seleção de genótipos.** Disponível em <http://www.scielo.br/scielo.php?script=sci_arttex&pid=S0006-87051997000200019>. Acesso em: Jan 2009.

COSTA, J. R. **Análise espacial e de vizinhança no melhoramento genético de plantas**. Lavras: ESAL, 2003. 90p. (Dissertação - Mestrado em Fitotecnia).

DUARTE, J. B. **Sobre o emprego e a análise estatística do delineamento em blocos aumentados no melhoramento genético vegetal.** 2000.293p. Tese (Doutorado em Agronomia/Genética e Melhoramento de Plantas)-Escola Superior de Agricultura Luiz de Queiroz, Piracicaba, SP.

EMBRAPA, 2003. **Agência de informação Embrapa feijão**. Disponível em: <http://sistemasdeproducao.cnptia.embrapa.br/FontesHTML/Feijao/CultivodoFeijoeiro/cultivares.htm> Acesso em: jan 2009.

EMBRAPA, 2008. **Agência de informação Embrapa feijão**. Disponível em: < http://www.agencia.cnptia.embrapa.br/Agencia4/AG01/Abertura.html> Acesso em: Nov 2008.

GUIMARÃES, E. C. **Geoestatística básica e aplicada.** Disponível em: < http://www.famat.ufu.br/prof/ednaldo/ednaldo.htm > Acesso em: jan 2009.

PAPADAKIS, J. S. **Méthode statistique pour des expériences sur champ.** Thessalonike: Institut d'Amélioration des Plantes à Salonique, 1937. 30p. (Bulletin, 23).

PIMENTEL GOMES, F. **Curso de estatística experimental.** 13 ed. Piracicaba: Nobel/USP-ESALQ, 1990. 468P.

RAMALHO, M. A. P.; ABREU, A. de F. B.; CARNEIRO, J. E. S. Cultivares. **Informe Agropecuário,** EPAMIG, V.25, N. 223, P. 21-32, 2004.

RIBEIRO JÚNIOR, P. J. **Métodos geoestatístico no estudo da variabilidade espacial de parâmetro s do solo**. Piracicaba, 1995.99p. Dissertação (Mestrado)-ESALQ/USP.

RICKEY, F.D. Adjusting yields to their regression on moving average, as a means of correction for soil heterogeneity. **Journal of Agricultural Research**, v. 27, 79-90, 1924.

SILVA, L. C. **Recomendações de cultivares de feijão-vermelho para o estado de Minas Gerais**. 2005. 77p. Dissertação (Mestrado) – Universidade Federal de Viçosa, Viçosa, MG, 2005.

VIEIRA, S. R. **Curso de atualização em conservação do solo. Uso de geoestatística.** Campinas, IAC, v.1e2, 1995.

CAPÍTULO 1

CARACTERIZAÇÃO DA DEPENDÊNCIA ESPACIAL EM EXPERIMENTOS PARA AVALIAÇÃO DA PRODUTIVIDADE NO FEIJOEIRO (*Phaseolus vulgaris L.*)

RESUMO

O uso do solo, com o passar do tempo, conduz a aumento na sua heterogeneidade pelos processos de desmatamento, seu preparo, rotação de cultura, localização de aplicação de fertilizantes etc. No caso dos fertilizantes, como são aplicados em faixas ou em linhas, isso faz com que o sistema de amostragem varie consideravelmente, podendo gerar uma dependência espacial entre parcelas (autocorrelação espacial), que afeta um pressuposto da análise de variância clássica que é a independência dos erros experimentais. Normalmente, em experimentos de campo, pressupõe-se a independência entre resíduos, mas nem sempre esta condição é atendida, comprometendo os resultados obtidos. Uma escolha equivocada do modelo de análise pode comprometer a eficiência do programa de melhoramento no sentido de os genótipos pouco promissores poderem ser avaliados e mantidos em próximos ciclos seletivos acarretando desperdício de tempo e recursos. No presente trabalho, avaliou-se a presença da dependência espacial em experimentos de avaliação de genótipos de feijoeiro. Os dados utilizados nesse estudo foram obtidos de experimentos instalados em látice e analisados como látice e blocos casualizados. O teste de Durbin- Watson e o semivariograma empírico foram usados para verificar a presença de autocorrelação espacial, além do semivariograma teórico com o ajustamento de modelos geoestatísticos para resíduos para verificar a coerência dos dados. Dos oito experimentos, cinco apresentaram dependência espacial para análise em blocos e um para análise em látice e em blocos, com alcance variando de 2,6 a 33,3 m. O tamanho dos experimentos não foi fator determinante da dependência espacial. Os modelos geoestatísticos esférico, exponencial e gaussiano com efeito pepita foram adequados para representar a estrutura espacial na análise em blocos. E a análise em látice foi eficiente para garantir a independência entre resíduos.

CHARACTERIZING THE SPATIAL DEPENDENCE IN EXPERIMENTS SET TO EVALUATE THE BEANS PRODUCTIVITY (*Phaseolus vulgaris L.*)

ABSTRACT

The land cover leads the increase of its heterogeneity during some time because of the deforestation, soil tillage, crop rotation, local fertilizer application, etc. Considering the fertilizers, as they are applied in stripe or lines, this makes sampling system significantly varies, creating a spatial dependence among plots (spatial autocorrelation) which affects a presupposed classical analyses of variance which is the independence of experimental errors. Normally, in experimental field, presuppose the independence among residues, but this condition is not always attended, affecting the experiment's outcome. Choosing a wrong model for analyses can compromise the program's improving efficiency in a way that genotypes less promising could be evaluated and maintained for next selective cycles, causing time and fond waste. In this present work was evaluated the existence of spatial dependence in experiments of genotypes evaluation of beans plantation. The data set used in this study was acquired from experiments set in lattice and analyzed as lattice and random blocks. The Durbin-Watson's test and the empirical semivariogram were used to look for the presence of spatial autocorrelation, besides the theoretical semivariogram with geostatistic fitting model for residues to check the data reliability. Five out of eight experiments showed spatial dependence for analyses in blocks and one for both, lattice and block analyses, reaching 2.6 to 33.3m. The experiment's size was not a factor that determined the spatial dependence. The geostatistic models, such as spherical, exponential and Gaussian with nugget effect, were the right ones to represent spatial structure for the analyses in blocks. And the analyses in lattice were efficient to guarantee the residues' independence.

1. INTRODUÇÃO

Daniel Krige, em 1951, trabalhava com dados de concentração de ouro na África do Sul e concluiu que a variância dos dados possuía uma estruturação que dependia da distância de amostragem. A partir desta constatação, surgiram os conceitos básicos de geoestatística (GUIMARÃES, 2004)

A geoestatística é baseada fundamentalmente em apenas dois conceitos: semivariograma e a Krigagem. O primeiro tem o papel de descrever a estrutura da variabilidade espacial e o segundo possui a função de predizer, não tendenciosamente e com variância mínima, os valores não medidos (CRESSIE, 1993).

A utilização da análise espacial busca aumentar a eficiência do controle local e a precisão experimental. Os métodos de análise espacial são métodos de análise estatística que levam em conta a posição das parcelas na área experimental. Neste tipo de análise, considera-se que observações localizadas próximas são mais semelhantes entre si do que as que se encontram mais distantes, caracterizando a dependência espacial entre os resíduos.

Fenômenos naturais, principalmente aqueles observados na agricultura, tais como produtividade de grãos e incidência de doenças, apresentam-se espacialmente correlacionados (SILVA, 2004).

Segundo Oda (2005), deixar de considerar a dependência espacial não apenas viola o pressuposto, mas também significa deixar de observar importantes aspectos que determinam a ocorrência do fenômeno estudado. Considerar a estrutura de variabilidade espacial contribui para ampliar a fração explicada da variabilidade dos dados.

Martinez (1994) argumentou que a violação das pressuposições assumidas no modelo de análise não espacial compromete a eficiência da análise de variância tradicional, mesmo que se cumpram os princípios básicos da experimentação.

O objetivo do presente trabalho foi avaliar a dependência espacial entre resíduos, em experimentos de avaliação de produtividade de famílias do feijoeiro, para análises em látice e em blocos. E também avaliar a eficiência de modelos geoestatísticos para caracterização da estrutura de variabilidade espacial entre resíduos de experimentos instalados em látice e analisados em blocos.

2. MATERIAL E MÉTODO

Neste trabalho, foram utilizados dados de experimentos de avaliação de progênies, famílias e linhagens de feijão do programa de melhoramento da Universidade Federal de Viçosa, conduzidos nas safras do inverno 2006/2007 e seca 2007/2008 na estação experimental de Coimbra, município de Coimbra, MG (690m de altitude, 20º45' S e 42º51' W) pertencente ao Departamento de Fitotecnia da Universidade Federal de Viçosa (UFV).

Foram analisados dados de oito experimentos em látice quadrado, conforme detalhes apresentados na Tabela 1.

Tabela 1 - Características dos experimentos avaliados: tipo de látice, número de repetições (r), número de famílias (F), número total de parcelas (NP), espaçamento (E), número de linhas por parcela (L), comprimento das linhas (C) em metros

Exp	Tipo de látice	r	F	NP	E	Tamanho da parcela	
						L	C
1	7x7	3	49	147	0,45	2	2
2	5x5	3	25	75	0,45	2	2
3	20x20	2	400	800	0,45	1	2
4	10x10	3	100	300	0,45	2	2
5	7x7	3	49	147	0,45	2	2
6	6x6	3	36	108	0,45	3	2
7	15x15	3	225	675	0,50	2	2
8	10x10	3	100	300	0,45	2	1,5

Os dados referentes ao espaçamento e tamanho da parcela foram utilizados para a obtenção das coordenadas relativas ao centro de cada parcela dentro da área experimental. A posição e os valores observados são informações necessárias para a utilização da análise estatística espacial. A posição de cada parcela na grade experimental foi especificada pelas coordenadas L e C, relativas ao centro das parcelas: L a coordenada no sentido da largura da parcela e C a coordenada no sentido do seu comprimento. Assim, a distância (d) entre as parcelas i e j foi obtida por $d = [(L_j - L_i)^2 + (C_j - C_i)^2]^{0,5}$

L_J = ordenada relacionada à largura na parcela J;

L_i = ordenada relacionada à largura na parcela i;

C_j= abscissa relacionada ao comprimento na parcela j; e

C_i= abscissa relacionada ao comprimento na parcela i.

Os dados experimentais foram analisados estatisticamente de acordo com os seguintes modelos estatísticos:

i) modelos com erros independentes

ii) modelo com erros dependentes

Para a análise de dados com erros independentes, foram adotados os seguintes modelos:

Modelo 1 (análise em látice): $y_{ijk} = \mu + r_k + b_{j(k)} + f_i + e_{ijk}$, em que

y_{ijk} é o valor observado da família i, no bloco j dentro da repetição k ;

μ é a constante associada a todas as observações;

r_k é o efeito fixo da repetição k;

$b_{j(k)}$ é o efeito fixo do bloco j dentro da repetição k;

f_i é o efeito fixo de família i;

e_{ijk} são os erros aleatórios associados às observações, assumindo independência entre os erros.

Modelo 2 (análise em blocos ao acaso): $y_{ik} = \mu + r_k + f_i + e_{ijk,}$ em que y_{ik} é o valor observado para a família i, na repetição k.

E para análise com erros dependentes, a análise dos dados adotou o seguinte modelo:

Modelo 3 (erros dependentes e análise em blocos ao acaso): $y_{ij} = \mu + r_k + f_i + e_{ijk,}$ em que e_{ijk} são os erros aleatórios associados às observações, assumindo a dependência entre os erros.

As seguintes estruturas para matriz de variâncias e covariâncias (R) foram consideradas:

i) R= I C (modelo com erros independentes)

ii) R= I C_0 + F C (modelo com erros dependentes), em que I é a matriz identidade, C é a variância residual no modelo com erros independentes e contribuição no modelo com erros dependentes, C_0 é o efeito pepita e F é a matriz formada pelos elementos da função de distância f(h). Esta função corresponde aos modelos geoestatísticos usados para representar a variabilidade espacial.

Os experimentos foram implantados utilizando o delineamento em látice, porém foram realizados dois tipos de análises: uma em látice e outra em blocos. Os resíduos estimados nas análises em látice e em blocos foram submetidos aos testes para verificar a existência da dependência espacial.

Inicialmente a dependência espacial foi avaliada usando o teste de Durbin-Watson (DW), que testa a hipótese de autocorrelação espacial nula (H_0: ρ = 0). A estatística do teste DW é definida como:

$$d = \frac{\sum_{s=2}^{n} (\hat{e}_s - \hat{e}_{s-1})^2}{\sum_{l=1}^{n} \hat{e}_s^2}$$

sendo s=1, 2, 3,..., n, a ordem de posicionamento da parcela no experimento, associada ao resíduo $\hat{e}_s$, tendo essa ordem obedecido à numeração sucessiva das parcelas, de modo que $\hat{e}_s$ e $\hat{e}_{s-1}$ indicam resíduos de parcelas adjacentes.

A relação entre d e ρ é aproximadamente d=2(1- ρ). Assim, caso não exista autocorrelação, o valor esperado de d é 2: valores significativamente inferiores a 2 indicam autocorrelação positiva; e valores significativamente superiores a 2 indicam autocorrelação negativa (REIS E MIRANDA FILHO, 2003).

O grau de dependência espacial (GD) foi estimado em função dos parâmetros do semivariograma, efeito pepita (C_0) e patamar (C + C_0), dado por:

$$GD = \frac{C}{C + Co},$$

Os seguintes intervalos para o grau de dependência espacial foram adotados:

i) Forte dependência espacial se 0,75<GD < 1;

ii) Moderada dependência espacial se 0,25 < GD<0,75; e

iii) Fraca dependência espacial se GD<0,25 (GUIMARÃES, 2004).

A dependência espacial também foi analisada por meio de semivariogramas empíricos que mostram os comportamentos das semivariâncias em função da distância (h). A semivariância foi estimada pela expressão a seguir:

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

sendo N(h) o número de pares de resíduos separados por um vetor de distância h; Z(i) e Z(i + h), valores do i-ésimo resíduo, relativos às parcelas i e i +h separados pelo vetor h.

A partir dos dados do semivariograma empírico, foram ajustados os modelos geoestatísticos, com e sem efeito pepita, para descrever a variabilidade espacial dos resíduos nas análises em látice e em blocos, estimando os parâmetros da contribuição (C), alcance (a) e efeito pepita (C_0). Os principais modelos ajustados, com efeito pepita, são descritos abaixo:

Modelo exponencial (Exp)

$$S(h) = C_0 + C\,[1 - e^{(-3h/a)}],$$

para $0 < h < d$, em que d é a distância máxima na qual o semivariograma é definido.

Modelo esférico (Esf)

$$S(h) = C_0 + C\,[1{,}5\,(h/a) - 0{,}5\,(h/a)^3],$$

para $0 < h < a$.

Modelo Gaussiano (Gau)

$$S(h) = C_0 + C\,[1 - e^{(3h^2/a^2)}],$$

para $0 \leq h \leq d$, sendo d a distância máxima na qual o semivariograma é definido, pois neste caso o patamar é atingido apenas assintoticamente (GUIMARÃES, 2004).

Nos modelos exponencial e gaussiano, o valor do alcance efetivo é obtido multiplicando o alcance estimado no SAS por 3 e $\sqrt{3}$, respectivamente (SAS INSTITUTE, 2006).

De acordo com Duarte (2000), é importante ressaltar que a vantagem de avaliar a dependência espacial pelo semivariograma é que, sob

estacionariedade, ele tem uma relação direta e simples com a função de autocovariância (C(h)) (Figura 3):

$$S(h) = \sigma^2 - C(h),$$

sendo σ^2 =C(h=0).

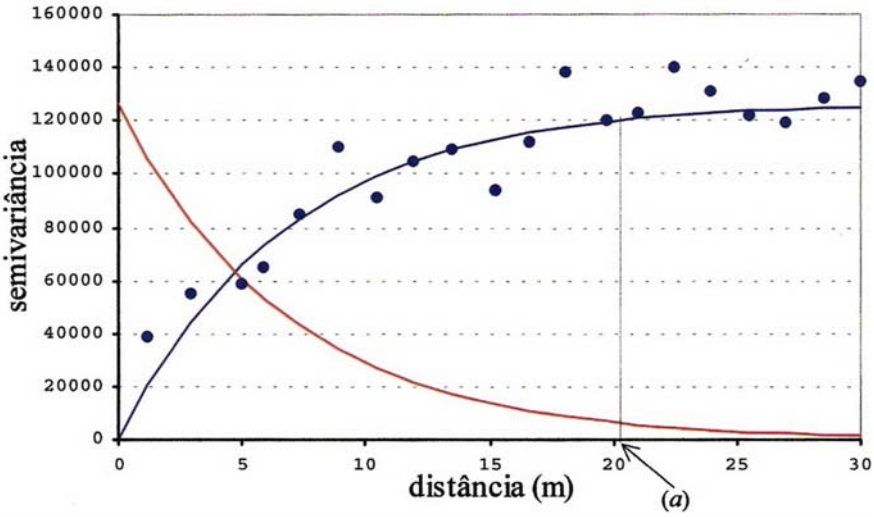

Figura 3 - Representação ilustrativa das funções de semivâriancias e covariâncias (linha azul e vermelha, respectivamente). Fonte: Duarte (2000).

O teste da razão de verossimilhança restrita foi aplicado para testar estatisticamente a diferença de adequação dos modelos com erros independentes e dependentes. Esta comparação consiste em testar a significância do alcance (a):

$$\begin{cases} H_0 : a = 0 \\ H_a : a > 0 \end{cases}$$

Sob normalidade dos dados, a estatística do teste tem distribuição de qui-quadrado com v graus de liberdade dada por:

$$X^2_{v=} [-2l_{REML}(R)_I] - [-2l_{REML}(R)_D],$$

em que I e D indicam os modelos com erros independentes e dependentes, respectivamente. Os valores $[-2l_{REML}(R)]$ são listados nas respectivas tabelas de informação das duas análises e v é a diferença entre o número de parâmetros do modelo completo (D) e do reduzido (I) (DUARTE, 2000).

O mesmo teste foi utilizado para verificar a adequação dos modelos com erros dependentes em relação à significância do efeito pepita, com a hipótese

$$\begin{cases} \mathrm{H}_0 : \mathrm{C}_0 = 0 \\ H_a : \mathrm{C}_0 > 0 \end{cases}$$

Os critérios de ajuste utilizados foram o valor da verossimilhança restrita (-2RLL) e o critério de Informação de Akaike (AIC). Foi considerado o melhor modelo, o mais adequado, aquele que apresentou o menor valor absoluto para estes critérios, como indicado por Littel et al. (1996).

O critério de AIC é baseado no valor que maximiza o logaritmo da verossimilhança restrita [$l_{REML}(R)$]. Estes critérios são estimados com a seguinte expressão:

$$AIC = RLL - p,$$

em que RLL é o valor da máxima verossimilhança e p o número de parâmetros do modelo

Segundo Akaike (1973), citado por Burnham e Anderson (2002), o modelo com menor valor de AIC é aquele que pode ser considerado o mais próximo do modelo real.

O valor da verossimilhança restrita permite testar estatisticamente a diferença de adequação de dois modelos como citado acima.

Os procedimentos do SAS (versão 8.0), utilizados para o teste de Durbin-Watson, foram o proc autoreg para gerar os semivariogramas empíricos, o proc variogram, utilizado para ajustes dos modelos geoestatísticos e o proc mixed, utilizado para avaliar os modelos com erros independentes.

3. RESULTADOS E DISCUSSÃO

Nas figuras e tabelas subsequentes são apresentados os resultados da avaliação da estrutura de dependência espacial entre resíduos estimados na análise em blocos e em látice. Dos oito experimentos analisados, seis apresentaram dependência espacial: um com dependência fraca e cinco com dependência moderada, de acordo com os intervalos de grau de dependência estimado (Tabela 2).

Tabela 2 - Graus de dependência espacial obtidos nos experimentos, tipo de látice e dimensões dos experimentos

Dependência espacial	GD	Experimento	Tipo de látice	Dimensões do experimento
Nula	-	1	7x7	20,25 x 16,80 = 340,20 m^2
	-	2	5x5	12,83 x 28,10 = 360,52 m^2
Fraca	0,09	3	20x20	21,38 x 45,80 = 979,20 m^2
Moderada	0,35	4	10x10	21,15 x34, 60 = 731,79 m^2
	0,34	5	7x7	19,35 x 20,60 = 398,61 m^2
	0,52	6	6x6	20,93 x 17,80 = 372,55 m^2
	0,25	7	15x15	23,50 x 79,40 = 1865,9 m^2
	0,43	8	10x10	25,65 x 30,65 = 786,17 m^2

* ne = não estimado, classificação com base no alto valor para autocorrelação espacial ($\hat{\rho}$ = 0,324)

No experimento 1, látice 7x7, as estatísticas do teste de Durbin-Watson (d) para autocorrelação espacial entre resíduos, de 1ª a 3ª ordem, apresentaram valores bem próximos de 2 para análise em látice com p-valores iguais a 0,0728, 0,2236 e 0,2590 (Tabela 3) não significativos para a hipótese de autocorrelação nula (P>0,01) e entre 1,6727 e 1,8608 para a análise em bloco com p-valores iguais a 0,1987, 0,0432 e 0,0334, não rejeitando a hipótese de nulidade da autocorrelação espacial (p>0,01). Estes resultados indicam que para o experimento 1, os resíduos estimados na análise em látice e em blocos são independentes, ou seja, não apresentam dependência espacial.

Tabela 3 - Estatística do teste de Durbin-Watson (d) para autocorrelação espacial entre resíduos estimados nas análises em látice e em blocos do experimento 1

Tipo de análise	Ordem	d	$\hat{\rho}^{(1)}$	p-valor
Látice	1	2, 2388	-0, 1194	0, 0728
	2	2,1113	-0,0556	0,2236
	3	2, 0789	-0, 0394	0, 2590
Blocos	1	1,8608	0,0696	0,1987
	2	1, 7054	0, 1473	0, 0432
	3	1, 6727	0, 1636	0, 0334

[1] $\hat{\rho}$ = autocorrelação espacial estimada

A independência espacial não está representada de forma bem caracterizada na análise em blocos (Figura 4), em que as semivariâncias apresentam um comportamento oscilatório.

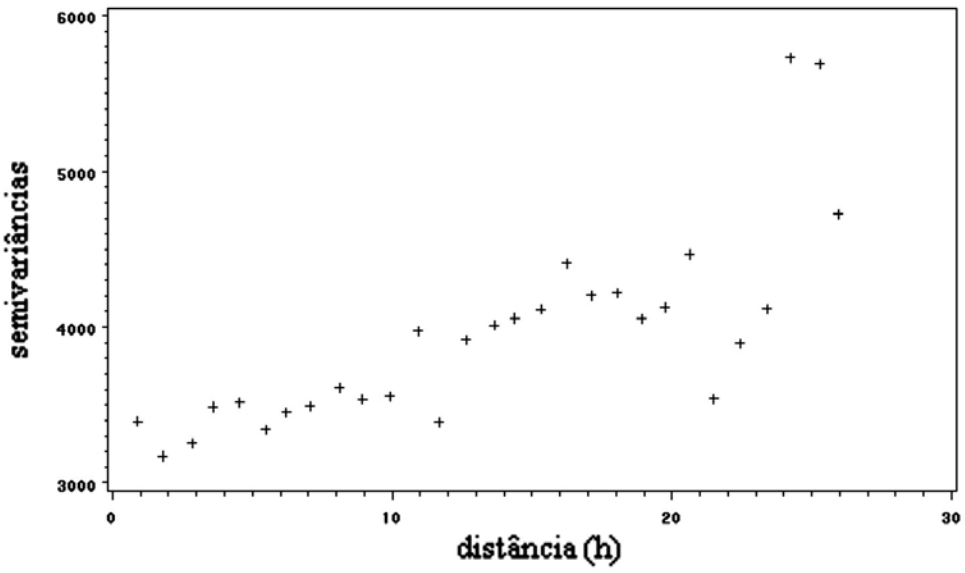

Figura 4 - Semivariograma empírico para resíduos estimados na análise em blocos, relativo ao experimento 1. (Semivâriancias em função da distância entre parcelas)

Entretanto, o látice se adapta ao modelo com efeito pepita puro que indica a independência entre resíduos (Figura 5), tornando-se, assim, caracterizada a independência espacial na análise em látice

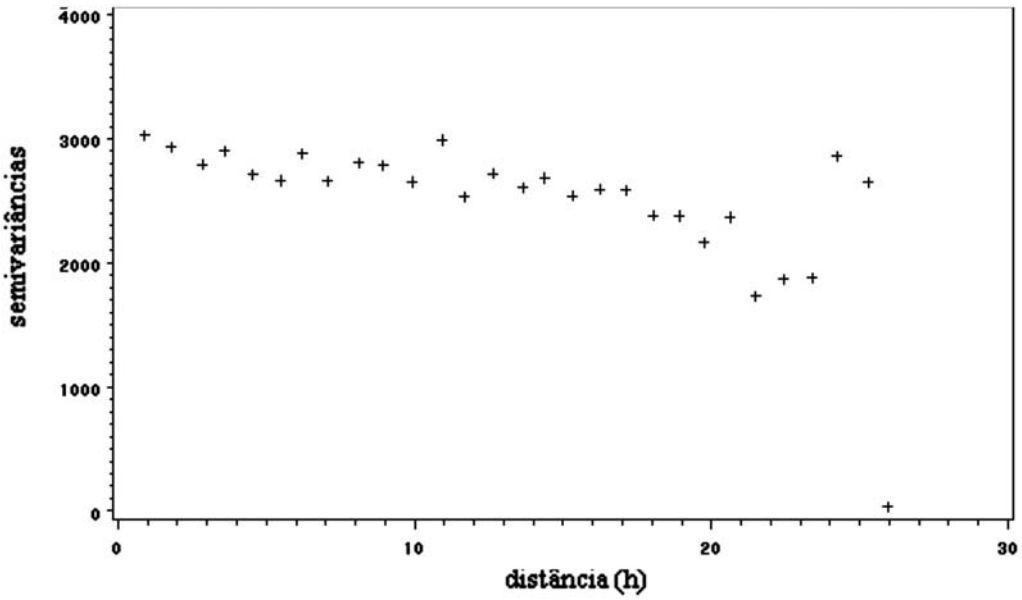

Figura 5 - Semivariograma empírico para resíduos estimados na análise em látice relativo ao experimento 1. (Semivâriancias em função da distância entre parcelas)

Embora o teste de DW e o semivariograma empírico indicassem ausência de dependência espacial, tentou-se ajustar um modelo geoestatístico para as semivariâncias apenas para verificar a coerência entre os resultados dos métodos para avaliação de dependência espacial.

Nas Tabelas 4 e 5, encontram-se as estimativas para os parâmetros contribuição (C), alcance (A) e efeito pepita (C_0), bem como os critérios de ajustamento adotados para escolha do modelo geoestatístico, além do teste da razão de verossimilhança (LRT) que compara os modelos geoestatísticos com o modelo de erros independentes (Experimento 1). Todos os modelos apresentaram valores próximos para os critérios de ajuste adotados. Para a análise em blocos com efeito pepita (Tabela 4), o ajuste do modelo exponencial não atingiu a convergência. Pelo teste da razão de verossimilhança (LRT), os modelos esférico e gaussiano, com efeito pepita, não diferiram do modelo com erros independentes (P>0,01), embora os p-valores tenham sidos baixos, 0,0820 e 0,0608, respectivamente. Para a análise sem efeito pepita, os

modelos exponencial e gaussiano apresentaram alcance estimado nulo, e o modelo esférico não diferiu do modelo com erros independentes (p-valor = 0,4795) (Tabela 4). Logo, considerando os dados do experimento 1, não houve dependência espacial entre os resíduos estimados na análise em blocos.

Tabela 4 - Estimativas dos parâmetros contribuição (C), alcance (a) e efeito pepita (C_0), valor da verossimilhança restrita (-2RLL), Critério de Informação de Akaike (AIC), estatística do teste da razão da verossimilhança (LRT) para modelos geoestatístico com e sem efeito pepita em relação ao modelo com erros independentes na análise em bloco do experimento 1

	C	a	C_0	-2 RLL	AIC	LRT	p-valor
Modelos com efeito pepita							
Exp[(1)]	Não	Convergiu	-	-		-	-
Esf	438,16	9,7	3.280,08	1.614,1	-810,1	5,0	0, 0820
Gau	358,39	9,6	3.362,49	1.613,5	-809,9	5,6	0, 0608
Modelos sem efeito pepita							
Exp	3.704,61	0	-	1.619,1	-810,5	-	-
Esf	3.705,34	1,1	-	1.618,6	-811,3	0,5	0, 4795
Gau	3.704,61	0	-	1.619,1	-811,5	-	-
Modelo com erros independentes							
	3.704,61	-	-	1.619,1	-810,5	-	-

[1]Modelos geoestatísticos: Exp= Exponencial, Esf= Esférico, Gau = Gaussiano.

Na análise em látice (Tabela 5), os modelos com erros dependentes não diferiram do modelo com erros independentes, e os critérios de ajuste foram todos idênticos, tendo sido desnecessário aplicar o LRT, ficando caracterizada a independência dos resíduos. Desta maneira, os resultados relativos ao ajuste de modelos geoestatísticos para o experimento 1 foram coerentes com aqueles obtidos no teste de DW e na interpretação do semivariograma empírico.

Tabela 5 - Estimativas dos parâmetros contribuição (C), alcance (a) e efeito pepita (C_0), valor da verossimilhança restrita (-2RLL), Critério de Informação de Akaike (AIC), estatística do teste da razão da verossimilhança (LRT) para modelos geoestatístico com e sem efeito pepita em relação ao modelo com erros independentes na análise em látice do experimento 1

	C	a	C_0	-2 RLL	AIC
Modelos com efeito pepita					
[(1)] Exp	0	4	2.708,59	1.573,3	-788,7
Esf	0	4	2.708,59	1.573,3	-788,7
Gau	0	4	2.708,59	1.573,3	-788,7
Modelos sem efeito pepita					
Exp	2.708,59	0	-	1.573,3	-788,7
Esf	2.708,59	0	-	1.573,3	-788,7
Gau	2.708,59	0	-	1.573,3	-788,7
Modelo com erros independentes					
	2.708,59	-	-	1.573,3	-788,7

[1]Modelos geoestatísticos: Exp= Exponencial, Esf= Esférico, Gau = Gaussiano.

O experimento 2 é um látice 5x5, e seus resultados relativos à estrutura de dependência espacial foram semelhantes aos obtidos para o experimento 1. Na análise em bloco, o teste LRT para os modelos exponencial, esférico e gaussiano, com efeito pepita, apresentou p-valores iguais a 0,1572, 0,1353 e 0,1422 e as análises sem efeito pepita, p-valores iguais a 0,0483, 0,0614 e 0,0613, respectivamente, todos não significativos em relação ao modelo com erros independentes. Na análise em látice, os modelos com erros dependentes não diferiram do modelo com erros independentes, o critério de ajuste AIC assumiu valores todos próximos e os valores da verossimilhança (-2RLL) foram iguais, sendo desnecessário aplicar o teste LRT, ficando caracterizada a independência dos resíduos. Desta maneira, os resultados relativos ao ajuste de modelos geoestatísticos para o experimento 2 também foram coerentes com os resultados obtidos no teste de DW e na interpretação do semivariograma empírico.

No experimento 3, látice 20x20, os resíduos estimados na análise em látice se distribuem de forma aleatória no campo experimental, de acordo com o teste de Durbin-Watson para autocorrelação espacial. As autocorrelações espaciais não foram significativas com p-valores iguais a 0,2164 , 0,1909 e 0,4502 para 1ª, 2ª e 3ª ordens, respectivamente (Tabela 6).

Tabela 6 - Estatística do teste de Durbin-Watson (d) para autocorrelação espacial entre resíduos estimados nas análises em látice e em blocos do experimento 3.

Tipo de análise	Ordem	D	$\hat{\rho}^{(1)}$	p-valor
Látice	1	1,9446	0,0277	0,2164
	2	2,0593	-0,0296	0,1909
	3	2,0038	-0,0019	0,4502
Blocos	1	1,8451	0,0774	0,0141
	2	1,9217	0,0391	0,1418
	3	1,8857	0,0500	0,0609

[1] $\hat{\rho}$ = autocorrelação espacial estimada

Na análise em blocos, as autocorrelações espaciais também não foram significativas a 1% de probabilidade com p-valores iguais a 0,0141, 0,1418 e 0, 0609, indicando a independência entre resíduos, porém os p-valores da análise em blocos foram menores que os encontrados na análise em látice e para p- valores iguais ou superiores a 0,0141, fica caracterizada a independência espacial de 1ª ordem, embora com baixa autocorrelação, igual a 0,0774 (Tabela 6).

O semivariograma empírico para resíduos estimados na análise em blocos (Figura 6) indica a presença de dependência espacial, com alcance estimado entre 2 e 10 m , efeito pepita entre 1.100 a 1.300, contribuição entre 100 e 300 e patamar entre 1250 e 1400, enquanto na análise em látice (Figura 7), as semivariâncias estimadas se encontram praticamente em torno de uma reta paralela até uma distancia de 30m, indicando que o modelo mais adequado é o modelo com efeito pepita puro que caracteriza a independência espacial entre os resíduos.

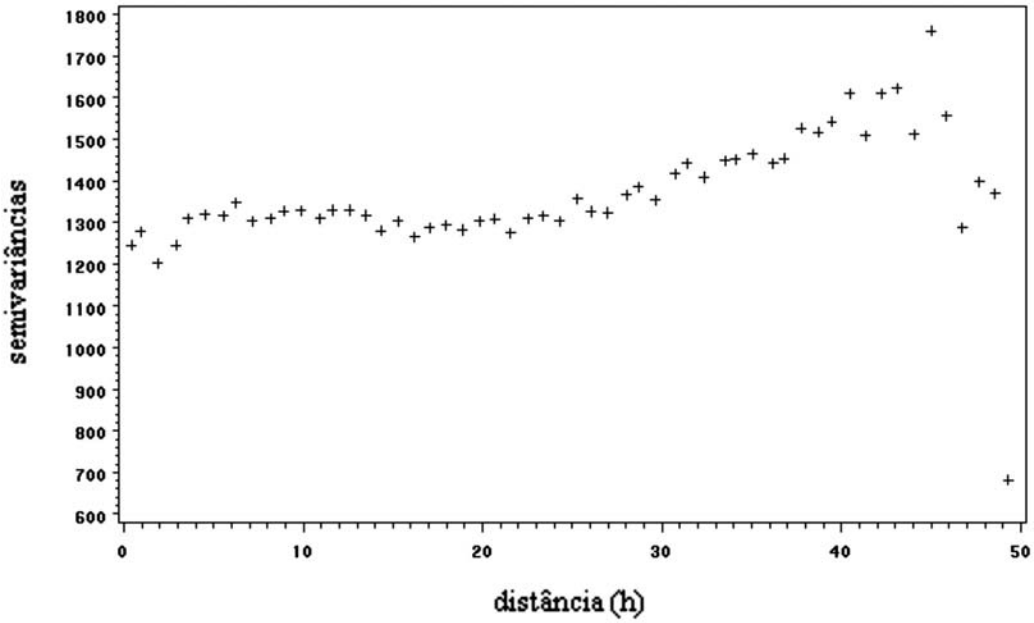

Figura 6 - Semivariograma empírico para resíduos estimados na análise em blocos relativo ao experimento 3. (Semivâriancias em função da distância entre parcelas)

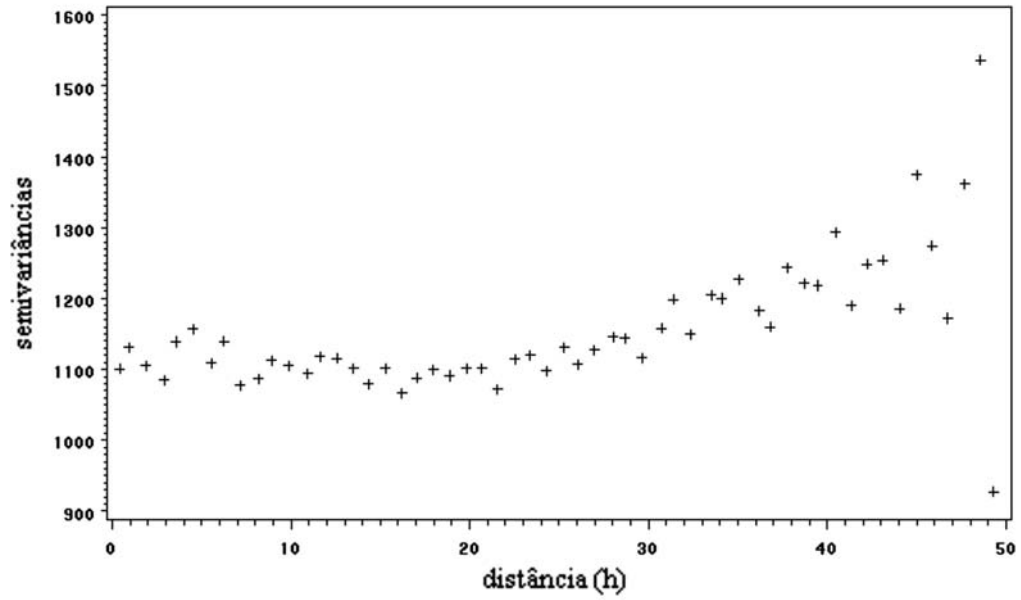

Figura 7 - Semivariograma empírico para resíduos estimados na análise em látice relativo ao experimento 3. (Semivâriancias em função da distância entre parcelas)

Pelo teste da razão de verossimilhança (Tabela 7), percebe-se que para análise em blocos com efeito pepita, os modelos exponencial, esférico e gaussiano, apresentaram melhor ajuste que o modelo com erros independentes($p \leq 0,0001$). Logo, o alcance difere significativamente de zero, indicando que os resíduos apresentaram dependência espacial, enquanto no ajuste sem efeito pepita todos os modelos apresentaram alcance igual a zero.

Tabela 7 - Estimativas dos parâmetros contribuição (C), alcance (a) e efeito pepita (C_0), valor da verossimilhança restrita (-2RLL), Critério de Informação de Akaike (AIC), estatística do teste da razão da verossimilhança (LRT) para modelos geoestatístico com e sem efeito pepita em relação ao modelo com erros independentes na análise em bloco do experimento 3

	C	a	C_0	-2 RLL	AIC	LRT[(2)]	p-valor
Modelos com efeito pepita							
[(1)]Exp	147,67	9,1	1.198,23	8.007,2	-4.006,6	18,4	0,0001
						(18,4)	(0,0001)
Esf	118,25	10,3	1233,26	8.007,6	-4.006,8	18	0,0001
						(18)	(0,0001)
Gau	125,98	5,9	1.216,61	8.003,7	-4.004,9	21,9	<0,0001
						(21,9)	(<0,0001)
Modelos sem efeito pepita							
Exp	1.337,21	0	-	8.025,6	-4.014,8	-	-
Esf	1.337,21	0	-	8.025,6	-4.014,8	-	-
Gau	1.337,21	0	-	8.025,6	-4.014,8	-	-
Modelo com erros independentes							
	1.337,21	-	-	8.025,6	-4.014,8	-	-

[1]Modelos geoestatísticos: Exp= Exponencial, Esf= Esférico, Gau = Gaussiano.[2]Entre parênteses:LRT para significância do efeito pepita (H_o: Co=0), comparação entre os modelos geoestatísticos com e sem efeito pepita

Pelo LRT, para todos os modelos, rejeita-se a hipótese de efeito pepita nulo com ($p<0,0001$)(Tabela 7). Assim, os modelos com efeito pepita foram mais adequados. De acordo com o critério AIC e com os resultados anteriores, o modelo mais adequado, considerando a análise em blocos, é o gaussiano com efeito pepita com parâmetros estimados C= 125,98, a=5,9m e C_0 =1.216,61. Assim, todos os erros de parcelas selecionada a uma distância

inferior a 5,9m estão correlacionados entre si (Tabela 7). Os parâmetros do Gaussiano foram usados para cálculo do GD, indicando dependência espacial fraca, com GD=0,09.

Entretanto, na análise em látice do experimento 3, para todos os modelos geoestatísticos com efeito pepita (p> 0,81) não se rejeita a hipótese de alcance nulo (Tabela 8) e para todos os modelos sem efeito pepita o alcance é nulo. Tais resultados indicam que os modelos com erros dependentes não diferiram do modelo com erros independentes, ficando caracterizada a independência dos resíduos quando estes foram estimados na análise em látice.

Tabela 8 - Estimativas dos parâmetros contribuição (C), alcance (a) e efeito pepita (C_0), valor da verossimilhança restrita (-2RLL),Critérios de Informação de Akaike (AIC), estatística do teste da razão da verossimilhança (LRT) para modelos geoestatístico com e sem efeito pepita em relação ao modelo com erros independentes na análise em látice do experimento 3

	C	a	C_0	-2 RLL	AIC	LRT	p-valor
Modelos com efeito pepita							
[1] Exp	762,04	0,4	358,90	7.884,4	-3.945,2	0,3	0,8607
Esf	1.106,08	0,5	14,96	7.884,3	-3.945,2	0,4	0,8187
Gau	210,10	0,5	910,94	7.884,3	-3.945,2	0,4	0,8187
Modelos sem efeito pepita							
Exp	1.120,94	0	-	7.884,7	-3.944,3	-	-
Esf	1.120,94	0	-	7.884,7	-3.944,3	-	-
Gau	1.120,94	0	-	7.884,7	-3.944,3	-	-
Modelo com erros independentes							
	1.120,94	0	-	7.884,7	-3.944,3	-	-

[1]Modelos geoestatísticos: Exp= Exponencial, Esf= Esférico, Gau = Gaussiano.

No experimento 4 (látice 10x10), os valores da estatística do teste de DW(d) para autocorrelação espacial entre resíduos na análise em látice estão próximos a 2 e seus p-valores iguais a 0, 4385, 0,3503 e 0,0757, indicando que os resíduos são independentes (Tabela 9).

Tabela 9 - Estatística do teste de Durbin-Watson (d) para autocorrelação espacial entre resíduos estimados nas análises em látice e em blocos casualizados para o experimento 4

Tipo de análise	Ordem	D	$\hat{\rho}^{(1)}$	p-valor
Látice	1	2,0179	-0,0089	0,4385
	2	2,0377	-0,0188	0,3503
	3	2,1519	0,0757	0,0757
Blocos	1	1,3959	0,3020	<0,0001
	2	1,4305	0,2847	<0,0001
	3	1,5860	0,2070	0,0002

[1] $\hat{\rho}$ = autocorrelação espacial estimada

Porém, na análise em blocos, a autocorrelação foi significativa (p<0,0002), indicando que os resíduos estão correlacionados até a 3ª ordem, com valores de autocorrelação iguais a 0,3020 , 0,2847 e 0,2070 para 1ª , 2ª e 3ª ordens, respectivamente, e com dependência espacial moderada. Desta maneira, fica caracterizada a independência dos resíduos na análise em látice e a dependência espacial moderada na análise em blocos para o experimento 4.

A dependência espacial se apresentou de forma bem definida no semivariograma empírico na análise em blocos dos dados do experimento 4 (Figura 8), com alcance entre 6 e 18 m, efeito pepita entre 4500 e 5500, contribuição entre 2000 e 3000 e patamar entre 7000 e 8000. Entretanto, na análise em látice (Figura 9), o modelo mais adequado para descrever o comportamento dos resíduos foi o modelo com efeito pepita puro indicando que os resíduos são independentes .

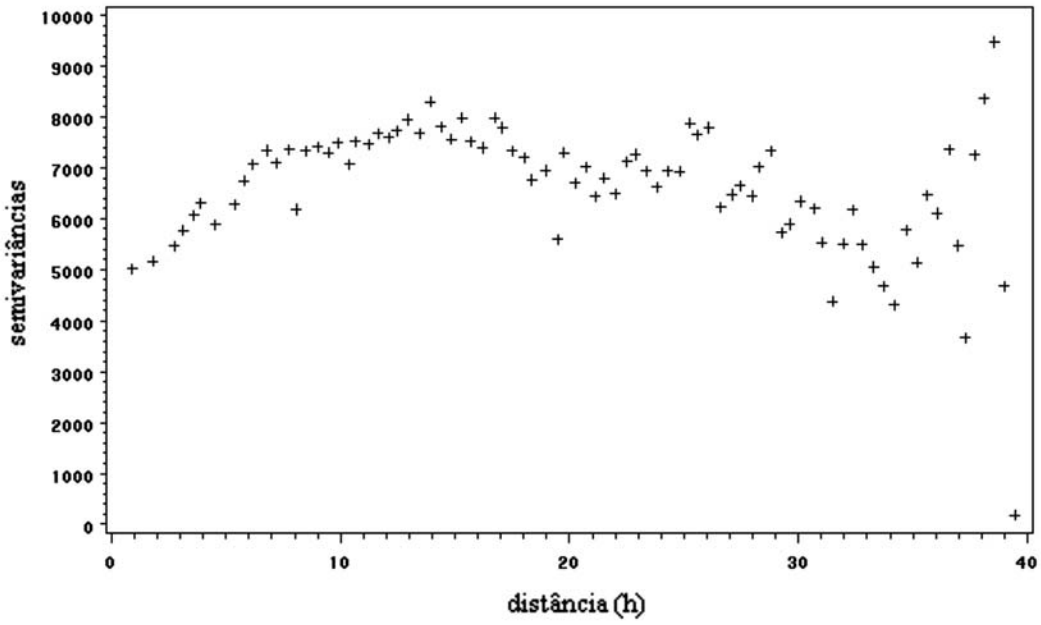

Figura 8 - Semivariograma empírico para resíduos estimados na análise em blocos relativo ao experimento 4. (Semivâriancias em função da distância entre parcelas)

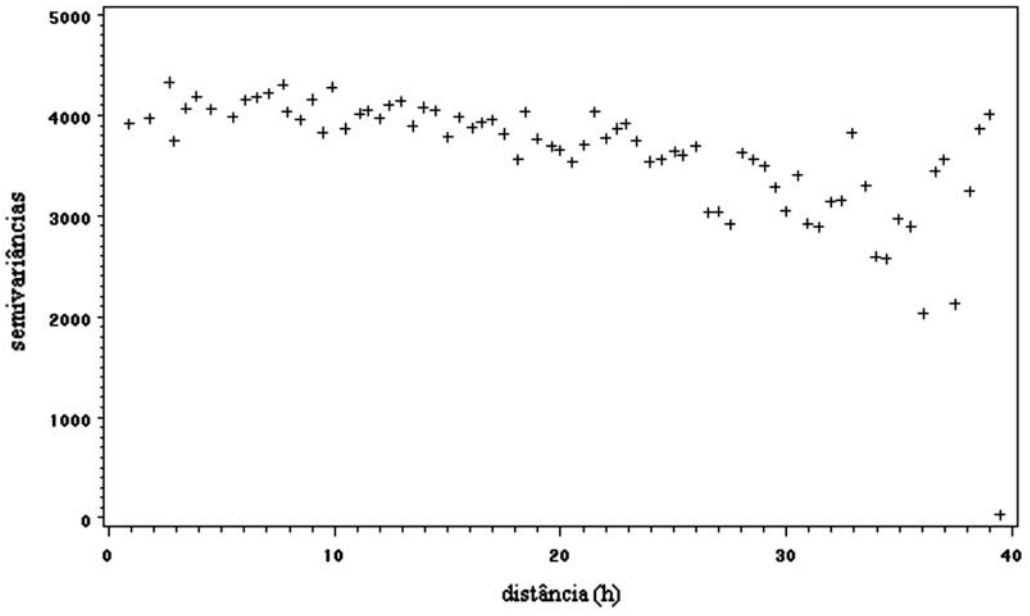

Figura 9 - Semivariograma empírico para resíduos estimados na análise em látice relativo ao experimento 4. (Semivâriancias em função da distância entre parcelas)

O teste da razão de verossimilhança (tabela 10) mostrou que na análise em blocos os modelos com erros dependentes (com ou sem efeito pepita) diferiram do modelo com erros independentes ($p < 0,0001$), com

exceção do modelo gaussiano sem efeito pepita. Pelo mesmo teste (LRT), comparando os modelos com e sem efeito pepita, concluiu-se que os modelos com efeito pepita foram os mais adequados($p< 0,0001$).

Tabela 10 - Estimativas dos parâmetros contribuição (C), alcance (a) e efeito pepita(C_0), valor da verossimilhança restrita (-2RLL), Critérios de Informação de Akaike(AIC), estatística do teste da razão da verossimilhança para modelos geoestatístico com e sem efeito pepita em relação ao modelo com erros independentes na análise em bloco do experimento 4

	C	a	C_0	-2 RLL	AIC	LRT[(2)]	p-valor
Modelos com efeito pepita							
[(1)]Exp	3.419,84	18,1	4.343,85	3.448,8	-1.727,4	58,1	<0,0001
						(28,3)	<0,0001
Esf	3.782,70	15,8	4.532,08	3.445,8	-1.725,9	61,1	<0,0001
						(43,0)	<0,0001
Gau	2.657,42	9,6	4.899,60	3.455,2	-1.725,6	61,7	<0,0001
						(61,7)	<0,0001
Modelos sem efeito pepita							
Exp	7.267,03	2,5	-	3.477,1	-1.740,5	29,8	<0,0001
Esf	7.014,47	1,5	-	3.488,8	-1.746,4	18,1	<0,0001
Gau	7.127,58	1,7	-	3.506,9	-1.755,4	-	-
Modelo com erros independentes							
	7.127,58	-	-	3.506,9	-1.755,4	-	-

[1]Modelos geoestatísticos: Exp = Exponencial, Esf = Esférico, Gau = Gaussiano.[2]Entre parênteses:LRT para significância do efeito pepita (Ho: Co=0), comparação entre os modelos geoestatísticos com e sem efeito pepita.

As estimativas de alcance para análise em blocos casualizados variaram entre 1,5 e 18,1 metros (Tabela 10). Segundo o critério AIC, o modelo gaussiano foi o mais adequado para a análise em blocos, e a função contínua que melhor se ajustou aos dados tem parâmetros C =2.657,42, a=9,6m e C_0= 4.899,60 (Tabela 10), com GD=0,35.

Na análise em látice (Tabela 11), todos os modelos, com e sem efeito pepita, não diferiram do modelo com erros independentes, indicando que a análise em látice garantiu a independência entre resíduos.

Tabela 11 - Estimativas dos parâmetros contribuição (C), alcance (a) e efeito pepita (C_0), Critério de Informação de Akaike (AIC), valor da verossimilhança restrita (-2RLL), estatística do teste da razão da verossimilhança para modelos geoestatístico com e sem efeito pepita em relação ao modelo com erros independentes na análise em látice do experimento 4

	C	a	C_0	-2 RLL	AIC
Modelos com efeito pepita					
Exp[(1)]	0	0	3.878,58	3.324,9	-1.664,5
Esf	0	0	3.878,58	3.324,9	-1.664,5
Gau	0	0	3.878,58	3.324,9	-1.664,5
Modelos sem efeito pepita					
Exp	3.878,58	-	-	3.324,9	-1.664,5
Esf	3.878,58	-	-	3.324,9	-1.664,5
Gau	3.878,58	-	-	3.324,9	-1.664,5
Modelo com erros independentes					
	3.878,58	-	-	3.324,9	-1.664,5

[1]Modelos geoestatísticos: Exp= Exponencial, Esf= Esférico, Gau = Gaussiano

Para o experimento 5 (látice7x7), os resultados para análise em blocos foram semelhantes aos do experimento 4, apresentando dependência espacial moderada (GD=0,34), porém o modelo exponencial não conseguiu modelar as semivariâncias com e sem efeito pepita. E de acordo com os critério AIC, o modelo gaussiano com efeito pepita foi o mais adequado para a análise em blocos, e a função contínua que melhor se ajustou aos dados tem parâmetros C = 1.415,70, A= 7,2m e C_0 = 2.629,69. Para a análise em látice, todos os modelos, com e sem efeito pepita, não diferiram do modelo com erros independentes, indicando que a análise em látice garantiu a independência entre resíduos.

Para o experimento 6 (látice 6x6), na análise em blocos ocorreu autocorrelação de 1ª ordem ($p<0,001$), com valor de 0,367; e de 2ª ordem, com valor 0,201 para p-valores superiores a 0,021 com dependência espacial moderada. Pelo teste da razão de verossimilhança, o alcance foi significativo ($p \leq 0,0004$) para todos os modelos com e sem efeito pepita, e o efeito pepita foi significativo para os modelos esféricos e gaussiano com p-valores 0,003 e 0,004, respectivamente, e não significativo para o modelo exponencial com p-

valor 0,2059. Os modelos mais adequados para análise em blocos foram: exponencial sem efeito pepita, esférico e gaussiano com efeito pepita. Todos os modelos apresentam valores muito próximos, mas, de acordo com o critérios AIC, o modelo esférico foi o mais adequado para a análise em blocos com parâmetros estimados C = 5.191, 76, a=6,8m e C_0 = 4.738,01 com GD=0,52. Para a análise em látice, todos os modelos com e sem efeito pepita foram não significativos ($p \cong 1$), rejeitando a hipótese de nulidade do alcance, indicando a independência entre os resíduos.

Para o experimento 7 (látice 15x15), a estatística de DW(d) na análise em látice variou de 1,9701 a 2,0407, com p-valores iguais 0,2985, 0,1385 e 0,3776 (Tabela 12) maiores que 0,01, indicando a independência entre resíduos. Porém, na análise em blocos, os valores para as estatísticas de DW(d) variaram entre 1,6462 a 1,6762, indicando autocorrelação espacial significativa (p<0,01) da 1ª à 3ª ordem, com valores iguais a 0,1619, 0,1928 e 0,1769, respectivamente, expressando a dependência espacial no bloco cujo experimento foi realizado com 225 tratamentos em uma área experimental de 1.848,6 m^2 (23,4 x 79), sendo o maior experimento analisado. A dependência espacial estimada no experimento 7 foi de fraca a moderada, com GD= 0,253.

Tabela 12 - Estatística do teste de Durbin-Watson (d) para autocorrelação espacial entre resíduos estimados nas análises em látice e em blocos casualizados para o experimento 7

Tipo de análise	Ordem	d	$\hat{\rho}^{(1)}$	p-valor
Látice	1	2,0407	-0,0203	0,2985
	2	1,9134	0,0043	0,1385
	3	1,9701	0,0149	0,3776
Blocos	1	1,6762	0,1619	<0,0001
	2	1,6144	0,1928	<0,0001
	3	1,6462	0,1769	<0,0001

[1] $\hat{\rho}$ = autocorrelação espacial estimada

A dependência espacial se apresenta de forma bem caracterizada no semivariograma da análise em bloco (Figura10) com alcance entre 10 e 30 m.

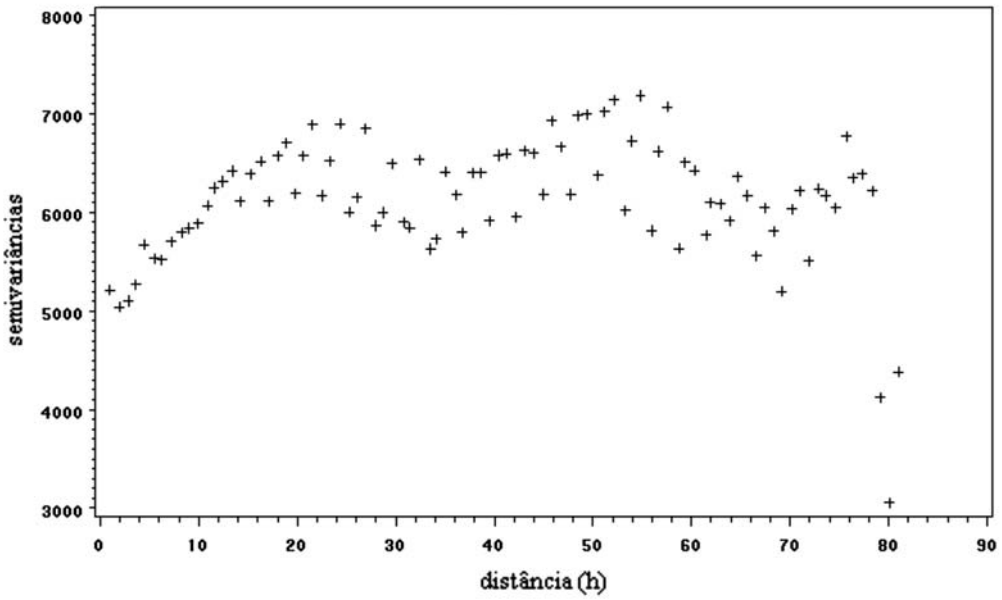

Figura 10 - Semivariograma enpírico para resíduos estimados na análise em blocos relativo ao experimento 7. (Semivâriancias em função da distância entre parcelas)

Porém, para os resíduos na análise em látice (Figura 11), o modelo mais adequado para descrever o comportamento das semivariâncias foi aquele com efeito pepita puro, indicando que a distribuição espacial dos resíduos na área experimental é aleatória e os resíduos são independentes.

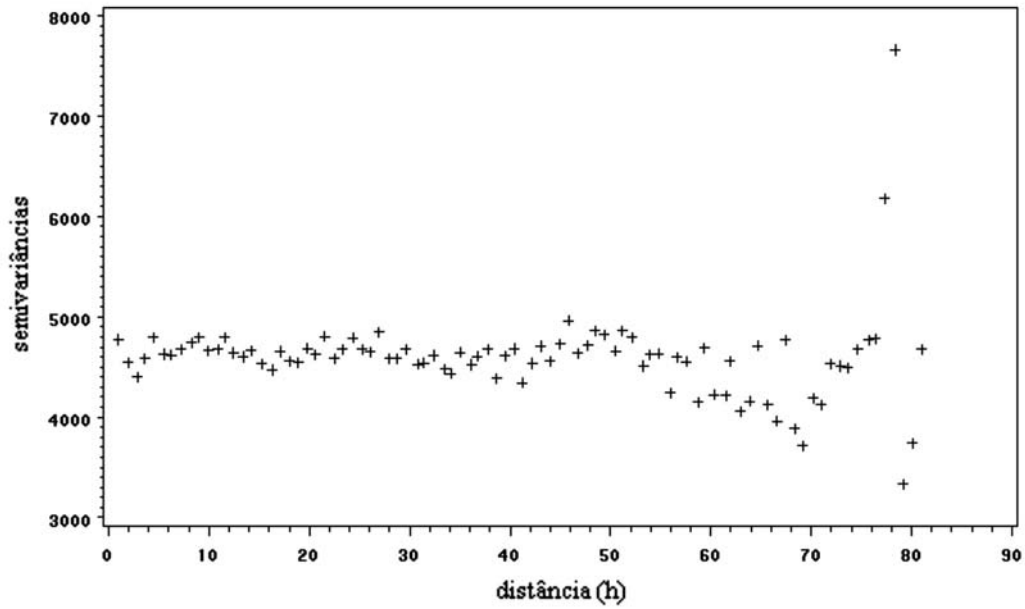

Figura 11 - Semivariograma empírico para resíduos estimados na análise em látice relativo ao experimento 7. (Semivâriancias em função da distância entre parcelas)

O teste da razão de verossimilhança (Tabela 13) mostrou que, para todos os modelos da análise em blocos com e sem efeito pepita ($p \leq 0,0004$), os modelos com erros dependentes diferiram do modelo com erros independentes, com exceção do modelo gaussiano sem efeito pepita que não diferiu do modelo com erros independentes. Pelo mesmo teste (LRT), comparando os modelos com e sem efeito pepita, concluiu-se que os modelos com efeito pepita foram os mais adequados ($p< 0,0001$).

As estimativas de alcance para análise em blocos casualizados variaram entre 0 e 43 metros (Tabela 13). Pelo mesmo teste da razão de verossimilhança, todos os modelos foram significativos ($p<0,0001$) para a hipótese de efeito pepita nulo (Tabela 13). Os modelos que melhor se adequaram às semivariâncias são o exponencial, o esférico e o gaussiano com efeito pepita. Segundo os critérios AIC, o modelo exponencial foi o mais adequado para a análise em blocos, e a função contínua que melhor se ajustou aos dados tem parâmetros estimados C = 1.646,86, a=33,3m e C_0 =4.858,93 (Tabela 13).

Tabela 13 - Estimativas dos parâmetros contribuição (C), alcance (a) e efeito pepita (C_0), valor da verossimilhança restrita (-2RLL), Critérios de Informação de Akaike (AIC), estatística do teste da razão da verossimilhança para modelos geoestatístico com e sem efeito pepita em relação ao modelo com erros independentes na análise em bloco do experimento 7

	C	a	C_0	-2 RLL	AIC	LRT[(2)]	p-valor
Modelos com efeito pepita							
Exp[(1)]	1.646,86	33,3	4.858,93	7.725,8	-3.865,9	75,9 (56,5)	<0,0001 (<0,0001)
Esf	2131.00	35	4989.55	7.727,2	-3.866,6	74,5 (62)	<0,0001 (<0,0001)
Gau	2.156,11	43	5.441,78	7.739,9	-3.873,0	61,8 (61,8)	<0,0001 (<0,0001)
Modelos sem efeito pepita							
Exp	6.213,54	1,8	-	7.782,3	-3.893,1	19,4	<0,0001
Esf	6.146,14	1,4	-	7.789,2	-3.896,6	12,5	0,0004
Gau	6.171,81	0	-	7.801,7	-3.902,9	-	-
Modelo com erros independentes							
	6.171,81			7.801,7	-3.901,9		

[1]Modelos geoestatísticos: Exp= Exponencial, Esf= Esférico, Gau = Gaussiano.[2]Entre parênteses:LRT para significância do efeito pepita (Ho: Co=0), comparação entre os modelos geoestatísticos com e sem efeito pepita.

Em todos os modelos, na análise em látice com efeito pepita, os p-valores para o LRT foram não significativos(p>0,5), indicando que os modelos com erros dependentes não diferiram do modelo com erros independente (Tabela 14). E para todos os modelos sem efeito pepita, os critérios de ajuste foram todos idênticos, tendo sido desnecessário aplicar o LRT, pois os modelos com erros dependentes não diferiram do modelo com erros independentes, ficando assim caracterizada a independência entre resíduos na análise em látice.

Tabela 14 - Estimativas dos parâmetros contribuição (C), alcance (a) e efeito pepita (C_0) valor da verossimilhança restrita(-2RLL), Critérios de Informação de Akaike (AIC), estatística do teste da razão da verossimilhança para modelos geoestatístico com e sem efeito pepita em relação ao modelo com erros independentes na análise em látice do experimento 7

	C	a	C_0	-2 RLL	AIC	LRT	p-valor
Modelos com efeito pepita							
[(1)] Exp	141,43	6,1	4.481,79	7.606,3	-3.806,2	0,5	0,7788
Esf	173,36	5,9	4.449,64	7.605,7	-3.805,8	1,1	0,5769
Gau	134,82	5,9	4.489,01	7.605,6	-3.805,8	1,2	0,5488
Modelos sem efeito pepita							
Exp	4.621,47	0	-	7.606,8	-3.805,4	-	-
Esf	4.621,47	0,1	-	7.606,8	-3.805,4	-	-
Gau	4.621,47	0	-	7.606,8	-3.805,4	-	-
Modelo com erros independentes							
	4.621.47	-	-	7.606,8	-3.804.4	-	-

[1]Modelos geoestatísticos: Exp= Exponencial, Esf= Esférico, Gau = Gaussiano

No experimento 8 (látice 10x10), o teste de DW(d) para a análise em blocos indicou que houve autocorrelação espacial de 1ª ordem significativa (p<0,001), com valor estimado de 0,324, gerando uma moderada dependência espacial com GD = 0,43. O teste da razão de verossimilhança mostrou que na análise em blocos, os modelos com erros dependentes (com ou sem efeito pepita) diferiram do modelo com erros independente (p< 0,0001). Pelo mesmo teste (LRT), comparando os modelos com e sem efeito pepita, concluiu-se que os modelos gaussiano com efeito pepita e exponencial e esférico sem efeito

pepita foram os mais adequados. Segundo o critério AIC, o modelo gaussiano foi o mais adequado, e a função contínua que melhor se ajustou tem parâmetros C=2.494,34, a=3m e C_0= 3.213.23.

Na análise em látice, pelo teste de Durbin-watson para autocorrelação espacial entre resíduos, verificou-se autocorrelação espacial de 1ª ordem com valor igual a 0, 170. Pelo teste da razão de verossimilhança, verificou-se que todos os modelos diferiram do modelo com erros independentes (p<0,008). Para o efeito pepita, não foi necessário fazer o teste da verossimilhança, pois os modelos não diferiram entre si. Assim, o modelo que melhor se adequou aos dados foi o sem efeito pepita. De acordo com os critérios AIC, tanto o modelo exponencial com C = 4.347,65 e a=1,7m, esférico com C = 4.330,34 e a=1,4 m ou gaussiano com C = 4.330,66 e a=1,2m pode ser escolhido, pois todos se adequaram às semivariâncias. Desta maneira, ficou caracterizada a dependência espacial no experimento 8, tanto para a análise em bloco quanto em látice.

4. CONCLUSÃO

Em experimentos instalados em látice para avaliação genética da produtividade em feijoeiro, foi identificada dependência espacial entre resíduos de fraca a forte, quando estes experimentos foram analisados em blocos. No entanto, a análise em látice foi eficiente para garantir a independência entre resíduos.

O tamanho dos experimentos não foi fator determinante da dependência espacial, uma vez que alguns experimentos com grande número de famílias apresentaram dependência fraca, enquanto experimentos com menor número de famílias apresentaram dependência moderada.

Os modelos geoestatísticos esférico, exponencial e gaussiano com efeito pepita foram adequados para representar a estrutura de dependência espacial entre resíduos estimados na análise em blocos. Quanto ao ajuste, os três modelos foram bem similares, destacando-se que o alcance estimado no modelo gaussiano foi inferior aos demais.

5. REFERÊNCIAS BIBLIOGRÁFICAS

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

DUARTE, J. B. **Sobre o emprego e a análise estatística do delineamento em blocos aumentados no melhoramento genético vegetal.** 2000. 293p. Tese (Doutorado em Agronomia/Genética e Melhoramento de Plantas) - Escola Superior de Agricultura Luiz de Queiroz, Piracicaba, SP.

GUIMARÃES, E. C. **Geoestatística básica e aplicada.** Disponível em: < http://www.famat.ufu.br/prof/ednaldo/ednaldo.htm > Acesso em: jan 2009.

LITTELL, R. C.; MILLIKEN, G. A.; STROUP, W. W.; WOLFINGER; RUSSELL, D. **SAS System for mixed models.** Cary: NC: SAS Institute INC, 2006. 814 p.

MARTÌNEZ, R. 1994. **Control de La correlacion espacial em experimentos de campo em El sector agrícola.** Agronomia colombiana, 11(1): 83-89

PONTES, J. M.; OLIVEIRA, M. S. **Uma proposta alternativa para a análise de experimentos de campo, utilizando a geoestatística.** Disponível em: <http://www.editora.ufla.br/revista/28_1/art18.PDF>. Acesso em: Nov 2008.

ODA, M. L. **Aplicação de métodos geoestatísticos para identificação de dependência espacial na análise de dados de um experimento em delineamento sistemático tipo "LEQUE".** Disponível em: <http://www.teses.usp.br/teses/disponiveis/11/11134/tde-22082005-135811/>. Acessado em: Jan 2009.

REIS, A. J. dos S.; MIRANDA FILHO,J. B. de. **Autocorrelação espacial na avaliação de composto de milho para resistência à largata do cartucho (*Spodoptera frugiperda*).** Pesquisa Agropecuária Tropical. v.33,nº2,p.65-72,2003.

SILVA, H. D.; GUIMARÃES, E. C.; PEDROSA, M. G. **Incorporação da dependência espacial na análise de um experimento de avaliação de progênies de milho quanto à resistência à ferrugem comum.** Disponível em: <http://www.editora.ufla.br/revista/28_5/art24.PDF>. Acessado em jan/2009.

CAPÍTULO 2

EFICIÊNCIA DE MÉTODOS DE ANÁLISE ESTATÍSTICA ESPACIAL PARA AVALIAÇÃO DA PRODUTIVIDADE DE FAMÍLIAS DO FEIJOEIRO

RESUMO

Nos experimentos de campo, a dependência espacial se caracteriza por parcelas localizadas mais próximas umas das outras apresentarem comportamentos semelhantes às parcelas localizadas mais distantes. Dos oito experimentos de avaliação de produtividade de famílias do feijoeiro considerados no capítulo 1, seis apresentaram dependência espacial, com alcances variando de 2,6 a 33,3 metros. No capítulo 2, foram utilizados o teste de Durbin-Watson para verificar a dependência espacial entre resíduos para as diferentes metodologias, além do critério de informação de Akaike (AIC), e o teste da razão da verossimilhança (LRT) para comparar o ajuste dos modelos. Foram utilizados o coeficiente de variação e a eficiência relativa para comparar a eficiência das metodologias testadas em relação à análise em látice. Também foram utilizados o coeficiente de correlação de Spearman e a eficiência de seleção para comparar as classificações de famílias do feijoeiro obtidas na análise em látice e por meio de métodos de análise espacial. Verificou-se que os métodos de análise espacial ajustaram-se melhor aos dados em relação aos métodos que assumem erros independentes. Quanto à precisão experimental, a eficiência da análise em látice foi maior em relação à análise em blocos quando houve maior dependência espacial. Os métodos de análise espacial, médias móveis, Papadakis e o método com erros dependentes apresentaram razoável eficiência, enquanto as metodologias propostas, reaplicações das médias móveis e Papadakis foram altamente eficientes em relação à análise em látice, sendo mais eficiente o método reaplicação de Papadakis cuja vizinhança é composta por uma parcela superior e uma inferior em relação à parcela referência. Os métodos testados apresentaram razoável nível de concordância em relação às famílias selecionadas na análise em látice. As metodologias de reaplicação de médias móveis e reaplicação de Papadakis mostram-se altamente eficientes para

melhorar a precisão experimental, além de conseguir garantir a independência entre resíduos, sendo uma excelente alternativa de análise em relação à análise em látice.

ANALYSES OF THE EFFICIENCE OF GEOSTATISCTIC METHODS TO EVALUATE THE PRODUCTIVITY OF BEANS FAMILIES

ABSTRACT

In the experimental fields, the spatial dependence characterizes by plots located much closer one of another showed behavior similar to the plots located far from one of another. Six out of eight experiments set to analyze the productivity of beans families, considered in chapter one, showed spatial dependence, reaching 2.6 to 33.3m. In chapter 2, it was used the Durbin-Watson's test to observe the spatial dependence among residues for different methodologies, besides the Akaike Information Criteria (AIC) and the Maximum Likelihood Method to compare the fitting model. The variation coefficient and the relative efficiency were used to compare the efficiency of tested methodologies related to the analyses in lattice. The Spearman's Correlation Coefficient and the efficiency of selection were also used to compare the classification of beans plantation families acquired through the analyses in lattice and through the spatial analyses method. It was noticed that the spatial analyses method had a better fitting to the data than the methods that assume independent errors. The experiments' accuracy, the analyses in lattice efficiency was higher compared to the analyses in blocks when there is higher spatial dependence. The spatial analyses method, Moving Means Method, Papadakis Method and the method with independent errors showed reasonable efficiency, while the proposed methodologies, Moving Means Replication and Papadakis were highly efficient than the analyses in lattice. Papadakis replication Method showed to be more efficient once its neighborhood is constituted by a superior and inferior plot compared to the reference plot. The tested methods showed reasonable concordance level compared to the selected families for analyses in lattice. The Moving Means Replication and Papadakis Replication methodologies showed highly efficient to improve the experiments accuracy, besides to guarantee the independence among residues, being an excellent alternative for analyses compared to the analyses in lattice.

1. INTRODUÇÃO

A dependência espacial entre parcelas afeta a pressuposição de independência dos erros experimentais da análise de variância clássica. Martinez (1994) argumentou que a violação das pressuposições assumidas no modelo de análise não espacial compromete a eficiência da análise, mesmo que se cumpram os princípios básicos da experimentação agrícola. Duarte (2000) comenta que a autocorrelação espacial pode comprometer a comparação entre os tratamentos genéticos. O autor ressalta que, na presença de autocorrelação, a qualidade de ajustamento de um modelo de análise espacial pode ser bastante superior à dos modelos tradicionais e que os genótipos selecionados pelas duas abordagens (espacial e não espacial) podem diferir substancialmente em decorrência do ajuste para os efeitos de posição. Considerando que tais efeitos são de natureza puramente ambiental, pode-se concluir que, nestas condições, a seleção de genótipo baseada na análise espacial é seguramente superior.

A aplicação de análise estatística espacial tem sido mais estudada recentemente devido à disponibilidade de técnicas computacionais que auxiliam no grande volume de cálculos necessário. Estudos envolvendo métodos de análise espacial foram feitos por Vivaldi (1990), Souza (2000), Reis & Miranda Filho (2003) e Costa (2005), entre outros.

Algumas formas de modelagens espaciais têm sido sugeridas atualmente por serem mais apropriadas para experimentos que apresentam dependência espacial entre os erros experimentais. Entre elas, a análise de vizinhança, métodos médias móveis e Papadakis e a análise usando modelo com erros dependentes.

Os métodos médias móveis e Papadakis modelam a dependência espacial na forma de análise de covariância. Já o modelo com erros dependentes modela a dependência na forma de matriz de variâncias e covariâncias definida a partir de uma função de distância entre parcelas.

Souza (2000), trabalhando com dados de produtividade de feijão, concluiu que os métodos de análise de vizinhança, Papadakis e médias móveis, foram eficientes no controle da heterogeneidade dentro dos blocos, sendo esta eficiência similar àquela proporcionada pela análise em látice,

Costa (2003), trabalhando com dados de experimentos com feijão e milho, afirma que os métodos de análise de vizinhança, Papadakis e médias móveis, melhoraram a precisão experimental dos delineamentos com controle local. Porém, o mesmo não ocorreu quando se utilizou o método de erros dependentes.

O objetivo deste trabalho foi avaliar a eficiência da análise estatística espacial em relação à análise em látice, na avaliação da produtividade de famílias em experimentos do programa de melhoramento genético do feijoeiro da UFV.

2. MATERIAL E MÉTODO

Neste trabalho, foram utilizados dados de produtividade de grãos de famílias de feijão, safras do Inverno 2006/2007 e seca 2007/2008, em experimentos conduzidos na estação experimental de Coimbra, município de Coimbra, MG (690m de altitude, 20º45' S e 42º51' W) pertencente ao Departamento de Fitotecnia da Universidade Federal de Viçosa (UFV). Estes ensaios fazem parte do programa de melhoramento genético do feijoeiro da Universidade Federal de Viçosa.

Foram obtidos e analisados dados de oito experimentos instalados em látice quadrado simples para avaliação da produtividade de grãos (g/parcela) em famílias de feijão.

A Tabela 15 apresenta a relação dos experimentos com suas respectivas características.

Tabela 15 - Características dos experimentos avaliados: tipo de látice, número de repetições (r), número de famílias (F), número total de parcelas (NP), espaçamento (E), número de linhas por parcela (L), comprimento das linhas (C) em metros

Exp	Tipo de látice	r	F	NP	E	Tamanho da parcela	
						L	C
1	7x7	3	49	147	0,45	2	2
2	5x5	3	25	75	0,45	2	2
3	20x20	2	400	800	0,45	1	2
4	10x10	3	100	300	0,45	2	2
5	7x7	3	49	147	0,45	2	2
6	6x6	3	36	108	0,45	3	2
7	15x15	3	225	675	0,50	2	2
8	10x10	3	100	300	0,45	2	1,5

Para aplicação das metodologias de análise espacial, a posição de cada parcela na grade experimental foi especificada por duas coordenadas (L, C) relativas ao centro das parcelas: L a coordenada no sentido da largura da parcela e C a coordenada no sentido do comprimento da parcela. Assim, a distância (d) entre as parcelas i e j foi obtida por $d = [(L_j - L_i)^2 + (C_j - C_i)^2]^{0,5}$.

Os experimentos foram instalados em látice e analisados em látice e em blocos de acordo com os modelos descritos abaixo. Nestas análises, foi assumido que os erros aleatórios associados a cada observação são independentes, com distribuição normal, média zero e variâncias homogêneas: $e \sim NID(0, \sigma^2)$.

Os dados de produtividade também foram analisados utilizando as seguintes metodologias de análise estatística espacial em conjunto com o controle local para blocos: método das Médias Móveis (M), método de Papadakis (P) e o modelo com erros dependentes (ED). Também foram propostas e avaliadas duas modificações nos métodos M e P: a reaplicação do método médias móveis (RM) e reaplicação do método Papadakis (RP).

Os seguintes modelos estatísticos foram utilizados, assumindo efeito fixo para todos os fatores:

Modelo 1 (látice): $y_{ijk} = \mu + f_i + r_k + b_{j(k)} + e_{ijk,}$,

em que:

y_{ijk} é o valor observado da família i, no bloco j dentro da repetição k;

μ é a constante associada a todas as observações;

fi é o efeito do tratamento i;

r_k é o efeito fixo da repetição k;

$b_{j(k)}$ é o efeito do bloco j dentro da repetição k; e

e_{ijk} é o erro aleatório associado às observações assumindo independência entre resíduos.

Modelo 2 (blocos ao acaso): $y_{ik} = \mu + fi + r_k + e_{ijk,}$

em que y_{ik} é o valor observado da família i na repetição k.

Modelo 3 (M): $y_{ik} = \mu + fi + r_k + \beta_1 I_{Mik} + e_{ijk}$

em que β_1 é o coeficiente de regressão linear entre produtividade de grãos e a covariável índice ambiental; e I_{Mik} é o índice ambiental para os modelos M.

Modelo 4 (RM): $y_{ik} = \mu + fi + r_k + \beta_1 I_{RMik} + e_{ijk,}$

em que I_{Pik} é o índice ambiental para os modelos P.

Modelo 5 (P): $y_{ij} = \mu + fi + r_k + \beta_1 I_{Pik} + e_{ijk,}$

em que I_{Pik} é o índice ambiental para os modelos P.

Modelo 6 (RP): $y_{ik} = \mu + fi + r_k + \beta_1 I_{RPik} + e_{ijk,}$

em que I_{RPik} é o índice ambiental para as reaplicações dos modelos P.

Modelo 7 (ED): $y_{ik} = \mu + fi + r_k + \varepsilon_{ij,}$

em que ε_{ij} é o erro aleatório associado às observações assumindo dependência entre os erros.

O modelo 1 para análise em látice foi o modelo utilizado como modelo referência para comparação com os demais modelos.

As metodologias M, P, RM e RP (modelos 3, 4,5 e 6) correspondem a uma análise de covariância em que o índice ambiental é utilizado como covariável, ajustando a produção de cada parcela pela média das parcelas vizinhas. Na metodologia M, o índice ambiental I_{Mik} é calculado como a média simples dos valores observados das parcelas vizinhas. Na metodologia P, o índice ambiental I_{Pik} é calculado como a média dos erros aleatórios das parcelas vizinhas, assim, o índice ambiental é ajustado para os demais efeitos do modelo.

Para o cálculo do índice ambiental, foram quatro formas de vizinhança (V) para as metodologias M e P, com variações na posição e número de parcelas vizinhas: V_1 = 2 parcelas vizinhas, uma superior e uma inferior (V_1= V_{SI}) (Figura 12); V_2 = 4 parcelas vizinhas, duas acima e duas abaixo ($V_2 = V_{2SI}$) (Figura 13); V_3 = 4 parcelas vizinhas, acima, abaixo à direita e à esquerda (V_3= V_{4L}) (Figura 14); V_4 = 8 parcelas vizinhas, considerando todas as parcelas que fazem “fronteira” com a parcela referência (R) (V_4= V_{8L}) (Figura 15). Para as parcelas das extremidades da grade experimental, foram consideradas as parcelas vizinhas que compõem a vizinhança analisada.

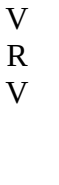

Figura 12. V_1

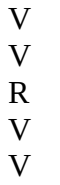

Figura 13. V_2

Figura 14. V_3

Figura 15. V_4

Nas modificações propostas para aplicação dos métodos M e P, calculou-se o índice ambiental duas vezes. Na metodologia Reaplicação do M (RM), o índice ambiental I_{RMik} foi calculado como a média simples dos valores dos índices ambientais (I_{Mik}) das parcelas vizinhas (Figura 16); e para a metodologia Reaplicação do P (RP), o índice ambiental I_{RPik} foi calculado como a média dos índices ambientais(I_{Pik}) das parcelas vizinhas (Figura 17) .

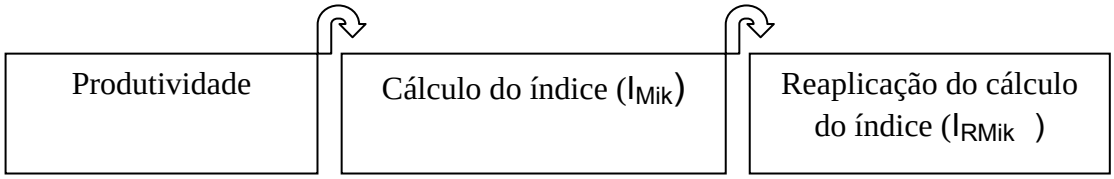

Figura 16 - Representação da metodologia reaplicação do método Médias Móveis.

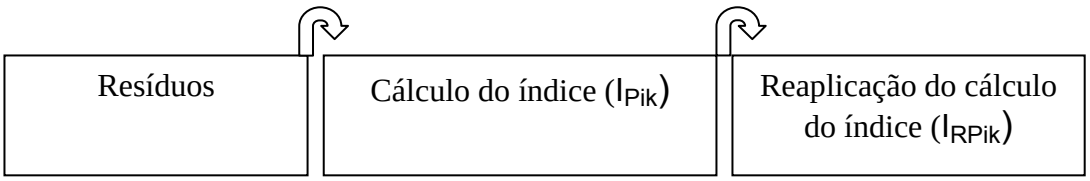

Figura 17 - Representação da metodologia reaplicação do método Papadakis.

Além das quatro formas de vizinhança descritas anteriormente, para o método médias móveis (M) também foi proposto e testado um índice ambiental que considera todas as parcelas da grade experimental. O índice ambiental proposto (I_{TP}) foi calculado como a média ponderada da produtividade de todas as parcelas, em que os pesos são dados pelo inverso da distância de cada parcela até a parcela referência, ou seja,

$$I_{Pi} = \frac{1}{n-1} \sum prod_i \frac{1}{dij},$$

em que n-1 é o número de termos; $prod_i$ é a produtividade de cada parcela; e d_{ij} é a distância de cada parcela até a parcela referência i .

No modelo de análise com erros dependentes - ED (modelo 7), considerou-se a dependência espacial entre os erros aleatórios associados às observações e que eles têm distribuição normal com média zero e matriz de

variâncias e covariâncias residuais não diagonal (R), ε_{ij} ~N(0,R). A seguinte estrutura foi considerada para matriz de variâncias e covariâncias R.

i) R= I C (modelo com erros independentes) e

ii) R= I C_0 + F C (modelo com erros dependentes),

em que I é a matriz identidade, C a variância residual no modelo com erros independentes e contribuição no modelo com erros dependentes, C_0 o efeito pepita e F a matriz formada pelos elementos da função de distância f(h). Esta função corresponde a modelos geoestatísticos usados para representar a variabilidade espacial. Os principais modelos, com efeito pepita, são descritos abaixo:

Modelo exponencial (Exp)

$$S(h) = C_0 + C\,[1 - e^{(-3h/a)}],$$

para $0 < h < d$, em que d é a distância máxima na qual o semivariograma é definido.

Modelo esférico (Esf)

$$S(h) = C_0 + C\,[1{,}5\,(h/a) - 0{,}5\,(h/a)^3], \text{ para } 0 < h < a.$$

Modelo Gaussiano (Gau)

$$S(h) = C_0 + C\,[1 - \exp(3h^2/a^2)],$$

para $0 \leq h \leq d$, em que d é a distância máxima na qual o semivariograma é definido, pois neste caso o patamar é atingido apenas assintoticamente (GUIMARÃES, 2004).

Nos modelos exponencial e gaussiano, o valor do alcance efetivo é obtido multiplicando o alcance estimado no SAS por 3 e $\sqrt{3}$, respectivamente (SAS INSTITUTE, 2006).

De acordo com os resultados do capítulo 1, para os experimentos 3,4,5 e 8, o modelo que melhor se adequou foi o gaussiano, para o experimento 6 foi o esférico e para o experimento 7 foi o exponencial.

A análise, usando modelo com erros dependentes, basicamente, foi compreendida pelas seguintes etapas:

i) Construção do semivariograma empírico que mostra o comportamento das semivariâncias S(h) em função da distância entre parcelas (h);

ii) Ajustamento do modelo geoestatístico que explica a dependência espacial dos dados;

iii) Estimação das médias das famílias considerando a dependência espacial.

Para comparar a eficiência de ajuste dos modelos testados, foi utilizado o critério de Informação de Akaike (AIC), baseado em verossimilhança.

O critério AIC se baseia no valor que maximiza o logaritmo da verossimilhança restrita ajustado para uma função do número de parâmetros. Modelos com menores valores absolutos de AIC são preferidos.

A autocorrelação espacial entre resíduos estimados nas diferentes análises foi avaliada usando-se a estatística de Durbin-Watson(d), que testa a hipótese de autocorrelação linear (H_0: ρ = 0), sendo definida como

$$d = \frac{\sum_{l=2}^{n} (\hat{e}_l - \hat{e}_{l-1})^2}{\sum_{l=1}^{n} \hat{e}_l^2},$$

sendo l=1, 2, 3,..., n a ordem de posicionamento da parcela no experimento, associada ao resíduo $\hat{e}_l$ (essa ordem obedeceu à numeração sucessiva das parcelas, de modo que $\hat{e}_l$ e $\hat{e}_{l-1}$ indicam resíduos de parcelas adjacentes).

A relação entre d e ρ é aproximadamente d=2(1-ρ). Assim, caso não exista autocorrelação, o valor esperado de d é 2. E valores significativamente inferiores a 2 indicam autocorrelação positiva; e valores significativamente superiores a 2 indicam autocorrelação negativa (SAS INSTITUTE, 1993).

A eficiência relativa (ER) foi utilizada para comparar a eficiência das metodologias de análise espacial em relação ao látice e foi calculada pela razão entre o quadrado médio da metodologia testada e o quadrado médio do látice.

O coeficiente de variação experimental foi utilizado para comparar a eficiência das metodologias testadas quanto à precisão experimental.

A correlação de Spearman e a eficiência de seleção (ES), essa última apresentado por Souza et al. (2000), foram utilizadas para verificar a coincidência na ordem de classificação das famílias, utilizando metodologias de análise espacial (M, P, RM, RP e ED) com a análise em látice. As comparações

para ES foram feitas para os níveis de seleção 10, 20 e 30% dos melhores genótipos.

A eficiência de seleção (ES) foi proposta por Hamblin & Zimmermann (1986) e se refere à proporção de genótipos selecionados em cada metodologia de análise, que são coincidentes com os genótipos selecionados na metodologia referência, neste caso, a análise em látice. A ES é dada por ES (%) = 100 (A – C)/(B-C), em que A é o número de genótipos selecionados tanto na metodologia testada como na metodologia referência; B, o número de genótipos selecionados; e C, o número de genótipos esperado, que, por simples coincidência, unicamente pelo acaso, é igual a 10% de B.

Os procedimentos do SAS utlizados foi proc glm para análise em látice e blocos e prox mixed para os critérios AIC e -2RELL para os métodos de análise em blocos, látice e metodologias testadas.

3. RESULTADOS E DISCUSSÃO

De acordo com os resultados obtidos no capítulo 1, os experimentos 1 e 2 não apresentaram autocorrelação espacial significativa (p-valor>0,01) pelo teste de Durbin-Watson. Assim, não foi utilizada nenhuma metodologia de análise espacial para estes dois experimentos.

O experimento 3 apresentou fraca dependência espacial entre resíduos estimados na análise em blocos com alcance igual a 5,9m, de acordo com os resultados apresentados no capítulo 1.

Para as metodologias testadas no experimento 3, de acordo com o teste de Durbin-Watson, houve pequena redução nas estimativas da autocorrelação espacial, de 1ª ordem em relação à análise em blocos ($\hat{\rho}$=0,0775), devido ao fato de a análise em blocos apresentar fraca dependência espacial com p-valor igual a 0,0141. Todos os métodos conseguiram garantir a independência de 1ª e 2ª ordem entre resíduos, com exceção do método RM_{SI}, que apresentou autocorrelação fraca de 2ª ordem (p-valor = 0,0070). Os métodos M_4 (p-valor= 0, 1120 e p-valor= 0, 4015), M_{8L}(p-valor= 0, 1766 e p-valor=0,4978), RM_{4L}(p-valor= 0, 0,1874 e p= 0,1234), RM_{8L}(p-valor= 0, 3899 e p-valor= 0,1233), RP_{4L} (p-valor=0,1022 e p-valor= 0,2751) e RP_{8L} (p-valor= 0,3136 e p-valor = 0,2698) se destacam por apresentar independência de 1ª e 2ª ordem, respectivamente, com altos p-valores (Tabela 16).

De acordo com os critérios de ajuste AIC e -2RLL, todas as metodologias testadas tiveram melhor ajuste que as análises em látice e em blocos (Tabela 16). A utilização de metodologias espaciais promoveu grande redução no valor da função de verossimilhança, de 7.884,7 na análise em látice para valores inferiores a 4.665,4, indicando melhor ajuste destas metodologias, destacando-se as reaplicações de médias móveis (RM) e a reaplicação de Papadakis (RP) (Tabela 16).

Tabela 16 - Autocorrelação espacial estimada ($\hat{\rho}$) de primeira e segunda ordem e probabilidade do erro tipo I (H_0: ρ =0) para o teste de Durbin-Watson, valor da verossimilhança restrita (-2RLL) e Critério de Informação de Akaike (AIC), para diferentes tipos de análises aplicadas ao experimento 3

Tipo de análise	DW 1ª ordem		DW 2ª ordem		-2RLL	AIC
	$\hat{\rho}$	p - valor	$\hat{\rho}$	p - valor		
Látice	0, 0277	0, 2164 ns	-0, 0297	0, 1909 ns	7.884,7	-3.943,3
Blocos	0, 0775	0, 0141 ns	0, 0392	0, 1418 ns	8.025,6	-4.014,8
[1]M_{SI} [6]	0, 0504	0, 0770 ns	0, 0312	0, 1989 ns	4.543,6	-2.272,8
M_{2SI}	0, 0782	0, 0134 ns	0, 0338	0, 1786 ns	4.567,2	-2.284,6
M_{4L}	-0, 0430	0, 1120 ns	0, 0101	0, 4015 ns	4.534,4	-2.268,2
M_{8L}	-0, 0328	0, 1766 ns	0, 0011	0, 4978 ns	4.528,5	-2.265,3
M_{TP}	0, 0775	0, 0141 ns	0, 0391	0, 1421 ns	4.563,0	-2.282,5
[2]RM_{SI}	0, 0254	0, 2362 ns	0, 0880	0, 0070 **	4.269,3	-2.135,6
RM_{2SI}	0, 0496	0, 0804 ns	0, 0470	0, 0976 ns	4.387,9	-2.195,0
RM_{4L}	-0, 0314	0, 1874 ns	-0, 0397	0, 1234 ns	4.357,9	-2.180,0
RM_{8L}	-0, 0099	0, 3899 ns	-0, 0397	0, 1233 ns	4.400,1	-2.201,0
[3]P_{SI}	0, 0609	0, 0421 ns	0, 0357	0, 1652 ns	4.554,6	-2.278,3
P_{2SI}	0, 0546	0, 0612 ns	0, 0377	0, 1511 ns	4.525,7	-2.263,9
P_{4L}	-0, 0536	0, 0646 ns	0, 0292	0, 2151 ns	4.549,4	-2.275,7
P_{8L}	-0, 0461	0, 0958 ns	0, 0176	0, 3222 ns	4.535,2	-2.268,6
[4]RP_{SI}	0, 0544	0, 0619 ns	0, 0327	0, 1867 ns	4.045,1	-2.023,5
RP_{2SI}	0, 0740	0, 0180 ns	0, 0534	0, 0700 ns	4.203,6	-2.102,8
RP_{4L}	0, 0448	0, 1022 ns	-0, 0198	0, 2751 ns	4.478,7	-2.240,3
RP_{8L}	-0, 0172	0, 3136 ns	-0, 0205	0, 2698 ns	4.325,6	-2.163,8
[5]ED	-	-	-	-	4.665,4	-2.335,7

** ns Autocorrelação espacial significativa e não significativa, respectivamente, pelo teste de DW a 1% de probabilidade. [1] M= método médias móveis, [2] RM= reaplicação médias móveis, [3] P= método Papadakis, [4] RP= reaplicação do método Papadakis, [5] ED= método com erros dependentes, [6] SI, 2SI, 4L, 8L e TP = posição das parcelas que compõem as vizinhanças no cálculo do índice ambiental.

A análise em látice foi consideravelmente menos eficiente que as reaplicações de médias móveis e Papadakis com valores de eficiência relativa variando entre 29,15 e 86,59%. O coeficiente de variação nas análises com reaplicação de médias móveis e reaplicação de Papadakis variou entre 7,51 e 12,95, indicando de alta a média precisão experimental (Tabela 17).

Tabela 17 - Coeficientes de variação experimental (CV), Eficiência relativa (ER), Correlação de Spearman (rs) e Eficiência de seleção (ES) da análise em látice em relação às análises em blocos e às análises utilizando métodos de análise espacial no experimento 3

Tipo de análise	CV(%)	ER (%)	[1]rs	Eficiência de seleção (%)		
				10(40)[2]	20(80)	30(120)
Látice	13,92	-	-	-	-	-
Blocos	14,46	107,93	0,96	80,55(33)[3]	84,72(69)	85,18(104)
[4]M_{SI} [9]	14,04	101,71	0,95	75,00(31)	88,89(72)	83,33(102)
M_{2SI}	14,43	107,47	0,96	80,55(33)	84,72(69)	85,18(104)
M_{4L}	13,89	99,56	0,96	77,77(32)	88,89(72)	85,18(104)
M_{8L}	13,80	98,22	0,96	75,00(31)	84,72(69)	83,33(102)
M_{TP}	14,48	108,2	0,96	80,55(33)	84,72(69)	85,18(104)
[5]RM_{SI}	9,95	51,12	0,81	52,78(23)	59,72(51)	69,44(87)
RM_{2SI}	11,56	68,94	0,88	61,11(26)	72,22(60)	76,85(95)
RM_{4L}	11,13	63,98	0,89	63,89(27)	72,22(60)	75,00(93)
RM_{8L}	11,75	71,24	0,90	75,00(31)	77,77(64)	77,77(96)
[6]P_{SI}	14,26	104,87	0,96	77,77(32)	88,89(72)	83,33(102)
P_{2SI}	13,96	100,52	0,94	75,00(31)	83,33(68)	82,40(101)
P_{4L}	14,18	103,7	0,96	80,55(33)	86,11(70)	85,18(104)
P_{8L}	13,93	100,19	0,96	86,11(35)	83,33(68)	84,26(103)
[7]RP_{SI}	7,51	29,15	0,91	69,44(29)	73,61(61)	75,92(94)
RP_{2SI}	9,18	43,5	0,91	66,66(28)	73,61(61)	79,63(98)
RP_{4L}	12,95	86,59	0,95	83,33(34)	83,33(68)	84,26(103)
RP_{8L}	10,71	59,22	0,94	77,77(32)	76,39(63)	79,63(98)
[8]ED	-	54,12	0,98	88,89(36)	90,28(73)	88,89(108)

[1] rs para classificação das médias das famílias usando a análise em látice como referência . [2] ES para 10, 20 e 30% das melhores famílias (número de famílias selecionadas). [3] Entre parênteses – número de famílias comuns em relação a análise em látice. [4] M= método médias móveis, [5] RM= reaplicação médias móveis, [6] P= método Papadakis, [7] RP= reaplicação do método Papadakis, [8] ED= modelo com erros dependentes, [9] SI,2SI,4L,8L e TP = posição das parcelas que compõem as vizinhanças no cálculo do índice ambiental.

As estimativas para os coeficientes de correlação de Spearman para classificação de famílias foram altas entre as análises testadas e análise em látice usada como referência, com valores acima de 0,90, exceto para os métodos RM_{SI}, RM_{2SI} e RM_{4L} que apresentaram valores de correlação de Spearman iguais a 0,81, 0,88 e 0,89 respectivamente (Tabela 17). De acordo com estes resultados, a classificação das famílias de feijoeiro avaliadas com

base na análise em látice é bastante semelhante às classificações obtidas a partir das análises que utilizam técnicas espaciais, bem como da análise em blocos. Porém, a eficiência de seleção do látice para 10, 20 e 30% das melhores famílias do experimento 3 apresentou valores entre 52 e 91%, mostrando que houve grande variação na concordância de seleção das diferentes metodologias testadas em relação à análise em látice (Tabela 17).

Entre todas as metodologias testadas, destacaram-se $M_{4L,}$ M_{8L}, $P_{4L,}$ $P_{8L,}$ RP_{4L} e $RP_{8L.}$ Estas metodologias conseguiram garantir a independência de 1ª e 2ª ordem com p-valores superiores a 0,01, pelo teste de Durbin-Watson. Os métodos $M_{4L,}$ M_{8L}, RP_{4L} e RP_{8L} tiveram eficiência relativa melhor e coeficiente de variação menor que a análise em látice. Porém, o método RP_{SI} conseguiu se sobressair entre todos, garantindo a independência entre resíduos de 1ª e 2ª ordem, o melhor ajuste, menor coeficiente de variação, melhor eficiência relativa entre todas as metodologias testadas e alto valor para correlação de Spearman com eficiência de seleção superior a 75% em todos os níveis de seleção testados.

As metodologias citadas acima se mostram mais eficientes que a análise em látice, tendo a vantagem de poder ser aplicadas diretamente no delineamento em blocos, pois o delineamento em látice quadrado, bastante utilizado em programa de melhoramento de feijão, apresenta algumas desvantagens em sua utilização. Para satisfazer a exigência de ser quadrado, o pesquisador pode ser levado a reduzir o número de famílias a serem testadas, podendo ocasionar o descarte de uma família promissora ou ter que completar com testemunhas que não têm nenhuma ligação com o experimento, ocasionando nesta situação a necessidade de uma maior área experimental, além de aumento nos custos.

De acordo com o capítulo 1, o experimento 4 (látice 10x10) apresentou moderada dependência espacial entre resíduos estimados na análise em blocos com valores de autocorrelação espacial iguais a 0,3021 e 0,2848 para 1ª e 2ª ordens, respectivamente, e alcance igual a 9,6m.

Quanto à autocorrelação espacial, houve razoável redução dos valores estimados para as metodologias espaciais em relação à análise em blocos, porém para a maioria das análises a autocorrelação foi significativa ($p<0,01$) (Tabela 18).

De acordo com os critérios de ajuste AIC e -2RLL, todas as metodologias testadas tiveram melhor ajuste que as análises em látice e em blocos (Tabela 18).

A reaplicação de Papadakis se destacou apresentando os melhores ajustes entre todas as metodologias testadas (Tabela 18), alta eficiência relativa em relação às análises em látice e em blocos com coeficiente de variação variando entre 6,74 a 9,82%, indicando alta precisão (Tabela 19).

Tabela 18 - Autocorrelação espacial estimada ($\hat{\rho}$) de primeira e segunda ordem e probabilidade do erro tipo I (H_0: ρ =0) para o teste de Durbin-Watson, valor da verossimilhança restrita (-2RLL) e Critério de Informação de Akaike (AIC) para diferentes tipos de análises aplicadas ao experimento 4

Tipo de análise	DW 1ª ordem		DW 2ª ordem		-2RLL	AIC
	$\hat{\rho}$	p - valor	$\hat{\rho}$	p - valor		
Látice	-0, 0090	0, 4385^{ns}	-0, 0189	0, 3503^{ns}	3.324,9	- 1.663,5
Blocos	0, 3021	<0, 0001**	0, 2848	<0, 0001**	3.506,9	-1.7544
[1]M_{SI} [6]	0, 1674	0, 0017**	0, 2362	<0, 0001**	2.480,0	-1.241,0
M_{2SI}	0, 2292	<0, 0001**	0, 2455	<0, 0001**	2.492,2	-1.247,1
M_{4L}	-0, 0531	0, 1787^{ns}	0, 1980	0, 0003**	2.460,0	-1.231,0
M_{8L}	-0, 0206	0, 3605^{ns}	0, 1713	0, 0017**	2.454,5	-1.228,3
M_{TP}	0, 3013	<0, 0001**	0, 2827	<0, 0001**	2.511,0	-1.256,5
[2]RM_{SI}	0, 2438	<0, 0001**	0, 2180	<0, 0001**	2.357,7	-1.179,9
RM_{2SI}	0, 2205	<0, 0001**	0, 2464	<0, 0001**	2.414,0	-1.208,0
RM_{4L}	-0, 0352	0, 2712^{ns}	0, 1121	0, 0293^{ns}	2.369,4	-1.185,7
RM_{8L}	0, 0415	0, 2359^{ns}	0, 1164	0, 0247^{ns}	2.405,8	-1.203,9
[3]P_{SI}	0, 1756	0, 0011**	0, 2393	<0, 0001**	2.485,0	-1.243,5
P_{2SI}	0, 2359	<0, 0001**	0, 2588	<0, 0001**	2.489,0	-1.245,5
P_{4L}	-0, 1609	0, 0025**	0, 2045	0, 0002**	2.458,5	-1.230,3
P_{8L}	-0, 1290	0, 0124^{ns}	0, 1635	0, 0026**	2.446,5	-1.224,2
[4]RP_{SI}	0, 2448	<0, 0001**	0, 1770	0, 0012**	2.198,0	-1.100,0
RP_{2SI}	0, 2595	<0, 0001**	0, 2872	<0, 0001**	2.345,5	-1.173,8
RP_{4L}	-0, 0550	0, 1700^{ns}	0, 0433	0, 2444^{ns}	2.291,7	-1.146,8
RP_{8L}	0, 0137	0, 4066^{ns}	0, 1095	0, 0325^{ns}	2.336,2	-1.169,1
[5]ED	-	-	-	-	2.451,1	-1.228,6

** ns Autocorrelação espacial significativa e não significativa, respectivamente, pelo teste de DW a 1% de probabilidade. [1] M= método médias móveis, [2] RM= reaplicação médias móveis [3] P= método Papadakis, [4] RP= reaplicação do método Papadakis, [5] ED= método com erros dependentes, [6] SI, 2SI, 4L, 8L e TP = posição das parcelas que compõem as vizinhanças no cálculo do índice ambiental.

As estimativas para os coeficientes de correlação de Spearman, usando a análise em látice como referência, foram altas para todas as

metodologias testadas, com valores próximos a 0,93, exceto RM_{SI} e $RM_{2SI,}$ com coeficientes iguais a 0,85 e 0,88, respectivamente. Porém a eficiência de seleção mostrou que houve alta concordância quando selecionadas 20 e 30% das famílias do experimento 4 e que, para 10%, os valores estimados para eficiência de seleção variaram entre 30 e 70 %.

As metodologias de análise RM4L, RM8L, RP4L e RP8L garantiram a independência espacial de 1ª e 2ª ordens (Tabela 18), tendo sido mais eficientes que a análise em látice e apresentado coeficientes de variação variando entre 8,56 a 11,43 (Tabela 19), indicando de alta a média precisão experimental. Percebe-se que os métodos com parcelas laterais foram eficientes para tornar os resíduos independentes nos métodos de reaplicação de médias móveis e Papadakis (Tabela 19).

Tabela 19 - Coeficientes de variação experimental (CV), Eficiência relativa (ER), Correlação de Spearman (rs) e Eficiência de seleção (ES) da análise em látice em relação às análises em blocos e às análises utilizando métodos de análise espacial no experimento 4

Tipo de análise	CV(%)	ER (%)	[1]rs	Eficiência de seleção (%) 10(10)[2]	20(20)	30(30)
Látice	12,10	-	-	-	-	-
Blocos	15,25	158,7	0,93	66,66(7)[3]	83,33(17)	88,89(27)
[4]M_{SI} [9]	13,75	129	0,93	44,44(5)	77,77(16)	88,89(27)
M_{2SI}	14,19	137,54	0,93	44,44(5)	77,77(16)	88,89(27)
M_{4L}	13,08	116,84	0,95	66,66(7)	83,33(17)	88,89(27)
M_{8L}	12,92	113,86	0,93	55,55(6)	83,33(17)	88,89(27)
M_{TP}	15,09	155,42	0,93	77,78(8)	83,33(17)	88,89(27)
[5]RM_{SI}	10,10	69,61	0,85	22,22(3)	61,11(13)	74,07(23)
RM_{2SI}	11,66	92,75	0,88	22,22(3)	61,11(13)	81,48(25)
RM_{4L}	10,41	74	0,90	44,44(5)	77,77(16)	81,48(25)
RM_{8L}	11,43	89,19	0,95	44,44(5)	72,22(15)	85,18(26)
[6]P_{SI}	13,97	133,11	0,93	55,55(6)	77,77(16)	88,89(27)
P_{2SI}	14,13	136,23	0,92	55,55(6)	77,77(16)	88,89(27)
P_{4L}	13,07	116,64	0,95	66,66(7)	77,77(16)	92,59(28)
P_{8L}	12,69	109,86	0,95	44,44(5)	83,33(17)	92,59(28)
[7]RP_{SI}	6,74	31,01	0,90	55,55(6)	72,22(15)	88,89(27)
RP_{2SI}	9,82	65,81	0,90	44,44(5)	66,66(14)	85,18(26)
RP_{4L}	8,56	50,02	0,93	66,66(7)	83,33(17)	88,89(27)
RP_{8L}	9,59	62,82	0,94	55,55(6)	77,77(16)	92,59(28)
[8]ED	-	-	0,97	77,78(8)	83,33(17)	96,29(29)

[1] rs para classificação das médias das famílias usando a análise em látice como referência . [2] ES para 10, 20 e 30% das melhores famílias (número de famílias selecionadas). [3] Entre parênteses – número de famílias comuns em relação a análise em látice. [4] M= método médias móveis, [5] RM= reaplicação médias móveis, [6] P= método Papadakis, [7] RP= reaplicação do método Papadakis, [8] ED= modelo com erros dependentes, [9] SI,2SI,4L,8L e TP = posição das parcelas que compõem as vizinhanças no cálculo do índice ambiental.

O experimento 5 (látice 7x7) apresentou moderada dependência espacial (capítulo 1) entre resíduos estimados na análise em blocos com autocorrelações iguais $\hat{\rho}$ =0,2077 e $\hat{\rho}$= 0,2100, para 1ª e 2ª ordem , respectivamente (Tabela 20).

Tabela 20 - Autocorrelação espacial estimada ($\hat{\rho}$) de primeira e segunda ordem e probabilidade do erro tipo I (H_0: ρ =0) para o teste de Durbin-Watson, valor da verossimilhança restrita (-2RLL) e Critério de Informação de Akaike (AIC) para diferentes tipos de análises aplicadas ao experimento 5

Tipo de análise	DW 1ª ordem		DW 2ª ordem		-2RLL	AIC
	$\hat{\rho}$	p - valor	$\hat{\rho}$	p - valor		
Látice	-0, 0801	0, 1649^{ns}	-0, 0123	0, 4081^{ns}	3.324,9	- 1.663,5
Blocos	0, 2077	0, 0054**	0, 2100	0, 0063**	3.506.9	-1.754.4
[1]M_{SI} [6]	0, 0944	0, 1254^{ns}	0, 1141	0, 0957^{ns}	1.139,5	-570,7
M_{2SI}	0, 0769	0, 1749^{ns}	0, 1721	0, 0216^{ns}	1.138,5	-570,2
M_{4L}	-0, 0462	0, 2871^{ns}	0, 1194	0, 0851^{ns}	1.142,7	-572,4
M_{8L}	0, 0378	0, 3229^{ns}	0, 0094	0, 4878^{ns}	1.140,0	-571,0
M_{TP}	0, 1966	0, 0080**	0, 1891	0, 0128^{ns}	1.156,0	-579,0
[2]RM_{SI}	-0, 0214	0, 3974^{ns}	0, 2049	0, 0075**	1.049,6	-525,8
RM_{2SI}	0, 0540	0, 2559^{ns}	0, 1389	0, 0536^{ns}	1.109,7	-555,8
RM_{4L}	0, 0086	0, 4583^{ns}	0, 0307	0, 3863^{ns}	1.100,7	-551,4
RM_{8L}	0, 0378	0, 3229^{ns}	0, 0094	0, 4878^{ns}	1.126,1	-564,1
[3]P_{SI}	0, 1979	0, 0076**	0, 1907	0, 0121^{ns}	1.158,9	-580,5
P_{2SI}	0, 0873	0, 1441^{ns}	0, 1772	0, 0185^{ns}	1.147,4	-574,7
P_{4L}	-0, 1964	0, 0080**	0, 1243	0, 0762^{ns}	1.142,8	-572,4
P_{8L}	-0, 0755	0, 1792^{ns}	0, 0815	0, 1821^{ns}	1.146,9	-574,4
[4]RP_{SI}	-0, 1126	0, 0851^{ns}	0, 1684	0, 0241^{ns}	1.009,4	-505,7
RP_{2SI}	-0, 0638	0, 2191^{ns}	0, 0384	0, 3509^{ns}	1041,8	-521,9
RP_{4L}	-0, 0215	0, 3969^{ns}	-0, 0225	0, 3608^{ns}	1.056,0	-529,0
RP_{8L}	-0, 0401	0, 3131^{ns}	-0, 0067	0, 4346^{ns}	1.080,7	-541,3
[5]ED	-	-	-	-	1.144,9	-575,5

** ns Autocorrelação espacial significativa e não significativa, respectivamente, pelo teste de DW a 1% de probabilidade. [1] M= método médias móveis, [2] RM= reaplicação médias móveis [3] P= método Papadakis, [4] RP= reaplicação do método Papadakis, [5] ED= método com erros dependentes, [6] SI, 2SI, 4L, 8L e TP = posição das parcelas que compõem as vizinhanças no cálculo do índice ambiental.

Entre as metodologias testadas, houve considerável redução nas estimativas da autocorrelação espacial em relação à análise em blocos (Tabela 20). Pelos resultados obtidos, todas as metodologias conseguiram garantir a independência de 1ª e 2ª ordem, com exceção de M_{TP} e RM_{SI}. Os métodos M_{8L}, RM_{4L}, RM_{8L}, P_{8L}, $RP_{2SI,}$ RP_{4L} e RP_{8L} destacaram-se por garantir independência entre resíduos para 1ª e 2ª ordem com altos p-valores.

De acordo com os critérios de ajuste AIC e -2RLL, todas as metodologias testadas tiveram melhor ajuste que as análises em látice e em blocos (Tabela 20). E as metodologias reaplicação de médias móveis e reaplicação de Papadakis se destacaram por apresentar maior eficiência em relação à análise em látice com valores variando entre 28,85 a 98,39 e coeficiente de variação com valores entre 4,37 a 8,08, indicando alta precisão experimental.

Tabela 21 - Coeficientes de variação experimental (CV), Eficiência relativa (ER), Correlação de Spearman (rs) e Eficiência de seleção (ES) da análise em látice em relação às análises em blocos e às análises utilizando métodos de análise espacial no experimento 5

Tipo de análise	CV(%)	ER (%)	[1]rs	Eficiência de seleção (%)		
				10(5)[2]	20(10)	30(15)
Látice	8,14	-	-	-	-	-
Blocos	9,64	140,15	0,88	77,78(4)[3]	66,66(7)	85,18(13)
[4]M_{SI} [9]	8,63	112,34	0,88	77,78(4)	66,66(7)	85,18(13)
M_{2SI}	8,60	111,61	0,87	77,78(4)	66,66(7)	85,18(13)
M_{4L}	8,80	116,84	0,91	77,78(4)	66,66(7)	85,18(13)
M_{8L}	8,08	98,39	0,91	77,78(4)	55,55(6)	85,18(13)
M_{TP}	9,60	138,89	0,88	77,78(4)	66,66(7)	77,77(12)
[5]RM_{SI}	5,40	43,91	0,63	33,33(2)	33,33(4)	70,37(11)
RM_{2SI}	7,39	82,44	0,83	77,78(4)	66,66(7)	70,37(11)
RM_{4L}	7,07	75,47	0,89	55,55(3)	77,77(8)	85,18(13)
RM_{8L}	8,08	98,39	0,93	77,78(4)	66,66(7)	92,59(14)
[6]P_{SI}	9,59	138,6	0,88	55,55(3)	66,66(7)	85,18(13)
P_{2SI}	9,07	124,1	0,89	77,78(4)	66,66(7)	85,18(13)
P_{4L}	8,86	118,33	0,92	77,78(4)	66,66(7)	92,59(14)
P_{8L}	9,05	123,63	0,90	77,78(4)	66,66(7)	92,59(14)
[7]RP_{SI}	4,37	28,85	0,87	77,78(4)	66,66(7)	77,77(12)
RP_{2SI}	5,20	40,84	0,85	77,78(4)	66,66(7)	85,18(13)
RP_{4L}	5,61	47,48	0,89	77,78(4)	66,66(7)	77,77(12)
RP_{8L}	6,40	61,88	0,93	77,78(4)	66,66(7)	85,18(13)
[8]ED	-	104,4	0,94	77,78(4)	77,77(8)	92,59(14)

[1] rs para classificação das médias das famílias usando a análise em látice como referência . [2] ES para 10, 20 e 30% das melhores famílias (número de famílias selecionadas). [3] Entre parênteses – número de famílias comuns em relação a análise em látice. [4] M= método médias móveis, [5] RM= reaplicação médias móveis, [6] P= método Papadakis, [7] RP= reaplicação do método Papadakis, [8] ED= modelo com erros dependentes, [9] SI, 2SI, 4L, 8L e TP = posição das parcelas que compõem as vizinhanças no cálculo do índice ambiental.

As estimativas para os coeficientes de correlação de Spearman, usando a análise em látice como referência, variaram de 0,63 a 0,94, tendo apresentado valores superiores a 0,82 para todas as metodologias testadas, com exceção do método RM_{SI} (rs =0,63) (Tabela 21). Porém a eficiência de seleção mostrou que houve concordância inferior a 80% para todas as

metodologias para seleção de 10 e 20% das famílias do experimento 5, indicando razoável coincidência entre as famílias de feijoeiro selecionadas na análise em látice e nas metodologias espaciais testadas.

A metodologia RP_{SI} se destacou entre todas testadas garantindo a independência de 1ª e 2ª ordem entre resíduos, melhor ajuste, maior precisão com menor CV, alta eficiência em relação ao látice, correlação de Spearman superior a 0,80, porém a eficiência de seleção foi inferior a 80% em todos os níveis de seleção.

De acordo com o capítulo 1, o experimento 6 (látice 6x6) apresentou moderada dependência espacial entre resíduos estimados na análise em blocos com valores de autocorrelação espacial iguais a 0,3674 e 0,2011 para 1ª e 2ª ordens, respectivamente, e alcance igual a 6,8 m (Tabela 22).

Tabela 22 - Autocorrelação espacial estimada ($\hat{\rho}$) de primeira e segunda ordem e probabilidade do erro tipo I (H_0: ρ =0) para o teste de Durbin-Watson, valor da verossimilhança restrita (-2RLL) e Critério de Informação de Akaike (AIC) para diferentes tipos de análises aplicadas ao experimento 6

Tipo de análise	DW 1ª ordem		DW 2ª ordem		-2RLL	AIC
	$\hat{\rho}$	p - valor	$\hat{\rho}$	p - valor		
Látice	-0, 0442	0, 3224 ns	-0, 1218	0, 0852 ns	1181,3	-591,7
Blocos	0, 3674	<0, 0001**	0, 2011	0, 0218 ns	1292,4	-647,2
[1]M_{SI} [6]	0, 3053	0, 0006**	0, 1302	0, 1032 ns	911,8	-456,9
M_{2SI}	0, 3631	<0, 0001**	0, 1783	0, 0381 ns	914,4	-458,2
M_{4L}	0, 0160	0, 4339 ns	0, 0053	0, 4832 ns	887,2	-444,6
M_{8L}	0, 1279	0, 0906 ns	0, 0639	0, 2851 ns	893,3	-447,6
M_{TP}	0, 3455	0, 0001**	0, 2225	0, 0122 ns	914,1	-458,0
[2]RM_{SI}	0, 4241	<0, 0001**	0, 2198	0, 0132 ns	868,2	-435,1
RM_{2SI}	0, 3660	<0, 0001**	0, 0968	0, 1807 ns	891,0	-446,5
RM_{4L}	0, 1573	0, 0495 ns	0, 0858	0, 2126 ns	856,7	-429,3
RM_{8L}	0, 2379	0, 0060**	0, 1221	0, 1193 ns	877,3	-439,7
[3]P_{SI}	0, 3321	0, 0002**	0, 1960	0, 0248 ns	915,1	-458,6
P_{2SI}	0, 3365	0, 0002**	0, 1927	0, 0269 ns	914,0	-458,0
P_{4L}	0, 1872	0, 0246 ns	0, 1704	0, 0455 ns	909,2	-455,6
P_{8L}	0, 1097	0, 1259 ns	0, 1097	0, 1474 ns	899,2	-450,6
[4]RP_{SI}	0, 2520	0, 0038**	0, 1196	0, 1245 ns	876,6	-439,3
RP_{2SI}	0, 3307	0, 0002**	0, 1270	0, 1093 ns	883,8	-442,9
RP_{4L}	0, 1450	0, 0644 ns	0, 1246	0, 1141 ns	846,8	-424,4
RP_{8L}	0, 2113	0, 0130 ns	0, 1434	0, 0802 ns	865,2	-433,6
[5]ED	-	-	-	-	895,5	-450,8

** ns Autocorrelação espacial significativa e não significativa, respectivamente, pelo teste de DW a 1% de probabilidade. [1] M= método médias móveis, [2] RM= reaplicação médias móveis, [3] P= método Papadakis, [4] RP= reaplicação do método Papadakis, [5] ED= método com erros dependentes, [6] SI, 2SI, 4L, 8L e TP = posição das parcelas que compõem as vizinhanças no cálculo do índice ambiental.

Entre as metodologias testadas, somente os métodos $M_{4L,}$ $M_{8L,}$ $RM_{4L,}$ $P_{4L,}$ $P_{8L,}$ RP_{4L} e RP_{8L} reduziram os valores das estimativas da autocorrelação espacial em relação à análise em blocos, conseguindo garantir a independência entre resíduos de 1ª e 2ª ordem. Os métodos M_{4L} (p-valor = 0,4339 e p-valor= 0,4832) e P_{8L} (p-valor=0,1259 e p-valor=0,1474) se destacaram por apresentar independência de 1ª e 2ª ordem com altos p-valores (Tabela 22).

De acordo com os critérios de ajuste AIC e -2RLL, todas as metodologias testadas tiveram melhor ajuste que as análises em látice e em blocos (Tabela 22).

As metodologias RM_{4L} e RP_{4L} se destacaram por serem mais eficientes que a análise em látice com valores de eficiência relativa iguais a 90,86 e 79,55 e coeficiente de variação igual a 10,14 e 9,49, respectivamente, indicando alta precisão experimental (Tabela 23).

Tabela 23 - Coeficientes de variação experimental (CV), Eficiência relativa (ER), Correlação de Spearman (rs) e Eficiência de seleção (ES) da análise em látice em relação às análises em blocos e às análises utilizando métodos de análise espacial no experimento 6

Tipo de análise	CV(%)	ER (%)	[1]rs	Eficiência de seleção (%)		
				10(4)[2]	20(7)	30(11)
Látice	10,64	-				
Blocos	15,86	221,96	0,80	50,62(2)[3]	66,04(5)	40,32(5)
[4]M_{SI} [9]	15,04	199,55	0,79	50,62(2)	66,04(5)	40,32(5)
M_{2SI}	15,35	207,9	0,80	81,46(3)	66,04(5)	40,32(5)
M_{4L}	12,61	140,31	0,86	50,62(2)	66,04(5)	50,62(6)
M_{8L}	13,21	153,96	0,82	50,62(2)	66,04(5)	50,62(6)
M_{TP}	15,83	221,14	0,82	50,62(2)	66,04(5)	40,32(5)
[5]RM_{SI}	10,99	106,6	0,70	81,46(3)	50,62(4)	30,04(4)
RM_{2SI}	12,99	148,97	0,78	81,46(3)	66,04(5)	40,32(5)
RM_{4L}	10,14	90,86	0,79	50,62(2)	50,62(4)	60,90(7)
RM_{8L}	11,81	123,13	0,83	81,46(3)	66,04(5)	50,62(6)
[6]P_{SI}	15,49	211,84	0,82	50,62(2)	66,04(5)	50,62(6)
P_{2SI}	15,42	209,99	0,83	50,62(2)	66,04(5)	40,32(5)
P_{4L}	14,90	195,91	0,85	81,46(3)	66,04(5)	60,90(7)
P_{8L}	13,88	170,16	0,84	50,62(2)	66,04(5)	50,62(6)
[7]RP_{SI}	11,73	121,43	0,80	50,62(2)	66,04(5)	50,62(6)
RP_{2SI}	12,41	135,89	0,80	81,46(3)	66,04(5)	50,62(6)
RP_{4L}	9,49	79,55	0,79	50,62(2)	66,04(5)	40,32(5)
RP_{8L}	10,88	104,53	0,80	81,46(3)	66,04(5)	50,62(6)
[8]ED	-	146	0,88	50,62(2)	66,04(5)	60,90(7)

[1] rs para classificação das médias das famílias usando a análise em látice como referência . [2] ES para 10, 20 e 30% das melhores famílias (número de famílias selecionadas). [3] Entre parênteses – número de famílias comuns em relação a análise em látice, [4] M = método médias móveis, [5] RM= reaplicação médias móveis, [6] P = método Papadakis, [7] RP= reaplicação do método Papadakis, [8] ED= modelo com erros dependentes, [(9)] SI, 2SI, 4L, 8L e TP = posição das parcelas que compõem as vizinhanças no cálculo do índice ambiental.

As estimativas para os coeficientes de correlação de Spearman, usando a análise em látice como referência, variaram de 0,70 a 0,86. A eficiência de seleção mostrou que houve baixa concordância entre as diferentes metodologias testadas no experimento 6 e na análise em látice. Todos os métodos apresentaram valores de eficiência de seleção inferiores a 80%, exceto $M_{2SI,}$ $RM_{SI,}$ $RM_{2SI,}$ RM_{8L}, $P_{4L,}$ RP_{2SI} e $RP_{8L,}$ todos com eficiência de seleção igual a 81,46 % para 10% das famílias selecionadas.

Pelos resultados obtidos, várias metodologias não conseguiram garantir a independência entre resíduos, tendo sido menos eficientes que a análise em látice e apresentado maiores coeficientes de variação, indicando que a análise em látice se mostrou superior às metodologias espaciais testadas, com exceção dos métodos RM_{4L} e RP_{4L}.

O experimento 7 (látice 15x15) apresentou de fraca a moderada dependência espacial entre resíduos estimados na análise em blocos, com alcance igual a 33,3m (de acordo com o resultado do capítulo 1). Com a aplicação dos métodos de análise espacial, houve considerável redução nas estimativas da autocorrelação espacial em relação à análise em blocos, com exceção dos métodos $M_{SI,}$ M_{TP} e P_{4L} para autocorrelação de 1ª ordem (Tabela 24)

Tabela 24 - Autocorrelação espacial estimada ($\hat{\rho}$) de primeira e segunda ordem e probabilidade do erro tipo I (H_0: ρ =0) para o teste de Durbin-Watson, valor da verossimilhança restrita (-2RLL) e Critério de Informação de Akaike (AIC) para diferentes tipos de análises aplicadas ao experimento 7

Tipo de análise	DW 1ª ordem		DW 2ª ordem		-2RLL	AIC
	$\hat{\rho}$	p - valor	$\hat{\rho}$	p - valor		
Látice	-0,0203	0, 2985 ns	0, 0433	0, 1385 ns	7.606,8	-3.804.4
Blocos	0, 1619	<0, 0001**	0, 1928	<0, 0001**	7.801,7	-3.901,9
[1]M_{SI} [6]	0, 1613	<0, 0001**	0,1900	<0, 0001**	5.618,4	-2.810,2
M_{2SI}	0, 1148	0, 0001**	0, 1253	<0, 0001**	5.557,1	-2.778,6
M_{4L}	-0, 0181	0, 6814 ns	0, 1513	<0, 0001**	5.569,9	-2.785,9
M_{8L}	-0, 0222	0, 2818 ns	0, 1540	<0, 0001**	5.560,4	-2.781,2
M_{TP}	0, 1442	<0, 0001**	0, 1769	<0, 0001**	5.612,4	-2.807,2
[2]RM_{SI}	0, 0198	0, 3032 ns	0, 0344	0, 1959 ns	5.210,6	-2.606,3
RM_{2SI}	0, 0951	0, 0006**	0, 0952	0, 0073**	5.349,1	-2.675,6
RM_{4L}	-0, 0172	0, 3271 ns	0, 0962	0, 0068**	5.402,5	-2.702,2
RM_{8L}	0, 0203	0, 2982 ns	0, 0686	0, 0040**	5.419,3	-2.710,6
[3]P_{SI}	0, 1008	0, 0043**	0, 1065	0, 0031**	5.518,5	-2.760,3
P_{2SI}	0, 0814	0, 0170 ns	0, 1076	0, 0028**	5.536,4	-2.769,2
P_{4L}	-0, 2442	<0, 0001**	0, 1097	0, 0024**	5.516,6	-2.759,3
P_{8L}	-0, 1373	0, 0020**	0, 1364	0, 0020**	5.550,7	-2.776,3
[4]RP_{SI}	0, 0317	0, 2046 ns	0, 0610	0, 0606 ns	4.907,3	-2.454,7
RP_{2SI}	0, 1147	0, 0114 ns	0, 1080	0, 0027**	5.158,6	-2.580,3
RP_{4L}	-0, 0816	0, 0168 ns	0, 0331	0, 2054 ns	5.192,7	-2.597,4
RP_{8L}	-0, 0649	0, 0457 ns	0, 0065	0, 4475 ns	5.294,6	-2.648,3
[5]ED	-	-	-	-	5.587,1	-2.796.6

** ns Autocorrelação espacial significativa e não significativa, respectivamente, pelo teste de DW a 1% de probabilidade. (1) M= método médias móveis, (2) RM= reaplicação médias móveis (3) P= método Papadakis, (4) RP= reaplicação do método Papadakis, (5) ED= método com erros dependentes, (6) SI, 2SI, 4L, 8L e TP = posição das parcelas que compõem as vizinhanças no cálculo do índice ambiental.

Os métodos RM_{SI} (p-valor = 0,3032 e p-valor = 0,1959), RP_{SI} (p-valor = 0,2046 e p-valor=0,0606), RP_{4L}(p-valor = 0,0168 e p-valor =0,2054) e RP_{8L}(p=0,0457 e p=0,4475) conseguiram garantir a independência entre resíduos de 1ª e 2ª ordem (Tabela 24).

De acordo com os critérios de ajuste AIC e -2RLL, todas as metodologias testadas tiveram melhor ajuste que as análises em látice e em blocos, destacando-se as reaplicações do método de Papadakis com melhores resultados entre todas as metodologias testadas (Tabela 24).

Nos diferentes tipos de análises relativas ao experimento 7, as metodologias de análise espacial foram mais eficientes que a análise em látice, com exceção dos métodos médias móveis, destacando-se as reaplicações

médias móveis e reaplicações de Papadakis, com valores de eficiência relativa variando de 48 a 76,65 e 24,7 a 58,20 e coeficiente de variação com valores entre 11,14 a 14,08 e 7,94 a 12,26, respectivamente, indicando de alta a média precisão experimental (Tabela 25).

Entre as metodologias testadas, destacou-se a reaplicação do método de Papadakis $RP_{SI,}$ RP_{2SI} e RP_{4L}, e, em relação a estes métodos, a análise em látice foi 75, 57,19 e 53,77%, respectivamente, menos eficiente.

As estimativas para os coeficientes de correlação de Spearman e eficiência de seleção, usando a análise em látice como referência, foram altas para todas as análises realizadas no experimento 7, exceto para a $RM_{SI,}$ que apresentou menor correlação de Spearman (rs =0,88) e eficiência de seleção menor que 70% para 10% e 20% de seleção. Observou-se que o aumento do número de famílias selecionadas ocasionou maior concordância entre a análise em látice e as demais metodologias testadas.

Pelos resultados obtidos, os métodos RM_{SI}, RP_{SI}, RP_{4L} e RP_{8L} mostram-se iguais ou superiores à análise em látice, pois, além de garantirem a independência entre resíduos, ainda apresentaram melhores ajustes, maior eficiência e melhor precisão experimental em relação à análise em látice.

Tabela 25 - Coeficientes de variação experimental (CV), Eficiência relativa (ER), Correlação de Spearman (rs) e Eficiência de seleção (ES) da análise em látice em relação às análises em blocos e às análises utilizando métodos de análise espacial no experimento 7

Tipo de análise	CV(%)	ER (%)	[1]rs	Eficiência de seleção (%) 10(23)[2]	20(45)	30(68)
Látice	16,08	-	-	-	-	-
Blocos	17,69	121,00	0,97	84,85(19) [3]	80,25(37)	90,05(61)
[4]M_{SI} [9]	17,52	118,71	0,97	84,85(19)	80,25(37)	91,71(62)
M_{2SI}	16,41	104,13	0,96	79,80(18)	80,25(37)	91,71(62)
M_{4L}	16,64	107,13	0,97	84,85(19)	82,72(38)	88,39(60)
M_{8L}	16,48	104,99	0,97	79,80(18)	80,25(37)	90,05(61)
M_{TP}	17,66	120,52	0,97	84,85(19)	82,72(38	91,71(62)
[5]RM_{SI}	11,14	48,00	0,88	59,60(14)	67,90(32)	83,42(57)
RM_{2SI}	13,01	65,48	0,93	74,75(17)	77,78(36)	88,39(60)
RM_{4L}	13,81	73,77	0,95	69,70(16)	80,25(37)	88,39(60)
RM_{8L}	14,08	76,65	0,97	79,80(18)	85,19(39)	90,05(61)
[6]P_{SI}	15,73	95,66	0,96	84,85(19)	85,19(39)	90,05(61)
P_{2SI}	16,05	99,70	096	74,75(17)	80,25(37)	91,71(62)
P_{4L}	15,70	95,36	0,97	89,90(20)	87,65(40)	90,05(61)
P_{8L}	15,73	95,66	0,97	84,85(19)	82,72(38)	91,71(62)
[7]RP_{SI}	7,94	24,37	0,96	74,75(17)	85,19(39)	90,05(61)
RP_{2SI}	10,52	42,81	0,95	79,80(18)	75,31(35)	90,05(61)
RP_{4L}	10,93	46,23	0,97	79,80(18)	82,72(38)	88,39(60)
RP_{8L}	12,26	58,20	0,97	79,80(18)	82,72(38)	91,71(62
[8]ED	-	84,80	0,98	84,85(19)	90,12(41)	95,02(64)

[1] rs para classificação das médias das famílias usando a análise em látice como referência . [2] ES para 10, 20 e 30% das melhores famílias (número de famílias selecionadas). [3] Entre parênteses – número de famílias comuns em relação a análise em látice. [4] M= método médias móveis, [5] RM= reaplicação médias móveis, [6] P= método Papadakis, [7] RP= reaplicação do método Papadakis, [8] ED= modelo com erros dependentes, [(9)] SI, 2SI, 4L, 8L e TP = posição das parcelas que compõem as vizinhanças no cálculo do índice ambiental.

De acordo com o capítulo 1, o experimento 8 (látice 10x10) apresentou moderada dependência espacial entre resíduos para as análises em látice e em blocos, com alcance estimado em 2,6 m .

Houve considerável redução nas estimativas da autocorrelação espacial em relação às análises em blocos e látice para 1ª e 2ª ordens com a utilização dos métodos de análise espacial que garantiram a independência entre resíduos ($p<0,01$), com exceção dos métodos M_{SI}, M_{2SI}, P_{2SI}, RP_{2SI}, M_{TP} (Tabela 26).

Tabela 26 - Autocorrelação espacial estimada ($\hat{\rho}$) de primeira e segunda ordem e probabilidade do erro tipo I (H0: ρ=0) para o teste de Durbin-Watson, valor da verossimilhança restrita (-2RLL) e Critério de Informação de Akaike (AIC) para diferentes tipos de análises aplicadas ao experimento 8

Tipo de análise	DW 1ª ordem		DW 2ª ordem		-2RLL	AIC
	$\hat{\rho}$	p - valor	$\hat{\rho}$	p - valor		
Látice	0, 1707	0.0014**	-0, 0913	0, 0500 ns	3.356,8	-1.679,4
Blocos	0, 3246	<0, 0001**	0, 0784	0, 0961 ns	3.439,2	-1. 720,6
[1]M_{SI} [6]	0, 2432	<0, 0001**	0, 0095	0, 4578 ns	2.441,5	-1.221,8
M_{2SI}	0, 1859	0, 0006**	0, 0165	0, 4098 ns	2.442,7	-1.222,4
M_{4L}	0, 0238	0, 3400 ns	0, 0145	0, 2311 ns	2.435,5	-1.218,8
M_{8L}	0, 0774	0, 0894 ns	0, 0445	0, 2373 ns	2.433,4	-1.217,7
M_{TP}	0, 3177	<0, 0001**	0, 0680	0, 1308 ns	2.471,0	-1.236,5
[2]RM_{SI}	0, 0195	0, 3677 ns	0, 1082	0, 0341 ns	2.232,7	-1.117,4
RM_{2SI}	0, 1169	0, 0210 ns	0, 0375	0, 2769 ns	2.322,6	-1.162,3
RM_{4L}	0, 0561	0, 1649 ns	0, 0541	0, 1892 ns	2.322,5	-1.162,3
RM_{8L}	0, 1094	0, 0285 ns	-0, 0336	0, 2606 ns	2.348,1	-1.175,1
[3]P_{SI}	0,0774	<0, 0001 ns	0,0445	0, 4887 ns	2.460,3	-1.231,1
P_{2SI}	0, 2085	0, 0001**	0, 0293	0, 3260 ns	2.451,0	-1.226,5
P_{4L}	-0, 0648	0, 1307 ns	-0, 0308	0, 2770 ns	2.377,9	-1.189,9
P_{8L}	0, 1148	0, 0229 ns	0, 0337	0, 2990 ns	2.438,3	-1.220,2
[4]RP_{SI}	0, 0949	0, 0496 ns	0, 0190	0, 3929 ns	2.144,0	-1.073,0
RP_{2SI}	0, 1703	0, 0015**	-0, 0015	0, 4663 ns	2.227,0	-1.139,5
RP_{4L}	0, 1117	0, 0261 ns	0, 0016	0, 4882 ns	2.255,7	-1.128,8
RP_{8L}	0, 0553	0, 1685 ns	-0, 0757	0, 0849 ns	2.312,8	-1.157,4
[5]ED	-	-	-	-	2555.6	-1279.8

** ns Autocorrelação espacial significativa e não significativa, respectivamente, pelo teste de DW a 1% de probabilidade. (1) M= método médias móveis, (2) RM= reaplicação médias móveis (3) P= método Papadakis, (4) RP= reaplicação do método Papadakis, (5) ED= método com erros dependentes, (6) SI, 2SI, 4L, 8L e TP = posição das parcelas que compõem as vizinhanças no cálculo do índice ambiental.

De acordo com os critérios AIC e -2RLL, todas as metodologias testadas apresentaram melhor ajuste que as análises em látice (-2RLL = 3.356,8) e em blocos (-2RLL = 3439,2) (Tabela 26). As reaplicações de médias móveis e reaplicações de Papadakis destacaram-se por apresentar melhores ajustes entre todas as metodologias testadas.

A análise em látice foi menos eficiente em relação às metodologias reaplicação de médias móveis, reaplicação de Papadakis, P_{4L}, P_{8L}, M_{4L}, M_{8L} e $M_{2SI,}$ com valores de eficiência relativa variando entre 20,80 e 96,50 e coeficiente de variação variando entre 5,82 a 12,53, indicando de alta a média precisão experimental (Tabela 27).

Tabela 27 - Coeficientes de variação experimental (CV), Eficiência relativa (ER), Correlação de Spearman (rs) e Eficiência de seleção (ES) para análise em látice em relação às análises em blocos e às análises utilizando métodos de análise espacial para o experimento 8

Tipo de análise	CV(%)	ER (%)	[1]rs	Eficiência de seleção (%) 10(10)[2]	20(20)	30(30)
Látice	12,75	-	-	-	-	-
Blocos	13,61	113,7	0,88	66,66(7)[3]	61,11(13)	66,66(21)
[4]M_{SI} [9]	12,87	101,8	0,86	66,66(7)	66,66(14)	74,07(23)
M_{2SI}	12,53	96,5	0,84	66,66(7)	66,66(14)	66,66(21)
M_{4L}	12,31	93,1	0,89	77,77(8)	72,22(15)	70,37(22)
M_{8L}	12,25	92,2	0,87	77,77(8)	72,22(15)	70,37(22)
M_{TP}	13,61	113,7	0,87	77,77(8)	61,11(13)	66,66(21)
[5]RM_{SI}	7,35	33,3	0,65	33,33(4)	38,86(9)	48,14(16)
RM_{2SI}	9,25	52,6	0,72	55,55(6)	55,55(12)	55,55(18)
RM_{4L}	9,25	52,6	0,77	66,66(7)	55,55(12)	59,26(19)
RM_{8L}	9,88	60	0,82	66,66(7)	66,66(14)	62,96(20)
[6]P_{SI}	12,99	103,7	0,86	66,66(7)	72,22(15)	66,66(21)
P_{2SI}	12,71	109,9	0,83	66,66(7)	66,66(14)	62,96(20)
P_{4L}	10,89	72,8	0,83	66,66(7)	77,77(16)	62,96(20)
P_{8L}	12,30	93	0,87	77,77(8)	77,77(16)	66,66(21)
[7]RP_{SI}	5,82	20,8	0,81	66,66(7)	61,11(13)	66,66(21)
RP_{2SI}	8,17	41	0,79	66,66(7)	72,22(15)	62,96(20)
RP_{4L}	7,75	36,9	0,81	66,66(7)	66,66(14)	62,96(20)
RP_{8L}	8,96	49,4	0,85	66,66(7)	72,22(15)	66,66(21)
[8]ED	-	-	0,89	77,77(8)	77,77(16)	70,37(22)

[1] rs para classificação das médias das famílias usando a análise em látice como referência . [2] ES para 10, 20 e 30% das melhores famílias (número de famílias selecionadas). [3] Entre parênteses – número de famílias comuns em relação a análise em látice. [4] M= método médias móveis, [5] RM= reaplicação médias móveis, [6] P= método Papadakis, [7] RP= reaplicação do método Papadakis, [8] ED= modelo com erros dependentes, [(9)] SI, 2SI, 4L, 8L e TP = posição das parcelas que compõem as vizinhanças no cálculo do índice ambiental.

As estimativas para os coeficientes de Correlação de Spearman variaram entre 0,65 e 0,89 com menores valores para os métodos RM_{SI}(rs = 0,65), RM_{2SI} (rs = 0,77), RM_{4L}(rs = 0,77) e RP_{2SI} (rs = 0,79). As demais metodologias apresentam valores próximos a 0,84 (Tabela 27). E a eficiência de seleção mostrou que houve baixa concordância entre os vários métodos testados e a análise em látice e que o aumento no número de famílias não ocasionou maior concordância entre as famílias selecionadas na análise em látice e os demais métodos testados.

As metodologias baseadas na reaplicação de médias móveis e reaplicação de Papadakis foram superiores à análise em látice, conseguindo

garantir independência entre resíduos, maior eficiência relativa e menor precisão experimental.

Costa et al. (2005), trabalhando com progênies de milho e feijão, concluíram que os métodos médias móveis e Papadakis foram mais eficientes que a análise em látice**,** e a eficiência relativa do látice em relação a estes dois métodos foi de 94,5 e 93,8%, respectivamente. Souza et al. (2000), trabalhando com famílias provenientes do programa de melhoramento do feijoeiro, concluíram que os métodos de análise de vizinhança Papadakis e médias móveis são eficientes no controle da heterogeneidade dos blocos, sendo esta eficiência similar à proporcionada pela análise em látice, resultados semelhantes aos obtidos nesta dissertação onde as metodologias de análise espacial testadas também se mostraram eficientes em relação à análise em látice. De acordo com Costa et al. (2005), o método com erros dependentes (ER=100) foi ineficiente em relação à análise em látice, contrário ao obtido neste trabalho, em que este método apresentou razoável eficiência em relação à análise em látice. Ainda de acordo com Souza et al. (2000), os métodos médias móveis e Papadakis proporcionaram uma razoável redução no valor do CV para o método Papadakis (CV=30,28) e para o método médias móveis (CV= 29,94) em relação à análise em látice (CV= 33,93) , enquanto neste trabalho estes métodos foram pouco eficientes para reduzir o valor do coeficiente de variação.

Em relação às estimativas do coeficiente de Spearman, Costa et al. (2005) e Souza et al. (2000) obtiveram elevados valores indicando que não houve grandes diferenças na classificação encontrada entre a análise em látice e as metodologias de análise espacial testada, resultados semelhantes aos encontrados neste trabalho. Souza et al. (2000) avaliaram a eficiência de seleção para as porcentagens 16, 33 e 50% das famílias superiores em relação aos métodos de análise de vizinhança médias móveis e Papadakis e constataram que houve boa coincidência entre as metodologias médias móveis (ES= 68,75%) e Papadakis (ES=68,75%), coincidindo com os resultados encontrados neste trabalho. Duarte (2000), em um ensaio de competição de soja, comparou a eficiência de seleção entre o método de análise com erros dependentes espacialmente e blocos aumentados, e, considerando uma seleção hipotética de 25% das linhagens mais produtivas, observou uma

coincidência de apenas 46% entre as duas seleções. Estes resultados mostraram que ocorreu grande diferença no ordenamento das linhagens entre a análise clássica e o método de análise espacial. Resultados semelhantes aos de Duarte (2000) não foram observados no presente estudo, em que dos seis experimentos testados, cinco apresentaram coincidência acima de 70% entre as metodologias látice e método com erros dependentes.

4. CONCLUSÕES

Em relação à precisão experimental, a eficiência da análise em látice foi maior que a eficiência da análise em blocos para experimentos com alta dependência espacial entre resíduos.

Comparados com a análise em látice, os métodos de análise espacial médias móveis, Papadakis e o método com erros dependentes apresentaram razoável eficiência e as reaplicações das médias móveis e Papadakis foram altamente eficientes, destacando-se o método reaplicação de Papadakis cuja vizinhança é uma parcela superior e uma inferior em relação à parcela referência.

Os métodos de análise estatística espacial avaliados, com controle local para blocos, apresentaram melhor ajuste que a análise em látice.

Os métodos testados apresentaram razoável nível de concordância em relação às famílias selecionada na análise em látice.

As novas metodologias propostas, reaplicação de médias móveis e reaplicação de Papadakis, mostram-se altamente eficientes para melhorar a precisão experimental, além de conseguir garantir a independência entre resíduos, sendo uma excelente alternativa de análise em relação à análise em látice.

5. REFERÊNCIAS BIBLIOGRÁFICAS

COSTA, J. R., BUENO FILHO, J. S., RAMALHO, M. A. P. **Análise espacial e de vizinhança no melhoramento genético de plantas**. Pesq. agropec. bras., Brasília, v. 40, n. 11, 1073-1079, 2005.

DUARTE, J. B. **Sobre o emprego e a análise estatística do delineamento em blocos aumentados no melhoramento genético vegetal.** 2000.293p. Tese (Doutorado em Agronomia/Genética e Melhoramento de Plantas)-Escola Superior de Agricultura Luiz de Queiroz, Piracicaba, SP.

DUARTE, J. B. e VENCOVSKY, R. **Spatial statistical analysis and selection of genotypes in plant breeding**. Pesq. agropec. bras., Brasília, v. 40, n. 2, 107-114, 2005.

MARTÌNEZ, R. **Control de La correlacion espacial em experimentos de campo em El sector agrícola.** Agronomia colombiana, 11(1): 83-89. 1994.

REIS, J. A. S.; MIRANDA FILHO, J. B. **Autocorrelação espacial na avaliação de compostos de milho para resistência à lagarta do cartucho (*Spodoptera frugiperda*)**. Pesquisa Agropecuária Tropical, 33 (2): 65-72, 2003.

SAS INSTITUTE. **SAS/STAT User´s Guide 8.0**, Cary N. C.: SAS Institute Inc., 1999.

SOUZA, E. A.; GERALDI, I. O.; RAMALHO, M. A. P. **Alternativas experimentais na avaliação de famílias em programas de melhoramento genético do feijoeiro**. Pesq. agropec. bras., Brasília, v. 35, n. 9, 1765-1771, 2000.

VIVALDI, L. J. **Comparação entre métodos de análise espacial de experimentos de campo.** Pesquisa Agropecuária Brasileira, Brasília, v.25, n.1, p.77-84, jan.1990.

HAMBLIN, J.; ZIMMERMANN, M. J. O. **Breeding common bean for yield mixtures. Plant Breeding Reviews**, New York, v.4, 245-272, 1986.

SAS Online Documentation. Disponível em:<http://www.v8doc.sas.com/sashtlm>. Acesso em dez. 2007.

ANEXOS

Anexo 1

Rotinas utilizadas para verificar a existência de dependência espacial (SAS 8.0). Os modelos foram ajustados utilizando-se o método da máxima verossimilhança restrita, do proc mixed do SAS.

Os proc do SAS utilizados foram:
-Teste de Durbin Watson

```
proc autoreg data=residuos;
model reslat = / dw=10 dwprob;
run;
```

-semivariograma empírico

```
proc variogram data=residuos outvar=plot_variogram;
coordinates xc=xmm yc=ymm; /*xc e yc são os valores das coordenada em
milímetro  */
compute lagd=675 maxlags=28000;
var resid;
run;
proc gplot data=plot_variogram;
plot variog*distance;
run;
```

- Critérios de ajuste para análise com erros independentes

```
proc mixed data=residuos scoring=5 convh=1e-6;
title "ajuste dos modelos de semivariograma";
model resid= ;
run;
```

- Ajuste dos modelos de semivariograma com efeito pepita

```
proc mixed data=residuos convh=1e-6;
title "ajuste dos modelos de semivariograma";
model resid= ;
/* local especifica que o modelo e com efeito pepita*/
repeated / sub=intercept local type=sp (sph) (xmm ymm);
/*entrar com valores iniciais para busca. sequência: (alcance)(contribuiçao)
(efeito pepita)*/
parms (0 to 1000 by 500) (2000 to 3000 by 500) (500 to 1000 by 500);
run;
```

- Ajuste dos modelos de semivariograma sem efeito pepita

```
proc mixed data=residuos /*scoring=5*/ ;
title "ajuste dos modelos de semivariograma";
model resid = ;
repeated / sub=intercept type=sp (sph) (xmm ymm);
/*entrar com valores iniciais para busca. sequência: (alcance)(contribuiçao ou
sill) */
parms (0 to 7000 by 100) (100 to 10000 by 500);
```

run;
Os proc do SAS utilizados foram:

-Teste de Durbin Watson

```
proc autoreg data=residuos;
model reslat = / dw=10 dwprob;
run;
```

-semivariograma empírico

```
proc variogram data=residuos outvar=plot_variogram;
coordinates xc=xmm yc=ymm; /*xc e yc são os valores das coordenada em milímetro */
compute lagd=675 maxlags=28000;
var resid;
run;
proc gplot data=plot_variogram;
plot variog*distance;
run;
```

- Critérios de ajuste para análise com erros independentes

```
proc mixed data=residuos scoring=5 convh=1e-6;
title "ajuste dos modelos de semivariograma";
model resid= ;
run;
```

- Ajuste dos modelos de semivariograma com efeito pepita

```
proc mixed data=residuos convh=1e-6;
title "ajuste dos modelos de semivariograma";
model resid= ;
/* local especifica que o modelo e com efeito pepita*/
repeated / sub=intercept local type=sp (sph) (xmm ymm);
/*entrar com valores iniciais para busca. sequência: (alcance)(contribuiçao) (efeito pepita)*/
parms (0 to 1000 by 500) (2000 to 3000 by 500) (500 to 1000 by 500);
run;
```

- Ajuste dos modelos de semivariograma sem efeito pepita

```
proc mixed data=residuos /*scoring=5*/ ;
title "ajuste dos modelos de semivariograma";
model resid = ;
repeated / sub=intercept type=sp (sph) (xmm ymm);
/*entrar com valores iniciais para busca. sequência: (alcance)(contribuiçao ou sill) */
parms (0 to 7000 by 100) (100 to 10000 by 500);
```

```
run;
```

Rotinas utilizadas no capítulo 2

- **Análise do látice tradicional**

```
proc glm data=;
title  "Analise em latice tradicional (modelo fixo)";
class trat rep bloco;
model prodpar = trat rep bloco(rep);
lsmeans trat /out=saida2;
run;
Resíduos
output out=residuos r=reslat;
run;
proc print data=residuos;
run;

Médias
data mlat;
set saida2;
rename lsmean=mlat;
proc print data=mlat;
run;
```

- A rotina para análise em blocos é a mesma da rotina do látice, basta retirar o termo bloco na frase class trat rep bloco;

- Para as metodologias MM, PP, RMM, RPP, é a mesma rotina da análise em blocos basta substituir o termo bloco por índice/solution na frase model prodpar = trat rep bloco (rep);

- Rotina das análises em látice e em blocos para análise espacial

```
proc mixed data=   ;
title  "Analise EM LATICE (modelo fixo)";
class rep trat bloco;
model prodpar=rep trat bloco(rep)/ outp=residuos;
run;
proc variogram data=residuos outvar=plot_variogram;
coordinates xc=xmm yc=ymm;
compute lagd=   maxlags=  ;
var resid;
run;

proc print data=plot_variogram;
run;
proc gplot data=plot_variogram;
plot variog*distance;
run;
```

- A rotina para análise em blocos é a mesma da rotina do látice, basta retirar o termo bloco na frase class trat rep bloco;

- Critérios AIC,BIC e -2RELL para os métodos de análise de vizinhança

```
/*Analise em bloco, usando indice como covariavel*/
proc mixed data=gilmar1;
title "Analise em bloco, usando distancia como covariavel";
class rep trat ;
model prodpar = indice rep trat /solution;
lsmeans trat;
run;
```

- Critérios AIC, BIC E -2RELL para o método de erros dependentes

```
proc mixed data=      noprofile;
title  "Analise do latice em bloco(modelo fixo)";
class rep trat;
model prodpar=rep trat/ outp=residuos;
repeated / subject=intercept local type=sp (sph) (xmm ymm);
   /*     ( alcance,  contribuição sill) */
parms ( ---------       ------          ------) / noiter;
run;
```

*** O cálculo do índice ambiental foi realizada no programa Excel. Foi reproduzido o croqui na planilha e executados os cálculos.**

*** as distâncias entre parcelas utilizadas no índice médias móveis (M_{TP}) foi calculada no programa Statistic.**