



**AVALIAÇÃO E APLICAÇÃO DE TESTES
PARA A DETECÇÃO DE
AUTOCORRELAÇÃO ESPACIAL
UTILIZANDO DADOS DE MARCADORES
GENÉTICOS**

JOÃO MARCOS LOUZADA

2003



João Marcos Louzada

**AVALIAÇÃO E APLICAÇÃO DE TESTES PARA A
DETECÇÃO DE AUTOCORRELAÇÃO ESPACIAL
UTILIZANDO DADOS DE MARCADORES
GENÉTICOS**

Dissertação apresentada à Universidade Federal de Lavras como parte das exigências do Curso de Mestrado em Agronomia, área de concentração em Estatística e Experimentação Agropecuária, para obtenção do título de "Mestre".

Orientador:

Eduardo Bearzoti

Co-orientador:

Profa. Dulcinéia de Carvalho

LAVRAS - MINAS GERAIS - BRASIL

2003

**Ficha Catalográfica Preparada pela Divisão de Processos Técnicos da
Biblioteca Central da UFLA**

Louzada, João Marcos

Avaliação e aplicação de testes para a detecção de autocorrelação espacial
utilizando dados de marcadores genéticos / João Marcos Louzada -- Lavras :
UFLA, 2003.

103 p. : il.

Orientador: Eduardo Bearzoti.
Dissertação (Mestrado) – UFLA.
Bibliografia.

1. Estatística aplicada. 2. Marcador genético. I. Universidade Federal de
Lavras. II. Título.

CDD-519.537
-630.2195

João Marcos Louzada

**AVALIAÇÃO E APLICAÇÃO DE TESTES PARA A
DETECÇÃO DE AUTOCORRELAÇÃO ESPACIAL
UTILIZANDO DADOS DE MARCADORES
GENÉTICOS**

Dissertação apresentada à Universidade Federal de Lavras como parte das exigências do Curso de Mestrado em Agronomia, área de concentração em Estatística e Experimentação Agropecuária, para obtenção do título de “Mestre”.

APROVADA em 27 de fevereiro de 2003

Prof. Dr. Marcelo Silva de Oliveira

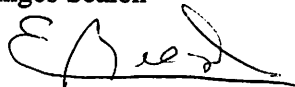
UFLA

Profa. Dra. Dulcinéia de Carvalho

UFLA

Prof. Dr. João Domingos Scalon

UFSJ



Prof. Dr. Eduardo Bearzoti

UFLA

(Orientador)

LAVRAS
MINAS GERAIS - BRASIL
2003

Cântico de vitória

Mas em todas estas coisas somos mais que vencedores, por aquele que nos amou.

Pois estou certo de que, nem a morte, nem a vida, nem os anjos, nem os principados, nem as potestades, nem o porvir, nem a altura, nem a profundidade, nem alguma outra criatura nos poderá separar do amor de Deus, que está em Cristo Jesus nosso Senhor.

Romanos 8: 37-39

À minha esposa, Andreia, pelo apoio e carinho; meus queridos filhos Tarcylla e Talles, meus irmãos e aos meus pais,

DEDICO.

Agradecimentos

A Deus, por ter me sustentado até aqui e por mais esta benção em minha vida.

Ao meu pai, Veredino e à minha mãe, Jandyra, pelos ensinamentos, apoio, imensurável amor e por suas orações.

À minha esposa, Andreia, pelo seu amor e dedicação para comigo e por ter sido minha companheira em todos os momentos do curso.

Aos meus queridos filhos, Tarcylla e Thales, por serem minha fonte de inspiração e motivação para superar os momentos de dificuldades.

Ao Professor Eduardo Bearzoti, pela competência, seriedade, amizade e orientação.

Ao Professor Joel, por sua incansável dedicação em aprimorar o Curso e à Capes, pela concessão de uma bolsa de estudos, por meio do programa PCIDTec.

À Universidade Federal de Lavras e ao Departamento de Ciências Exatas, pela oportunidade de aprendizado.

Aos professores do Mestrado em Agronomia/Estatística e Experimentação Agropecuária, pelos ensinamentos e amizade.

Ao grande amigo e irmão Veredino e minha cunhada Adriana, pelo incentivo e apoio.

À minha cunhada Márcia, pelo grande apoio.

Aos colegas de curso pela convivência, em especial ao grande amigo Washington, pela sinceridade e aprendizado.

Aos funcionários do Departamento de Ciências Exatas e da biblioteca, pela eficiência e boa vontade.

À Escola Agrotécnica Federal de Colatina, por ter me concedido a oportuni-

Ao casal Nilson e Sandra, pela inestimável amizade e apoio moral.

Ao Prof. Dr. Marcelo Silva de Oliveira, Profa. Dra. Dulcinéia de Carvalho e ao Prof. Dr. João Domingos Scalon, pela grande contribuição para o aperfeiçoamento deste trabalho.

SUMÁRIO

RESUMO	i
ABSTRACT	iii
1. INTRODUÇÃO	1
2. REFERENCIAL TEÓRICO	4
2.1 Padrão espacial.....	5
2.2 O índice I de Moran.....	14
2.3 Construção da matriz de continuidade W.....	16
2.4 Distribuição amostral da estatística I.....	19
2.4.1 Momentos do índice I de Moran.....	20
2.5 Teste de aleatorização para dados espaciais.....	29
2.5.1 Teste de Mantel.....	30
2.6 Avaliação de teste para autocorrelação espacial via simulação Monte Carlo.....	32
2.6.1 Região crítica, poder e o melhor teste.....	32
2.6.2 A eficiência relativa assintótica de testes para autocorrelação espacial.....	36
2.7 Valor P.....	37
3. MATERIAL DE MÉTODOS	39
3.1 Simulação Monte Carlo na avaliação de critérios de testes para dados com autocorrelação espacial.....	39
3.2 Procedimento da análise de autocorrelação em uma população de <i>Copaifera langsdorffii</i> Desf.....	43
4. RESULTADOS E DISCUSSÕES	46
4.1 Sob a hipótese H_0	46
4.2 Sob dependência espacial.....	49
4.3 Uma aplicação com dados reais.....	79
5. CONCLUSÕES	85
REFERÊNCIAS BIBLIOGRÁFICAS	88

RESUMO

LOUZADA, João Marcos. **Avaliação e aplicação de testes para a detecção da autocorrelação espacial usando marcadores genéticos** Lavras: UFLA, 2003. 101 p. (Dissertação - Mestrado em Estatística e Experimentação Agropecuária) *

O índice I de Moran, o qual tem a forma equivalente a de um coeficiente de correlação, é a ferramenta usual para se medir a intensidade da autocorrelação espacial em dados de características genéticas. A estatística I é assintoticamente normalmente distribuída quando o tamanho da amostra cresce, podendo, portanto, ser avaliada como desvios da normal padrão (pressuposto-N). Porém, para pequenos números de populações ($n < 8$), deve-se aplicar o teste de aleatorização (pressuposto-R). Em genética de populações, o teste de aleatorização comumente utilizado foi desenvolvido por Mantel (1967). Assim, por meio do presente trabalho, buscou-se avaliar o desempenho de ambos os testes no que tange às taxas de erro tipo I e o poder dos mesmos. Estes foram avaliados via simulação de Monte Carlo, em que analisaram-se sob a H_0 , as situações $\{p = 0,1, p = 0,25$ e $p = 0,5\}$ de frequências alélicas médias, para as quais consideraram-se os números $\{n = 5, n = 10, m = 25$ e $n = 50\}$ de populações, sendo que variou-se, para cada população, o número de indivíduos em $\{m = 1, m = 2, m = 5$ e $m = 10\}$. Em relação à hipótese alternativa (com padrão espacial), além das mesmas situações simuladas em H_0 , avaliou-se o comportamento destes critérios de testes segundo a variação da amplitude na frequência alélica média local em $\{A = 0,1, A = 0,2, A = 0,5, A = 0,8$ e $A = 1,0\}$. Desta forma, pôde-se analisar a performance dos testes em questão, quanto ao grau de variabilidade das frequências alélicas médias geradas sobre uma superfície linear, em função do espaço geográfico, por meio de diferentes inclinações da mesma. Observou-se que a magnitude de n diminuiu a variância do I quando combinada com as ponderações inverso da distância e inverso da distância ao quadrado. Contrariamente, a amplitude (A) tendeu a aumentá-la. O teste da aproximação normal foi considerado melhor com $\{n \leq 10\}$ populações, quando combinado com as ponderações inverso da distância e inverso da distância ao quadrado, com os níveis

*Comitê Orientador: Eduardo Bearzoti - UFLA (Orientador), Dulcinéia de Carvalho.

de significância 1% e 5%. Porém, o mesmo não deve ser aplicado associado com a ponderação vizinho mais próximo. Com $\{n \geq 25\}$, deve-se aplicar o teste de Mantel em qualquer das situações simuladas. Aplicou-se o índice I em dados de frequências alélicas (obtidos por sistema de isoenzimas), com a finalidade de se detectar a presença de padrão espacial entre árvores de *Copaifera langsdorffii* (copaíba). Para tal, utilizaram-se os critérios de testes de Mantel e o da aproximação normal, segundo as ponderações *inverso da distância*, *inverso da distância ao quadrado* e *vizinho mais próximo*, para avaliar a significância da estatística I . Entretanto, por meio dessa metodologia, não se detectou a existência de autocorrelação, a 5% de significância.

ABSTRACT

LOUZADA, João Marcos. **Evaluation and application of tests for the detection of spatial autocorrelation using data of genetics markers.** Lavras: UFLA, 2003. 101 p. (Dissertation - Master Program in Statistics and Agricultural Experimentation) *

The *index of Moran I* is the usual tool to measure the intensity of the spatial autocorrelation in genetic data; it has a form similar to a correlation coefficient. The statistics *I* is asymptotically normally distributed when the size of the sample increases, and it can be compared to standard normal deviates (assumption-N). However, for small numbers of populations ($n < 8$), the randomization test should be applied (assumption-R). In population genetics, the randomization test commonly used was developed by Mantel (1967). This study evaluated the performance of both tests with regard to type *I* error rates and power, using Monte Carlo simulation, under H_0 , considering the situations $\{p=0,1, p=0,25 \text{ and } p=0,5\}$ of mean gene frequencies, and the amounts $\{n=5, n=10, n=25 \text{ and } n=50\}$ of populations, with $\{m=1, m=2, m=5 \text{ and } m=10\}$ individuals per population. The alternative hypothesis (spatial pattern) was also simulated, with the same conditions of H_0 , and also varying the range in the mean gene frequencies in a given population $\{R=0,1, R=0,2, R=0,5, R=0,8 \text{ and } R=1,0\}$. Thus, the performance of the tests under spatial correlation could be evaluated according to the degree of variability of the frequencies as points of a linear surface on the geographical space, that is, varying its inclination. It was observed that the magnitude of *m* reduced the variance of the *I* when combined with the weights inverse of the distance and inverse of the squared distance. The range (*R*) tended to increase it. The normal approximation behaved better with $\{n \leq 10\}$ populations, when combined to the weights inverse of the distance and inverse of the squared distance, with significance levels of 0.01 and 0.05. However, this was not observed with the weight nearest neighbour. With $\{n \geq 25\}$, the Mantel test should be used in any of the simulated situations. The index *I* was applied to data of isoenzymatic marker data, with the purpose of detecting the presence of spatial pattern among trees of *Co-*

*Guidance Committee: Eduardo Bearzoti - UFLA (Major Professor), Dulcinéia de Carvalho.

paifera langsdorffii. The Mantel test and the normal approximation were used, with weights inverse of the distance, inverse of the squared distance and nearest neighbour, to test significance of statistic I . No autocorrelation was detected, at the level 0.05 of significance.

A principal justificativa do uso dos métodos da estatística espacial está na problemática da independência. Dados independentes são uma suposição conveniente, a qual torna mais simples a inferência. Porém, modelos que envolvem dependência estatística são freqüentemente mais realistas (Cressie (1993)). A idéia de que, em fenômenos estruturados espacialmente, as observações próximas tendem a ser mais semelhantes que aquelas mais distantes parece bem razoável, embora, em estudos de propriedades genéticas, possam ocorrer situações em que populações distantes serem mais semelhantes que as mais próximas.

Quando um indivíduo apresenta alguma qualidade que, provavelmente, se faz presente em indivíduos vizinhos, diz-se que o fenômeno exhibe *autocorrelação espacial*. Assim, o conceito de *autocorrelação espacial* está estruturado sobre a idéia de região em geografia, que trata um conjunto de valores $\{x_i\}$ de dados espacialmente localizados como sendo provavelmente relacionados sobre o espaço. Esta técnica foi principalmente utilizada em estudos de geologia por pesquisadores ingleses no início da década de 1970 (Cliff & Ord (1969, 1970, 1971, 1973)), e posteriormente aplicada também em biologia de populações a partir dos trabalhos de Sokal & Jumars, Thistle & Jones (1977) e Sokal & Oden (1978a, 1978b). Em genética de populações, o instrumento que usualmente se utiliza para medir a intensidade da autocorrelação espacial é o *índice I de Moran* (Cliff & Ord (1973)). O *I* de Moran é similar ao coeficiente de correlação de Pearson.

Em estudos de genética de populações é de suma importância detectar padrões espaciais entre populações ou entre indivíduos de uma mesma espécie pertencente a uma mesma comunidade. Segundo Slatkin & Maruyama (1975), padrões es-

paciais de frequências de genes em populações naturais são determinados pela combinação de deriva genética, seleção natural variando espacialmente e fluxo gênico entre populações locais. Esses padrões espaciais poderão subsidiar precisas informações científicas que possivelmente irão auxiliar em pesquisas futuras sobre os processos biológicos, tais como deriva genética, mutação, dispersão e o conhecimento da variabilidade em estruturas genéticas. Esses conhecimentos são úteis na adoção de técnicas de conservação do patrimônio gênico e seleção de indivíduos para fins de melhoramento de uma particular espécie de interesse do pesquisador.

Na abordagem espacial da genética de populações, a análise de autocorrelação espacial é recomendada para detectar processos microevolutivos em populações naturais, conforme demonstraram Sokal & Jacquez (1991). Estes processos são seleção, migração, mutação e isolamento por distância. Isto posto, a análise de autocorrelação espacial tem sido aplicada como ferramenta para se fazer inferências de processos que estão subjacentes nos padrões espaciais observados de variação em frequências alélicas. Esta metodologia foi primeiramente descrita para estudos sobre evolução por Sokal & Oden (1978a, 1978b) e, baseados nestes estudos, Sokal (1979, 1983, 1986) e Sokal & Wartenberg (1981) desenvolveram os procedimentos inferenciais.

Geralmente, os dados de frequências alélicas apresentam as seguintes características, descritas por Piazza, Menozzi & Cavalli-Sforza (1981): (i) possuem distribuição geográfica altamente irregular, sendo algumas áreas significativamente mais estudadas que outras; (ii) estão sujeitos a mensurações inacuradas, principalmente por causa de erros de amostragem; (iii) referem-se a uma população cobrindo uma área normalmente mal definida e não somente um ponto na superfície; (iv) "outliers" estão presentes e sua identificação pode ser de interesse.

Dada a grande importância de se investigar os processos biológicos que atuam em populações naturais, conforme já mencionado, teve-se por objetivo com este

trabalho avaliar o desempenho (poder) dos testes de Mantel e da aproximação normal para detectar a presença de autocorrelação espacial, visto que os mesmos são usualmente utilizados para se testar a significância do I de Moran.

REFERENCIAL TEÓRICO

Everything is related to everything else, but near things are more related than distant things.

Tobler (1970)

Why we should expect independence in spatial observations which are of the slightest intellectual interest or importance in geographic research I cannot imagine. All our efforts to understand spatial pattern, structure, and process have indicated that it is precisely the lack of independence—the interdependence—of spatial phenomena that allows us to substitute pattern, and therefore predictability and order, for chaos and apparent lack of interdependence—of things in time and space.

Gould (1970)

Data of geographic units are tied together like bunches of grapes, not separate like balls in an urn.

Hepple (1978)

2.1 Padrão espacial

Análises de *padrões espaciais* em fenômenos ambientais, mais precisamente em estudos ecológicos e estudos em genética de populações, são realizadas por três métodos básicos: padrões espaciais, variação espacial discreta (látice) e variação espacial contínua (superfície).

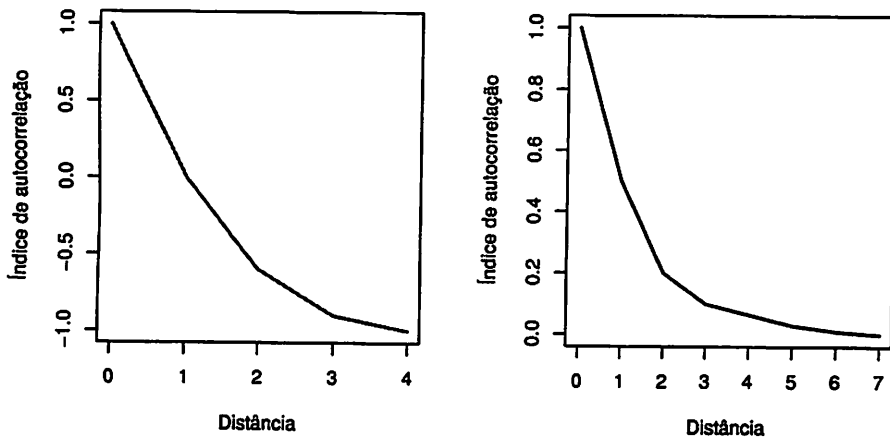
Primeiramente, o estudo de *padrões pontuais* preocupa-se com a distribuição geográfica dos indivíduos, ou seja, padrões pontuais que ocorrem quando a variável de interesse é o local do evento. Este tipo de fenômeno sugere questões básicas, tais como se o padrão apresenta aleatorização espacial, regularidade ou agrupamentos. Assim, o objetivo principal deste tipo de análise é o de descrever o tipo de padrão que o fenômeno apresenta e, então, tentar inferir qual o tipo de processo que pode ter gerado a estrutura observada (Cliff & Ord (1981a), Legendre (1993), Cressie (1993)). Este tipo de fenômeno é mais comumente estudado por ecologistas, como enfatiza Pielou (1977), devido à importância de se vincular os locais relativos dos indivíduos aos estudos sobre agregação entre indivíduos de uma espécie, bem como competição entre espécies.

A análise de *padrão de superfície*, por outro lado, está ligada à descrição de fenômenos que se distribuem de forma contínua sobre o espaço. Aqui, a distribuição espacial da variável de interesse é obtida, usualmente, por amostragem discreta, sendo que cada ponto observado (univariado ou multivariado) representa sua particular porção do espaço geográfico. A análise de superfície possui uma família de métodos que são utilizados para *descrever, estimar e mapear* o padrão

especial existente no fenômeno sob estudo. Assim, a estrutura espacial é primeiramente descrita por *funções de estruturas*, que permitem quantificar a dependência (autocorrelação) em função da distância ou classe de distância. As funções de estrutura mais comumente utilizadas são *semivariograma*, *correlogramas* e o *periodograma* (Legendre & Fortin (1989)). Neste trabalho, será dado enfoque somente em correlogramas e, para informações sobre as outras funções, pode-se consultar Cressie (1993), Isaaks & Srivastava (1989) e Legendre & Fortin (1989).

Define-se correlograma como sendo um gráfico cujos valores dos coeficientes de autocorrelação são plotados na ordenada contra suas respectivas classes de distâncias entre os locais amostrados. Ou, ainda, de modo geral, o correlograma é um espectro que sumaria o grau de relacionamento, que é medido por meio de coeficientes de correlação ou autocorrelação, entre a mesma variável ou variáveis distintas em função da distância; isto permite então descrever como varia a continuidade espacial em relação à distância (Isaaks & Srivastava (1989)). Em áreas contíguas, geralmente, há uma forte associação, esperando-se que esta associação varie de algum modo com a distância (Cliff & Ord (1973)). Então, por meio de um correlograma pode-se ter uma boa noção de como se processa este padrão de continuidade espacial do fenômeno em estudo. A Figura 1 mostra um exemplo de correlograma em forma de gradiente, que comumente se obtém em estudos de genética de populações (Sokal & Oden (1978b), Barbujani (1987), Telles (2001)).

A autocorrelação em uma variável gera um problema estatístico relacionado à habilidade de se realizar testes de hipóteses. O motivo pode ser melhor ilustrado pelo coeficiente de correlação. O problema ocorre quando, por exemplo, ambas as variáveis são positivamente correlacionadas e o intervalo de confiança em torno do coeficiente de Pearson, estimado pelos métodos convencionais são mais estreitos que realmente deveriam ser quando estimado corretamente. Assim, declara-se muito freqüentemente que o coeficiente de correlação seria significativamente diferente de zero sob a hipótese da nulidade, o que poderia não ser bem verdade, conforme ilustra a Figura 2 adaptada de Bivand (1980). Para mais detalhes sobre



(a) Correlograma em plano clinal

(b) Correlograma em variação local

FIGURA 1: Correlogramas espaciais hipotéticos

este assunto ver Cliff & Ord (1981a).

Segundo Legendre (1993), as funções de estruturas (semivariograma, correlograma e periodograma) precisam ser interpretadas e, para tal, cada particular estrutura deve ser *mapeada* de modo a representar a variação espacial da variável em questão ou, em outros casos, representar a variância da variável. Ripley (1981) relata que o modo mais fácil de se obter um mapa de contorno de uma única variável, por exemplo, é utilizar o método de interpolação por médias móveis, denominado inverso da distância ao quadrado. Pode-se utilizar qualquer outro método de interpolação. É importante ressaltar que os antigos métodos de análise de superfície de tendência (Student (1914)), nos quais a variação da variável de interesse é expressada em função das coordenadas geográficas dos pontos amostrados, não produzem mapas precisos, exceto nos casos mais simples. Este método é útil quando, em uma análise, se deseja remover dos seus dados um padrão espacial simples, como, por exemplo, *tendência espacial* ou “*manchas*” (patches) de larga escala, ou porque se procura estudar estruturas espaciais de escala mais

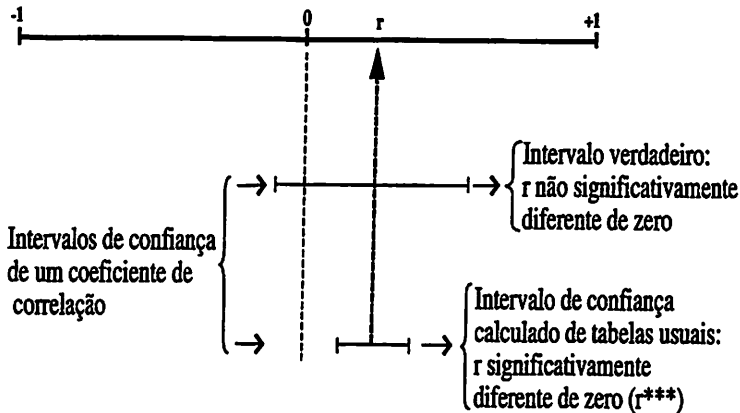


FIGURA 2: Ilustração de testes de coeficientes de correlação para dados positivamente autocorrelacionados, adaptado de Bivand (1980).

reduzida ou porque não há realmente nenhum interesse em se manter estruturas espaciais significativas nos dados.

Segundo Legendre (1993), *gradientes* podem ser vistos como uma estrutura espacial que pode facilmente ser expressa como uma simples função das coordenadas geográficas (x, y) dos pontos amostrados e extraídas dos dados antes da análise propriamente dita ser realizada, satisfazendo então à condição de *estacionariedade de segunda ordem*. A hipótese de estacionariedade de segunda ordem é uma restrição imposta à variável aleatória regionalizada $Z(s)$, em que s representa uma posição (em uma, duas, ou três dimensões), dentro de uma área, S , considerada. Assim, assume-se que o valor esperado (média) da variável aleatória sob estudo deve ter valor finito e constante, conforme definido a seguir.

$$E[Z(s)] = m(s) = m. \quad (2.1)$$

Admite-se também que a função de covariância (ou a função de autocorrelação) deve depender somente do comprimento e orientação do vetor distância (h) entre

quaisquer dois pontos amostrados sobre a área S , de acordo com a seguinte expressão:

$$C(h) = Cov[Z(s), Z(s+h)] = E[Z(s) \cdot Z(s+h)] - m^2, \quad (2.2)$$

para qualquer que seja o valor de s .

Em outras palavras, diz-se que uma função aleatória é estacionária quando não houver tendência na região (detalhes em Cressie (1993)). Este procedimento é requerido em muitos modelos espaciais, principalmente nos modelos de geoestatística, em que a estacionariedade de segunda ordem é tudo que se deseja para que se proceda adequadamente à análise. Porém, em fenômenos nos quais a *variância* não converge, a geoestatística, pode adotar uma hipótese de estacionariedade menos restritiva denominada *intrínseca*. Neste caso, admite-se que $E[Z(s)] = m(s) = m$, para qualquer s e que a variância das diferenças depende somente do vetor distância (h), isto é:

$$Var[Z(s) - Z(s+h)] = E\{[Z(s) - Z(s+h)]^2\} = 2\gamma(h), \quad (2.3)$$

o que implica, supondo $m(s)$ constante, que a variação local das amostras e seu relacionamento espacial podem ser caracterizados pelo semivariograma $2\gamma(h)$.

Legendre (1990) fez a distinção entre gradiente “verdadeiro” ou “falso”, conforme definição a seguir.

Em um gradiente “verdadeiro”, o valor p_{ij} , medido em qualquer posição s no espaço bi-dimensional, pode ser expresso como função de suas coordenadas geográficas x_i e y_j , com $\{i, j = 1, 2, 3, \dots, n\}$, sendo ainda associado o termo erro (ϵ_{ij}), que varia independentemente de um ponto amostral para outro. Assim, pode-se então modelar um gradiente linear por:

$$p_{ij} = \beta_0 + \beta_1 x_i + \beta_2 y_j + \epsilon_{ij}, \quad (2.4)$$

sendo que o termo do erro (ϵ_{ij}) em locais vizinhos não são correlacionados.

Neste tipo de estrutura há violação do pressuposto de estacionariedade na maioria dos métodos de análise porque o valor esperado, dado por $(\beta_0 + \beta_1 x_{ij} + \beta_2 y_{ij})$, varia de local para local como uma função das coordenadas geográficas. Portanto, ele deveria ser removido dos dados antes de proceder, por exemplo, à análise de superfície de tendência.

Legendre (1993) também relata que em um *gradiente "falso"*, a correlação está presente entre locais vizinhos, onde o gradiente como estrutura foi gerado. Assim, o valor esperado (β_0) permanece constante ao longo da superfície, embora o valor observado em cada localidade seja particularmente determinado por valores vizinhos:

$$p_{ij} = \beta_0 + \sum f(p_{i'j'}) + \varepsilon_{ij}. \quad (2.5)$$

Neste modelo, a expressão $\sum f(p_{i'j'})$ representa a soma dos efeitos dos pontos localizados dentro de alguma distância d do valor p_{ij} que se pretende descrever. O somatório é feito sobre todas as classes de distâncias, tal que o valor de interesse é modelado, de certo modo, por alguma função f de todos os outros pontos sobre a superfície. Por simplicidade, deve-se assumir que os pontos estão distribuídos sobre um gride espacial regular. Se um ponto é determinado somente por seus primeiros vizinhos, então a distância d refere-se a uma classe apenas e o processo é chamado "autorregressivo de 1ª ordem". Em situações em que a autocorrelação é positiva, a superfície tem a aparência de uma paisagem oscilando suavemente. Os modelos de isolamento-por-distância, que têm sido amplamente estudados por geneticistas de populações, pertencem a esta classe de modelos autorregressivos de baixa ordem. Se a influência de pontos de dados sobre os outros pontos permanece por muitas classes de distâncias d , em um modo mais ou menos linear (em outras palavras, sendo a função f , por exemplo, o inverso da distância geográfica vezes o valor nestes pontos vizinhos), então a superfície parecerá mais como uma paisagem de montanha.

O termo "*clines*", muito empregado em estudos de genética de populações,

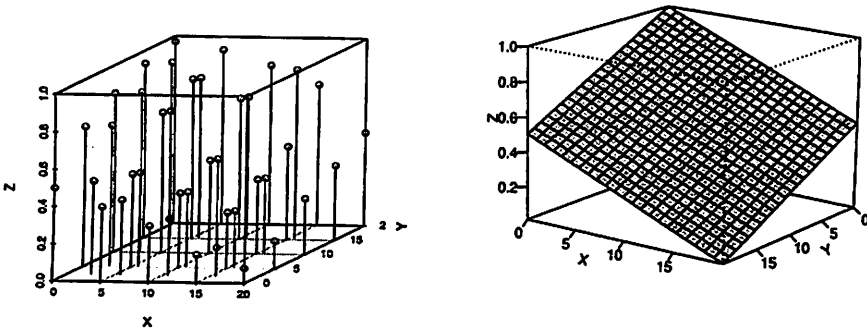


pode ser definido, segundo Endler (1977), como sendo gradientes controlados por frequências de genes ou de qualquer outra característica, que produzirão correlogramas monotonicamente decrescentes a partir de autocorrelações positivas a negativas (Sokal & Oden (1978a)). Embora não se possa garantir que sempre se obterá correlogramas que implique estritamente em um “cline”, deve-se, mesmo assim, descrevê-lo como correlogramas “cliniais”. O termo “cline”, neste contexto, refere-se somente ao padrão observado para as frequências de genes ou deduzido a partir de um correlograma, decrescendo monotonicamente, conforme apresentado na Figura 1a. O que não implica a presença de qualquer processo causativo, tal como *migração* ou *seleção*, por exemplo. Semelhantemente, correlogramas que decrescem a partir de autocorrelações positivas, tendendo assintoticamente à autocorrelação zero, recebem o adjetivo “*patchy*”(manchas), visto que estes descrevem áreas homogêneas cujo tamanho médio é indicado pelo intercepto-x (Figura 1b).

Estudos feitos por Sokal (1979), especificamente em genética de populações, mostraram que o ponto onde o correlograma intercepta o eixo das distâncias (índice de autocorrelação zero) estima a dimensão da mancha (Bjørnstad & Falck (2002)). É sabido que, no que tange à genética de populações, o processo de diferenciação entre populações locais por isolamento-por-distância tende a gerar correlogramas com esse perfil ou seja, autocorrelação positiva, decaindo suavemente em função da distância, até que deixe de ser significativa (Diniz-Filho (1998)).

Conforme Adler & Lauenroth (2000), no contexto de padrões de superfícies, a heterogeneidade espacial refere-se à variação sobre o espaço geográfico dos valores observados de um descritor qualitativo ou quantitativo. Desde que a estrutura e a composição de comunidades de plantas são mensuradas por variáveis contínuas, a heterogeneidade espacial será usada no sentido de padrões de superfícies. Esta heterogeneidade de padrão de superfície pode ser medida pela dependência espacial, o relacionamento entre os valores de uma variável como uma função da distância entre as observações. Dependência espacial mais intensa implica hete-

componentes espaciais, sobre as quais associa-se o eixo Z das frequências alélicas (Figura 3a). Dessa forma, define-se o tipo básico de dados para a análise espacial de diferenciação de populações locais, em que a variável (frequência alélica ou qualquer outra característica de interesse) mensurada em diversos pontos amostrais, cuja posição geográfica é conhecida, pode ser representada por meio de um mapa (Diniz-Filho (1998)). A Figura 3a sugere uma superfície contínua, em forma de um *gradiente de variação*, que pode ser interpolada em relação aos valores de Z , como mostra a Figura 3b. Esta superfície ajustada ilustra razoavelmente bem o comportamento real da variação das frequências alélicas, de modo que este fenômeno dependa apenas da posição geográfica s_i em qualquer parte da região D , tal que $\{s_i \in D\}$.



(a) Variação de frequência alélica

(b) Superfície ajustada

FIGURA 3: Superfície de frequência alélica simulada sobre espaço geográfico com variação clinal.

Diniz-Filho (1998) afirma que quanto maior for o número de amostras coletadas (número de populações locais), torna-se mais fácil perceber que uma superfície de variação pode ser interpolada sobre os valores do eixo- Z . Se essa superfície for bem ajustada, reflete a existência de um *padrão espacial de variação*. Certamente, quanto maior for o número de populações locais, mais robusta será a estimativa do

padrão. É claro que nem sempre esta superfície é necessariamente uma estrutura real, pois as populações locais podem não se distribuir de forma contínua sobre o espaço; porém, deve refletir a *dinâmica espacial de um processo microevolutivo* que atua sobre as frequências alélicas. De fato, a existência de um padrão espacial implica no fato de que é possível encontrar uma explicação biológica única para a variação (ou pelo menos grande parte dessa) nas diversas populações locais.

2.2 O índice *I* de Moran

A estatística espacial desempenha um importante papel na análise de processos biológicos de populações que, geralmente, ocorrem em um contexto espacial (Sokal & Wartenberg (1981)). Em genética de populações, estudos de padrões espaciais em estruturas genéticas são analisados geralmente por meio de ferramentas específicas da estatística espacial. Usualmente, utiliza-se um caso especial da estatística de produto cruzado geral denominado *índice I de Moran* (também conhecido como razão de continuidade de Moran), para quantificar esses padrões. Em outras palavras, o índice *I* de Moran mede a intensidade da *autocorrelação espacial* em dados de variáveis contínuas, baseada em *vizinhança de populações locais* (Moran (1950), Sokal & Oden (1978a, 1978b), Cliff & Ord (1973, 1981b)). Para desenvolver este tipo de estudo, é ainda fundamental que se obtenham as posições geográficas de cada indivíduo ou de cada subpopulação local de interesse. Estas posições geográficas podem ser obtidas por meio de sistema de coordenadas cartesianas, medidas de latitude-longitude, ou qualquer sistema de referência arbitrário eficaz. Define-se esta estatística, segundo Moran (1950), como sendo:

$$I = \frac{n \sum_{i=1}^n \sum_{j=1}^n z_i z_j w_{ij}}{S_0 \sum_{i=1}^n z_i^2}, \quad (2.6)$$

em que n é o tamanho da amostra (populações); $z_i = (p_i - \bar{p})$ e $z_j = (p_j - \bar{p})$ são as variáveis frequências alélicas das populações i e j centradas na média; w_{ij} é o elemento da matriz quadrada e simétrica W , com dimensão $n \times n$, a qual expressa

a relação espacial entre as n populações e S_0 é o somatório dos elementos w_{ij} da matriz simétrica de pesos espaciais W .

É notório que a expressão (2.6) tem a forma clássica de qualquer coeficiente de autocorrelação, com o numerador expressando o termo do produto cruzado (covariância) entre quaisquer dois valores p_i e p_j , enquanto o termo do denominador é a medida da variância (Cliff & Ord (1981a)). Assim, a interpretação do índice I é análoga ao coeficiente de correlação de Pearson, sendo, então, suas estatísticas, na maior parte dos casos, restritas ao intervalo $[-1,+1]$ (o índice I de Moran ocasionalmente receberá valores fora do intervalo $[-1,+1]$, mas, estes são conseqüências de uma matriz de ponderação não usual ou de calcular a média no interior de classes de distâncias). No contexto de genética de populações, valores de I positivos e próximos de $\{1\}$ indicam que há uma forte autocorrelação positiva e conclui-se que localidades vizinhas tendem a ser mais semelhantes que as mais distantes. Contrariamente, para valores de I negativos e próximos de $\{-1\}$, isto indica elevada autocorrelação negativa e que os valores da variável de interesse amostradas em localidades próximas tendem a ser mais distintas, enquanto que as mais distantes são mais similares entre si. Finalmente, valores em torno de zero indicam que não há autocorrelação espacial.

Conforme Diniz-Filho (1998), é importante ressaltar que o índice I , por ser estandardizado, não depende da variância de p_i , sendo função apenas da forma da superfície de frequências alélicas sobre o espaço. Isto implica que superfícies de frequências alélicas em gradientes lineares possuem o mesmo valor do I de Moran, independente de sua inclinação. Essa propriedade do I produz, em dados reais, uma variação residual em p_i , devido a erros de amostragem ou a efeitos locais independentes do gradiente. Além disso, a estandardização do valor do I faz com o mesmo também não dependa da magnitude dos valores dos dados, tornando possível a comparação dos índices encontrados, livre da escala em que foram mensurados.

Embora o índice I assemelhe-se ao índice de Pearson, não deve ser inter-

pretado como uma estimativa direta do parâmetro de correlação estimado; antes, deve-se especificar um particular modelo estocástico, tal como um modelo autorregressivo (Markov espacial) condicional de primeira ordem para habilitar o parâmetro estimado. A estimação de parâmetros para processos espaciais não é trivial e envolve solução de equações não lineares.

De modo geral, o procedimento de quantificação da autocorrelação de unidades espaciais requer que se façam algumas suposições sobre a estrutura subjacente dos dados. Dentre estas suposições, aquela que requer o mínimo do nosso conhecimento a priori do relacionamento entre as observações, está baseada em *continuidade* (Bivand, 1999). Cliff & Ord (1973) descrevem detalhadamente a formalização básica dos relacionamentos entre vizinhos como uma matriz de pesos, usualmente denominada $W \equiv \{w_{ij}\}$, em que denota-se o efeito de um dado objeto espacial i sobre o objeto j pelo peso w_{ij} .

Henebry (1995) comenta que o I de Moran é calculado diretamente em dados de gride regular (látice) para o qual existem sessões padronizadas de uma vizinhança de células. Entretanto, em uma grande maioria de dados reais, não há um sistema de vizinhança padronizado e bem definido. Portanto, faz-se necessário definir os vizinhos e compor a matriz de pesos. O modo como se define esta vizinhança é subjetivo e depende do fenômeno em estudo. A seguir, definir-se-ão os procedimentos da construção da matriz W .

2.3 Construção da matriz de continuidade W

É fundamental, nos procedimentos da análise espacial, definir o *tipo de vizinhos* que será adotado para o conjunto de dados a ser analisado. A composição da matriz de continuidade, W , é bastante flexível, constando de esquemas mais simples, como a medida de conectividade (vizinhança), bem como por meio de esquema de pesos generalizados, tais como o recíproco da distância, o recíproco da distância ao quadrado, ou alguma outra potência: $\{\frac{1}{d}, \frac{1}{d^2}, \dots, \frac{1}{d^r}\}$. Este último es-

quema de peso é mais adequado porque, deste modo, obtêm-se os maiores valores para w_{ij} , no numerador do I , para locais espaciais mais relacionados e, contrariamente, menores valores para w_{ij} , para locais menos relacionados. A matriz de conectividade é composta de pesos binários, com $w_{ij} = 1$, se duas populações estão conectadas espacialmente e $w_{ij} = 0$, caso contrário (Moran (1948)). A segunda alternativa citada, contrariamente à matriz de pesos binários, permite ao pesquisador escolher o esquema de ponderação mais apropriado a partir de informações anteriores. Assim, por exemplo, o pesquisador é capaz de definir esquemas de pesos adequados, a fim de evitar conexão de objetos espaciais que estão isolados por barreiras naturais. Cliff & Ord (1973, p. 12) enfatizam o cuidado que se deve ter na escolha dos pesos, com a finalidade de evitar autocorrelações espúrias. Especificamente em genética de populações, não parece razoável definir populações locais como conectadas em situações em que as mesmas se encontram isoladas ou parcialmente isoladas reprodutivamente por barreiras físicas (montanhas, rios, etc.), embora estejam próximas. Como visto anteriormente, o esquema de pesos binários tem algumas limitações, ou seja, ignora alguns fatos de relevante importância, o que pode acarretar perda de informações preciosas para a análise de fenômenos espaciais. Quando utiliza-se o esquema de pesos generalizados como forma de compor subjetivamente a matriz de vizinhança, o pesquisador tem como incorporar informações não só de vizinhos de primeira ordem (vizinhos mais próximos), como também de segunda ordem ou de ordem maior. Com isso, pode-se analisar a função de autocorrelação decaindo continuamente sobre o espaço, sumariada em espectro (correlogramas), conforme já mencionado anteriormente, o que não se consegue fazer com vizinhos de primeira ordem; às vezes o que interessa é a relação dos vizinho de ordem superior com os de primeira ordem.

Sobre um sistema de célula regular, tipicamente dados “látice”, podem-se definir três formas básicas de influência de vizinhos que, por analogia com o jogo de xadrez, são: “torre”, “bispo” e “rainha”. A influência de vizinhos tipo “torre” ocorre nos eixos principais do látice regular, como nos pontos cardeais – N, S, L,

O; já a influência tipo “bispo” se dá pela rotação de 45° dos eixos principais, ou seja, ND, SU, SO, NO e, finalmente, a influência “rainha”, que é a composição de “torre” e “bispo”, simultaneamente (Cliff & Ord (1973), Henebry (1995)).

Especificamente em genética de populações, em que as amostras se distribuem de forma irregular, é muito comum utilizar a rede de Gabriel para definir os pesos binários $\{w_{ij}\}$. Entende-se que duas populações A e B estão conectadas por meio da rede de Gabriel, se for traçado um círculo hipotético de diâmetro AB e nenhuma outra população pode estar contida dentro deste círculo hipotético (Gabriel & Sokal (1969)). Matematicamente, pode-se dizer que duas populações A e B estão conectadas se:

$$d_{AB}^2 < d_{AC}^2 + d_{BC}^2, \quad (2.7)$$

em que d é a distância e C é uma terceira população local qualquer. A Figura 4 mostra o procedimento da utilização da rede de Gabriel para verificar se duas populações locais quaisquer estão conectadas. Pode-se notar, por exemplo, que as populações $\{C$ e H , B e D , E e F , O e P , ...} estão conectadas.

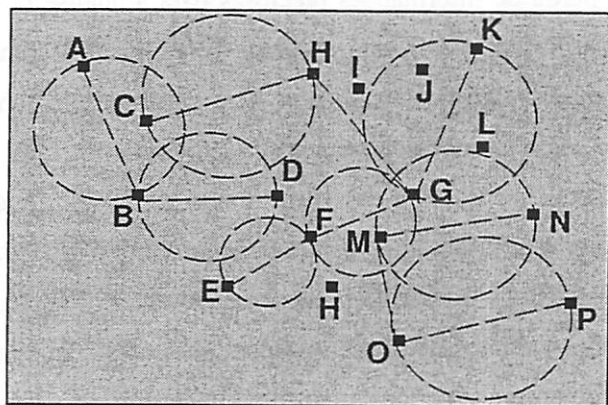


FIGURA 4: Utilização da rede de Gabriel para definir se duas populações locais estão conectadas ou não.

2.4 Distribuição amostral da estatística I

Há duas suposições básicas para avaliar, sob H_0 , a significância dos momentos de I , das quais pode-se assumir:

1. *suposto N* - assume-se que a estrutura dos dados (amostra) observados são realizações de n amostras independentes e identicamente distribuídas, provenientes de uma ou mais populações normais;
2. *suposto R* - admite-se que as observações são realizações aleatórias de n amostras independentes, tomadas de uma ou mais populações cuja função de distribuição é desconhecida. Quando se desconhece a função de distribuição, pode-se considerar que uma particular estrutura de dados é uma permutação aleatória equiprovável de $n!$ possíveis arranjos dessa estrutura de dados. Assim, podem-se avaliar as estatísticas de I sob a hipótese de nulidade de não autocorrelação espacial nas n amostras observadas.

Em ambas as suposições acima, as distribuições amostrais de I têm propriedades *assintoticamente* normais. Portanto, seus parâmetros podem ser avaliados como desvios de uma distribuição normal padrão, contanto que se tomem valores amostrais m grandes. Nos estudos de genética de populações, geralmente considera-se $\{m > 8\}$ um tamanho razoável de amostra para se analisar autocorrelação espacial, admitindo-se que a população de origem segue um distribuição aproximadamente normal – suposto N. Em pesquisas sobre propriedades genéticas, em populações naturais, a estimação dos *momentos de I* é usualmente realizada sob a suposição de aleatorização – suposto R, por ser menos restritivo.

Aleatorização é útil quando se pensa em medidas de autocorrelação espacial, devido ao fato de que, muitas vezes, ou não se obtém a distribuição exata da população sob estudo, ou têm-se problemas quanto ao tamanho da amostra

disponível; isto, conseqüentemente, força o pesquisador a utilizar quaisquer distribuições teóricas aproximadas cabíveis para estimar os parâmetros de interesse. Por essa razão, esse procedimento conduz a estimativas sem acurácia, o que não ocorre com o princípio da aleatorização.

2.4.1 Momentos do índice I de Moran

Segundo Cliff & Ord (1973, p. 29-33), a formalização das estatísticas do índice I de Moran, sob ambas as suposições citadas, é como definidas a seguir:

- seja a variável aleatória normal X com $N(\mu, \sigma^2)$; então pode-se definir $E[X_i] = \mu$, $E[(X_i - \mu)^2] = E[X_i^2] - \{E[X_i]\}^2 = \sigma^2$, $E[(X_i - \mu)^3] = 0$ e $E[(X_i - \mu)^4] = 3\sigma^4$. Desta forma, toma-se uma amostra de tamanho n cujos valores observados foram $\{x_1, x_2, \dots, x_n\}$, e cuja média amostral é:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} \quad (2.8)$$

- sendo ainda a variável Z_i correspondente aos desvios dos valores amostrais observados $z_i = x_i - \bar{x}$, tem-se:

$$\begin{aligned} E[Z_i] &= 0 \\ E[Z_i^2] &= \left(\frac{n-1}{n}\right) \sigma^2 \\ E[Z_i Z_j] &= -\frac{\sigma^2}{n} \\ E[Z_i^2 Z_j^2] &= \frac{(n^2 - 2n + 3)\sigma^4}{n^2} \\ E[Z_i^2 Z_j Z_k] &= -\frac{(n-3)\sigma^4}{n^2} \\ E[Z_i Z_j Z_k Z_l] &= \frac{3\sigma^4}{n^2} \end{aligned} \quad (2.9)$$

Os momentos de I , sob o *suposto* N , está fundamentado no seguinte teorema de Pitman (1937) e Koopmans (1942):

Teorema 1 Se (X_1, X_2, \dots, X_n) são variáveis normais padronizadas independentes e identicamente distribuídas, então qualquer função delas independente de escala, $f(X_1, X_2, \dots, X_n)$, é distribuída independentemente de $Q = \sum_{i=1}^n X_i^2$, a qual tem a forma de uma χ_n^2 .

Prova: seja a função geradora de momentos,

$$\begin{aligned}
 M_X(\theta_1, \theta_2) &= E[e^{\theta_1 f + \theta_2 Q}] \\
 &= E[e^{\theta_1 f + \theta_2 \sum_{i=1}^n X_i^2}] \\
 &= E[e^{\theta_1 f}] E[e^{\theta_2 \sum_{i=1}^n X_i^2}] \\
 &= E[e^{\theta_1 f}] E \left[\prod_{i=1}^n e^{\theta_2 X_i^2} \right]. \tag{2.10}
 \end{aligned}$$

Assim, tem-se que

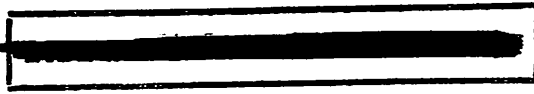
$$\begin{aligned}
 \prod_{i=1}^n E[e^{\theta_2 X_i^2}] &= \int_1^n e^{\theta_2 x^2} dx \\
 &= \int_1^n e^{\theta_2 x^2} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx \\
 &= \int_1^n e^{-\frac{1}{2} \frac{x^2}{(1-2\theta_2)^{-1}}} dx.
 \end{aligned}$$

Deve-se multiplicar pelo seguinte termo para chegar à normal:

$$\frac{(1-2\theta_2)^{-\frac{1}{2}}}{(1-2\theta_2)^{-\frac{1}{2}}}.$$

Então

$$\begin{aligned}
 \prod_{i=1}^n E[e^{\theta_2 x^2}] &= (1-2\theta_2)^{-\frac{1}{2}} \int_1^n \frac{1}{\sqrt{2\pi}(1-2\theta_2)^{-\frac{1}{2}}} e^{-\frac{1}{2} \frac{x^2}{(1-2\theta_2)^{-1}}} dx \\
 &= \left(\frac{1}{1-2\theta_2} \right)^{\frac{1}{2}}. \tag{2.11}
 \end{aligned}$$



Substituindo-se a expressão (2.11) na equação (2.10), obtém-se:

$$M_X(\theta_1, \theta_2) = E[e^{\theta_1 f}] \prod_{i=1}^n \left(\frac{1}{1 - 2\theta_2} \right)^{\frac{1}{2}}$$
$$M_X(\theta_1, \theta_2) = E[e^{\theta_1 f}] \left(\frac{1}{1 - 2\theta_2} \right)^{\frac{n}{2}}. \quad (2.12)$$

Portanto, f e Q são independentes e Q tem uma distribuição $\chi_{(n)}^2$ com n graus de liberdade.

Mesmo que as realizações de Z_i não sejam independentes, elas podem ser transformadas em $(n - 1)$ variáveis normais idênticas e independentes, tal que, pelo teorema de Cochran, $\sum_{i=1}^n Z_i^2$ é distribuído como qui-quadrado com $(n - 1)$ graus de liberdade (Casella & Berger (1990), Kendall & Stuart (1973), Searle (1971)). Assim, do Teorema de Pitman-Koopmans, o I e Q ($=\sum_{i=1}^n Z_i^2$), são *independentes*, de modo que pode-se escrever:

$$E[I^r Q^r] = E[I^r] E[Q^r], \quad (\forall r > 0). \quad (2.13)$$

Então, os momentos de I podem ser encontrados como a seguir.

De acordo com Teorema (1) e utilizando-se a expressão

$$Q = \sum_{i=1}^n z_i^2,$$

pode-se deduzir a seguinte identidade:

$$I \cdot Q = \left(\frac{n}{S_0} \right) \sum_{i=1}^n \sum_{j=1}^n z_i z_j w_{ij}.$$

A partir da qual pode-se, então, estabelecer

$$\begin{aligned}
 E[I^r] &= \frac{E[I^r Q^r]}{E[Q^r]} \\
 &= \left(\frac{n}{S_0}\right)^r E\left[\left(\sum_{i=1}^n \sum_{j=1}^n n Z_i Z_j w_{ij}\right)^r\right].
 \end{aligned} \tag{2.14}$$

Fixando-se $r = 1$ na equação (2.14), pode-se encontrar o 1º momento de I

$$E[I] = \left(\frac{n}{S_0}\right) \frac{E[\sum_{i=1}^n \sum_{j=1}^n n Z_i Z_j w_{ij}]}{E[\sum_{i=1}^n Z_i^2]}. \tag{2.15}$$

Da equação (2.9), temos que o denominador da equação (2.15) segue uma distribuição $\chi_{(n-1)}^2$, com esperança $(n-1)\sigma^2$. E do numerador obtém-se

$$\sum_{i=1}^n \sum_{j=1}^n w_{ij} E[Z_i Z_j] = -\frac{w\sigma^2}{n}.$$

Logo,

$$\begin{aligned}
 E[I] &= \left(\frac{n}{S_0}\right) \frac{-S_0\sigma^2/n}{(n-1)\sigma^2} \\
 E[I] &= -\frac{1}{(n-1)}.
 \end{aligned} \tag{2.16}$$

O estudo do segundo momento de I , sob normalidade, pode ser definido como

$$E[I^2] = \frac{n^2 E[(\sum_{i=1}^n \sum_{j=1}^n z_i z_j w_{ij})^2]}{S_0^2 E[(\sum_i Z_i^2)^2]}. \tag{2.17}$$

A esperança da expressão no denominador na equação (2.17) é

$$(n-1)(n+1)\sigma^4, \tag{2.18}$$

porque, como no primeiro momento, o denominador segue uma distribuição qui-quadrado.

A esperança do termo do numerador, torna-se, então

$$\begin{aligned}
 E \left[\left(\sum_{i=1}^n \sum_{i=1}^n z_i z_j w_{ij} \right)^2 \right] &= \sum_{i=1}^n \sum_{j=1}^n w_{ij} (w_{ij} + w_{ji}) Z_i^2 Z_j^2 \\
 &+ \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n (w_{ij} + w_{ji}) (w_{ik} + w_{ki}) Z_i^2 Z_j Z_k \\
 &+ \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \sum_{l=1}^n w_{ij} w_{kl} Z_i Z_j Z_k Z_l. \quad (2.19)
 \end{aligned}$$

Substituindo-se adequadamente os termos da equação (2.9), na expressão (2.19) tem-se

$$\begin{aligned}
 \left(\sum_{i=1}^n \sum_{i=1}^n z_i z_j w_{ij} \right)^2 &= \frac{\sigma^4}{n^2} [(n^2 - 2n + 3) \sum_{i=1}^n \sum_{j=1}^n w_{ij} (w_{ij} + w_{ji}) \\
 &- (n - 3) \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n (w_{ij} + w_{ji}) (w_{ik} + w_{ki}) \\
 &+ 3 \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \sum_{l=1}^n w_{ij} w_{kl}]. \quad (2.20)
 \end{aligned}$$

Os termos da equação (2.20) podem ser reescritos de uma forma conveniente por meio de manipulação algébrica, em que são definidas algumas estatísticas de resumo com o intuito de simplificar a expressão (detalhes em Cliff & Ord (1973, p. 24–26)). Assim,

$$S_1 = \sum_{i=1}^n \sum_{j=1}^n w_{ij} (w_{ij} + w_{ji}) = \frac{1}{2} \sum_{i=1}^n (w_{ij} + w_{ji})^2, \quad (2.21)$$

$$\begin{aligned}
\sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n (w_{ij} + w_{ji})(w_{ik} + w_{ki}) &= \sum_{i=1}^n \sum_{j=1}^n (w_{ij} + w_{ji})(w_{i.} + w_{.j} - 2w_{ij}) \\
&= \sum_{i=1}^n (w_{i.} + w_{.i})^2 - \sum_{i=1}^n \sum_{j=1}^n (w_{ij} + w_{ji})^2 \\
&= S_2 - 2S_1, \tag{2.22}
\end{aligned}$$

sendo:

$$\begin{aligned}
S_2 &= \sum_{i=1}^n (w_{i.} + w_{.i})^2, \\
w_{i.} &= \sum_{j=1}^n w_{ij} \text{ a soma da linha da matriz } W \text{ de pesos,} \\
w_{.j} &= \sum_{i=1}^n w_{ij} \text{ a soma da coluna de } W.
\end{aligned} \tag{2.23}$$

E, finalmente, define-se a seguinte identidade:

$$\begin{aligned}
\sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \sum_{l=1}^n w_{ij} w_{kl} &= \left(\sum_{i=1}^n \sum_{j=1}^n w_{ij} \right)^2 - \sum_{i=1}^n (w_{i.} + w_{.i})^2 \\
&= + \sum_{i=1}^n \sum_{j=1}^n w_{ij} (w_{ij} (w_{ij} + w_{ji})) \\
&= S_0^2 - S_2 + S_1, \tag{2.24}
\end{aligned}$$

tal que:

$$S_0 = \sum_{i=1}^n \sum_{j=1}^n w_{ij}. \tag{2.25}$$

Retomando-se à expressão (2.20) pela substituição adequada das estatísticas de resumo definidas nas equações (2.21), (2.22) e (2.24), a expressão é reduzida a

$$\frac{\sigma^4}{n^2} [(n^2 - 2n + 3)S_1 - (n - 3)(S_2 - 2S_1) + 3(S_0^2 - 2S_2 + S_1)]$$

$$\frac{\sigma^4}{n^2} [n^2 S_1 - nS_2 + 3S_0^2] \quad (2.26)$$

Finalmente, substituindo-se ambas as equações (2.18) e (2.26) na expressão (2.17), conclui-se que

$$E[I^2] = \frac{n^2 \frac{\sigma^4}{n^2} [n^2 S_1 - nS_2 + 3S_0^2]}{(n - 1)(n + 1)\sigma^4}$$

$$Var[I] = \frac{n^2 S_1 - nS_2 + 3S_0^2}{(n - 1)(n + 1)S_0^2}. \quad (2.27)$$

Os outros momentos de ordem superior podem ser definidos de forma semelhante; contudo, o grau de complexidade aumenta muito rapidamente conforme mostra a expressão do terceiro momento de I :

$$E[I^3] = \frac{-3(n - 1)(n - 5)S_1 + 2n^3 S_3 - 3n^2(S_0 S_1 + S_4) + 3S_0(3nS_2 - 5S_0^2)}{2(n - 1)(n + 1)(n + 3)S_0^3}, \quad (2.28)$$

em que

$$S_3 = \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \sum_{l=1}^n (w_{ij} + w_{ji})(w_{ik} + w_{ki})(w_{jk} + w_{kj})$$

$$S_4 = \sum_{i=1}^n \sum_{j=1}^n (w_{ij} + w_{ji})(w_{i.} + w_{.i})(w_{jk} + w_{kj}). \quad (2.29)$$

Quanto ao *suposto-R*, os momentos de I podem ser avaliados sob a hipótese de nulidade, de não autocorrelação espacial, condicional nos n valores amostrais observados, considerando $n!$ permutações aleatórias igualmente prováveis. Assim, a probabilidade de um determinado evento z_i ocorrer em uma dada unidade

espacial k é dada por

$$\text{prob}(k = z_i) = \frac{1}{n}, \quad (2.30)$$

para todo i e k .

A média observada de Z é

$$\frac{z_1 + z_2 + \dots + z_n}{n} = \sum_i^n z_i = \bar{z}, \quad (2.31)$$

em que, por definição, $E[Z_i] = 0$, sendo esta esperança calculada sobre a extensão das $n!$ permutações aleatórias. De modo semelhante, podem-se definir outras esperanças que serão importantes no processo de avaliação dos momentos de I sobre a hipótese de aleatorização. Entretanto, se uma particular unidade espacial, k , receber a realização z_i com probabilidade $\{\frac{1}{n}\}$, a unidade espacial vizinha pode receber qualquer z_j , com uma probabilidade $\{\frac{1}{n-1}\}$. Então, as seguintes esperanças são definidas:

$$\begin{aligned} E[Z_i^2] &= \frac{1}{n} \sum_{i=1}^n z_i^2 = m_2 \\ E[Z_i Z_j] &= \frac{1}{n-1} \sum_{i=1}^n z_i^2 = -\frac{m_2}{n-1} \\ E[Z_i^2 Z_j^2] &= \frac{2m_2^2 - m_4}{n-1} \\ E[Z_i^2 Z_j Z_k] &= \frac{2m_4 - nm_2^2}{(n-1)(n-2)} \\ E[Z_i Z_j Z_k Z_l] &= \frac{3nm_2^2 - 6m_4}{(n-1)(n-2)(n-3)}. \end{aligned} \quad (2.32)$$

Utilizando-se as esperanças apropriadas do item (2.32), o primeiro e o segundo momentos de I sob a aleatorização (suposto-R) podem ser avaliados como segue:

$$\begin{aligned}
E[I] &= \frac{E[n \sum_{i=1}^n \sum_{j=1}^n Z_i Z_j w_{ij}]}{E[S_0 \sum_{j=1}^n Z_j^2]} \\
&= \frac{n \sum_{i=1}^n \sum_{j=1}^n w_{ij} E[Z_i Z_j]}{S_0 \sum_{j=1}^n E[Z_j^2]} \\
&= \frac{n}{S_0} \left(-\frac{m_2 S_0}{n-1} \right) \frac{1}{nm_2} \\
&= \left(-\frac{1}{n-1} \right), \tag{2.33}
\end{aligned}$$

o que é igual ao do suposto-N. Aplicando-se o operador E na expressão em (2.19) e substituindo-se adequadamente as esperanças determinadas em (2.32), obtém-se

$$\begin{aligned}
E\left[\left(\sum_{i=1}^n \sum_{i=1}^n z_i z_j w_{ij}\right)^2\right] &= \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} (w_{ij} + w_{ji})(nm_2^2 - m_4)}{n-1} \\
&+ \frac{\sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n (w_{ij} + w_{ji})(w_{ik} + w_{ki})(2m_4 - nm_2^2)}{(n-1)(n-2)} \\
&+ \frac{\sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \sum_{l=1}^n w_{ij} w_{kl} (3nm_2^2 - 6m_4)}{(n-1)(n-2)(n-3)}. \tag{2.34}
\end{aligned}$$

A seguir, tomando-se as estatísticas de resumo em (2.21), (2.22) e (2.24), e reorganizando-se os termos convenientemente do lado direito da expressão (2.34), chega-se em

$$\begin{aligned}
&\frac{S_1(nm_2^2 - m_4)}{(n-1)} + \frac{(S_2 - 2S_1)(2m_4 - nm_2^2)}{(n-1)(n-2)} + \frac{(S_0^2 - S_2 + S_1)(3nm_2^2 - 6m_4)}{(n-1)(n-2)(n-3)} \\
Var[I] &= \frac{n[(n^2 - 3n + 3)S_1 - nS_2 + 3S_0^2] - k[(n^2 - n)S_1 - 2nS_2 + 6S_0^2]}{(n-1)(n-2)(n-3)S_0^2}. \tag{2.35}
\end{aligned}$$

Em que

$$k = \frac{m_4}{m_2^2} \text{ é o coeficiente de curtose.} \tag{2.36}$$

2.5 Teste de aleatorização para dados espaciais

Devido a problemas de amostragem, como já mencionado, o *teste de aleatorização ou permutação* tem papel importante na avaliação de fenômenos biológicos, principalmente no contexto espacial, oferecendo mais qualidade na análise e, conseqüentemente, estimativas mais confiáveis. Um outro motivo que envolve a utilização do teste de aleatorização é o da natureza das variáveis que são analisadas no contexto ambiental, situação bem diferente dos experimentos devidamente planejados em que o pesquisador tem mais domínio sobre o fenômeno. Geralmente, estas variáveis apresentam um sincronismo espacial, o que compromete a acurácia dos testes tradicionais, por violação de um dos seus pressupostos básicos, o de *independência* entre as amostras. Neste tipo de análise, o teste de permutação é uma alternativa bem eficaz.

Aleatorização é um princípio muito simples e, para idealizá-la melhor, imagine-se uma dada região, em que são tomados n locais georreferenciados. Em cada local quantifica-se uma dada variável, tal que o conjunto de valores obtidos são $\{z_1, z_2, \dots, z_n\}$. Com a finalidade de ilustrar este fato, considere-se uma amostra com $n = 4$ locais ordenados em linhas e colunas, como descrito a seguir.

$$\begin{array}{|c|c|} \hline z_1 & z_2 \\ \hline z_3 & z_4 \\ \hline \end{array} \equiv \begin{array}{|c|c|} \hline 1 & 2 \\ \hline 3 & 4 \\ \hline \end{array}$$

Estes valores podem ser arranjados de $4! = 24$ maneiras diferentes, conforme mostra tabela a seguir, adaptada de Gumpertz (1999):

1 2 3 4	1 2 4 3	1 3 4 2	1 3 2 4	1 4 2 3	1 4 3 2	2 1 3 4	2 1 4 3
2 3 1 4	2 3 4 1	2 4 1 3	2 4 3 1	3 1 2 4	3 1 4 2	3 2 1 4	3 2 4 1

3 4	3 4	4 1	4 1	4 2	4 2	4 3	4 3
1 2	2 1	2 3	3 2	1 3	3 1	1 2	2 1

Uma forma bem simples de verificar se um particular arranjo dos valores $\{1, 2, 3, 4\}$ acima foi provavelmente designado ao acaso, seria a de calcular, para cada possível arranjo, uma *estatística de resumo* qualquer, como por exemplo, a *semivariância*, o *coeficiente de Pearson*, a *estatística C de Geary*, ou o *índice I de Moran*. Desse modo, compara-se o valor “real” calculado a partir da amostra observada com a distribuição, aleatorizada, de uma dessas estatísticas de resumo de interesse do pesquisador. Se o valor observado for extremo, então não há razões para duvidar de que é improvável que estes valores estejam dispostos de forma aleatória. Em outras palavras, há motivos para crer que exista um sincronismo espacial (autocorrelação espacial).

2.5.1 Teste de Mantel

O teste de Mantel (1967) é um procedimento de aleatorização que visa avaliar o relacionamento geral entre a *distância* e a *similaridade* entre os n locais amostrados (Koenig (1999)), em que ambas as medidas de distância e similaridade são representadas por meio de duas matrizes distintas. Este relacionamento é medido por uma estatística satisfatória, tal como o *I de Moran* ou qualquer outra medida de correlação, espacial ou temporal, e a significância desta é estimada pela comparação com a distribuição de todas as estatísticas encontradas por aleatorização dos elementos de uma das matrizes (Manly (1991)). Este método se inicia pela construção de duas matrizes, da seguinte maneira: (i) constrói-se a matriz de distâncias, a qual contém informações de todos os $\{\frac{n(n-1)}{2}\}$ pares de distâncias entre os locais amostrados, e (ii) compõe-se a matriz de autocorrelação, consistindo dos índices que medem o grau de similaridade em todos os pares de locais. Estas matrizes obviamente são quadradas de ordem $(n \times n)$ e simétricas porque as distâncias são simétricas. Devido à simetria, os cálculos podem ser

feitos utilizando-se somente os vetores colunas da parte triangular inferior das matrizes, facilitando, assim, o procedimento de cálculo. A estatística λ de Mantel é formalizada como

$$\Lambda = \sum_i \sum_j d_{ij} s_{ij}, \quad (2.37)$$

sendo d_{ij} e s_{ij} , respectivamente, os elementos das matrizes de distância (D) e similaridade (S) a serem comparadas. No contexto matemático, o teste de Mantel nada mais é do que o produto interno dos vetores colunas da parte triangular inferior das matrizes simétricas D e S .

A idéia do teste de Mantel (1967) é embaralhar os elementos nas linhas e nas colunas de uma destas matrizes, com a finalidade de se quebrar, caso houver, a associação entre elas. Este procedimento deve ser realizado um grande número de vezes, por exemplo 5.000 vezes e, para cada embaralhamento feito, calcula-se uma nova estatística λ de Mantel. Deve-se, primeiramente, ordenar os valores para, então, prosseguir com o processo de permutação aleatória e, posteriormente, os cálculos das estatísticas. Com este procedimento obtém-se uma distribuição empírica destas estatísticas, que são usadas para avaliar a significância das matrizes em questão. Deste modo, se as matrizes tiveram alguma relação, as estimativas de λ obtidas deverão ser menores que o valor “real” λ_0 observado no conjunto de dados.

Assim, por exemplo, ao se fazer P permutações das quais se obtém-se λ_k estatísticas maiores ou iguais ao valor λ_0 , pode-se então estimar a proporção de valores λ_k maiores ou iguais a λ_0 , em relação a todas as possíveis permutações. Com este procedimento gera-se o nível de significância do teste, ou seja, a probabilidade de a estimativa de Mantel ser um produto do acaso.

O uso do teste de Mantel é justificado pelo fato de os dados geralmente serem de natureza observacional. Então, o seu fundamento é baseado na teoria de *amostragem aleatória*, em que os n elementos para os quais as distâncias são calculadas, são amostras aleatórias provenientes de uma grande população de n possíveis valo-

res que podem ser avaliados. Pode-se então entender que cada matriz de distância é uma particular amostra aleatória independente tomada desta população. Neste caso, o processo de amostragem assegura que o total de permutações dos elementos de uma das matrizes sejam equiprováveis em relação ao ordenamento da outra matriz a ser comparada e, portanto, justificando o teste de aleatorização (Manly (1991)). Finalmente, uma outra justificativa para a utilização do teste de Mantel está relacionada ao fato de os elementos de uma matriz de distâncias não serem observações independentes e, por isso, não se podem aplicar os métodos paramétricos tradicionais.

2.6 Avaliação de teste para autocorrelação espacial via simulação Monte Carlo

2.6.1 Região crítica, poder e o melhor teste

Sabe-se que as hipóteses estatísticas são julgadas com base no comportamento de variáveis aleatórias observáveis. Assim, define-se um conjunto de variáveis aleatórias $\{x_1, x_2, \dots, x_n\}$ de um espaço amostral n -dimensional que podem ser representadas como coordenadas de um ponto x . Visto que x é uma variável aleatória, ele possui uma *distribuição de probabilidade* e, portanto, se alguém selecionar uma região qualquer, ω , no espaço amostral Ω , pode-se calcular a probabilidade de o ponto amostral x cair na região ω ($P\{x \in \omega\}$). Dessa forma, diz-se que qualquer hipótese concernente a $P\{x \in \omega\}$ é uma hipótese estatística (Kendall & Stuart (1973)). Alguém pode desejar saber, por exemplo, se: (i) duas distribuições contínuas não especificadas são idênticas; (ii) a autocorrelação presente em um dado fenômeno é igual a zero; (iii) duas populações estão geneticamente isoladas pela distância ou (iv) duas ou mais populações são diferentes geneticamente ou se são meros desvios aleatórios de uma população panmítica. Cada um dos quatro exemplos são hipóteses que envolvem certas propriedades do espaço amostral. Conseqüentemente, cada uma destas se resume em afirmações a respeito do espaço

amostral, as quais podem ser testadas, via comparação, por meio de observação.

O procedimento para testar qualquer hipótese sobre os fundamentos de uma *amostra aleatória* de observações está baseado no fato de se dividir o espaço amostral em duas regiões. Se o ponto amostral x cair em uma dessas regiões, por exemplo ω , deve-se rejeitar a hipótese e, se x cair na região complementar, $\{\Omega - \omega\}$, não se rejeita a hipótese. ω é conhecido como *região crítica* do teste e $\{\Omega - \omega\}$ é denominada região de *não-rejeição*.

Quando se conhece a distribuição de probabilidade da amostra sob a hipótese a ser testada, a qual é denominada hipótese nula, H_0 , é possível selecionar uma dada região ω , dentre uma infinidade delas, de modo que, dada H_0 , a probabilidade de x cair em ω (rejeitar H_0) é dada pelo valor pré-estabelecido α , que pode ser simbolizada por

$$Prob\{x \in \omega \mid H_0\} = \alpha. \quad (2.38)$$

α é denominado o tamanho do teste ou, em outras palavras, o *nível de significância* do teste. É importante observar que, no caso de uma distribuição ser *descontínua*, nem sempre é possível satisfazer (2.38), para qualquer α em um intervalo (0, 1).

Devido ao fato de haver uma infinidade de sub-regiões ω que satisfazem a (2.38), gera-se a problemática, que envolve a teoria de testes de hipóteses, que resume-se em qual dessas regiões deveria ser preferida. Colocando de maneira prática, quais conjuntos de observações são considerados como *favorecendo*, e quais como *desfavorecendo*, uma determinada hipótese. Por causa disso, não basta apenas conhecer as propriedades inerentes à região crítica quando H_0 ocorre. É imprescindível saber o que acontece quando uma outra hipótese também ocorre. Isso implica que, ao se afirmar que uma massa de dados favorece uma dada hipótese, deve-se, primeiro, compará-la a uma hipótese alternativa. Assim, é perfeitamente possível uma amostra de observações ser mais provável para a hipótese original do que para a hipótese alternativa ou, ainda, podem haver fortes evidências que apóiam a hipótese alternativa. Desta forma, o problema dos testes de hipótese está

relacionado com a escolha de uma dessas hipóteses formuladas e a decisão de aceitar ou não a hipótese original dependerá essencialmente das alternativas contra as quais a mesma está sendo testada.

A idéia discutida anteriormente está diretamente ligada com o *poder do teste*. Um teste de hipótese consiste, basicamente, em elaborar a hipótese nula (H_0), a hipótese alternativa, geralmente denominada H_a , com a qual a hipótese original deve ser comparada e a escolha de uma região crítica com um certo α . A seguir, mediante a distribuição de probabilidades amostral sob H_0 , o teste deve ser julgado quanto às propriedades de cada hipótese testada, *falsa* ou *verdadeira*. Assim, ao se decidir entre uma ou outra hipótese estatística pode-se ter tomado uma decisão incorreta. Este fato induz a dois tipos de erros mais comumente encontrados em testes estatísticos e podem ser de dois tipos, conforme definição a seguir.

- o *erro tipo I* é aquele que se comete quando, erradamente, rejeita-se uma hipótese que é verdadeira;
- pode-se erradamente não-rejeitar a hipótese H_0 , quando ela é falsa e incorrer no *erro tipo II*.

A probabilidade de se incorrer no erro tipo *I* é igual ao tamanho da região crítica escolhida, α . E a probabilidade de um erro tipo *II*, depende da hipótese alternativa considerada, e é usualmente denotada por β . Então,

$$\text{Prob}\{\mathbf{x} \in \Omega - \omega \setminus H_a\} = \beta. \quad (2.39)$$

É quase sempre conveniente trabalhar exclusivamente com a região crítica. Por essa razão, é normal se calcular $1-\beta$. Portanto, reescrevendo (2.39) em relação à probabilidade complementar, tem-se

$$\text{Prob}\{\mathbf{x} \in \omega \setminus H_a\} = 1 - \beta. \quad (2.40)$$

Esta probabilidade complementar, $1 - \beta$, é denominada de *poder* do teste da hipótese H_0 contra a hipótese alternativa H_a , a qual é definida como a probabilidade de se rejeitar uma hipótese nula, H_0 , quando de fato ela é falsa. É essencial que se especifique H_a em (2.40), visto que o poder é uma função de H_a .

Sabe-se que não é possível evitar o erro tipo I , mas deve-se mantê-lo em níveis os mais baixos quanto possível (Hoel (1980)). Assim, dado um tamanho amostral (n) fixo, o estatístico frequentemente determina antes o razoável erro que deseja cometer (erro tipo I) e, a partir daí, ele tenta construir seu teste de forma a minimizar a taxa de erro tipo II . Pois, sendo a amostra de tamanho fixo, o erro tipo II tende a aumentar quando se reduz o erro tipo I ; por isso, não se pode fazer o erro tipo I tão pequeno quanto se deseja sem que o erro tipo II sofra proporcionalmente um grande aumento. O que se faz, na prática, é ajustar o erro tipo I até que se encontre satisfatoriamente um equilíbrio entre os dois tipos de erros.

Uma discussão comum seria a de se questionar qual teste ou combinação de testes é melhor. Claro que este questionamento deve estar fundamentado em um critério de avaliação para se definir o que é ser melhor neste contexto. A resposta a este questionamento está relacionada com a escolha de uma região crítica ótima que satisfaça à expressão em (2.38). Então, um critério padrão atribuído a Neyman e Pearson considera uma região ω , tal que seu poder seja tão maior quanto possível. Desta forma, além de se ter controlado a taxa de erro tipo I em α , consegue-se também minimizar o erro tipo II , β . Isso implica em dizer que a região crítica, cujo poder não é menor que aquele de qualquer outra região do mesmo tamanho para testar a hipótese H_0 contra a hipótese alternativa, é chamada de melhor região crítica (MRC), e um teste baseado em uma MRC é denominado um teste mais poderoso (Kendall & Stuart (1973)). Segundo Costa (1992), o poder

de um teste representa a *sensibilidade* da região crítica para perceber e rejeitar uma hipótese falsa.

E, finalmente, é considerado um bom teste aquele que rejeita a hipótese H_0 com uma alta probabilidade, em favor da hipótese alternativa (H_a) quando a hipótese original (H_0) é falsa e, do mesmo modo, não a rejeita também com alta probabilidade quando a mesma é verdadeira.

2.6.2 A eficiência relativa assintótica de testes para autocorrelação espacial

Primeiramente, o modo mais simples de se comparar dois testes é por meio de exame direto de sua *função poder*. Quando um estatístico dispõe de um tamanho amostral razoável, é natural que ele busque um critério de teste eficiente da mesma forma como se busca um estimador eficiente. Se um teste mais poderoso (eficiente) de tamanho α requer n_1 observações para adquirir um certo poder, e um segundo teste, também de tamanho α , necessita de n_2 observações para atingir o mesmo poder contra a hipótese alternativa, pode-se definir a *eficiência relativa* do segundo teste em atingir aquele poder contra aquela hipótese alternativa como sendo $\left\{\frac{n_1}{n_2}\right\}$. Esta medida é, como nos casos de estimação, o recíproco da razão dos tamanhos amostrais exigidos para um dado desempenho, mas, além de esta definição não ser a de uma eficiência relativa assintótica, a mesma não impõe restrições na forma de distribuições amostrais dos testes a serem comparados. Podem-se comparar quaisquer dois testes desta maneira, porque a função poder dos testes leva em conta a distribuição da estatística de teste.

O conceito de eficiência relativa aqui definido não é conciso. É função de três parâmetros, o tamanho do teste (α), a “distância”(em torno de algum parâmetro) entre a hipótese testada e a alternativa, e a amostra (n_1) requerida pela eficiência do teste. Então, seria interessante encontrar uma medida de eficiência sumária. E é claro que se pode alcançar isto impondo algum processo de limitação. Assim,

vê-se a necessidade de impor restrições para os resultados assintóticos.

Uma sugestão para se fazer isso seria a de permitir que os tamanhos amostrais tendam ao infinito e tomar a razão dos poderes dos testes como uma medida de teste eficiente. Porém, imediatamente, encontra-se uma dificuldade, pois se, na avaliação considerada, ambos são testes consistentes de tamanho α contra a classe hipóteses alternativas no problema, segue, por definição, que a função poder de cada um tende a 1 quando o tamanho amostral aumenta. Assim, ao se compararam os testes contra algum valor alternativo fixo de um parâmetro qualquer, θ , por exemplo, ter-se-á sempre a eficiência assim definida tendendo a 1, quando o tamanho da amostra cresce. Esta medida sugerida de eficiência é, portanto, bastante inútil. Geralmente, é fácil verificar-se que a função poder de testes assintoticamente consistentes em n é de valor limitado. Por exemplo, Wald (1941) definiu um teste assintoticamente mais poderoso como aquele cuja função poder não pode ser melhorada quando o tamanho da amostra tende ao infinito. Isto é, o teste é assintoticamente uniformemente mais poderoso (Kendall & Stuart (1973)).

Para se obter uma medida assintoticamente útil da eficiência do teste a partir da eficiência relativa, deve-se considerar a eficiência relativa limite dos testes contra a sequência de hipóteses alternativas, na qual o parâmetro (θ) aproxima-se do valor testado, θ_0 , quando n cresce. Faz-se isto para evitar forçar o poder dos testes a estar próximo de 1.

2.7 Valor P

O valor p é a probabilidade de uma variável (estatística) assumir um valor maior ou igual ao valor observado estritamente ao acaso: $P(z \geq z_{\text{observado}})$. Portanto, no teste de aleatorização em que N valores da estatística de interesse são gerados sobre H_0 e sendo k desses valores, maiores ou iguais à estatística “real” calculada (estatística calculada sem aleatorização), deve-se rejeitar H_0 a um nível percentual $\hat{p} = \{[(k + 1)/(N + 1)] \times 100\}$ se o valor observado da estatística

“real” superar, pelo menos, os $\{N - k\}$ quantis da distribuição de reamostragem ou distribuição empírica amostral dos dados (Cliff & Ord (1973)), sendo este um teste unilateral com $H_0 : \rho = 0$ contra $H_1 : \rho > 0$.

Se $\{(N + 1)\alpha\}$ for um número inteiro, o valor de α será exato. Um ponto crucial é saber quantas permutações devem ser feitas para se obter um nível α de significância razoável. Um consenso geral é o de que, para se obter um $\alpha = 0,05$, são necessárias 1.000 permutações e 5.000 permutações para um nível $\alpha = 0,01$ (Manly (1991)).

3.1 Simulação Monte Carlo na avaliação de critérios de testes para dados com autocorrelação espacial

Avaliou-se a performance dos critérios de teste de *Mantel* e da *aproximação normal* por meio de simulação Monte Carlo, em aplicações simuladas de estudos de genética de populações. A finalidade foi verificar em qual ou em quais situações os testes em questão seriam mais eficientes para detectar a presença de autocorrelação espacial, medida pelo I de Moran, em variáveis de frequências alélicas, bem como avaliar o comportamento dos testes sob a hipótese de nulidade.

Ambos os testes foram avaliados em diferentes condições de padrão espacial, considerando diferentes números de populações $\{n = 5, n = 10, n = 25 \text{ e } n = 50\}$, variando os tamanhos (m) de amostras, com $\{1, 2, 5 \text{ e } 10\}$ indivíduos por população; estabeleçam-se, ainda os valores nominais de significância $\alpha = 5\%$ e $\alpha = 1\%$ para os testes. Deve-se observar que, com $\{m = 1\}$, pode-se interpretar a situação como a de uma única população com m indivíduos. Assim, por exemplo, uma situação com $\{n = 10\}$ populações de tamanho $\{m = 1\}$ indivíduo por população, equivale a uma única população com tamanho $\{m = 10\}$ indivíduos.

A simulação dos dados foi feita como descrito a seguir. Uma superfície linear foi concebida sobre um plano cartesiano, de maneira a representar uma variação clinal ao longo de populações hipotéticas dispostas no espaço geográfico, sendo este, por sua vez, representado pelo plano cartesiano. Sem perda de generalidade, foram admitidos como 0 e 10 os valores mínimo e máximo de cada uma das coordenadas deste plano.

A superfície foi construída sobre este plano de maneira a ter o valor mínimo no ponto do domínio $(0, 0)$ e o valor máximo em $(10, 10)$. Dessa forma, o valor médio da superfície era encontrado nos pontos pertencentes à reta ligando as coordenadas $(0, 10)$ e $(10, 0)$.

Assim, definida uma frequência alélica média μ ao longo das populações, e uma amplitude de variação A (diferença entre o máximo e o mínimo), tem-se a equação da superfície dada por:

$$E(p_{ij}) = \mu - \frac{A}{2} + \frac{A}{20}x_i + \frac{A}{20}y_j, \quad (3.1)$$

com $0 \leq p_{ij} \leq 1$ e sendo x_i e y_i as coordenadas da população i no plano cartesiano. Conforme sugere a notação $E(p_{ij})$, a superfície linear buscou justamente representar a frequência alélica média de uma população na posição (x_i, y_i) do espaço.

Dessa forma, na simulação, cada uma das n populações tinha sua posição aleatoriamente sorteada no plano cartesiano. Substituindo-se as coordenadas da população em questão, obtinha-se $E(p_{ij})$, o qual era utilizado para simular os m indivíduos desta população. Para tanto, admitiu-se uma espécie diplóide (dois alelos por indivíduo). Um número aleatório era gerado para cada alelo de cada indivíduo. Se este número fosse igual ou menor que $E(p_{ij})$, o alelo em questão era considerado presente e, em caso contrário, ausente. Estes números aleatórios eram gerados independentemente, de alelo para alelo, nos diferentes indivíduos. A consequência disto foi a geração de populações panmíticas que guardavam, no entanto, uma dependência espacial entre si.

Para gerar populações sob a hipótese de nulidade (ausência de padrão espacial), bastava atribuir o valor $A = 0$ em (3.1). Dessa forma, obtém-se uma superfície paralela ao plano cartesiano, com uma dada frequência alélica média μ .

Em presença de dependência espacial, as seguintes condições foram avaliadas:

$\mu = 0,5; A = 1,0; 0,8; 0,5; 0,2$ e $0,1$

$\mu = 0,25; A = 0,5; 0,2$ e $0,1$

$\mu = 0,1; A = 0,2$ e $0,1$

Sob H_0 , considerou-se $\mu = 0,5; 0,25$ e $0,1$.

Exclusivamente para os casos $\{n = 5\}$ populações, analisou-se também o tamanho amostral $\{m = 30\}$ indivíduos por população, somente para as situações com frequência alélica média $p = 0,5$ e amplitude de variação na frequência média local $\{A = 1\}$. Assim, considerando-se todas estas combinações entre parâmetros, um total de 211 situações foram abordadas. Para cada uma destas situações, um total de 5.000 amostras diferentes eram simuladas. Eventualmente, se uma amostra aleatória era monomórfica (todos os alelos de todos os indivíduos iguais), esta era descartada, ficando-se apenas com as amostras válidas (polimórficas).

Com o objetivo de se verificar o melhor esquema de peso para compor a matriz de distâncias, W , combinou-se, com cada situação citada, três esquemas de pesos: *inverso da distância (ID)*, *inverso da distância ao quadrado (IDQ)* e o esquema *vizinho mais próximo (VMP)*.

Com cada uma dessas ponderações, em cada uma das 5.000 amostras, o teste de *Mantel* e o teste baseado na *aproximação normal* do I de Moran eram feitos para julgar a hipótese H_0 de ausência de padrão espacial. No caso do primeiro, um total de 2.000 permutações eram feitas. O teste baseado na aproximação normal era realizado utilizando-se o pressuposto R , conforme apresentado no Referencial Teórico, empregando-se a expressão (2.35) para a variância de I . Assim, em cada amostra, e para cada matriz de ponderação, a seguinte estatística era calculada:

$$Z_I = \frac{I - E(I)}{\sqrt{\text{var}(I)}}, \quad (3.2)$$

na qual $E(I)$ é dada por (2.33).

Cliff & Ord (1973) sugerem um fator de correção para $E(I)$, em (3.2), para quando:

$$\frac{2\sqrt{n} - 1}{\sqrt{n}} < \frac{W}{2n} < \frac{2(\sqrt{n} - 1)(2\sqrt{n} - 1)}{n}. \quad (3.3)$$

Embora esta possibilidade estivesse prevista nas rotinas computacionais, a condição (3.3) nunca foi verificada, a qual dificilmente acontece em disposições espaciais aleatórias, como nas simulações realizadas.

Tanto no critério de Mantel quanto na aproximação normal, testes unilaterais à direita eram feitos para julgar a hipótese H_0 de inexistência de padrão espacial.

Ao longo das 5.000 amostras, calculava-se a proporção de amostras com as quais H_0 era rejeitada. Intervalos de confiança exatos também eram calculados para tais proporções. Tais intervalos exatos têm como limite inferior e superior, respectivamente (adaptado de Leemis & Trivedi (1996)):

$$LI = \frac{1}{1 + \frac{(N-y+1)F_{2(N-y+1);2y}}{y}} \quad (3.4)$$

$$LS = \frac{1}{1 + \frac{(N-y)}{(y+1)F_{2(y+1);2(N-y)}}}, \quad (3.5)$$

nos quais $N = 5000$, y é o número de amostras, em uma dada condição, com as quais H_0 foi rejeitada, e F_{v_1, v_2} refere-se ao valor da distribuição F , com v_1 e v_2 graus de liberdade, que deixa um área acima dele igual a $\{\frac{1-\gamma}{2}\}$, para intervalos com coeficiente de confiança igual a γ . Neste estudo, utilizou-se $\gamma = 0,95$.

Tais intervalos de confiança eram especialmente úteis nas condições sob H_0 , nas quais é interessante verificar se a taxa de rejeição de H_0 era estatisticamente igual ao valor nominal de significância, ou seja, se este valor nominal estava contido nos intervalos de confiança. Como níveis de significância, foram utilizados os valores 0,05 e 0,01.

Além das taxas de rejeição de H_0 , o valor médio e a variância do índice I de Moran, ao longo das 5.000 amostras, também eram calculados para cada um dos

três critérios de ponderação.

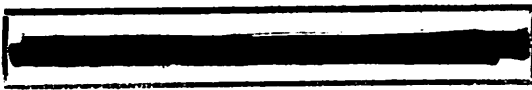
As rotinas computacionais foram desenvolvidas utilizando a linguagem Delphi, versão 6.

3.2 Procedimento da análise de autocorrelação em uma população de *Copaifera langsdorffii* Desf.

Dados de marcadores genéticos de árvores de copaíba foram apurados (Tabela 1) e submetidos à análise de *autocorrelação espacial* por meio do usual índice *I* de Moran. A significância do mesmo foi avaliada paralelamente pelos critérios teste de Mantel e o da aproximação normal, cujas matrizes de distância foram compostas distintamente pelos esquemas de peso inverso da distância, inverso da distância ao quadrado e o vizinho mais próximo. Toda a análise foi realizada por um programa de computador, elaborado em linguagem pascal e implementado em Delphi 6.0.

Apresentaram-se neste trabalho resultados referentes aos marcadores $\alpha - Est - 1$ e $Po - 2$, os quais foram analisados. Os resultados foram dispostos em tabelas distintas para cada particular critério de teste adotado e também ajustaram-se aos dados de frequências alélicas, superfícies interpoladas, as quais foram geradas por meio do software S-plus 2000.

O conjunto de dados para análise foi uma amostra aleatória obtida de uma população de *Copaifera langsdorffii* Desf., popularmente conhecida por copaíba, localizada no campus da Universidade Federal de Lavras (UFLA), Lavras, MG. Trata-se de um fragmento florestal que possui uma área de 5,83 ha, o qual é conhecido como "Matinha da UFLA". A vegetação é classificada como Floresta Estacional Semidecidual Montana (floresta tropical subcadúcifolia) com dossel emergente (Fme), de acordo com o sistema de classificação da Fundação Instituto Brasileiro de Geografia e Estatística - FIBGE (1993). Este fragmento está localizado nas



coordenadas 21°14'40"S. e 44°57'50"W., em uma altitude média de 925 m.

Os dados foram elaborados com base em folhas coletadas de 20 indivíduos da população de copaíba, na qual os indivíduos amostrados foram considerados adultos. Mediram-se também as posições geográficas destes indivíduos por meio de bússola e a distância entre os mesmos por trena, com a finalidade de obter as coordenadas cartesianas destes indivíduos. Este procedimento é conhecido como caminhamento. Então, a amostra coletada foi submetida à caracterização genética da espécie, utilizando-se a técnica de eletroforese de isoenzimas, conforme a metodologia descrita por Alfenas (1991, 1998). Utilizou-se a eletroforese de isoenzima vertical, na qual os extratos foram submetidos à separação em gel de poliacrilamida no sistema descontínuo, concentrado a 12,5% e o gel de concentração a 4,0%.

Toda a análise laboratorial foi realizada no Laboratório de Melhoramento Florestal e Biologia Molecular do Departamento de Ciências Florestais da Universidade Federal de Lavras (DCF-UFLA), Lavras, MG. Assim, o formato do conjunto de dados consta das coordenadas cartesianas (x, y) de cada indivíduo e de oito locos polimorfos do sistema enzimático que apresentaram atividades com possibilidades de interpretação expressas no zimograma. A variável de interesse foi a frequência alélica, sendo que, como se tratava de um estudo entre indivíduos dentro de uma dada população, atribuem-se os valores categóricos 0, 0,5 ou 1,0 para indicar a frequência de um particular alelo de um dado loco em estudo. Assim, por exemplo, suponha que no loco α -Est-1 existam três alelos denominados A_1 , A_2 e A_3 . Então, a frequência alélica relativa ao alelo A_1 , neste loco em particular, para o genótipo A_1A_2 , é 0,5; para o genótipo A_2A_3 é 0(zero) e para o genótipo A_1A_1 é 1,0. Note-se que as frequências são calculadas para cada indivíduo, sendo que considera-se um alelo de cada vez. O conjunto de dados estão disponíveis na Tabela 1.

TABELA 1: Planilha de dados mostrando os genótipos de cada indivíduo quanto aos locos, sendo cada indivíduo associado as coordenadas cartesianas (x, y) .

Ind.	X	Y	$\alpha - Est$		Got-1	Po		Acp-1	Mdh	
			1	2		1	2		1	2
1	163,45	155,47	13	11	22	22	13	11	12	22
2	125,27	157,26	00	23	22	12	23	11	12	22
3	77,48	141,78	12	22	22	00	12	11	22	00
4	77,48	146,08	00	12	22	00	13	11	12	12
5	95,17	133,69	22	22	22	12	13	11	22	00
6	90,55	131,54	00	12	22	00	23	11	12	00
7	100,00	142,80	00	00	11	11	13	11	12	12
8	47,56	109,07	13	22	11	12	33	11	12	12
9	59,76	106,87	00	11	22	11	22	11	12	00
10	56,83	103,72	22	11	22	00	23	11	12	22
11	110,59	102,52	11	12	00	00	33	11	12	12
12	132,70	112,83	00	00	22	11	33	11	12	12
13	135,21	106,86	00	23	00	00	13	11	12	12
14	154,70	81,39	00	00	00	00	23	00	12	12
15	182,28	95,28	11	00	00	11	33	11	12	00
16	194,69	105,69	00	00	00	00	12	11	22	00
17	278,42	78,02	12	22	22	00	23	11	12	22
18	274,53	74,13	22	12	22	00	23	00	12	22
19	266,27	64,04	11	14	00	11	23	00	12	22
20	306,72	2,20	00	33	22	00	23	00	12	12

4.1 Sob a hipótese H_0

As tabelas relativas às taxas de erro tipo I e ao poder dos testes de Mantel e da aproximação normal encontram-se nas Tabelas 2, 3, 4, 5, 6 e 7.

A Tabela 2 apresenta os resultados da simulação de Monte Carlo de amostras em diferentes condições de H_0 , ou seja, ausência de dependência espacial, tendo como fator de ponderação o inverso da distância. Um primeiro aspecto que chama a atenção é o de que os valores médios (ao longo das 5.000 amostras) do índice I de Moran estão muito próximos daqueles esperados conforme a expressão (2.33). Assim, para os valores considerados $\{n=5, 10, 25 \text{ e } 50\}$ tem-se as esperanças de I , sob H_0 , respectivamente iguais a $-0,2500$, $-0,1111$, $-0,0417$ e $-0,0204$, os quais estão relativamente próximos dos valores correspondentes apresentados na Tabela 2.

A magnitude de n teve um efeito observável na variância de I , diminuindo-a com seu aumento. Isto foi um comportamento coerente, pois espera-se que, com maior número de pontos tomados da superfície, o índice seja estimado com maior precisão. Por outro lado, o número de indivíduos por população não apresentou um efeito marcante na variância de I .

Conforme pode-se observar nesta mesma Tabela 2, as taxas de rejeição de H_0 , segundo o critério de Mantel, tenderam a se apresentar estatisticamente diferentes dos níveis nominais de significância; notadamente com $\{n = 5\}$, as taxas foram estatisticamente menores que 0,05, quando este era o nível nominal e, quando o nível nominal era 0,01, as taxas foram um pouco superiores. Nos casos de $\{n=10\}$,

quando significativas, as taxas estiveram abaixo dos níveis nominais. Este fato denota um certo conservadorismo do critério de teste.

Se os pontos amostrais fossem gerados diretamente da superfície clinal de referência, sem nenhuma flutuação aleatória, esperar-se-ia que o teste de Mantel fosse fiel aos níveis nominais. No entanto, como os pontos da superfície correspondiam, na realidade, a frequências *médias* de cada população, as flutuações aleatórias em relação a estas médias devem ter contribuído para realçar o padrão aleatório (sem dependência espacial), dificultando a rejeição de H_0 , e levando a este conservadorismo do teste. Deve-se notar, contudo, que a partir de $\{n = 25\}$, as taxas de rejeição de H_0 foram quase sempre estatisticamente iguais aos valores nominais, indicando um controle adequado do erro tipo I .

Na Tabela 2 estão também apresentadas as taxas de rejeição de H_0 para os testes baseados na aproximação normal. Com $n = 5$ populações, tais testes foram mais conservadores, tanto com níveis nominais iguais a 0,05 como 0,01. Ou seja, neste último nível, o teste promoveu um controle do erro tipo I melhor do que o critério de Mantel, que apresentou taxas superiores a 0,01 com $\{n = 5\}$.

Com valores de n superiores a 5, as taxas dos testes baseadas na normalidade apresentaram uma tendência de serem superiores aos níveis de significância, embora não muito distantes destes. Especialmente nos casos com $\{n = 10\}$ e $\alpha = 0,05$, o teste baseado na aproximação normal foi claramente superior ao critério de Mantel.

Estes resultados sugerem que, com amostras pequenas (iguais ou menores que 10), utilizando-se o inverso da distância como ponderação e com nível de significância 0,05, a aproximação normal é superior ao teste de Mantel, a julgar exclusivamente pelo controle de erro tipo I . Com $\{n = 25\}$ ou mais, o critério de Mantel já apresenta um controle satisfatório deste tipo de erro, devendo ser preferido. A aproximação normal só apresentou, consistentemente, taxas estatisticamente iguais aos valores nominais com $\{n=50\}$ e, mesmo assim com $\alpha = 0,05$.

Independente do critério de teste, foi marcante o efeito pouco pronunciado da frequência alélica média no controle do erro tipo I . Ou seja, as tendências apontadas acima foram muito semelhantes, independentemente da magnitude de μ .

A Tabela 3 apresenta os resultados da simulação de Monte Carlo nas mesmas condições da Tabela 2, só que tendo o inverso da distância ao quadrado como fator de ponderação. Uma vez que tais condições também representavam situações sob H_0 , não se esperavam grandes mudanças de tendências em relação à Tabela 2.

De fato, a diferença mais marcante talvez seja em relação às variâncias do índice I de Moran, que foram expressivamente mais altas, embora os valores de I tenham sido semelhantes. Conforme se nota na expressão (2.6), que define este índice, a soma dos pesos aparece no denominador e, portanto, a variância de I é função do inverso desta soma ao quadrado. Com a ponderação dada pelo inverso da distância ao quadrado, esta soma é menor que aquela correspondente à ponderação dada pelo inverso da distância. Assim, pode ter sido um fator que contribuiu para este aumento na variância.

Outra diferença entre as Tabelas 2 e 3 se deu com o critério de Mantel com $\{n = 5\}$ e $\alpha = 0,05$. Na Tabela 3 houve uma tendência de as taxas serem estatisticamente superiores ao valor nominal 0,05, ao contrário do observado na Tabela 2. Com n igual ou superior a 10, houve uma maior semelhança de tendências, tendo o teste de Mantel também apresentado um controle satisfatório do erro tipo I com n igual ou superior a 25.

Em relação ao teste baseado na aproximação normal, o comportamento também foi semelhante, tendo-se apenas observado, na Tabela 3, um controle estatisticamente diferente dos níveis de significância.

Na Tabela 4 encontram-se os resultados tendo-se utilizado o vizinho mais próximo como fator de ponderação. Pode-se notar que as variâncias do índice I de Moran apresentaram magnitudes intermediárias entre aquelas referentes às

ponderações inverso da distância e inverso da distância ao quadrado.

Quanto ao teste de Mantel, houve uma semelhança grande com as tendências observadas na Tabela 3, ou seja, com $\{n = 5\}$ as taxas de rejeição de H_0 foram superiores aos valores nominais de nível de significância. Com $\{n = 10\}$, tais taxas foram, quando significativas, em geral, inferiores, e, a partir de $\{n = 25\}$, pode-se dizer que o teste de Mantel promoveu um controle satisfatório do erro tipo I .

O mesmo não pode ser dito no tocante à aproximação normal, apresentando taxas muito elevadas de rejeição de H_0 , na grande maioria das situações. Ou seja, a julgar pelos resultados aqui encontrados, se o pesquisador optar pelo vizinho mais próximo como fator de ponderação, então não deve utilizar a aproximação normal como critério de teste, em função de seu desempenho muito insatisfatório de controle do erro tipo I .

É interessante destacar que, nas Tabelas 3 e 4, semelhantemente ao observado na Tabela 2, a frequência alélica média não foi um fator que marcadamente evidenciou alguma tendência digna de nota.

4.2 Sob dependência espacial

A Tabela 5 apresenta os resultados da simulação de Monte Carlo, com ponderação dada pelo inverso da distância, em condições semelhantes àsquelas do item anterior, mas sob dependência espacial, com a amplitude de variação A do plano clinal variando de 0,1 a 1. De imediato, pode-se verificar que, com $A = 0,1$, o índice I de Moran apresentou valores médios próximo aos esperados sob H_0 em que, coerentemente, com o aumento de A houve sempre um aumento correspondente em I . A amplitude também afetou a variância de I , tendendo a aumentá-la, mas também aqui o número de populações n foi o fator mais influente sobre tal variância.

No tocante ao teste de Mantel, houve também uma clara tendência de as taxas de rejeição de H_0 aumentarem com amplitude A . Ou seja, quanto maior a inclinação do plano de variação clinal, maior o poder do teste. Esta foi uma tendência observada com os níveis de significância 0,05 e 0,01.

Para se ter uma idéia das influências relativas que exercem o número de populações n e o número de indivíduos m por população no poder do teste de Mantel, é ilustrativo considerar situações de mesmo tamanho total de amostra, ou seja, nas quais o produto $n \cdot m$ é constante. Assim, tomando-se as situações $\{n = 5, m = 10\}$, $\{n = 10, m = 5\}$, $\{n = 25, m = 2\}$ e $\{n = 50, m = 1\}$, todas com amostra total igual a 50, verifica-se que as taxas de rejeição de H_0 , com $\alpha = 0,05$, e na condição mais extrema de $A = 1$, foram, respectivamente, iguais a 0,1508; 0,3904; 0,6348 e 0,7006. Tendência semelhante foi observada com outros valores de A . Estes resultados mostram claramente que, para se amostrar um mesmo plano de variação clinal em um dado espaço geográfico, é mais vantajoso, com a mesma quantidade de recursos, tomar o maior número possível de pontos, sendo um indivíduo por ponto. Evidentemente, em uma situação real, o pesquisador não estará interessado apenas na detecção de padrões espaciais, mas provavelmente estará interessado em outras inferências intrapopulacionais acerca da estrutura genética. Assim um equilíbrio deve ser buscado entre n e m , conforme os objetivos do estudo.

É relativamente comum que o pesquisador disponha, contudo, de um número pequeno de populações (tipicamente de 1 a 3) e, dessa forma, ele poderia estar interessado na detecção de padrões espaciais intrapopulacionais. Isto corresponde às situações em que se têm n populações com $m = 1$ indivíduos. Em tais casos, pode-se interpretar o valor de m como sendo o número de indivíduos em uma única população.

Assim, na Tabela 5, chamam a atenção os valores de poder muito baixos, com amostras pequenas ($n < 10$). Tomando-se o teste de Mantel com $\alpha = 0,05$, observa-se que, na maioria dos casos, o poder esteve abaixo de 0,1; tendo atin-

gido o valor 0,1458 na situação de maior inclinação do plano de variação clinal. Mesmo com um tamanho de amostra dado por $n = 25$, o que pode ser considerado uma condição relativamente típica, o poder foi muito baixo com valores de A intermediários. Na situação mais extrema ($A = 1$), o poder chegou a 0,3999, com $\alpha = 0,05$. Estes resultados sugerem que, em situações típicas, mesmo que haja considerável dependência espacial entre os indivíduos, o pesquisador terá dificuldades de detectar tal dependência, utilizando o índice I de Moran associado ao teste de Mantel. Este aspecto é ainda agravado se for utilizado o nível de significância 0,01.

Valores desejáveis de poder, como aqueles acima de 0,9, só puderam ser observados com tamanhos totais de amostra grandes, como a partir das combinações $\{n = 25, m = 10\}$ e $\{n = 50, m = 5\}$. Tais tamanhos, no entanto, são pouco freqüentes na prática em genética.

Considerando que o teste baseado na aproximação normal apresentou, em geral, um melhor controle do erro tipo I sob amostras menores ($n \leq 10$), conforme discutido no item anterior, é de especial interesse compará-lo com o teste de Mantel em tais condições. De fato, pode-se observar, na Tabela 5, que o poder da aproximação normal foi quase sempre maior que o teste de Mantel, com $\{n = 5\}$ e $\{m = 10\}$. Exceções foram observadas apenas com $\{n = 5\}$ e $\alpha = 0,01$. Assim, sugere-se que, com amostras pequenas, a aproximação normal deva ser preferida ao critério de Mantel. No entanto, é importante salientar que, mesmo com a aproximação normal, os valores de poder foram marcadamente baixos, sob amostras pequenas.

Conforme pode-se observar na Tabela 5, a freqüência alélica média desempenhou grande influência sobre o poder. A razão disto é simples e se deve ao fato de que quanto mais a freqüência alélica média se aproximar de zero, menor será a amplitude de variação possível entre o maior e o menor valor do plano clinal. Considerando uma distribuição simétrica em torno da média, espera-se um valor de A máximo, com freqüência alélica $p = 0,5$, com a qual a amplitude A pode

chegar a 1. Com $p = 0,25$, o valor máximo de A passa a ser 0,5. Ou seja, se p aproxima-se de zero (ou, se p aproxima-se de 1), a inclinação do plano de variação clinal forçosamente diminui, admitindo-se que este de fato seja uma superfície linear. Considere-se, por exemplo, o teste de Mantel com $\alpha = 0,05$. À exceção de 3 situações, $\{p = 0,25; n = 50; m = 5; A = 0,5\}$, $\{p = 0,25; n = 50; m = 10; A = 0,5\}$ e $\{p = 0,1; n = 50; m = 10; A = 0,2\}$, todos os demais casos com $p=0,25$ ou $p = 0,1$ apresentaram valores de poder abaixo de 0,5. Este é um aspecto que dificulta ainda mais a detecção de dependência espacial, pois nem sempre os locos marcadores apresentam alelos com frequência próxima a 0,5.

A Tabela 6 apresenta os resultados de simulação de Monte Carlo sob dependência espacial com fator de ponderação dado *pelo inverso da distância ao quadrado*. É interessante salientar que, se a pressuposição de que a variação clinal se dá por meio de um plano, então espera-se que a dependência espacial seja uma função *linear* da distância. Assim, as situações apresentadas na Tabela 6 são interessantes para avaliar a robustez da detecção de padrão espacial em face de uma ponderação escolhida que não seja a mais apropriada.

Em termos gerais, tendências semelhantes aos dados da Tabela 5 foram observadas na Tabela 6. Houve também um aumento, tanto no índice I de Moran como nas taxas de poder, ao aumentar n e/ou A , com algumas exceções. Por exemplo, considerando novamente os casos com tamanho total de amostra igual a 50, correspondentes às combinações $\{n = 5; m = 10\}$, $\{n = 10; m = 5\}$, $\{n = 25; m = 2\}$ e $\{n = 50; m = 1\}$, têm-se os valores de poder do teste de Mantel, com $\alpha = 0,05$, $p = 0,5$, e com $A = 1$, respectivamente iguais a 0,1952; 0,2518; 0,3182 e 0,2724. Comparando-se estes valores com aqueles correspondentes da Tabela 5, tem-se que apenas o primeiro foi ligeiramente superior com a ponderação inverso da distância ao quadrado. Também pode-se notar nos valores acima que, com $\{n = 50$ e $m = 1\}$, houve uma queda no poder em relação ao valor com $\{n = 25; m = 2\}$. Estes aspectos provavelmente se devem à pior especificação do fator de ponderação. As variâncias do índice I de Moran também foram, em

geral, superiores sob esta nova ponderação e as razões provavelmente são também aquelas apontadas no item anterior, ou seja, um maior valor S_0 .

Considerando novamente as condições em que $\{m = 1\}$, que guardam uma analogia à situação de um estudo intrapopulacional, nota-se, em geral, um comportamento pior ao observado com a ponderação inverso da distância (Tabela 5), mais coerente com o plano clinal.

A aproximação normal teve uma sensível queda de desempenho com $\{n = 5\}$ com a nova ponderação. No entanto, com $\{n = 10\}$, as taxas de poder foram maiores que às do teste de Mantel e assim como também foi observado um melhor controle do erro tipo I com esta ponderação (item anterior), também aqui pode-se sugerir a aproximação normal com número pequeno de populações, em torno de 10.

É interessante notar que, com a nova ponderação, houve uma sensível redução de poder nos casos com $\{n = 50\}$. Este fato demonstra que, mesmo com as amostras grandes, é importante que haja uma especificação adequada da ponderação a ser utilizada. Esta não é uma questão simples, sugerindo que estudos devam ser conduzidos no sentido de se elucidar como se dá o padrão de variação clinal. Na prática, é pouco provável que ele se dê por meio de uma superfície estritamente linear e, assim, a escolha da ponderação mais adequada não é óbvia.

Da mesma forma que os dados apresentados na Tabela 5 e também na Tabela 6 houve uma forte influência da frequência alélica média p . Se esta se aproxima de zero, a amplitude de variação é forçosamente reduzida, diminuindo o poder. Embora não se tenha estudado situações com p acima de 0,5, espera-se um comportamento semelhante se p aproxima-se de 1, pelas mesmas razões.

A Tabela 7 contém os resultados da simulação sob dependência espacial, utilizando a ponderação *vizinho mais próximo*. Também aqui o índice I de Moran cresceu com o aumento de n e de A , em geral, bem como os valores de poder, mas também aqui houve exceções. Comparando-se as taxas de poder do teste de

Mantel com o produto $n \cdot m$ igual a 50, nota-se uma redução do poder com $n = 50$, da mesma forma como observado com a ponderação inverso da distância ao quadrado. Este é um aspecto estatístico indesejável, pois sugere inconsistência, ou seja, aumento de taxas de erro (no caso, o tipo II) com o aumento do tamanho da amostra, se a ponderação especificada não é mais apropriada.

De maneira geral, pode-se dizer que, no teste de Mantel, a ponderação vizinho mais próximo promoveu valores de poder intermediários entre as ponderações inverso da distância e inverso da distância ao quadrado. Este fato sugere que o vizinho mais próximo seja uma ponderação um pouco mais robusta.

No tocante à aproximação normal, os valores de poder sempre foram consideravelmente superiores com a ponderação vizinho mais próximo, comparando-se a Tabela 7 com as Tabelas 5 e 6. Isto pode sugerir superioridade deste critério, à primeira vista. Porém, uma vez que o controle do erro tipo I com a aproximação normal deixou muito a desejar, conforme discutido no item anterior, realmente não se deve recomendá-la, em absoluto, se o vizinho mais próximo for a ponderação escolhida.

Como considerações gerais, pode-se dizer que a detecção de dependência espacial na estrutura genética de populações, utilizando marcadores moleculares, é algo difícil, requerendo condições ideais que dificilmente são observadas simultaneamente na prática, como amostras muito grandes, frequências alélicas próximas de 0,5 e planos de variação clinal com inclinação pronunciada. Estes aspectos sugerem que, além do índice I de Moran, outros critérios mais poderosos devam ser investigados e utilizados. Uma estratégia complementar que tem sido utilizada por geneticistas de populações de plantas é o uso de correlogramas. Assim, como uma perspectiva de trabalho futuro, seria interessante investigar o controle de erro tipo I e as taxas de poder também sob variadas condições, utilizando simulação de Monte Carlo.

TABELA 2: Taxa de erro tipo I sob hipótese de ausência de autocorrelação espacial, para os testes de Mantel e aproximação normal, em função da frequência alélica (p), número de população (n) e número de indivíduos por população (m), considerando a ponderação *inverso da distância*, sendo os níveis nominais de significância $\alpha = 5\%$ e $\alpha = 1\%$.

Situação			I de Moran		Mantel		Aprox. normal	
p	n	m	Média	var.	$\alpha = 5\%$	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 1\%$
0,5	5	1	-0,2482	0,0463	0,0315*	0,0165*	0,0379*	0,0019*
0,5	5	2	-0,2492	0,0497	0,0280*	0,0164*	0,0329*	0,0008*
0,5	5	5	-0,2540	0,0476	0,0250*	0,0110	0,0314*	0,0000*
0,5	5	10	-0,2574	0,0481	0,0230*	0,0114	0,0276*	0,0006*
0,5	5	30	-0,2470	0,0490	0,0254*	0,0114	0,0308*	0,0004*
0,5	10	1	-0,1102	0,0237	0,0336*	0,0032*	0,0545	0,0186*
0,5	10	2	-0,1142	0,0237	0,0302*	0,0032*	0,0564*	0,0184*
0,5	10	5	-0,1121	0,0250	0,0308*	0,0014*	0,0556	0,0146*
0,5	10	10	-0,1125	0,0247	0,0298*	0,0028*	0,0504	0,0140*
0,5	25	1	-0,0414	0,0093	0,0526	0,0100	0,0624*	0,0230*
0,5	25	2	-0,0425	0,0072	0,0488	0,0094	0,0558	0,0186*
0,5	25	5	-0,0434	0,0068	0,0454	0,0086	0,0552	0,0168*
0,5	25	10	-0,0435	0,0069	0,0456	0,0076	0,0524	0,0158*
0,5	50	1	-0,0194	0,0028	0,0566*	0,0140*	0,0652*	0,0228*
0,5	50	2	-0,0203	0,0026	0,0494	0,0100	0,0532	0,0172*
0,5	50	5	-0,0216	0,0037	0,0512	0,0118	0,0552	0,0208*
0,5	50	10	-0,0209	0,0032	0,0530	0,0108	0,0580*	0,0184*
0,25	5	1	-0,2472	0,0506	0,0363*	0,0225*	0,0492	0,0032*
0,25	5	2	-0,2523	0,0521	0,0327*	0,0183*	0,0409*	0,0016*
0,25	5	5	-0,2507	0,0486	0,0275*	0,0130*	0,0331*	0,0008*
0,25	5	10	-0,2506	0,0499	0,0252*	0,0128	0,0296*	0,0010*
0,25	5	30	-0,2473	0,0487	0,0222*	0,0100	0,0290*	0,0004*

Continua ...

* significa que as taxas de erro tipo I são estatisticamente diferentes dos valores nominais estabelecidos, $\alpha = 5\%$ e $\alpha = 1\%$.

TABELA 2: (Continuação)

Situação			I de Moran		Mantel		Aprox. normal	
p	n	m	Média	var.	$\alpha = 5\%$	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 1\%$
0,25	10	1	-0,1108	0,0273	0,0327*	0,0028*	0,0611	0,0194*
0,25	10	2	-0,1119	0,0281	0,0344*	0,0028*	0,0606*	0,0196*
0,25	10	5	-0,1081	0,0252	0,0314*	0,0036*	0,0534	0,0160*
0,25	10	10	-0,1151	0,0294	0,0268*	0,0026*	0,0482	0,0134*
0,25	25	1	-0,0431	0,0074	0,0448	0,0056*	0,0568	0,0164*
0,25	25	2	-0,0432	0,0083	0,0464	0,0082	0,0562	0,0190*
0,25	25	5	-0,0422	0,0091	0,0516	0,0088	0,0602*	0,0210*
0,25	25	10	-0,0422	0,0080	0,0506	0,0100	0,0582*	0,0180*
0,25	50	1	-0,0198	0,0023	0,0486	0,0078	0,0536	0,0162*
0,25	50	2	-0,0208	0,0036	0,0480	0,0102	0,0548	0,0160*
0,25	50	5	-0,0204	0,0023	0,0470	0,0080	0,0544	0,0160*
0,25	50	10	-0,0214	0,0036	0,0488	0,0104	0,0556	0,0188*
0,1	5	1	-0,2550	0,0420	0,0312*	0,0167*	0,0302*	0,0015*
0,1	5	2	-0,2548	0,0469	0,0309*	0,0170*	0,0363*	0,0036*
0,1	5	5	-0,2566	0,0485	0,0314*	0,0168*	0,0340*	0,0014*
0,1	5	10	-0,2496	0,0494	0,0257*	0,0122	0,0309*	0,0010*
0,1	5	30	-0,2513	0,0506	0,0256*	0,0126	0,0294*	0,0002*
0,1	10	1	-0,1116	0,0183	0,0269*	0,0023*	0,0452	0,0199*
0,1	10	2	-0,1102	0,0244	0,0309*	0,0032*	0,0550	0,0189*
0,1	10	5	-0,1097	0,0247	0,0356*	0,0024*	0,0622*	0,0188*
0,1	10	10	-0,1072	0,0230	0,0398*	0,0044*	0,0622*	0,0196*
0,1	25	1	-0,0409	0,0072	0,0469	0,0090	0,0593*	0,0282*
0,1	25	2	-0,0410	0,0067	0,0442	0,0090	0,0534	0,0188*
0,1	25	5	-0,0389	0,0073	0,0514	0,0080	0,0594*	0,0178*
0,1	25	10	-0,0414	0,0068	0,0478	0,0108	0,0552	0,0206*
0,1	50	1	-0,0203	0,0022	0,0500	0,0114	0,0570	0,0284*
0,1	50	2	-0,0211	0,0024	0,0464	0,0098	0,0510	0,0190*
0,1	50	5	-0,0208	0,0026	0,0470	0,0104	0,0504	0,0174*
0,1	50	10	-0,0206	0,0025	0,0454	0,0098	0,0522	0,0180*
0,1	50	10	-0,0206	0,0025	0,0454	0,0098	0,0522	0,0180*

Final

TABELA 3: Taxa de erro tipo I sob hipótese de ausência de autocorrelação espacial, para os testes de Mantel e aproximação normal, em função da frequência alélica (p), número de população (n) e o número de indivíduos por população (m), considerando a ponderação *inverso da distância ao quadrado*, para os níveis nominais de significância $\alpha = 5\%$ e $\alpha = 1\%$.

Situação			I de Moran		Mantel		Aprox. normal	
p	n	m	Média	var.	$\alpha = 5\%$	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 1\%$
0,5	5	1	-0,2512	0,1812	0,0562	0,0307*	0,0476	0,0004*
0,5	5	2	-0,2479	0,1884	0,0521	0,0275*	0,0388*	0,0006*
0,5	5	5	-0,2570	0,1853	0,0426*	0,0202*	0,0346*	0,0000*
0,5	5	10	-0,2629	0,1855	0,0421*	0,0196*	0,0304*	0,0002*
0,5	5	30	-0,2442	0,1858	0,0456	0,0172	0,0352	0,0002
0,5	10	1	-0,1055	0,1643	0,0330*	0,0028*	0,0539	0,0168*
0,5	10	2	-0,1172	0,1672	0,0296*	0,0024*	0,0508	0,0140*
0,5	10	5	-0,1133	0,1730	0,0268*	0,0010*	0,0480	0,0124
0,5	10	10	-0,1168	0,1581	0,0266*	0,0020*	0,0452	0,0104
0,5	25	1	-0,0374	0,1530	0,0542	0,0096	0,0664*	0,0180*
0,5	25	2	-0,0457	0,1451	0,0450	0,0088	0,0516	0,0178*
0,5	25	5	-0,0489	0,1310	0,0414*	0,0066*	0,0442	0,0158*
0,5	25	10	-0,0484	0,1219	0,0446	0,0098	0,0490	0,0170*
0,5	50	1	-0,0215	0,1202	0,0524	0,0120	0,0652*	0,0180*
0,5	50	2	-0,0190	0,1206	0,0490	0,0094	0,0538	0,0146*
0,5	50	5	-0,0272	0,1161	0,0484	0,0096	0,0488	0,0170*
0,5	50	10	-0,0245	0,1221	0,0484	0,0122	0,0486	0,0186*
0,25	5	1	-0,2446	0,1900	0,0702*	0,0405*	0,0611*	0,0011*
0,25	5	2	-0,2538	0,1904	0,0588*	0,0348*	0,0468	0,0004*
0,25	5	5	-0,2529	0,1856	0,0477	0,0233*	0,0373*	0,0008*
0,25	5	10	-0,2520	0,1880	0,0444	0,0192*	0,0340*	0,0004*

Continua ...

* significa que as taxas de erro tipo I são estatisticamente diferentes os valores nominais estabelecidos, $\alpha = 5\%$ e $\alpha = 1\%$.

TABELA 3: (Continuação)

Situação			I de Moran		Mantel		Aprox. normal	
p	n	m	Média	var.	$\alpha = 5\%$	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 1\%$
0,5	5	30	-0,2418	0,1838	0,0436*	0,0190*	0,0334*	0,0002*
0,25	10	1	-0,1155	0,1760	0,0307*	0,0026*	0,0515	0,0150*
0,25	10	2	-0,1151	0,1848	0,0324*	0,0030*	0,0556	0,0172*
0,25	10	5	-0,1037	0,1635	0,0312*	0,0030*	0,0476	0,0128
0,25	10	10	-0,1225	0,1906	0,0232*	0,0024*	0,0440	0,0116
0,25	25	1	-0,0423	0,1372	0,0444	0,0062*	0,0482	0,0150*
0,25	25	2	-0,0468	0,1505	0,0466	0,0086	0,0494	0,0178*
0,25	25	5	-0,0418	0,1389	0,0504	0,0086	0,0522	0,0174*
0,25	25	10	-0,0421	0,1515	0,0494	0,0084	0,0520	0,0164*
0,25	50	1	-0,0124	0,1130	0,0486	0,0088	0,0430*	0,0132*
0,25	50	2	-0,0180	0,1290	0,0502	0,0104	0,0514	0,0166*
0,25	50	5	-0,0177	0,1180	0,0516	0,0084	0,0530	0,0182*
0,25	50	10	-0,0272	0,1268	0,0498	0,0096	0,0512	0,0188*
0,1	5	1	-0,2614	0,1562	0,0759*	0,0441*	0,0407*	0,0015*
0,1	5	2	-0,2591	0,1778	0,0737*	0,0408*	0,0495	0,0009*
0,1	5	5	-0,2637	0,1847	0,0534	0,0273*	0,0429*	0,0008*
0,1	5	10	-0,2471	0,1862	0,0540	0,0247*	0,0367*	0,0004*
0,5	5	30	-0,2545	0,1828	0,0830*	0,0510*	0,1380*	0,0724*
0,1	10	1	-0,1115	0,1293	0,0336*	0,0059*	0,0393*	0,0187*
0,1	10	2	-0,1066	0,1685	0,0302*	0,0049*	0,0505	0,0195*
0,1	10	5	-0,1125	0,1648	0,0362*	0,0030*	0,0536	0,0164*
0,1	10	10	-0,1008	0,1647	0,0370*	0,0044*	0,0568*	0,0192*
0,1	25	1	-0,0381	0,1409	0,0465	0,0074	0,0493	0,0302*
0,1	25	2	-0,0379	0,1243	0,0444	0,0074	0,0472	0,0190*
0,1	25	5	-0,0321	0,1292	0,0476	0,0068*	0,0492	0,0170*
0,1	25	10	-0,0401	0,1382	0,0462	0,0084	0,0466	0,0172*
0,1	50	1	-0,0188	0,1100	0,4862	0,0122	0,0516	0,0332*
0,1	50	2	-0,0312	0,1183	0,0440	0,0082	0,0442	0,0168*
0,1	50	5	-0,0243	0,1253	0,0474	0,0078	0,0438*	0,0142*
0,1	50	10	-0,0183	0,1087	0,0462	0,0086	0,0428*	0,0156*

Final

TABELA 4: Taxa de erro tipo I sob hipótese de ausência de autocorrelação espacial, para os testes de Mantel e aproximação normal, em função da frequência alélica (p), número de população (n) e número de indivíduos por população (m), considerando a ponderação *vizinho mais próximo*, para os níveis nominais de significância $\alpha = 5\%$ e $\alpha = 1\%$.

Situação			I de Moran		Mantel		Aprox. normal	
p	n	m	Média	var.	$\alpha = 5\%$	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 1\%$
0,5	5	1	-0,2467	0,1861	0,0858*	0,0583*	0,1346*	0,0441*
0,5	5	2	-0,2465	0,1819	0,0926*	0,0649*	0,1418*	0,0621*
0,5	5	5	-0,2503	0,1803	0,0875*	0,0559*	0,1414*	0,0721*
0,5	5	10	-0,2656	0,1772	0,0785*	0,0517*	0,1318*	0,0741*
0,5	5	30	-0,2513	0,0506	0,0256*	0,0126	0,0294*	0,0002*
0,5	10	1	-0,1074	0,1201	0,0671*	0,0164*	0,2477*	0,1865*
0,5	10	2	-0,1112	0,1204	0,0548	0,0106	0,2540*	0,1741*
0,5	10	5	-0,1129	0,1199	0,0338*	0,0072*	0,2672*	0,1892*
0,5	10	10	-0,1146	0,1169	0,0378*	0,0044*	0,2662*	0,1934*
0,5	25	1	-0,0442	0,0570	0,0562	0,0112	0,3352*	0,2936*
0,5	25	2	-0,0431	0,0528	0,0474	0,0090	0,3504*	0,2840*
0,5	25	5	-0,0433	0,0546	0,0474	0,0102	0,3468*	0,2928*
0,5	25	10	-0,0423	0,0549	0,0500	0,0122	0,3450*	0,2940*
0,5	50	1	-0,0156	0,0292	0,0572*	0,0146*	0,4012*	0,3692*
0,5	50	2	-0,0223	0,2822	0,0488	0,0106	0,3844*	0,3406*
0,5	50	5	-0,0206	0,0276	0,0520	0,0100	0,3898*	0,3500*
0,5	50	10	-0,0152	0,0283	0,0556	0,0102	0,4066*	0,3684*
0,25	5	1	-0,2503	0,1887	0,0903*	0,0662*	0,0996*	0,0349*
0,25	5	2	-0,2530	0,1788	0,0858*	0,0608*	0,1183*	0,0414*
0,25	5	5	-0,2481	0,1811	0,0909*	0,0596*	0,1450*	0,0673*
0,25	5	10	-0,2441	0,1878	0,0969*	0,0655*	0,1506*	0,0842*
0,5	5	30	-0,2522	0,1895	0,0442	0,0194*	0,0340*	0,0002

Continua ...

* significa que as taxas de erro tipo I são estatisticamente diferentes dos valores nominais estabelecidos, $\alpha = 5\%$ e $\alpha = 1\%$.

TABELA 4: (Continuação)

Situação			I de Moran		Mantel		Aprox. normal	
p	n	m	Média	var.	$\alpha = 5\%$	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 1\%$
0,25	10	1	-0,1048	0,1250	0,0783*	0,0210*	0,2565*	0,2042*
0,25	10	2	-0,1189	0,1240	0,0590*	0,0156*	0,2522*	0,1907*
0,25	10	5	-0,1063	0,1177	0,0434*	0,0090	0,2639*	0,1959*
0,25	10	10	-0,1105	0,1174	0,0392*	0,0046*	0,2670*	0,2004*
0,25	25	1	-0,0441	0,0546	0,0476	0,0082	0,3395*	0,2867*
0,25	25	2	-0,0449	0,0547	0,0504	0,0112	0,3401*	0,2841*
0,25	25	5	-0,0436	0,0537	0,0470	0,0118	0,3474*	0,2932*
0,25	25	10	-0,0414	0,0542	0,0490	0,0094	0,3484*	0,2908*
0,25	50	1	-0,0201	0,0285	0,0508	0,0108	0,3898*	0,3460*
0,25	50	2	-0,0258	0,0280	0,0456	0,0100	0,3792*	0,3324*
0,25	50	5	-0,0202	0,0270	0,0522	0,0104	0,3912*	0,3530*
0,25	50	10	-0,0214	0,0270	0,0478	0,0096	0,3860*	0,3420*
0,1	5	1	-0,2601	0,1227	0,0519	0,0380*	0,0730*	0,0117
0,1	5	2	-0,2603	0,1581	0,0744*	0,0479*	0,0831*	0,0293*
0,1	5	5	-0,2629	0,1794	0,0915*	0,0613*	0,1232*	0,0452*
0,1	5	10	-0,2423	0,1856	0,1034*	0,0660*	0,1572*	0,0755*
0,5	5	30	-0,2522	0,1811	0,0826*	0,0542*	0,1431*	0,0834*
0,1	10	1	-0,1104	0,0814	0,0550	0,0242*	0,2086*	0,1888*
0,1	10	2	-0,1124	0,1099	0,0668*	0,0219*	0,2357*	0,1889*
0,1	10	5	-0,1034	0,1182	0,0614*	0,0118	0,2664*	0,1933*
0,1	10	10	-0,1037	0,1202	0,0518	0,0084	0,2729*	0,1976*
0,1	25	1	-0,0458	0,0464	0,0549	0,0127	0,2987*	0,2725*
0,1	25	2	-0,0410	0,0497	0,0458	0,0112	0,3317*	0,2824*
0,1	25	5	-0,0392	0,0531	0,0532	0,0096	0,3566*	0,2991*
0,1	25	10	-0,0446	0,0557	0,0532	0,0120	0,3357*	0,2841*
0,1	50	1	-0,0235	0,0256	0,0470	0,0088	0,3613*	0,3214*
0,1	50	2	-0,0219	0,0279	0,0504	0,0116	0,3790*	0,3411*
0,1	50	5	-0,0210	0,0277	0,0500	0,0126	0,3916*	0,3424*
0,1	50	10	-0,0234	0,0276	0,0470	0,0098	0,3834*	0,3408*

Final

TABELA 5: Taxa de poder dos testes de Mantel e aproximação normal, em função de p , n e m , considerando a ponderação *inverso da distância*, sendo os níveis nominais de significância $\alpha = 5\%$ e $\alpha = 1\%$.

Situação				I de Moran		Mantel		Aprox. normal	
p	n	m	A	Média	var.	$\alpha = 5\%$	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 1\%$
0,5	5	1	0,1	-0,2531	0,0514	0,0257	0,0157	0,0350	0,0014
0,5	5	1	0,2	-0,2478	0,0470	0,0345	0,0178	0,0427	0,0023
0,5	5	1	0,5	-0,2264	0,0511	0,0364	0,0206	0,0459	0,0016
0,5	5	1	0,8	-0,1998	0,0515	0,0484	0,0277	0,0722	0,0027
0,5	5	1	1,0	-0,1740	0,0529	0,0623	0,0379	0,0847	0,0049
0,5	5	2	0,1	-0,2503	0,0507	0,0300	0,0165	0,0345	0,0008
0,5	5	2	0,2	-0,2457	0,0521	0,0349	0,0175	0,0436	0,0006
0,5	5	2	0,5	-0,2159	0,0548	0,0435	0,0242	0,0572	0,0016
0,5	5	2	0,8	-0,1757	0,0525	0,0616	0,0307	0,0760	0,0022
0,5	5	2	1,0	-0,1336	0,0544	0,0853	0,0476	0,1054	0,0030
0,5	5	5	0,1	-0,2458	0,0481	0,0309	0,0144	0,0311	0,0010
0,5	5	5	0,2	-0,2300	0,0464	0,0324	0,0154	0,0400	0,0006
0,5	5	5	0,5	-0,1747	0,0517	0,0554	0,0262	0,0754	0,0014
0,5	5	5	0,8	-0,1169	0,0485	0,0900	0,0446	0,1178	0,0030
0,5	5	5	1,0	-0,0794	0,0449	0,1200	0,0598	0,1580	0,0052
0,5	5	10	0,1	-0,2426	0,0485	0,0272	0,0138	0,0324	0,0006
0,5	5	10	0,2	-0,2224	0,0508	0,0324	0,0140	0,0432	0,0006
0,5	5	10	0,5	-0,1329	0,0478	0,0700	0,0344	0,0938	0,0022
0,5	5	10	0,8	-0,0689	0,0443	0,1190	0,0596	0,1674	0,0038
0,5	5	10	1,0	-0,0378	0,0458	0,1508	0,0762	0,2064	0,0066
0,5	10	1	0,1	-0,1093	0,0263	0,0316	0,0040	0,0572	0,0192
0,5	10	1	0,2	-0,1105	0,0259	0,0384	0,0050	0,0613	0,0220
0,5	10	1	0,5	-0,0863	0,0285	0,0500	0,0072	0,0843	0,0282
0,5	10	1	0,8	-0,0434	0,0302	0,0996	0,0142	0,1469	0,0606
0,5	10	1	1,0	-0,0168	0,0318	0,1458	0,0226	0,2114	0,0896
0,5	10	2	0,1	-0,1120	0,0257	0,0346	0,0034	0,0576	0,0180

Continua ...

TABELA 5: (Continuação)

Situação				I de Moran		Mantel		Aprox. normal	
p	n	m	A	Média	var.	$\alpha = 5\%$	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 1\%$
0,5	10	2	0,2	-0,1064	0,0248	0,0366	0,0038	0,0640	0,0190
0,5	10	2	0,5	-0,0630	0,0275	0,0732	0,0072	0,1156	0,0432
0,5	10	2	0,8	-0,0105	0,0331	0,1550	0,0222	0,2222	0,0942
0,5	10	2	1,0	0,0369	0,0301	0,2290	0,0444	0,3134	0,1500
0,5	10	5	0,1	-0,1084	0,0242	0,0376	0,0034	0,0626	0,0174
0,5	10	5	0,2	-0,0919	0,0268	0,0502	0,0060	0,0796	0,0284
0,5	10	5	0,5	-0,0152	0,0295	0,1470	0,0196	0,2086	0,0910
0,5	10	5	0,8	0,0650	0,0312	0,2970	0,0586	0,3798	0,1966
0,5	10	5	1,0	0,1077	0,0296	0,3904	0,0970	0,4874	0,2764
0,5	10	10	0,1	-0,1054	0,0253	0,0356	0,0034	0,0664	0,0190
0,5	10	10	0,2	-0,0800	0,0257	0,0632	0,0074	0,1028	0,0372
0,5	10	10	0,5	0,0338	0,0288	0,2380	0,0404	0,3178	0,1514
0,5	10	10	0,8	0,1215	0,0306	0,4414	0,1190	0,5276	0,3034
0,5	10	10	1,0	0,1513	0,0304	0,5394	0,1664	0,6060	0,3834
0,5	25	1	0,1	-0,0390	0,0078	0,0540	0,0098	0,0666	0,0240
0,5	25	1	0,2	-0,0370	0,0073	0,0586	0,0124	0,0712	0,0238
0,5	25	1	0,5	-0,0120	0,0097	0,1300	0,0366	0,1520	0,0598
0,5	25	1	0,8	0,0294	0,0112	0,2682	0,1088	0,2988	0,1580
0,5	25	1	1,0	0,0617	0,0126	0,3999	0,1976	0,4258	0,2654
0,5	25	2	0,1	-0,0388	0,0084	0,0534	0,0126	0,0652	0,0252
0,5	25	2	0,2	-0,0333	0,0077	0,0622	0,0160	0,0738	0,0296
0,5	25	2	0,5	0,0118	0,0089	0,2052	0,0730	0,2256	0,1138
0,5	25	2	0,8	0,0782	0,0129	0,4618	0,2444	0,4872	0,3198
0,5	25	2	1,0	0,1228	0,0134	0,6348	0,4014	0,6540	0,4850
0,5	25	5	0,1	-0,0351	0,0080	0,2606	0,0124	0,0706	0,0270
0,5	25	5	0,2	-0,0175	0,0081	0,1056	0,0268	0,1172	0,0510
0,5	25	5	0,5	0,0663	0,0120	0,4156	0,1976	0,4398	0,2686
0,5	25	5	0,8	0,1586	0,0137	0,7670	0,5344	0,7820	0,6148

Continua ...

TABELA 5: (Continuação)

Situação				I de Moran		Mantel		Aprox. normal	
p	n	m	A	Média	var.	$\alpha = 5\%$	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 1\%$
0,5	25	5	1,0	0,2030	0,0128	0,8752	0,7040	0,8812	0,7656
0,5	25	10	0,1	-0,0306	0,0081	0,0774	0,0166	0,0888	0,0296
0,5	25	10	0,2	-0,0015	0,0091	0,1574	0,0486	0,1758	0,0818
0,5	25	10	0,5	0,1271	0,0115	0,6692	0,4214	0,6868	0,5092
0,5	25	10	0,8	0,2149	0,0131	0,9008	0,7412	0,9018	0,7974
0,5	25	10	1,0	0,2558	0,0132	0,9498	0,8478	0,9512	0,8868
0,5	50	1	0,1	-0,0180	0,0024	0,0546	0,0104	0,0668	0,0190
0,5	50	1	0,2	-0,0163	0,0026	0,0662	0,0154	0,0778	0,0262
0,5	50	1	0,5	0,0109	0,0036	0,2016	0,0800	0,2212	0,1070
0,5	50	1	0,8	0,0516	0,0041	0,4868	0,2900	0,5090	0,3464
0,5	50	1	1,0	0,0904	0,0056	0,7006	0,5116	0,7186	0,5640
0,5	50	2	0,1	-0,0181	0,0031	0,0610	0,0154	0,0680	0,0226
0,5	50	2	0,2	-0,0102	0,0028	0,0972	0,0268	0,1082	0,0404
0,5	50	2	0,5	0,0362	0,0043	0,3872	0,1918	0,4068	0,2404
0,5	50	2	0,8	0,1025	0,0050	0,7778	0,5758	0,7904	0,6388
0,5	50	2	1,0	0,1491	0,0055	0,9118	0,8058	0,9178	0,8382
0,5	50	5	0,1	-0,0134	0,0026	0,0724	0,0176	0,0800	0,0304
0,5	50	5	0,2	0,0033	0,0030	0,1622	0,0556	0,1720	0,0786
0,5	50	5	0,5	0,0921	0,0050	0,7360	0,5236	0,7452	0,5950
0,5	50	5	0,8	0,1868	0,0058	0,9660	0,9022	0,9672	0,9268
0,5	50	5	1,0	0,2356	0,0063	0,9856	0,9582	0,9848	0,9674
0,5	50	10	0,1	-0,0082	0,0026	0,1040	0,0306	0,1120	0,0452
0,5	50	10	0,2	0,0252	0,0044	0,3068	0,1352	0,3192	0,1800
0,5	50	10	0,5	0,1565	0,0061	0,9336	0,8314	0,9348	0,8732
0,5	50	10	0,8	0,2500	0,0063	0,9916	0,9692	0,9910	0,9790
0,5	50	10	1,0	0,2879	0,0064	0,9928	0,9772	0,9928	0,9838
0,25	5	1	0,1	-0,2451	0,0498	0,0372	0,0216	0,0024	0,0024
0,25	5	1	0,2	-0,2515	0,0487	0,0336	0,0186	0,0021	0,0021
0,25	5	1	0,5	-0,2203	0,0528	0,0486	0,0260	0,0037	0,0037

Continua ...

TABELA 5: (Continuação)

Situação				I de Moran		Mantel		Aprox. normal	
p	n	m	A	Média	var.	$\alpha = 5\%$	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 1\%$
0,25	5	2	0,1	-0,2527	0,0525	0,0316	0,0181	0,0020	0,0020
0,25	5	2	0,2	-0,2388	0,0494	0,0327	0,0179	0,0008	0,0008
0,25	5	2	0,5	-0,2061	0,0524	0,0493	0,0286	0,0028	0,0028
0,25	5	5	0,1	-0,2429	0,0515	0,0321	0,0151	0,0010	0,0010
0,25	5	5	0,2	-0,2299	0,0509	0,0368	0,0182	0,0012	0,0012
0,25	5	5	0,5	-0,1611	0,0490	0,0575	0,0276	0,0022	0,0022
0,25	5	10	0,1	-0,2426	0,0501	0,0302	0,0170	0,0012	0,0012
0,25	5	10	0,2	-0,2140	0,0523	0,0438	0,0212	0,0538	0,0012
0,25	5	10	0,5	-0,1059	0,0473	0,0900	0,0388	0,1238	0,0036
0,25	10	1	0,1	-0,1127	0,0250	0,0341	0,0034	0,0588	0,0197
0,25	10	1	0,2	-0,1085	0,0251	0,0295	0,0022	0,0617	0,0196
0,25	10	1	0,5	-0,0776	0,0286	0,0691	0,0092	0,1106	0,0446
0,25	10	2	0,1	-0,1092	0,0264	0,0368	0,0052	0,0626	0,0204
0,25	10	2	0,2	-0,1001	0,0274	0,0428	0,0052	0,0692	0,0250
0,25	10	2	0,5	-0,0531	0,0297	0,0910	0,0132	0,1400	0,0534
0,25	10	5	0,1	-0,1025	0,0271	0,0404	0,0038	0,0696	0,0226
0,25	10	5	0,2	-0,0871	0,0266	0,0482	0,0040	0,0814	0,0238
0,25	10	5	0,5	0,0097	0,0288	0,1882	0,0326	0,2588	0,1132
0,25	10	10	0,1	-0,0976	0,0259	0,0426	0,0038	0,0706	0,0232
0,25	10	10	0,2	-0,0623	0,0282	0,0800	0,0070	0,1232	0,0468
0,25	10	10	0,5	0,0643	0,0314	0,2924	0,0576	0,3792	0,1896
0,25	25	1	0,1	-0,0386	0,0082	0,0552	0,0124	0,0642	0,0242
0,25	25	1	0,2	-0,0365	0,0076	0,0618	0,0152	0,0724	0,0272
0,25	25	1	0,5	0,0004	0,0092	0,1640	0,0482	0,1810	0,0832
0,25	25	2	0,1	-0,0387	0,0083	0,0594	0,0114	0,0674	0,0272
0,25	25	2	0,2	-0,0304	0,0079	0,0734	0,0176	0,0830	0,0346
0,25	25	2	0,5	0,0273	0,0111	0,2562	0,1006	0,2812	0,1510
0,25	25	5	0,1	-0,0321	0,0078	0,0718	0,0196	0,0800	0,0328
0,25	25	5	0,2	-0,0112	0,0010	0,1186	0,0332	0,1318	0,0614

Continua ...

TABELA 5: (Continuação)

Situação				I de Moran		Mantel		Aprox. normal	
p	n	m	A	Média	var.	$\alpha = 5\%$	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 1\%$
0,25	25	5	0,5	0,0929	0,0125	0,5320	0,2796	0,5552	0,3670
0,25	25	10	0,1	-0,0253	0,0081	0,0870	0,0214	0,1016	0,0350
0,25	25	10	0,2	0,0132	0,0101	0,2102	0,0656	0,2254	0,1122
0,25	25	10	0,5	0,1590	0,0124	0,7782	0,5416	0,7928	0,6258
0,25	50	1	0,1	-0,0197	0,0027	0,0628	0,0152	0,0672	0,0256
0,25	50	1	0,2	-0,0133	0,0028	0,0822	0,0214	0,0892	0,0352
0,25	50	1	0,5	0,0199	0,0039	0,2658	0,1062	0,2852	0,1506
0,25	50	2	0,1	-0,0173	0,0032	0,0584	0,0150	0,0634	0,0260
0,25	50	2	0,2	-0,0080	0,0031	0,1060	0,0334	0,1184	0,0524
0,25	50	2	0,5	0,0528	0,0044	0,4974	0,2604	0,5122	0,3378
0,25	50	5	0,1	-0,0105	0,0029	0,0900	0,0250	0,0980	0,4000
0,25	50	5	0,2	0,0114	0,0031	0,2148	0,0736	0,2290	0,1102
0,25	50	5	0,5	0,1203	0,0051	0,8502	0,6554	0,8550	0,7240
0,25	50	10	0,1	-0,0046	0,0032	0,1226	0,0368	0,1318	0,0560
0,25	50	10	0,2	0,0377	0,0045	0,3842	0,1924	0,3988	0,2464
0,25	50	10	0,5	0,1867	0,0056	0,9656	0,9024	0,9640	0,9282
0,1	5	1	0,1	-0,2466	0,0406	0,0290	0,0168	0,0290	0,0018
0,1	5	1	0,2	-0,2429	0,0412	0,0370	0,0220	0,0370	0,0009
0,1	5	2	0,1	-0,2399	0,0488	0,0371	0,0207	0,0454	0,0016
0,1	5	2	0,2	-0,2338	0,0468	0,0401	0,0218	0,0505	0,0025
0,1	5	5	0,1	-0,2450	0,0508	0,0319	0,0169	0,0404	0,0022
0,1	5	5	0,2	-0,2107	0,0492	0,0441	0,0261	0,0559	0,0018
0,1	5	10	0,1	-0,2257	0,0516	0,0357	0,0176	0,0433	0,0004
0,1	5	10	0,2	-0,1867	0,0524	0,0499	0,0250	0,0641	0,0014
0,1	10	1	0,1	-0,1079	0,0183	0,0332	0,0050	0,0516	0,0241
0,1	10	1	0,2	-0,1003	0,0188	0,0338	0,0054	0,0646	0,0311
0,1	10	2	0,1	-0,1032	0,0241	0,0386	0,0038	0,0646	0,0238
0,1	10	2	0,2	-0,0889	0,0256	0,0561	0,0055	0,0901	0,0360
0,1	10	5	0,1	-0,0944	0,0260	0,0406	0,0032	0,0746	0,0230

Continua ...



TABELA 5: (Continuação)

Situação				I de Moran		Mantel		Aprox. normal	
p	n	m	A	Média	var.	$\alpha = 5\%$	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 1\%$
0,1	10	5	0,2	-0,0574	0,0274	0,0818	0,0120	0,1270	0,0466
0,1	10	10	0,1	-0,0825	0,0261	0,0574	0,0050	0,0868	0,0300
0,1	10	10	0,2	-0,0270	0,0288	0,1288	0,0166	0,1852	0,0712
0,1	25	1	0,1	-0,0374	0,0066	0,0574	0,0092	0,0687	0,0326
0,1	25	1	0,2	-0,0298	0,0086	0,0735	0,0153	0,0844	0,0448
0,1	25	2	0,1	-0,0343	0,0079	0,0632	0,0120	0,0732	0,0284
0,1	25	2	0,2	-0,0161	0,0089	0,1112	0,0266	0,1300	0,0560
0,1	25	5	0,1	-0,0243	0,0083	0,0906	0,0212	0,1020	0,0422
0,1	25	5	0,2	0,0156	0,0101	0,2194	0,0710	0,2316	0,1158
0,1	25	10	0,1	-0,0108	0,0089	0,1316	0,0390	0,1452	0,0698
0,1	25	10	0,2	0,0582	0,0105	0,3932	0,1732	0,4106	0,2500
0,1	50	1	0,1	-0,0166	0,0024	0,0702	0,0170	0,0788	0,0422
0,1	50	1	0,2	-0,0072	0,0039	0,0908	0,0250	0,1060	0,0526
0,1	50	2	0,1	-0,0131	0,0025	0,0788	0,0160	0,0872	0,0310
0,1	50	2	0,2	0,0049	0,0034	0,1664	0,0514	0,1778	0,0884
0,1	50	5	0,1	-0,0039	0,0026	0,1280	0,0366	0,1372	0,0598
0,1	50	5	0,2	0,0400	0,0037	0,4056	0,1898	0,4186	0,2596
0,1	50	10	0,1	0,0114	0,0032	0,2138	0,0744	0,2244	0,1108
0,1	50	10	0,2	0,0844	0,0048	0,7052	0,4638	0,7088	0,5448

Final

TABELA 6: Taxa de poder dos testes de Mantel e aproximação normal, em função de p , n e m , considerando a ponderação *inverso da distância ao quadrado*, sendo os níveis nominais de significância $\alpha = 5\%$ e $\alpha = 1\%$.

Situação				I de Moran		Mantel		Aprox. normal	
p	n	m	A	Média	var.	$\alpha = 5\%$	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 1\%$
0,5	5	1	0,1	-0,2563	0,1937	0,0532	0,0314	0,0451	0,0012
0,5	5	1	0,2	-0,2469	0,1803	0,0601	0,0339	0,0514	0,0012
0,5	5	1	0,5	-0,2085	0,1858	0,0628	0,0348	0,0564	0,0010
0,5	5	1	0,8	-0,1613	0,1874	0,0849	0,0496	0,0791	0,0012
0,5	5	1	1	-0,1156	0,1893	0,0957	0,0556	0,0904	0,0020
0,5	5	2	0,1	-0,2478	0,1918	0,0547	0,0266	0,0401	0,0000
0,5	5	2	0,2	-0,2461	0,1997	0,0611	0,0329	0,0490	0,0004
0,5	5	2	0,5	-0,1911	0,1971	0,0794	0,0391	0,0606	0,0008
0,5	5	2	0,8	-0,1213	0,1887	0,0949	0,0488	0,0778	0,0016
0,5	5	2	1	-0,0467	0,1824	0,1256	0,0648	0,1062	0,0012
0,5	5	5	0,1	-0,2419	0,1821	0,0485	0,0225	0,0361	0,0004
0,5	5	5	0,2	-0,2107	0,1751	0,0583	0,0386	0,0443	0,0002
0,5	5	5	0,5	-0,1163	0,1817	0,0843	0,0394	0,0748	0,0006
0,5	5	5	0,8	-0,0173	0,1632	0,1284	0,0590	0,1114	0,0006
0,5	5	5	1	0,0508	0,1486	0,1578	0,0774	0,1510	0,0018
0,5	5	10	0,1	-0,2373	0,1894	0,0500	0,0212	0,0400	0,0000
0,5	5	10	0,2	-0,2010	0,1892	0,0534	0,0248	0,0468	0,0002
0,5	5	10	0,5	-0,0444	0,1665	0,1028	0,0510	0,0954	0,0012
0,5	5	10	0,8	0,0701	0,1411	0,1624	0,0794	0,1640	0,0016
0,5	5	10	1	0,1209	0,1398	0,1952	0,0972	0,1948	0,0028
0,5	10	1	0,1	-0,1095	0,1813	0,0338	0,0026	0,0522	0,0156
0,5	10	1	0,2	-0,1115	0,1752	0,0356	0,0042	0,0583	0,0182
0,5	10	1	0,5	-0,0658	0,1788	0,0456	0,0070	0,0687	0,0218
0,5	10	1	0,8	0,0181	0,1863	0,0802	0,0090	0,1142	0,0352
0,5	10	1	1	0,0632	0,1964	0,1120	0,0184	0,1448	0,0500
0,5	10	2	0,1	-0,1160	0,1797	0,0354	0,0026	0,0504	0,0148

Continua ...

TABELA 6: (Continuação)

Situação				I de Moran		Mantel		Aprox. normal	
p	n	m	A	Média	var.	$\alpha = 5\%$	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 1\%$
0,5	10	2	0,2	-0,1048	0,1664	0,0340	0,0036	0,0536	0,0132
0,5	10	2	0,5	-0,0189	0,1704	0,0570	0,0058	0,0858	0,0264
0,5	10	2	0,8	0,0736	0,1876	0,1056	0,0130	0,1440	0,0534
0,5	10	2	1	0,1636	0,1673	0,1544	0,0286	0,1990	0,0828
0,5	10	5	0,1	-0,1043	0,1663	0,0324	0,0028	0,0520	0,0162
0,5	10	5	0,2	-0,0801	0,1794	0,0434	0,0052	0,0642	0,0198
0,5	10	5	0,5	0,0653	0,1713	0,1006	0,0142	0,1370	0,0500
0,5	10	5	0,8	0,2130	0,1703	0,1974	0,0292	0,2404	0,0942
0,5	10	5	1	0,2955	0,1573	0,2518	0,0508	0,3078	0,1376
0,5	10	10	0,1	-0,1011	0,1732	0,0330	0,0028	0,0544	0,0158
0,5	10	10	0,2	-0,0554	0,1672	0,0516	0,0058	0,0760	0,0244
0,5	10	10	0,5	0,1511	0,1580	0,1588	0,0256	0,1982	0,0736
0,5	10	10	0,8	0,3212	0,1579	0,2812	0,0606	0,3338	0,1496
0,5	10	10	1	0,3737	0,1529	0,3480	0,0816	0,3898	0,1878
0,5	25	1	0,1	-0,0353	0,1520	0,0488	0,0106	0,0622	0,0216
0,5	25	1	0,2	-0,0275	0,1388	0,0524	0,0108	0,0672	0,0206
0,5	25	1	0,5	0,0220	0,1498	0,0796	0,0192	0,0930	0,0308
0,5	25	1	0,8	0,0996	0,1446	0,1318	0,0420	0,1484	0,0550
0,5	25	1	1	0,1602	0,1534	0,1974	0,0784	0,2122	0,0900
0,5	25	2	0,1	-0,0377	0,1487	0,0476	0,0082	0,0552	0,0168
0,5	25	2	0,2	-0,0270	0,1329	0,0504	0,0094	0,0574	0,0164
0,5	25	2	0,5	0,0635	0,1358	0,1100	0,0282	0,1156	0,0474
0,5	25	2	0,8	0,1972	0,1493	0,2294	0,0792	0,2380	0,1106
0,5	25	2	1	0,2801	0,1453	0,3182	0,1316	0,3224	0,1620
0,5	25	5	0,1	-0,0281	0,1333	0,0488	0,0084	0,0528	0,0198
0,5	25	5	0,2	0,0106	0,1348	0,0694	0,0146	0,0746	0,0256
0,5	25	5	0,5	0,1719	0,1511	0,1964	0,0594	0,1986	0,0882
0,5	25	5	0,8	0,3525	0,1575	0,3964	0,1886	0,3964	0,2338
0,5	25	5	1	0,4417	0,1461	0,4950	0,2714	0,5036	0,3136

Continua ...

TABELA 6: (Continuação)

Situação				I de Moran		Mantel		Aprox. normal	
p	n	m	A	Média	var.	$\alpha = 5\%$	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 1\%$
0,5	25	10	0,1	-0,0245	0,1400	0,0548	0,0120	0,0602	0,0216
0,5	25	10	0,2	0,0310	0,1367	0,0928	0,0180	0,0976	0,0344
0,5	25	10	0,5	0,2878	0,1427	0,3164	0,1282	0,3164	0,1684
0,5	25	10	0,8	0,4565	0,1381	0,5236	0,2826	0,5194	0,3286
0,5	25	10	1	0,5452	0,1473	0,6218	0,3790	0,6168	0,4228
0,5	50	1	0,1	-0,0088	0,1172	0,0540	0,0098	0,0684	0,0180
0,5	50	1	0,2	-0,0157	0,1129	0,0498	0,0084	0,0670	0,0166
0,5	50	1	0,5	0,0475	0,1185	0,0874	0,0256	0,1040	0,0338
0,5	50	1	0,8	0,1228	0,1117	0,1792	0,0660	0,1914	0,0758
0,5	50	1	1	0,2114	0,1267	0,2724	0,1276	0,2854	0,1258
0,5	50	2	0,1	-0,0211	0,1250	0,0524	0,0138	0,0542	0,0210
0,5	50	2	0,2	0,0039	0,1116	0,0650	0,0152	0,0660	0,0250
0,5	50	2	0,5	0,0961	0,1267	0,1364	0,0434	0,1376	0,0564
0,5	50	2	0,8	0,2328	0,1174	0,2958	0,1314	0,2960	0,1548
0,5	50	2	1	0,3286	0,1222	0,4154	0,2200	0,4190	0,2492
0,5	50	5	0,1	0,0035	0,1193	0,0572	0,0110	0,0594	0,0200
0,5	50	5	0,2	0,0287	0,1110	0,0738	0,0174	0,0758	0,0316
0,5	50	5	0,5	0,1982	0,1292	0,2592	0,1006	0,2596	0,1300
0,5	50	5	0,8	0,4019	0,1139	0,5294	0,3016	0,5250	0,3418
0,5	50	5	1	0,5079	0,1300	0,6256	0,4202	0,6238	0,4626
0,5	50	10	0,1	0,0058	0,1087	0,0658	0,0132	0,0686	0,0228
0,5	50	10	0,2	0,0103	0,1266	0,1146	0,0306	0,1130	0,0482
0,5	50	10	0,5	0,3424	0,1340	0,4456	0,2284	0,4444	0,2704
0,5	50	10	0,8	0,5408	0,1342	0,6750	0,4522	0,6686	0,5030
0,5	50	10	1	0,6171	0,1308	0,7458	0,5388	0,7430	0,5900
0,25	5	1	0,1	-0,2408	0,1888	0,0683	0,0364	0,0567	0,0011
0,25	5	1	0,2	-0,2505	0,1874	0,0667	0,0423	0,0547	0,0009
0,25	5	1	0,5	-0,2003	0,1926	0,0877	0,0537	0,0729	0,0011
0,25	5	2	0,1	-0,2544	0,1981	0,0604	0,0332	0,0445	0,0012

Continua ...

TABELA 6: (Continuação)

Situação				I de Moran		Mantel		Aprox. normal	
p	n	m	A	Média	var.	$\alpha = 5\%$	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 1\%$
0,25	5	2	0,2	-0,2287	0,1853	0,0608	0,0321	0,0484	0,0004
0,25	5	2	0,5	-0,1759	0,1926	0,0833	0,0438	0,0703	0,0010
0,25	5	5	0,1	-0,2391	0,1903	0,0512	0,0235	0,0366	0,0008
0,25	5	5	0,2	-0,2137	0,1914	0,0579	0,0266	0,0429	0,0006
0,25	5	5	0,5	-0,0965	0,1740	0,1019	0,0497	0,0873	0,0010
0,25	5	10	0,1	-0,2370	0,1892	0,0560	0,0280	0,0432	0,0004
0,25	5	10	0,2	-0,1868	0,1877	0,0626	0,0280	0,0558	0,0002
0,25	5	10	0,5	0,0051	0,1597	0,1266	0,0606	0,1216	0,0018
0,25	10	1	0,1	-0,1195	0,1726	0,0353	0,0050	0,0505	0,0164
0,25	10	1	0,2	-0,1073	0,1710	0,0309	0,0026	0,0525	0,0156
0,25	10	1	0,5	-0,0499	0,1851	0,0599	0,0072	0,0872	0,0297
0,25	10	2	0,1	-0,1079	0,1814	0,0352	0,0046	0,0574	0,0166
0,25	10	2	0,2	-0,0887	0,1811	0,0406	0,0028	0,0610	0,0190
0,25	10	2	0,5	-0,0068	0,1812	0,0718	0,0080	0,0954	0,0332
0,25	10	5	0,1	-0,0941	0,1734	0,0404	0,0024	0,0628	0,0186
0,25	10	5	0,2	-0,0685	0,1730	0,0386	0,0024	0,0656	0,0142
0,25	10	5	0,5	0,1148	0,1637	0,1304	0,0182	0,1634	0,0566
0,25	10	10	0,1	-0,0834	0,1711	0,0376	0,0032	0,0616	0,0168
0,25	10	10	0,2	-0,0187	0,1789	0,0636	0,0034	0,0946	0,0280
0,25	10	10	0,5	0,2163	0,1671	0,1868	0,0328	0,2348	0,0978
0,25	25	1	0,1	-0,0349	0,1451	0,0514	0,0090	0,0518	0,0164
0,25	25	1	0,2	-0,0341	0,1410	0,0496	0,0086	0,0500	0,0152
0,25	25	1	0,5	0,0501	0,1478	0,1004	0,0234	0,0960	0,0358
0,25	25	2	0,1	-0,0357	0,1485	0,0530	0,0100	0,0564	0,0192
0,25	25	2	0,2	-0,0167	0,1402	0,0582	0,0126	0,0590	0,0218
0,25	25	2	0,5	0,0915	0,1483	0,1364	0,0324	0,1306	0,0518
0,25	25	5	0,1	-0,0234	0,1346	0,0544	0,0106	0,0600	0,0184
0,25	25	5	0,2	0,0186	0,1302	0,0726	0,0152	0,0750	0,0274
0,25	25	5	0,5	0,2220	0,1389	0,2470	0,0832	0,2428	0,1140

Continua ...

TABELA 6: (Continuação)

Situação				I de Moran		Mantel		Aprox. normal	
p	n	m	A	Média	var.	$\alpha = 5\%$	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 1\%$
0,25	25	10	0,1	-0,0085	0,1421	0,0634	0,0104	0,0652	0,0228
0,25	25	10	0,2	0,0641	0,1474	0,1012	0,0232	0,1072	0,0406
0,25	25	10	0,5	0,3580	0,1387	0,4092	0,1830	0,4022	0,2334
0,25	50	1	0,1	-0,0172	0,1216	0,0550	0,0114	0,0516	0,0170
0,25	50	1	0,2	-0,0054	0,1188	0,0600	0,0116	0,0560	0,0164
0,25	50	1	0,5	0,0656	0,1306	0,1140	0,0262	0,1026	0,0390
0,25	50	2	0,1	-0,0175	0,1240	0,0528	0,0096	0,0532	0,0166
0,25	50	2	0,2	0,0025	0,1211	0,0606	0,0132	0,0592	0,0216
0,25	50	2	0,5	0,1323	0,1368	0,1770	0,0524	0,1676	0,0692
0,25	50	5	0,1	0,0071	0,1215	0,0622	0,0134	0,0620	0,0236
0,25	50	5	0,2	0,0465	0,1175	0,0892	0,0198	0,0910	0,0320
0,25	50	5	0,5	0,2688	0,1248	0,3486	0,1442	0,3394	0,1840
0,25	50	10	0,1	0,0087	0,1220	0,0676	0,0150	0,0668	0,0254
0,25	50	10	0,2	0,1008	0,1300	0,1354	0,0360	0,1314	0,0544
0,25	50	10	0,5	0,4029	0,1257	0,5232	0,2902	0,5170	0,3430
0,1	5	1	0,1	-0,2443	0,1558	0,0733	0,0418	0,0446	0,0012
0,1	5	1	0,2	-0,2391	0,1549	0,0779	0,0467	0,0449	0,0009
0,1	5	2	0,1	-0,2323	0,1810	0,0726	0,0394	0,0528	0,0012
0,1	5	2	0,2	0,2239	0,1742	0,0723	0,0427	0,0544	0,0014
0,1	5	5	0,1	-0,2394	0,1874	0,0603	0,0331	0,0496	0,0006
0,1	5	5	0,2	-0,1793	0,1819	0,0753	0,0423	0,0619	0,0014
0,1	5	10	0,1	-0,2068	0,1942	0,0628	0,0309	0,0493	0,0002
0,1	5	10	0,2	-0,1418	0,1875	0,0767	0,0373	0,0657	0,0004
0,1	10	1	0,1	-0,1052	0,1265	0,0410	0,0078	0,0440	0,0225
0,1	10	1	0,2	-0,0877	0,1313	0,0433	0,0095	0,0533	0,0261
0,1	10	2	0,1	-0,0908	0,1680	0,0394	0,0047	0,0559	0,0215
0,1	10	2	0,2	-0,0675	0,1723	0,0496	0,0085	0,0695	0,0279
0,1	10	5	0,1	-0,0788	0,1714	0,0368	0,0034	0,0612	0,0172
0,1	10	5	0,2	-0,0105	0,1701	0,0664	0,0110	0,0898	0,0312

Continua ...

TABELA 6: (Continuação)

Situação			I de Moran		Mantel		Aprox. normal			
p	n	m	A	Média	var.	$\alpha = 5\%$	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 1\%$	
0,1	10	10	0,1	-0,0578	0,1776	0,0514	0,0042	0,0684	0,0240	
0,1	10	10	0,2	0,0425	0,1686	0,0846	0,0118	0,1174	0,0418	
0,1	25	1	0,1	-0,0300	0,1281	0,0572	0,0081	0,0538	0,0322	
0,1	25	1	0,2	-0,0188	0,1564	0,0597	0,0086	0,0599	0,0366	
0,1	25	2	0,1	-0,0262	0,1403	0,0534	0,0092	0,0588	0,0246	
0,1	25	2	0,2	0,0102	0,1530	0,0742	0,0148	0,0762	0,0364	
0,1	25	5	0,1	0,0032	0,1502	0,0696	0,0134	0,0702	0,0272	
0,1	25	5	0,2	0,0670	0,1447	0,1162	0,0246	0,1014	0,0442	
0,1	25	10	0,1	0,0173	0,1400	0,0804	0,0166	0,0822	0,0332	
0,1	25	10	0,2	0,1556	0,1345	0,1884	0,0568	0,1836	0,0814	
0,1	50	1	0,1	-0,0116	0,1100	0,0586	0,0154	0,0592	0,0366	
0,1	50	1	0,2	0,0055	0,1223	0,0638	0,0136	0,0626	0,0376	
0,1	50	2	0,1	-0,0051	0,1109	0,0526	0,0104	0,0502	0,0242	
0,1	50	2	0,2	0,0266	0,1171	0,0784	0,0158	0,0734	0,0306	
0,1	50	5	0,1	0,0170	0,1163	0,0732	0,0152	0,0644	0,0252	
0,1	50	5	0,2	0,1100	0,1235	0,1508	0,0358	0,1320	0,0594	
0,1	50	10	0,1	0,0448	0,1189	0,0908	0,0226	0,0858	0,0352	
0,1	50	10	0,2	0,1940	0,1452	0,2558	0,0792	0,2406	0,1124	

Final

TABELA 7: Taxa de poder dos testes de Mantel e aproximação normal, em função de p , n e m , considerando a ponderação *vizinho mais próximo*, sendo os níveis nominais de significância $\alpha = 5\%$ e $\alpha = 1\%$.

Situação			I de Moran		Mantel		Aprox. normal		
p	n	m	A	Média	var.	$\alpha = 5\%$	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 1\%$
0,5	5	1	0,1	-0,2556	0,1800	0,0825	0,0542	0,1204	0,0443
0,5	5	1	0,2	-0,2439	0,1816	0,0855	0,0570	0,1367	0,0500
0,5	5	1	0,5	-0,2142	0,1873	0,0907	0,0650	0,1545	0,0461
0,5	5	1	0,8	-0,1487	0,2086	0,1330	0,1011	0,1943	0,0671
0,5	5	1	1	-0,1062	0,2121	0,1527	0,1132	0,2335	0,0845
0,5	5	2	0,1	-0,2547	0,1833	0,0878	0,0609	0,1451	0,0633
0,5	5	2	0,2	-0,2317	0,1935	0,1031	0,0736	0,1506	0,0671
0,5	5	2	0,5	-0,1825	0,2008	0,1233	0,0883	0,1919	0,0913
0,5	5	2	0,8	-0,0978	0,2035	0,1685	0,1178	0,2480	0,1139
0,5	5	2	1	-0,0218	0,2058	0,2061	0,1435	0,3108	0,1585
0,5	5	5	0,1	-0,2531	0,1783	0,0920	0,0583	0,1404	0,0748
0,5	5	5	0,2	-0,2110	0,1863	0,1096	0,0729	0,1606	0,0781
0,5	5	5	0,5	-0,1094	0,2001	0,1403	0,0895	0,2527	0,1381
0,5	5	5	0,8	0,0114	0,1964	0,2056	0,1272	0,3464	0,1811
0,5	5	5	1	0,0916	0,1904	0,2372	0,1424	0,4119	0,2203
0,5	5	10	0,1	-0,2348	0,1891	0,0943	0,0588	0,1585	0,0868
0,5	5	10	0,2	-0,1893	0,1938	0,1074	0,0660	0,1859	0,0975
0,5	5	10	0,5	-0,0201	0,1923	0,1760	0,0968	0,3152	0,1661
0,5	5	10	0,8	0,1131	0,1803	0,2316	0,1302	0,4519	0,2370
0,5	5	10	1	0,1752	0,1689	0,2654	0,1458	0,5118	0,2636
0,5	10	1	0,1	-0,1038	0,1235	0,0712	0,0162	0,2529	0,1880
0,5	10	1	0,2	-0,1101	0,1213	0,0619	0,0144	0,2490	0,1881
0,5	10	1	0,5	-0,0525	0,1301	0,0885	0,0258	0,3041	0,2311
0,5	10	1	0,8	0,0403	0,1422	0,1409	0,0362	0,4145	0,3223
0,5	10	1	1	0,1127	0,1438	0,1900	0,0560	0,4851	0,3981
0,5	10	2	0,1	-0,1065	0,1243	0,0562	0,0122	0,2690	0,1978

Continua ...

TABELA 7: (Continuação)

Situação				I de Moran		Mantel		Aprox. normal	
p	n	m	A	Média	var.	$\alpha = 5\%$	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 1\%$
0,5	10	2	0,2	-0,0994	0,1232	0,0550	0,0104	0,2761	0,1969
0,5	10	2	0,5	0,0024	0,1297	0,0898	0,0190	0,3843	0,2948
0,5	10	2	0,8	0,1216	0,1392	0,1650	0,0340	0,5167	0,4194
0,5	10	2	1	0,2322	0,1285	0,2356	0,0560	0,6286	0,5314
0,5	10	5	0,1	-0,1025	0,1236	0,0466	0,0062	0,2785	0,2015
0,5	10	5	0,2	-0,0607	0,1267	0,0586	0,0104	0,3204	0,2428
0,5	10	5	0,5	0,1076	0,1338	0,1384	0,0242	0,5129	0,4234
0,5	10	5	0,8	0,3048	0,1123	0,2844	0,0568	0,7255	0,6366
0,5	10	5	1	0,3979	0,1005	0,3742	0,0812	0,8195	0,7474
0,5	10	10	0,1	-0,0960	0,1200	0,0436	0,0046	0,2784	0,2015
0,5	10	10	0,2	-0,0403	0,1317	0,0650	0,0110	0,3525	0,2749
0,5	10	10	0,5	0,2369	0,1234	0,2242	0,0410	0,6594	0,5678
0,5	10	10	0,8	0,4378	0,0894	0,4204	0,0978	0,8638	0,7998
0,5	10	10	1	0,5052	0,0744	0,5180	0,1252	0,9145	0,8656
0,5	25	1	0,1	-0,0378	0,0566	0,0558	0,0134	0,3536	0,3076
0,5	25	1	0,2	-0,0275	0,0561	0,0570	0,0134	0,3724	0,3202
0,5	25	1	0,5	0,0283	0,0617	0,1084	0,0306	0,4618	0,4078
0,5	25	1	0,8	0,1307	0,0611	0,1978	0,0690	0,6286	0,5668
0,5	25	1	1	0,2144	0,0605	0,3062	0,1214	0,7500	0,6984
0,5	25	2	0,1	-0,0314	0,0563	0,0564	0,0132	0,3628	0,3076
0,5	25	2	0,2	-0,0194	0,0571	0,0686	0,0140	0,3878	0,3254
0,5	25	2	0,5	0,0891	0,0600	0,1514	0,0488	0,5616	0,5062
0,5	25	2	0,8	0,2489	0,0569	0,3566	0,1512	0,7884	0,7404
0,5	25	2	1	0,3630	0,0522	0,5450	0,2924	0,9010	0,8734
0,5	25	5	0,1	-0,0317	0,0561	0,0550	0,0124	0,3676	0,3074
0,5	25	5	0,2	0,0164	0,0554	0,0832	0,0192	0,4416	0,3822
0,5	25	5	0,5	0,2216	0,0559	0,3136	0,1274	0,7578	0,7130
0,5	25	5	0,8	0,4538	0,0419	0,7172	0,4476	0,9636	0,9500
0,5	25	5	1	0,5691	0,0324	0,8828	0,6818	0,9938	0,9898

Continua ...

TABELA 7: (Continuação)

Situação				I de Moran		Mantel		Aprox. normal	
p	n	m	A	Média	var.	$\alpha = 5\%$	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 1\%$
0,5	25	10	0,1	-0,0123	0,0565	0,0672	0,0146	0,3964	0,3406
0,5	25	10	0,2	0,0619	0,0591	0,1190	0,0330	0,5315	0,4671
0,5	25	10	0,5	0,3794	0,0472	0,5904	0,3112	0,9266	0,9020
0,5	25	10	0,8	0,6000	0,0304	0,9204	0,7552	0,9946	0,9924
0,5	25	10	1	0,6968	0,0209	0,9836	0,9158	0,9996	0,9994
0,5	50	1	0,1	-0,0143	0,0282	0,0534	0,0134	0,4054	0,3732
0,5	50	1	0,2	-0,0085	0,0291	0,0604	0,0144	0,4250	0,3892
0,5	50	1	0,5	0,0558	0,0303	0,1218	0,0384	0,5754	0,5364
0,5	50	1	0,8	0,1632	0,0304	0,3112	0,1224	0,7836	0,7540
0,5	50	1	1	0,2520	0,0302	0,4982	0,2680	0,9008	0,8758
0,5	50	2	0,1	-0,0142	0,0276	0,0538	0,0084	0,4020	0,3572
0,5	50	2	0,2	0,0070	0,0288	0,0710	0,0156	0,4634	0,4172
0,5	50	2	0,5	0,1201	0,0316	0,2284	0,0854	0,6968	0,6672
0,5	50	2	0,8	0,2853	0,0284	0,5820	0,3292	0,9308	0,9168
0,5	50	2	1	0,4026	0,0244	0,8258	0,6098	0,9860	0,9822
0,5	50	5	0,1	-0,0077	0,0270	0,0578	0,0118	0,4196	0,3702
0,5	50	5	0,2	0,0411	0,0298	0,1074	0,0320	0,5372	0,4920
0,5	50	5	0,5	0,2642	0,0286	0,5348	0,2914	0,9158	0,8962
0,5	50	5	0,8	0,4993	0,0198	0,9476	0,8296	0,9988	0,9982
0,5	50	5	1	0,6201	0,0152	0,9932	0,9684	1,0000	1,0000
0,5	50	10	0,1	0,0124	0,0282	0,0760	0,0192	0,4694	0,4236
0,5	50	10	0,2	0,0882	0,0309	0,1760	0,0570	0,6498	0,6054
0,5	50	10	0,5	0,4239	0,0231	0,8708	0,6686	0,9902	0,9874
0,5	50	10	0,8	0,6579	0,0121	0,9986	0,9898	1,0000	1,0000
0,5	50	10	1	0,7502	0,0088	1,0000	0,9990	1,0000	1,0000
0,25	5	1	0,1	-0,2498	0,1878	0,0867	0,0584	0,1070	0,0312
0,25	5	1	0,2	-0,2483	0,1789	0,0930	0,0641	0,0923	0,0290
0,25	5	1	0,5	-0,1927	0,1962	0,1219	0,0853	0,1454	0,0516
0,25	5	2	0,1	-0,2428	0,1817	0,0893	0,0624	0,1245	0,0483

Continua ...

TABELA 7: (Continuação)

Situação				I de Moran		Mantel		Aprox. normal	
p	n	m	A	Média	var.	$\alpha = 5\%$	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 1\%$
0,25	5	2	0,2	-0,2351	0,1874	0,0951	0,0685	0,1384	0,0587
0,25	5	2	0,5	-0,1569	0,2006	0,1409	0,1038	0,1979	0,0806
0,25	5	5	0,1	-0,2341	0,1833	0,0986	0,0631	0,1511	0,0741
0,25	5	5	0,2	-0,2056	0,1851	0,1075	0,0713	0,1661	0,0835
0,25	5	5	0,5	-0,0632	0,2004	0,1760	0,1219	0,2797	0,1481
0,25	5	10	0,1	-0,2332	0,1891	0,1012	0,0706	0,1589	0,0816
0,25	5	10	0,2	-0,1751	0,2003	0,1252	0,0798	0,1985	0,1109
0,25	5	10	0,5	0,0355	0,1897	0,1992	0,1204	0,3702	0,1921
0,25	10	1	0,1	-0,1094	0,1265	0,0726	0,0213	0,2452	0,1915
0,25	10	1	0,2	-0,1018	0,1258	0,0791	0,0234	0,2541	0,2054
0,25	10	1	0,5	-0,0371	0,1346	0,1146	0,0370	0,3149	0,2578
0,25	10	2	0,1	-0,1032	0,1264	0,0640	0,0154	0,2696	0,1979
0,25	10	2	0,2	-0,0822	0,1271	0,0748	0,0184	0,2874	0,2155
0,25	10	2	0,5	0,0263	0,1334	0,1218	0,0314	0,4090	0,3248
0,25	10	5	0,1	-0,0909	0,1260	0,0534	0,0094	0,2896	0,2168
0,25	10	5	0,2	-0,0591	0,1240	0,0606	0,0116	0,3132	0,2317
0,25	10	5	0,5	0,1702	0,1296	0,1750	0,0344	0,5781	0,4861
0,25	10	10	0,1	-0,0803	0,1210	0,0510	0,0070	0,2981	0,2249
0,25	10	10	0,2	0,0011	0,1289	0,0784	0,0132	0,3857	0,2974
0,25	10	10	0,5	0,2949	0,1132	0,2662	0,0576	0,7231	0,6415
0,25	25	1	0,1	-0,0347	0,0554	0,0572	0,0132	0,3511	0,2939
0,25	25	1	0,2	-0,0253	0,0559	0,0634	0,0142	0,3651	0,3133
0,25	25	1	0,5	0,0571	0,0632	0,1234	0,0414	0,4918	0,4430
0,25	25	2	0,1	-0,0346	0,0541	0,0508	0,0118	0,3608	0,2972
0,25	25	2	0,2	-0,0123	0,0573	0,0698	0,0168	0,3996	0,3338
0,25	25	2	0,5	0,1320	0,0597	0,2002	0,0658	0,6301	0,5669
0,25	25	5	0,1	-0,0180	0,0559	0,0626	0,0150	0,3924	0,3334
0,25	25	5	0,2	0,0336	0,0579	0,1020	0,0256	0,4769	0,4155
0,25	25	5	0,5	0,2905	0,0530	0,4298	0,1840	0,8444	0,8118

Continua ...

TABELA 7: (Continuação)

Situação				I de Moran		Mantel		Aprox. normal	
p	n	m	A	Média	var.	$\alpha = 5\%$	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 1\%$
0,25	25	10	0,1	0,0015	0,0575	0,0772	0,0194	0,4209	0,3653
0,25	25	10	0,2	0,0921	0,0592	0,1492	0,0506	0,5767	0,5139
0,25	25	10	0,5	0,4502	0,0429	0,7112	0,4412	0,9632	0,9474
0,25	50	1	0,1	-0,0203	0,0294	0,0566	0,0144	0,3870	0,3482
0,25	50	1	0,2	-0,0024	0,0295	0,0666	0,0156	0,4258	0,3852
0,25	50	1	0,5	0,0786	0,0319	0,1596	0,0490	0,6106	0,5688
0,25	50	2	0,1	-0,0137	0,0279	0,0510	0,0114	0,4106	0,3592
0,25	50	2	0,2	0,0143	0,0297	0,0806	0,0208	0,4682	0,4182
0,25	50	2	0,5	0,1615	0,0312	0,2944	0,1196	0,7814	0,7454
0,25	50	5	0,1	0,0019	0,0286	0,0626	0,0168	0,4432	0,4038
0,25	50	5	0,2	0,0570	0,0303	0,1306	0,0398	0,5676	0,5272
0,25	50	5	0,5	0,3309	0,0280	0,6810	0,4300	0,9582	0,9500
0,25	50	10	0,1	0,0218	0,0286	0,0876	0,0222	0,4904	0,4428
0,25	50	10	0,2	0,1242	0,0296	0,2278	0,0840	0,7144	0,6732
0,25	50	10	0,5	0,4995	0,0198	0,9504	0,8322	0,9988	0,9980
0,1	5	1	0,1	-0,2437	0,1241	0,0559	0,0388	0,0710	0,0131
0,1	5	1	0,2	-0,2388	0,1259	0,0593	0,0409	0,0889	0,0216
0,1	5	2	0,1	-0,2469	0,1619	0,0765	0,0569	0,0996	0,0333
0,1	5	2	0,2	-0,2218	0,1657	0,0925	0,0668	0,1207	0,0474
0,1	5	5	0,1	-0,2300	0,1846	0,1097	0,0764	0,1469	0,0539
0,1	5	5	0,2	-0,1768	0,1980	0,1332	0,0951	0,1829	0,0787
0,1	5	10	0,1	-0,2087	0,1891	0,1099	0,0780	0,1740	0,0803
0,1	5	10	0,2	-0,1255	0,1965	0,1424	0,0942	0,2223	0,1117
0,1	10	1	0,1	-0,1031	0,0838	0,0587	0,0266	0,2117	0,1976
0,1	10	1	0,2	-0,0838	0,0884	0,0735	0,0279	0,2405	0,2192
0,1	10	2	0,1	-0,0953	0,1127	0,0748	0,0232	0,2561	0,2133
0,1	10	2	0,2	-0,0590	0,1184	0,0948	0,0340	0,2954	0,2486
0,1	10	5	0,1	-0,0727	0,1233	0,0756	0,0174	0,2956	0,2195
0,1	10	5	0,2	0,0066	0,1314	0,1088	0,0276	0,3897	0,3111

Continua ...

TABELA 7: (Continuação)

Situação				I de Moran		Mantel		Aprox. normal	
p	n	m	A	Média	var.	$\alpha = 5\%$	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 1\%$
0,1	10	10	0,1	-0,0423	0,1258	0,0674	0,0112	0,3467	0,2601
0,1	10	10	0,2	0,0859	0,1313	0,1320	0,0250	0,4812	0,3852
0,1	25	1	0,1	-0,0302	0,0493	0,0620	0,0143	0,3327	0,3034
0,1	25	1	0,2	-0,0156	0,0531	0,0727	0,0233	0,3523	0,3290
0,1	25	2	0,1	-0,0238	0,0534	0,0620	0,0122	0,3626	0,3135
0,1	25	2	0,2	0,0146	0,0566	0,0872	0,0222	0,4332	0,3772
0,1	25	5	0,1	-0,0024	0,0580	0,0734	0,0170	0,4285	0,3627
0,1	25	5	0,2	0,1005	0,0607	0,1626	0,0510	0,5880	0,5274
0,1	25	10	0,1	0,0293	0,0596	0,0996	0,0276	0,4607	0,4018
0,1	25	10	0,2	0,2043	0,0568	0,2820	0,1094	0,7425	0,6925
0,1	50	1	0,1	-0,0112	0,0271	0,0592	0,0106	0,3882	0,3517
0,1	50	1	0,2	0,0108	0,0298	0,0810	0,0206	0,4282	0,3890
0,1	50	2	0,1	0,0012	0,0283	0,0676	0,0176	0,4297	0,3903
0,1	50	2	0,2	0,0402	0,0304	0,1072	0,0274	0,5193	0,4763
0,1	50	5	0,1	0,0174	0,0301	0,0866	0,0230	0,4784	0,4324
0,1	50	5	0,2	0,1276	0,0305	0,2330	0,0840	0,7168	0,6802
0,1	50	10	0,1	0,0602	0,0298	0,1310	0,0394	0,5796	0,5378
0,1	50	10	0,2	0,2406	0,0281	0,4710	0,2322	0,8958	0,8742
Final									

4.3 Uma aplicação com dados reais

Na Tabela 8 encontram-se os resultados obtidos da análise de autocorrelação espacial (índice I de Moran) referente ao critério de teste de Mantel, em que adotaram-se os três critérios de ponderação abordados neste trabalho. Os dados submetidos a esta análise são de características genéticas de uma população de árvores de copaíba conforme descrito na Seção 3.2. Apresentou-se aqui um exemplo ilustrativo apenas dos resultados das análises provenientes dos alelos dos locos marcadores $\alpha - Est - 1$ e $Po - 2$, pelo fato de os outros locos apresentarem resultados semelhantes.

Pode-se verificar, por meio da Tabela 8, que as estimativas dos índices I de Moran não foram estatisticamente significativas a 5%, em nenhum dos casos estudados, quando se adotou o critério de teste de Mantel. Ou seja, não houve evidências, neste nível nominal de significância, de que houvesse a existência de dependência espacial.

Quanto ao critério da aproximação normal, de modo geral, estas estatísticas também não foram significativas neste mesmo nível de significância, conforme mostra a Tabela 9. Porém, em especial para o loco $Po - 2$, apresentou dependência significativa a 5%, em relação ao alelo 1, com o critério de ponderação vizinho mais próximo.

É importante relatar que a situação estudada via simulação de Monte Carlo $\{n=25 \text{ e } m=1\}$ equivale a uma amostra de 25 indivíduos por população, conforme já comentado anteriormente. Este fato é condizente com a situação dos dados reais de copaíbas citados na Seção (3.2). Embora na simulação o tamanho da amostra fosse de 25 indivíduos, pôde-se comparar, sem maiores riscos, o desempenho de ambos os testes, com os resultados obtidos relativo ao tamanho amostral dos dados reais com 20 indivíduos por população. Assim, fazendo-se uma analogia entre a situação de dados reais e a simulação Monte Carlo apresentada nas Seções (4.1) e (4.2), pode-se afirmar que o teste de Mantel foi coerente com o desempenho

apresentado na simulação, no que tange ao seu poder de decisão, principalmente em relação aos baixos valores de frequências alélicas médias, como é o caso do loco $Po - 2$, conforme apresentado na Tabela 10.

Apesar de o teste da aproximação normal ter detectado a existência de uma baixa autocorrelação positiva, quanto ao alelo 1 do loco $Po - 2$, deve-se atentar que, para a ponderação vizinho mais próximo, este critério de teste não foi recomendado, porque o mesmo atingiu altas taxas de erro tipo I , conforme os resultados obtidos por meio da simulação. Portanto, deve-se considerar que os resultados apresentados pelo teste de Mantel são mais confiáveis que aqueles apresentados pela aproximação normal. Para as outras ponderações, pôde-se também verificar, com base na simulação, que o teste de Mantel deve ser preferido quando se utilizam tamanhos de amostras acima de 10 indivíduos, porque, neste caso, o mesmo controlou satisfatoriamente a taxa de erro tipo I .

É ainda importante observar que, para o loco $\alpha - Est - 1$, exceto o alelo 3, as frequências alélicas médias tiveram valores intermediários (Tabela 10). Este aspecto, como pôde ser comprovado pela simulação de Monte Carlo, implica em maior poder para os testes.

As Figuras 5 e 6 ilustram o comportamento das frequências alélicas dos dados de copaíbas, para os locos marcadores $\alpha - Est - 1$ e $Po - 2$, em função do espaço geográfico. Embora tenha-se um tamanho muito pequeno de amostra relativo ao loco $\alpha - Est - 1$, pode-se ver na Figura 5 que há apenas pontos flutuando aleatoriamente sobre o espaço geográfico. Ou seja, não há um gradiente de variação de frequências alélicas bem definido, indicando uma situação típica de *padrões aleatórios*, sem padrão espacial. Mas, pode-se notar a existência de alguns valores discrepantes, conforme mostram, principalmente, as Figuras 5b e 5c. Essas superfícies têm a aparência de uma paisagem de montanha, reforçando o indício de ausência de padrão espacial. Este fato confirma os resultados discutidos anteriormente e apresentados nas Tabelas 8 e 9.

A Figura 6 também não evidencia nenhum padrão espacial definido. No entanto, deve-se observar que, aparentemente, também aqui existe a presença de “outliers” na massa de dados, atribuído, principalmente, ao genótipo homozigoto composto pelo alelo 2 do loco $Po - 2$, conforme retrata a Figura 6. Figura 6a está mais em consonância com os resultados obtidos por meio do teste de Mantel do que com o teste da aproximação normal, de acordo com o que já foi discutido anteriormente.

TABELA 8: Resultados da análise espacial utilizando o teste de Mantel por meio do I de Moran para os marcadores $\alpha - Est - 1$ e $Po - 2$, segundo os pesos inverso da distância, inverso da distância ao quadrado e vizinho mais próximo, considerando um nível de significância $\alpha = 5\%$.

Inverso da distância				
<i>Loco</i>	<i>Alelo</i>	<i>I de Moran</i>	$Q_1(0,05)$	$Q_2(0,995)$
$\alpha - Est - 1$	1	-0,2052 ^{ns}	-0,5818	0,6900
	2	-0,2280 ^{ns}	-0,5113	0,5317
	3	-0,0817 ^{ns}	-0,4146	0,3966
$Po - 2$	1	0,1048 ^{ns}	-0,2888	0,3402
	2	0,0221 ^{ns}	-0,3047	0,2302
	3	-0,0930 ^{ns}	-0,3215	0,2373
Inverso da distância ao quadrado				
<i>Loco</i>	<i>Alelo</i>	<i>I de Moran</i>	$Q_1(0,05)$	$Q_2(0,995)$
$\alpha - Est - 1$	1	-0,1861 ^{ns}	-1,2605	1,2177
	2	-0,1486 ^{ns}	-1,0296	1,2318
	3	0,0228 ^{ns}	-0,7936	0,7451
$Po - 2$	1	0,3647 ^{ns}	-0,7259	0,8413
	2	0,1731 ^{ns}	-0,8154	0,6663
	3	-0,0465 ^{ns}	-0,8927	0,7853
Vizinho mais próximo				
<i>Loco</i>	<i>Alelo</i>	<i>I de Moran</i>	$Q_1(0,05)$	$Q_2(0,995)$
$\alpha - Est - 1$	1	-0,5556 ^{ns}	-0,8333	0,8333
	2	-0,0027 ^{ns}	-0,0063	0,0058
	3	-0,0000 ^{ns}	-0,0000	0,0000
$Po - 2$	1	0,1774 ^{ns}	-0,6703	0,6954
	2	0,0001 ^{ns}	-0,0008	0,0008
	3	-0,0000 ^{ns}	-0,0000	0,0000

$Q_1(0,05)$ representa o quantil 0,05 da distribuição empírica.

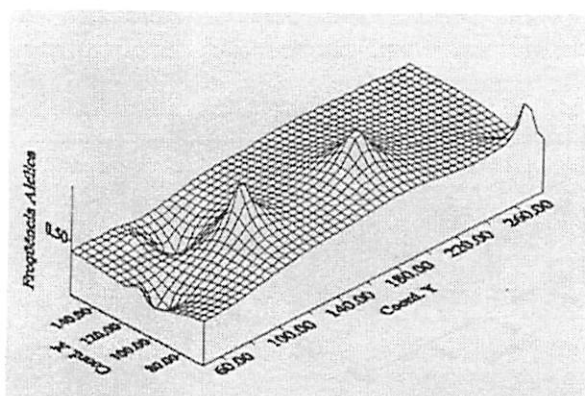
$Q_2(0,95)$ representa o quantil 0,95 da distribuição empírica.

TABELA 9: Resultados da análise de autocorrelação espacial utilizando o teste da aproximação normal por meio do I de Moran para os marcadores $\alpha - Est - 1$ e $Po - 2$, segundo os pesos inverso da distância, inverso da distância ao quadrado e vizinho mais próximo, considerando um nível de significância $\alpha = 5\%$.

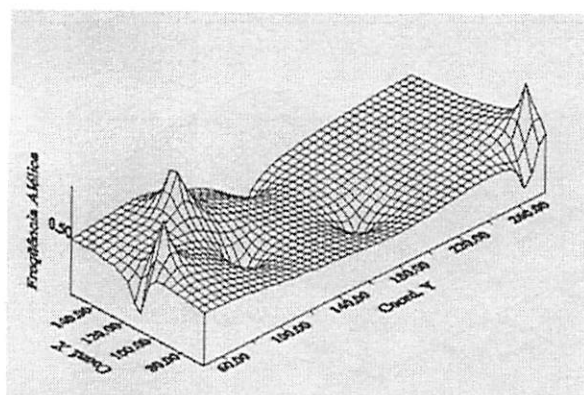
Inverso da distância			
<i>Loco</i>	<i>Alelo</i>	<i>I de Moran</i>	<i>valor-P</i>
$\alpha - Est - 1$	1	-0,2052 ^{ns}	0,6542
	2	-0,2280 ^{ns}	0,6863
	3	-0,0817 ^{ns}	0,4443
$Po - 2$	1	0,1048 ^{ns}	0,0840
	2	0,0221 ^{ns}	0,2513
	3	-0,0930 ^{ns}	0,6437
Inverso da distância ao quadrado			
<i>Loco</i>	<i>Alelo</i>	<i>I de Moran</i>	<i>valor-P</i>
$\alpha - Est - 1$	1	-0,1861 ^{ns}	0,5537
	2	-0,1486 ^{ns}	0,5265
	3	0,0228 ^{ns}	0,3941
$Po - 2$	1	0,3647 ^{ns}	0,0946
	2	0,1731 ^{ns}	0,2334
	3	-0,0465 ^{ns}	0,4919
Vizinho mais próximo			
<i>Loco</i>	<i>Alelo</i>	<i>I de Moran</i>	<i>valor-P</i>
$\alpha - Est - 1$	1	-0,5556 ^{ns}	0,9964
	2	-0,0027 ^{ns}	0,2661
	3	-0,0000 ^{ns}	0,1139
$Po - 2$	1	0,1774*	0,0016
	2	0,0001 ^{ns}	0,2322
	3	-0,0000 ^{ns}	0,2190

TABELA 10: Frequência alélica média dos dados de copaiba em relação aos locos $\alpha - Est - 1$ e $Po - 2$.

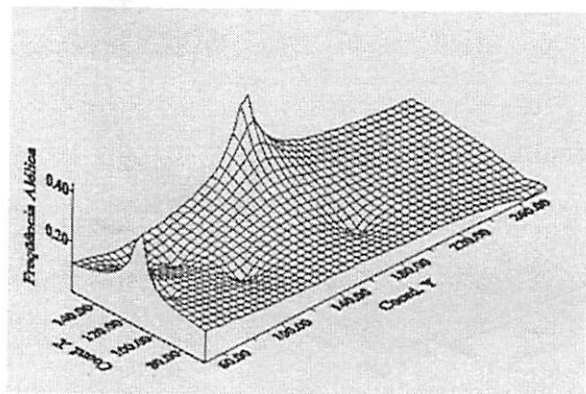
<i>Loco</i>	<i>Alelo</i>	Frequência alélica média
$\alpha - Est - 1$	1	0,5000
	2	0,4000
	3	0,1000
$Po - 2$	1	0,1750
	2	0,3000
	3	0,5250



(a) alelo 1

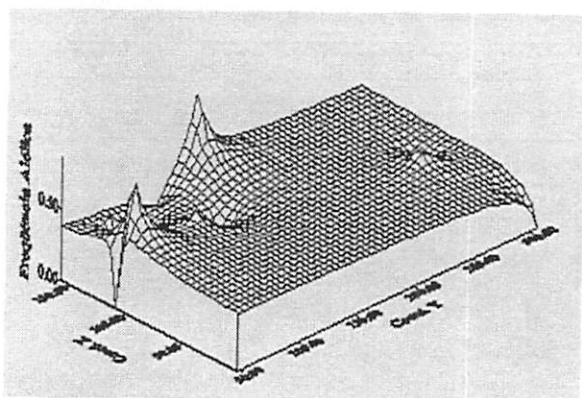


(b) alelo 2

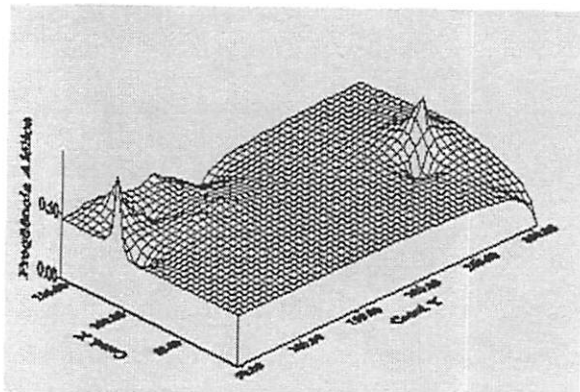


(c) alelo 3

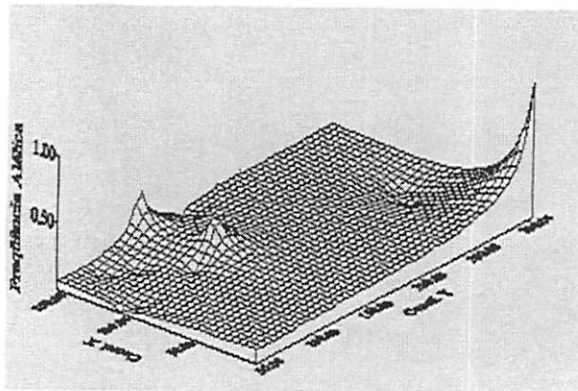
FIGURA 5: Superfície de frequências alélicas em relação ao loco $\alpha - Est - 1$



(a) alelo 1



(b) alelo 2



(c) alelo 3

FIGURA 6: Superfície de frequências alélicas em relação ao loco $Po - 2$

Para as condições analisadas via simulação de Monte Carlo, em relação às ponderações inverso da distância e inverso da distância ao quadrado, verificou-se que a variância do índice I de Moran foi diretamente afetada pela amplitude de variação média (A), tendendo a aumentá-la. Por outro lado, a magnitude de n , contrariamente, também influenciou esta mesma variância, tendendo a diminuí-la. Quanto ao esquema de ponderação vizinho mais próximo, as variâncias apresentaram magnitudes intermediárias quando comparadas com as duas outras ponderações adotadas.

Embora com baixo poder, o teste da aproximação normal deve ser preferido para as situações de pequenas amostras ($n \leq 10$), quando se utiliza os esquemas de ponderações inverso da distância e o inverso da distância ao quadrado, pois o mesmo controlou melhor as taxas de erro tipo I , para ambos os níveis de significância.

Para o tamanho populacional maior ou igual a 25, o critério de Mantel teve melhor desempenho, pois controlou satisfatoriamente as taxas de erros. Somente para o caso $\{n=50\}$ o teste da aproximação normal obteve taxas de erros consistentes e estatisticamente iguais aos valores nominais, com o valor nominal $\alpha = 0,05$.

Os poderes dos testes foram fortemente influenciados pelas frequências alélicas médias. De fato, quando se varia a frequência alélica média, ou aproximando-a de zero, ou de 1, isto faz com que a inclinação do plano de variação clinal sofra uma alteração, forçando-a a diminuir. Assim, para se obter uma amplitude razoável de variação deve-se ter frequências alélicas intermediárias, com as quais podem-se obter inclinações mais pronunciadas. Caso contrário, ter-se-ão mais dificuldades

em se detectar o padrão espacial, o que implica em menos poder para os testes em questão. O pesquisador deve estar atento a este aspecto, visto que nem sempre se encontram locos marcadores com frequência alélica média próxima de 0,5.

Quanto ao critério de teste de Mantel, ele controlou as taxas de erros tipo *I* a partir de $\{n=25\}$, com a mesma eficiência em todos os três esquemas de pesos adotados, para ambos os níveis de significância. Este critério de teste não controlou o erro tipo *I* para os menores valores populacionais ($n \leq 10$), sendo um pouco conservador quando foram utilizadas as ponderações inverso da distância, exceto para o tamanho $n = 5$ populações, em que o mesmo mostrou-se liberal quando combinado com a ponderação vizinho mais próximo, para $\alpha = 5\%$ e $\alpha = 1\%$.

Quanto ao fator de ponderação inverso da distância ao quadrado, pôde-se notar uma queda no poder dos testes nos casos com $\{n=50\}$. Isso é um forte indício de que, mesmo para grandes amostras, deve-se ter muita cautela na escolha adequada do esquema de ponderação. Este fato sugere a importância de estudos que expliquem como se dá o padrão de variação clinal. Na prática, é razoável supor que, provavelmente, este padrão espacial não se dê por meio de uma superfície rigorosamente linear. Assim, a escolha da ponderação adequada não é tão clara.

Em linhas gerais, o teste da aproximação normal não é recomendado em nenhuma das situações simuladas quando se utiliza o esquema vizinho mais próximo.

Quanto à aplicação com dados reais, conclui-se, baseado no índice *I*, que não houve autocorrelação espacial significativa ($p < 0,05$) entre as árvores de uma população de copaíba.

REFERÊNCIAS BIBLIOGRÁFICAS

ADLER, P. B.; LAUENROTH, W. K. Livestock exclusion increases the spatial heterogeneity of vegetation in colorado shortgrass steppe. **Applied Vegetation Science**, Dordrecht, v. 3, p. 313–222, 2000.

ALFENAS, A. C. **Eletroforese de isoenzimas e proteínas afims; fundamentos e aplicações em plantas e microrganismos**. Viçosa: Universidade Federal de Viçosa, 1998. 574p.

ALFENAS, A. C. et al. **Eletroforese de proteínas e isoenzimas de fungos e essências florestais**. Viçosa: Universidade Federal de Viçosa, 1991. 242p.

BARBUJANI, G. Diversity of some gene frequencies in european and asian populations III. Spatial Correlogram Analysis. **Annals of Human Genetics**, New York, v. 51, n. 4, p. 345–353, Oct. 1987.

BIVAND, R. A monte carlo study of correlation coeficient estimation with spatially autocorrelated observations. **Quaestiones Geographicae**, Poznan, v. 6, p. 5–10, 1980.

BJØRNSTAD, O. N.; FALCK, W. **An extension of the spatial correlogram and the x-intercept for genetic data**. (Working Papers from University of Oslo, Department of Biology), 2002. Disponível em: <<http://onb.ent.psu.edu/onb1/unpubl/blomster/blomster.pdf>>. Acesso em: 10 de set. 2002.

CASELLA, G.; BERGER, R. L. **Statistical inference**. California: Duxbury Press, 1990. 650p.

CLIFF, A. D.; ORD, J. K. Evaluating the percentage points of a spatial autocorrelation coeficient. **Geographical Analysis**, Columbus, v. 3, n. 1, p. 51–62, 1971.

CLIFF, A. D.; ORD, J. K. The problem of spatial autocorrelation. In: SCOTT, A. J. (Ed.). **Studies in regional science**. London: Pion Press, 1969. p. 25–55.

CLIFF, A. D.; ORD, J. K. **Spatial autocorrelation**. London: Pion Press, 1973. 178p.

- CLIFF, A. D.; ORD, J. K. Spatial autocorrelation: a review of existing measures with applications. *Econometric Geography*, Worcester, v. 46, n. 2, p. 269–292, 1970.
- CLIFF, A. D.; ORD, J. K. Spatial and temporal analysis: autocorrelation in space and time. In: WRIGLEY, N.; BENNET, R. J. (Ed.). **Quantitative geography: a british view**. London: Routledge and Kegan Paul, 1981a. p. 104–110.
- CLIFF, A. D.; ORD, J. K. **Spatial processes: models and applications**. London: Pion Press, 1981b. 266 p.
- COSTA, S. F. **Introdução ilustrada à estatística (com muito humor)**. 2 ed. São Paulo: Harbra, 1992.
- CRESSIE, N. A. C. **Statistics for spatial data**. Revised edition. Iowa State University, New York: A Wiley Interscience Publication, 1993. 900p.
- DINIZ-FILHO, J. A. F. Análise de autocorrelação espacial: inferências microevolutivas e aplicações em genética de populações. *Séries Monográficas*, Ribeirão Preto, n. 6, p. 89–138, 1998.
- ENDLER, J. A. **Geographic variation, speciation and clines**. Princeton: Princeton University Press, 1977.
- GABRIEL, K. R.; SOKAL, R. A new statistical approach to geographic variation analysis. *Systematic Zoology*, Washington, v. 18, n. 3, p. 259–278, 1969.
- GOULD, P. R. Is statistic inferens the geographical name for a wild goose? *Economic Geography Supplement*, Worcester, v. 46, p. 439–448, 1970.
- GUMPERTZ, M. **Applied spatial statistics**. (Working Papers from North Carolina State University, Department of Statistics), 1999. Disponível em: <<http://www4.stat.ncsu.edu/gumpertz/>>. Acesso em: 20 de mai. 2002.
- HENEGBRY, G. M. Spatial model error analysis using autocorrelation indices. *Ecological Modelling*, Amsterdam, v. 82, n. 1, p. 75–91, Sept. 1995.
- HEPPLE, L. W. The econometric specification and estimation of spatio-temporal models. In: T CARLSTEIN, D. P. e. N. T. (Ed.). **Timing space and spacing time**. London: Edward Arnold, 1978. p. 66–80.
- HOEL, P. G. **Estatística matemática**. 4 ed. Rio de Janeiro: Editora Guanabara Dois S. A., 1980. 373p.

- ISAAKS, E. H.; SRIVASTAVA, R. M. **Applied geostatistics**. New York: Oxford University Press, 1989. 561p.
- JUMARS, P. A.; THISTLE, D.; JONES, M. L. Detecting two-dimensional spatial structure in biological data. **Oecologia**, New York, v. 28, n. 2, p. 109–123, Feb. 1977.
- KENDALL, M. G.; STUART, A. **The advanced theory of statistics**. 3 ed. London: Charles Griffin and Company Ltd, 1973. 723 p.
- KOENIG, W. D. Spatial autocorrelation of ecological phenomena. **Trends in Ecology & Evolution**, London, v. 14, n. 1, p. 22–26, Jan. 1999.
- KOOPMANS, T. C. Serial correlation and quadratic forms in normal variables. **Annals of Mathematical Statistics**, Baltimore, v. 13, n. 1, p. 14–33, Mar. 1942.
- LEEMIS, L. M.; TRIVED, K. S. A comparison of approximate interval estimators for the bernoulli parameter. **The American Statistician**, Alexandria, v. 50, n. 1, p. 63–68, Feb. 1996.
- LEGENDRE, P. Spatial autocorrelation. **Ecology**, Washington, v. 74, n. 6, p. 1659–1673, Sept. 1993.
- LEGENDRE, P.; FORTIN, M. J. Spatial pattern and ecological analysis. **Vegetatio**, v. 80, n. 2, Jun. 1989.
- LEGENDRE, P. et al. Approximate analysis of variance of spatially autocorrelated regional data. **Journal of Classification**, v. 7, n. 1, p. 53–75, 1990.
- MANLY, B. F. J. **Randomization and Monte Carlo methods in biology**. London: Chapman and Hall, 1991. 399p.
- MANTEL, N. The detection of disease clustering and a generalized regression approach. **Cancer Research**, Birmingham, v. 27, n. 2, p. 209–202, 1967.
- MILNE, B. T. Heterogeneity as a multiscale characteristic of landscapes. In: PICKETT, J. K. . S. T. A. (Ed.). **Ecological heterogeneity**. New York: Springer-Verlag, 1991. p. 69–84.
- MORAN, P. A. F. The interpretation of statistical maps. **Journal of Royal Statistical Society, Series B**, London, v. 10, n. 2, p. 243–251, 1948.
- MORAN, P. A. F. Notes on continuous stochastic phenomena. **Biometrika**, London, v. 37, n. 1, p. 17–23, Jun. 1950.

- ODEN, N. L. Assessing the significance of a spatial correlogram. **Geographical Analysis**, Columbus, v. 16, n. 1, p. 1–16, 1984.
- PALMER, M. W. Fractal geometry: a tool for describing spatial patterns of plant communities. **Vegetatio**, Washington, v. 75, n. 1/2, p. 91–102, Dec. 1988.
- PIAZZA, A.; MENOZZI, P.; CAVALLI-SFORZA, L. The making and testing of geographic gene-frequency maps. **Biometrics**, Washington, v. 37, n. 4, p. 635–659, Dec. 1981.
- PIELOU, E. C. **Mathematical ecology**. New York: J. Wiley, 1977.
- PITTMAN, E. The closest estimates of statistical parameters. **Cambridge Philosophical Society**, Cambridge, v. 33, p. 212–222, 1937.
- RIPLEY, B. D. **Spatial statistics**. New York: J. Wiley, 1981.
- SEARLE, S. R. **Linear models**. New York: J. Wiley, 1971. 532p.
- SLATKIN, M.; MARUYAMA, T. Genetic drift in a cline. **Genetics**, v. 81, p. 209–222, 1975.
- SOKAL, R. R. Analyzing character variation in geographic space. In: FELSENSTEIN, J. (Ed.). **Numerical taxonomy**. New York, pp.: Springer-Verlag, 1983. p. 384–403.
- SOKAL, R. R. Spatial data analysis and historical processes. In: DIDAY, E. et al. (Ed.). **Data analysis and informatics IV**. Holland: Science publishers, 1986. p. 29–43.
- SOKAL, R. R. Testing statistical significance of geographic variation patterns. **Systematic Zoology**, Washington, v. 28, n. 2, p. 227–232, 1979.
- SOKAL, R. R.; JACQUEZ, G. M. Testing inferences about microevolutionary processes by means of spatial autocorrelation analysis. **Evolution**, Washington, v. 45, n. 1, p. 152–168, Feb. 1991.
- SOKAL, R. R.; ODEN, N. L. Spatial autocorrelation in biology 1. methodology. **Biology Journal of the Linnean Society**, London, v. 10, n. 1, p. 199–228, 1978a.
- SOKAL, R. R.; ODEN, N. L. Spatial autocorrelation in biology 2. some biological implications of evolutionary and ecological interest. **Biology Journal of the Linnean Society**, London, v. 10, n. 2, p. 229–249, 1978b.

SOKAL, R. R.; WARTENBERG, D. E. Space and population structure. In: GRIFFITH, D.; MCKINNON, R. (Ed.). **Dynamic spatial models**. Netherlands: Sijthoff and Noordhoff, Alphen aan den Rijn, 1981. p. 186–213.

STUDENT, W. S. G. The elimination of spurious correlation due to position in time or space. **Biometrika**, London, v. 10, n. 1/2, p. 179–180, Jun. 1914.

TELLES, M. P. C. et al. Autocorrelação espacial das frequências em superpulações de cagaiteira (*Eugenia dysentericadc.*, myrtaceae) no sudeste de goiás. **Revista Brasileira Botânica**, São Paulo, v. 24, n. 1, p. 145–154, Mar. 2001.

TOBLER, W. R. A computer movie simulating urban growth in the detroit region. **Economic Geography Supplement**, Worcester, v. 15, p. 134–240, 1970.

WALD, A. Asymptotically most powerful tests of statistical hypotheses. **Annals of Mathematical Statistics.**, Baltimore, v. 12, n. 1, p. 1, Mar. 1941.