

UMA COMPARAÇÃO DE TÉCNICAS BASEADAS EM *QUADRATS* PARA A CARACTERIZAÇÃO DE PADRÕES ESPACIAIS EM DOENÇAS DE PLANTAS

Renato Ribeiro de LIMA¹
Clarice Garcia Borges DEMÉTRIO²
Paulo Justiniano RIBEIRO JUNIOR³
Martin Spencer RIDOUT⁴

- RESUMO: A informação sobre a dinâmica espaço-temporal de doenças em plantas é relevante em estudos epidemiológicos, podendo ser utilizada para descrever e entender o desenvolvimento das doenças, desenvolver planos de amostragem, planejar experimentos controlados, avaliar o efeito de diferentes tratamentos e caracterizar perdas na produção ocasionadas pela doença. Métodos estatísticos, baseados em contagens por unidade amostral ou *quadrats*, têm sido utilizados na caracterização de padrões espaciais, tais como cálculo de índices de dispersão, de coeficiente de correlação intraclasse e técnicas de determinação de autocorrelação espacial. Tais métodos fornecem medidas diretas da heterogeneidade entre as unidades amostrais, que são funções da associação espacial dos indivíduos dentro dessas unidades. Neste trabalho, foi conduzido um estudo de simulação, considerando diferentes níveis de incidência da doença, diferentes tamanhos de *quadrats* e diferentes graus de agregação das plantas doentes. O estudo mostrou que existem problemas na caracterização de padrões espaciais, usando essas medidas de heterogeneidade, indicando que o uso e a interpretação de índices de agregação para caracterização de padrões espaciais devem ser vistos com cautela, em particular se apenas um deles é adotado em uma análise de dados. Os índices mostram discrepâncias nos seus resultados e o comportamento de cada um deles mostrou diferentes sensibilidades a efeitos de diferentes tamanhos de *quadrat* e níveis de incidência da doença.

¹Departamento de Ciências Exatas, Universidade Federal de Lavras, Caixa Postal 3037, CEP: 37200-000, Lavras, MG, Brasil, E-mail: rrlima@ufla.br

²Departamento de Ciências Exatas, Escola Superior de Agricultura "Luiz de Queiroz", ESALQ, Universidade de São Paulo, Caixa Postal 9, CEP: 13418-900, Piracicaba, SP, Brasil, E-mail: clarice@esalq.usp.br

³Laboratório de Estatística e Geoinformação, Universidade Federal do Paraná – UFPR, Caixa Postal 19,081, CEP: 81531-990, Curitiba, PR, Brasil, E-mail: paulojus@ufpr.br

⁴Institute of Mathematics, Statistics and Actuarial Science, University of Kent, CT2 7NF, Canterbury, Kent, United Kingdom, E-mail: M.S.Ridout@kent.ac.uk

- PALAVRAS-CHAVE: Autocorrelação espacial; correlação intraclasse; índice de dispersão; padrão espacial; simulação.

1 Introdução

A informação sobre a dinâmica espaço-temporal de doenças em plantas é de importância fundamental em estudos epidemiológicos, podendo ser utilizada para descrever e entender o desenvolvimento das doenças (Bassanezi et al., 2003; Gibson, 1997a, 1997b; Gibson e Austin, 1996; Gottwald, 1995; Gottwald et al., 1999; Turechek e Madden, 2000; Xu e Ridout, 2000, 2001), desenvolver planos de amostragem (Hughes e Madden, 1994; Madden e Hughes, 1999; Marion, Gibson e Renshaw, 2003; Turechek e Madden, 2001), planejar experimentos controlados (Hughes e Madden, 1995), avaliar o efeito de diferentes tratamentos (Madden, 1993; Madden e Hughes, 1995; Zadoks e Van Den Bosch, 1994) e caracterizar perdas na produção ocasionadas pela doença (Ferrandino, 1989; Hughes, 1988; Madden e Nutter Júnior, 1995). O conhecimento detalhado dessa dinâmica espaço-temporal possibilita uma visão mais completa da estrutura e do comportamento de patossistemas (Bergamin Filho et al., 2002) e auxilia no desenvolvimento de hipóteses em relação aos aspectos biológicos e ambientais associados à dispersão das doenças (Campbell e Madden, 1990).

Análises temporais de patossistemas vêm sendo discutidas na literatura há muitos anos, utilizando metodologias consolidadas e facilmente disponíveis em implementações computacionais. Nas duas últimas décadas, tem se voltado a atenção também para o estudo da caracterização de padrões espaciais de doenças em plantas, definido como o arranjo das plantas doentes, que é reflexo do processo de dispersão dos patógenos (Bergamin Filho et al., 2002; Hughes et al., 1997). Padrões espaciais podem ser estudados a partir de observações feitas em unidades que podem ser constituídas de uma única planta, o que comumente ocorre com culturas perenes (Amorin et al. 2007), ou grupos de plantas (Gumpertz, Graham e Ristaino, 1997). Neste trabalho, sem perda de generalidade, assume-se uma única planta como unidade, como por exemplo, ocorre em uma cultura de citros.

Tipicamente, um padrão espacial é, em geral, de início classificado como aleatório, agregado ou regular. O padrão espacial aleatório está relacionado a iguais oportunidades de infecção enquanto o agregado ocorre quando, em condições naturais, o patógeno dispersa-se apenas a curtas distâncias, como por exemplo, através de respingos da chuva e, assim, a probabilidade de que uma planta situada próxima à fonte de inóculo se torne infectada é maior do que a de plantas mais distantes. O padrão regular não ocorre de forma natural, mas pode ser provocado pelo plantio alternado de variedades resistentes e suscetíveis.

Segundo Madden e Hughes (1995), os primeiros autores que estudaram a caracterização de padrões espaciais, usando a distribuição binomial, foram Cochran (1936) e Todd (1940). Desde então, vários métodos estatísticos baseados em contagens por unidade amostral têm sido utilizados, tais como o cálculo de índices de

dispersão (Upton e Fingleton, 1985; Campbell e Madden, 1990; Madden e Hughes, 1995 e Bergamin Filho et al., 2002), do coeficiente de correlação intraclasse (Ridout e Xu, 2000; Ridout, Demétrio e Firth, 1999; Xu e Ridout, 1998, 2000, 2001) e de técnicas de determinação de autocorrelação espacial (Campbell e Madden, 1990; Gottwald, 1995; Madden e Hughes, 1995; Ridout e Xu, 2000; Turechek e Madden, 2001; Xu e Ridout, 1998, 2001). Esses fornecem medidas diretas da heterogeneidade entre as unidades amostrais, que são funções da associação espacial dos indivíduos nessas unidades.

Para a caracterização do padrão espacial, a área sob estudo é, em um método comumente utilizado na prática, dividida em subáreas chamadas *quadrats*, de uma forma sistemática e exaustiva, nas quais são determinados os números de plantas doentes (Campbell e Madden, 1990). Entretanto, dificuldades surgem na localização e na determinação da forma e do tamanho dos *quadrats*. Vários autores estudaram o efeito do tamanho e da forma do *quadrat* sobre as conclusões finais em relação ao padrão espacial da doença e sobre as estatísticas usadas. Assim, por exemplo, Ridout e Xu (2000) mostraram que o índice de dispersão D , definido como a razão entre a variância do número de plantas infectadas por *quadrat* e a variância teórica da distribuição binomial, é maior no caso de *quadrats* quadrados e se torna menor à medida que o *quadrat* se torna alongado ou retangular, recomendando que se evite o uso de *quadrats* extremamente alongados. Afirmaram, porém, que o efeito do tamanho do *quadrat* é mais pronunciado do que o da forma, sendo o índice de dispersão aumentado, aproximadamente de forma linear, com o aumento do tamanho do *quadrat*. Xu e Ridout (2000), considerando a correlação intraclasse e a forma binária da lei de potência e Xu e Ridout (2001), usando o coeficiente de autocorrelação espacial, também concluíram que o tamanho do *quadrat* tem influência maior do que a forma, sobre as estatísticas espaciais usadas.

O presente trabalho teve como objetivo a comparação de métodos baseados em contagem em *quadrats* que envolvem o cálculo de índices de dispersão, coeficientes de correlação intraclasse e coeficientes de autocorrelação espacial, por meio de um estudo de simulação. Os dados foram simulados considerando-se diferentes níveis de incidência da doença e intensidades de padrões espaciais. Na Seção 2, inicialmente são apresentados os métodos para detecção de padrões espaciais, e na sequência a metodologia utilizada para gerar simulações dos diferentes padrões, encerrando com a descrição do estudo de simulação. Na Seção 3, são apresentados e discutidos os resultados obtidos neste estudo e, finalmente, as conclusões encerram o texto.

2 Metodologia

2.1 Descrição dos métodos comparados

2.1.1 Índices de dispersão

Um primeiro índice de dispersão surgiu a partir da suposição de que a variável aleatória (v.a.) X , que representa o número de plantas doentes por *quadrat*, tem distribuição Poisson com intensidade de infecção λ , isto é, $X \sim P(\lambda)$ com $E(X) =$

$V(X) = \lambda$, se a disseminação da doença ocorresse de modo aleatório. Dessa forma, define-se o índice VM pela relação entre a variância e média estimada, isto é,

$$VM = \frac{s^2}{m}, \quad (1)$$

em que $m = \sum_{j=1}^N x_j/N$, $s^2 = \sum_{j=1}^N (x_j - m)^2/(N - 1)$ e x_1, x_2, \dots, x_N são as contagens de plantas doentes nos N quadrats.

O índice VM é usado como uma medida de agregação, em que $VM = 1$ representa um padrão espacial aleatório, enquanto $VM > 1$ e $VM < 1$ representam padrões espaciais agregado e regular, respectivamente. De acordo com Upton e Fingleton (1985), a estatística $(N - 1)VM$ tem aproximadamente distribuição χ_{N-1}^2 sob a pressuposição de aleatoriedade, o que permite testar a hipótese de que o índice de dispersão é igual a 1. Entretanto, VM é muito sensível à escolha do tamanho do *quadrat* e da incidência média da doença (Upton e Fingleton, 1985; Campbell e Madden, 1990).

Segundo Van Maanen e Xu (2003), com base na distribuição Poisson, VM pode ser visto como uma relação entre variância observada e esperada que pode ser generalizada como um índice de dispersão, D . Assim, supondo-se que a v.a. Y_i , $i = 1, 2, \dots, n$ representa se a planta i está doente ou não, tem-se que $Y_i \sim \text{Bernoulli}(\pi)$ com $E(Y_i) = \pi$ e $\text{Var}(\pi) = \pi(1 - \pi)$, sendo que π é a probabilidade constante de cada planta estar doente. Considerando a situação em que n plantas são avaliadas em um

quadrat, a v.a. número de plantas doentes por *quadrat* $X = \sum_{i=1}^n Y_i$ tem distribuição

binomial, isto é, $X \sim \text{Bin}(n, \pi)$ com $E(X) = n\pi$ e $\text{Var}(X) = n\pi(1 - \pi)$, se as v.a. Y_i 's forem independentes. Define-se D , então por,

$$D = \frac{s^2}{np(1 - p)} \quad (2)$$

em que $s^2 = \sum_{j=1}^N (x_j - m)^2/(N - 1)$ representa a estimativa da variância do número observado de plantas doentes por *quadrat* e $p = \sum_{j=1}^N x_j/(nN)$, a estimativa de π . Da mesma forma que VM , $D = 1$, $D > 1$ e $D < 1$ indicam padrões aleatório, agregado e regular, respectivamente (Madden e Hughes, 1995; Van Maanen e Xu, 2003). Sob a suposição de aleatoriedade, isto é, π é constante, $(N - 1)D \sim \chi_{N-1}^2$ e, portanto, para grandes valores de $(N - 1)D$, rejeita-se a hipótese de padrão aleatório indicando a evidência de superdispersão ou padrão espacial agregado.

Um outro índice encontrado na literatura é o de Morisita, I_δ (Upton e Fingleton, 1985; Campbell e Madden, 1990; Madden e Hughes, 1995; Bergamin Filho et al., 2002), definido por

$$I_\delta = \frac{N \left[\sum_{j=1}^N x_j(x_j - 1) \right]}{\sum_{j=1}^N x_j \left(\sum_{j=1}^N x_j - 1 \right)} = \frac{N \left[\sum_{j=1}^N x_j^2 - \sum_{j=1}^N x_j \right]}{\left(\sum_{j=1}^N x_j \right)^2 - \sum_{j=1}^N x_j}. \quad (3)$$

O padrão aleatório corresponde a $I_\delta = 0$ e quanto maior o valor de I_δ , maior a agregação das plantas doentes. Segundo Campbell e Madden (1990) a estatística

$$I_\delta^* = I_\delta \left(\sum_{j=1}^N x_j - 1 \right) + N - \sum_{j=1}^N x_j,$$

tem distribuição χ_{N-1}^2 .

2.1.2 Correlação intraclasse

Um modelo conceitualmente diferente assume, mais realisticamente, que as respostas individuais de plantas no mesmo *quadrat* não são independentes. Assumindo-se um coeficiente de correlação constante ρ entre as v.a. Y_i , isto é, que os pares de plantas dentro de um *quadrat* sejam igualmente correlacionados, tem-se que $Cov(Y_i, Y_{i'}) = \rho\pi(1 - \pi)$ e $Var(X = \sum_{i=1}^n Y_i) = n\pi(1 - \pi)[1 + \rho(n - 1)]$, para $-1/(n - 1) < \rho < 1$ (Hinde e Demétrio, 1998; Ridout e Xu, 2000). Esse tipo de variância pode ainda ser obtido de uma outra forma, a partir de uma extensão da distribuição beta-binomial (Prentice, 1986).

O parâmetro ρ é chamado coeficiente de correlação intraclasse e fornece uma medida de similaridade entre as plantas dentro dos *quadrats* (Ridout, Demétrio e Firth, 1999; Van Maanen e Xu, 2003). Assim, se $\rho = 0$, tem-se um padrão espacial aleatório e se $\rho > 0$, tem-se um padrão de agregação espacial, mostrando uma superdispersão em relação à distribuição binomial. Valores negativos para ρ raramente ocorrem na prática e correspondem a uma subdispersão em relação ao modelo binomial, indicando que a presença de uma planta doente diminui a chance de que plantas mais próximas também estejam doentes (Ridout e Xu, 2000).

Diferentes estimadores do coeficiente de correlação intraclasse têm sido propostos como mostram Ridout, Demétrio e Firth (1999). Fleiss (1971) e Gibson e Austin (1996) utilizaram o coeficiente de correlação intraclasse na caracterização de padrões espaciais, considerando dados provenientes de *quadrats*. Um estimador para o coeficiente de correlação intraclasse ρ , apresentado por Van Maanen e Xu (2003) é dado por

$$\hat{\rho}_F = 1 - \frac{1}{\hat{p}(1 - \hat{p})nN(n - 1)} \sum_{j=1}^N x_j(n - x_j), \quad (4)$$

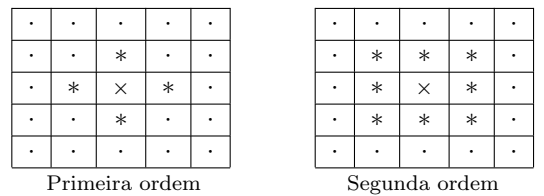
em que N é o número de *quadrats*, n é o total de plantas por *quadrat* e x_j é o número de plantas doentes no *quadrat* j . Um estimador alternativo que relaciona ρ com o índice de dispersão D , definido em (2), é dado por

$$\hat{\rho}_D = \frac{\hat{D} - 1}{n - 1}. \quad (5)$$

A relação com a beta-binomial estendida sugere ainda que se pode utilizar o teste da razão de verossimilhança para o teste da hipótese de padrão aleatório, isto é, $\rho = 0$.

2.1.3 Medidas de autocorrelação espacial

O estudo da autocorrelação espacial é utilizado para caracterizar a dependência espacial da incidência de uma doença, considerando diferentes distâncias e padrões de proximidade das unidades amostrais (Campbell e Madden, 1990; Cliff e Ord, 1981; Upton e Fingleton, 1985; Xu e Ridout, 2000). A idéia é comparar o número de plantas doentes de cada *quadrat* com o número de plantas doentes dos *quadrats* mais próximos, obedecendo a padrões de proximidade pré-estabelecidos, determinados pelo tipo de conexão e pela distância entre os *quadrats*. Para arranjos quadrangulares ou retangulares de plantas, em geral, são usados os padrões de vizinhança de primeira e segunda ordens (Gumpertz; Graham e Ristaino, 1997), sendo que a vizinhança de primeira ordem é composta pelas quatro plantas ou *quadrats* adjacentes na linha e entre linha da planta ou *quadrat* de referência e a vizinhança de segunda ordem inclui também vizinhos nas posições diagonais, como ilustrado no esquema a seguir,



em que $\boxed{\times}$ representa um *quadrat* de referência, $\boxed{*}$, os *quadrats* considerados na comparação e $\boxed{\cdot}$, os demais.

Um estimador para o coeficiente de autocorrelação espacial, denominado estatística I de Moran (Ripley, 1981; Upton e Fingleton, 1985), é dado por

$$\hat{I} = \frac{N}{\sum_i \sum_j W_{ij}} \frac{\sum_i \sum_j W_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sum_i (x_i - \bar{x})^2},$$

em que x_i e x_j referem-se ao número de plantas doentes nos *quadrats* i e j , respectivamente, $i, j = 1, 2, \dots, N$, $i \neq j$ e $W_{ij} = 1$, se os *quadrats* i e j são vizinhos, de acordo com o padrão de proximidade escolhido e $W_{ij} = 0$, caso contrário. Tem-se que $E(I) = -1/(N-1)$ e $E(I^2) = (N^2 S_1 - N S_2 + 3 S_0^2) / [(N^2 - 1) S_0^2]$, sob a suposição de padrão espacial aleatório, uma vez que S_0 , S_1 e S_2 , para o padrão de proximidade de primeira ordem, são dados por

$$S_0 = 2(2rc - r - c), \quad S_1 = 2S_0 \quad \text{e} \quad S_2 = 8(8rc - 7r - 7c + 4)$$

e, para o de segunda ordem, por

$$S_0 = 2(4rc - 3r - 3c + 2), \quad S_1 = 2S_0 \quad \text{e} \quad S_2 = 8(32rc - 39r - 39c + 46),$$

em que r e c são os números de linhas e de colunas de *quadrats*, respectivamente. Por exemplo, considerando um látice 32×32 e dividindo-o em *quadrats* 4×4 , isto

é, com quatro linhas e quatro colunas, tem-se um total de $8 \times 8=64$ *quadrats*, isto é, $r = 8$ e $c = 8$.

Valores de I próximos de 0 indicam padrões espaciais aleatórios, enquanto valores próximos de 1, indicam uma autocorrelação positiva entre as unidades amostrais, ou seja, padrões espaciais agregados. Segundo Ripley (1981), Upton e Fingleton (1985) e Campbell e Madden (1990), pode-se assumir que I tenha distribuição assintoticamente normal e testes de hipóteses podem ser utilizados para verificar se o padrão espacial é aleatório, ou não.

2.2 Modelo para simulação de padrões espaciais

As medidas descritas na seção anterior para detecção de padrões espaciais podem ser avaliadas e comparadas por estudos de simulação. Nesses estudos, são geradas replicações com padrões espaciais conhecidos, para as quais são obtidas as medidas. Para isso é necessário adotar um modelo que permita gerar simulações dos padrões desejados. A seguir descreve-se o modelo utilizado neste trabalho.

Para a simulação de doenças em plantas, classificadas como doentes ou sadias, são construídas grades regulares (*látices*) de dimensão $l \times c$, em que l é o número de linhas e c é o número de colunas ou de plantas por linha. Esses látices representam áreas ou talhões retangulares com espaçamentos entre linhas e entre plantas dentro das linhas fixados pelo pesquisador, sendo o espaçamento nas linhas menor do que ou igual aos das colunas. O modelo aqui adotado para gerar padrões espaciais é seqüencial, e em cada passo simula-se a posição de uma nova planta doente. Para isso, parte-se de uma infecção inicial e calculam-se as probabilidades de cada planta sadia tornar-se doente a partir da pressão infectiva sofrida pelas plantas sadias e, a partir dessas probabilidades, simula-se a posição de uma nova infecção. A seguir, todas as probabilidades são atualizadas levando-se agora em consideração também a posição da nova planta doente. A pressão de infecção é exercida pelas plantas doentes da mesma área e, às vezes, por outras fontes externas de infecção. Dessa forma, novas infecções são simuladas até que se atinja uma incidência desejada. O procedimento consiste, portanto, em uma simulação estocástica pelo método de Monte Carlo em que a simulação de observações é feita a partir de uma distribuição de probabilidades que reflete o padrão espacial especificado. Para a infecção inicial pode-se adotar uma única planta ou um grupo de plantas doentes, sendo neste trabalho, para cada simulação, escolhida aleatoriamente uma única planta entre todas as plantas do látice.

Supondo que y e x representam, respectivamente, uma planta doente e uma sadia, localizadas nas coordenadas (y_1, y_2) e (x_1, x_2) , o que significa que y está na linha y_1 e coluna y_2 e x , na linha x_1 e coluna x_2 , define-se a pressão infectiva sofrida por x como:

$$F_{\alpha}(y-x) = \alpha_1 + |y-x|^{-2\alpha_2}, \quad (6)$$

em que $|y-x| = \sqrt{(y_1-x_1)^2 + (y_2-x_2)^2}$ representa a distância entre y e x , α_1 é o parâmetro que representa uma fonte primária de infecção, ou seja, está relacionado à chance de que uma planta sadia seja infectada por fontes fora de sua área ou

talhão e α_2 é o parâmetro que está associado à transmissão local da doença, ou seja, de uma planta doente para uma sadia, presentes na mesma área ou talhão, sendo que $\alpha_1, \alpha_2 \geq 0$.

Portanto, a função $F_{\alpha}(y-x)$, apresentada em (6), descreve a pressão infectiva sofrida pelas plantas sadias, oriunda de fontes diferentes de infecção. Essa função é dependente da distância entre plantas doente e sadia e é determinada por um vetor com dois parâmetros diferentes, $\alpha' = [\alpha_1 \ \alpha_2]$. O termo $|y-x|^{-2\alpha_2}$ representa a chance de que uma planta sadia x seja infectada por uma planta doente y do mesmo talhão e, para um determinado valor de $\alpha_2 > 0$, esta chance diminui à medida que a distância entre x e y aumenta.

A probabilidade $P(x_i)$ da i -ésima planta ser infectada é dada por:

$$P(x_i) = \frac{r_{\alpha}(x_i)}{\sum_{i=1}^n r_{\alpha}(x_i)},$$

em que $n = lxc$, $r_{\alpha}(x_i) = \alpha_1 + \sum_{y \in D(t)} |y-x|^{-2\alpha_2}$ e $D(t)$ é o conjunto de todas as

plantas doentes em um determinado tempo t . O termo do somatório, em relação a todas as plantas doentes $y \in D(t)$, indica que as plantas doentes exercem influência sobre uma planta sadia x independentemente. Se uma determinada planta já está doente, a probabilidade que ela seja infectada é nula, ou seja, $P(x_i) = 0$, o que corresponde a $r_{\alpha}(x_i) = 0$.

Para simular uma nova infecção, basta gerar um valor u , a partir de uma distribuição $U(0,1)$, determinar a função de distribuição acumulada, dada por $F(x_i) = \sum_{i=1}^n P(x_i)$ e a primeira planta x_i , tal que $F(x_i) > u$, corresponde à planta que se tornará infectada.

2.3 O estudo de simulação

Neste trabalho, foram simulados dados considerando quatro tipos de padrões espaciais combinados com quatro níveis de incidência da doença, 5%, 10%, 20% e 30% de plantas doentes totalizando 16 situações diferentes. Para cada uma das 16 situações foram simuladas 1.000 repetições, considerando um látice 32×32 , com um total de 1.024 plantas, com espaçamento igual a 1, tanto entre linhas, quanto entre plantas dentro da linha.

Os quatro tipos de padrões espaciais diferem entre si, de acordo com os parâmetros α_1 e α_2 do modelo definido em (6), ou seja,

1. Padrão espacial aleatório ($\alpha_1 = 0, 1$ e $\alpha_2 = 0, 0$): quando $\alpha_2 = 0$ para cada tempo t , a pressão infectiva é $r_{\alpha}(x) = \alpha_1 + \sum_{y \in D(t)} |y-x|^0 = \alpha_1 + n_d$ em que n_d denota o número de plantas doentes e, para qualquer valor de α_1 , a

pressão infectiva é a mesma para toda planta sadia, definindo uma distribuição uniforme e, portanto, padrão espacial aleatório.

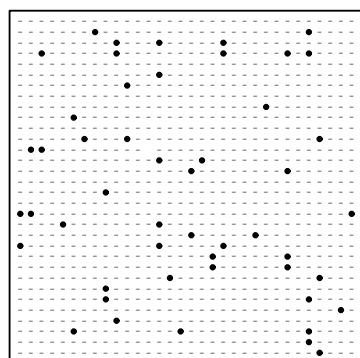
2. Padrão espacial moderadamente agregado com a presença de uma fonte externa de infecção ($\alpha_1 = 0, 1$ e $\alpha_2 = 1, 0$): a infecção ocorre das plantas doentes para as sadias, pois $\alpha_2 > 0$ e existe uma fonte externa de infecção, pois $\alpha_1 > 0$. Entretanto, como o valor de α_2 é pequeno, a influência de plantas doentes sobre as sadias se dá em distâncias maiores, o que implica um padrão moderadamente agregado;
3. Padrão espacial fortemente agregado com a presença de uma fonte externa de infecção ($\alpha_1 = 0, 1$ e $\alpha_2 = 2, 0$): α_2 é relativamente grande e, o efeito exercido por uma planta doente sobre suas vizinhas sadias diminui rapidamente com a distância, isto é, além da fonte externa, a infecção ocorre apenas entre plantas muito próximas;
4. Padrão fortemente agregado sem a presença de uma fonte externa de infecção ($\alpha_1 = 0, 0$ e $\alpha_2 = 2, 0$): tem-se apenas a transmissão da doença de plantas infectadas para plantas sadias muito próximas e dentro do mesmo látice.

A Figura 1 mostra exemplos de látices simulados para cada um dos quatro tipos de padrões espaciais considerados neste trabalho, em que são observados os diferentes padrões espaciais que podem ser obtidos considerando um modelo simples, como o apresentado em (6), apenas alterando os valores dos parâmetros α_1 e α_2 .

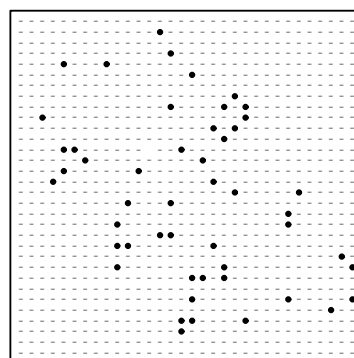
Na Figura 1A observa-se que as plantas doentes estão localizadas de uma maneira bem dispersa por toda a área, o que caracteriza um padrão espacial completamente aleatório. A não ocorrência de plantas doentes em uma determinada parte do látice ou o início de formação de algum agregado é meramente devido ao acaso. Assim, uma determinada planta sadia é infectada, aleatoriamente, por uma planta doente do próprio látice ou por uma fonte de infecção externa, sem apresentar um padrão espacial definido que possa estabelecer um padrão de disseminação da doença.

Um padrão espacial nitidamente diferente do anterior é obtido utilizando-se $\alpha_1 = 0, 0$ e $\alpha_2 = 2, 0$, como pode ser visto na Figura 1D. Esse é visualmente agregado pois para $\alpha_1 = 0$ não há fontes de infecção além das plantas doentes do próprio látice. Desse modo, a transmissão da doença ocorre apenas de uma planta doente para uma planta sadia com ambas localizadas na mesma área ou látice. Além disso, como o valor do parâmetro α_2 é grande, tem-se que a disseminação da doença ocorre apenas a curtas distâncias, ou seja, entre plantas vizinhas muito próximas. Assim, tem-se um padrão espacial agregado muito bem caracterizado pela presença de um único grupo de plantas doentes. Essa obtenção de padrões espaciais bastante agregados, com valor de $\alpha_2 = 2, 0$, confirma a afirmação de Pethybridge e Madden (2003), de que valores de α_2 em torno de 3,0 são considerados extremamente grandes.

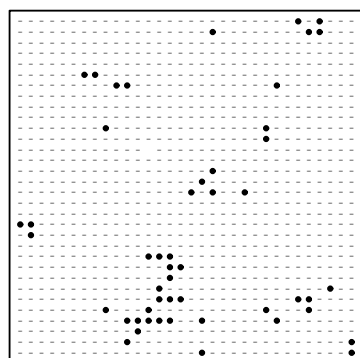
Os dois látices apresentados na Figura 1B e e 1C podem ser considerados como situações intermediárias entre as descritas anteriormente. Pode-se observar



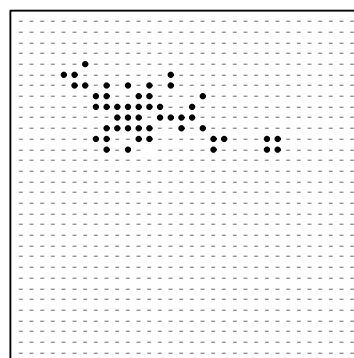
A. $\alpha_1 = 0,1; \alpha_2 = 0$



B. $\alpha_1 = 0,1; \alpha_2 = 1,0$



C. $\alpha_1 = 0,1; \alpha_2 = 2,0$



D. $\alpha_1 = 0; \alpha_2 = 2,0$

Figura 1 - Látices simulados a partir do modelo dado em (6), com diferentes valores para os parâmetros α_1 e α_2 , onde os símbolos - e ● representam plantas sadias e doentes, respectivamente.

que as plantas doentes tendem a se localizar mais próximas umas das outras e não estão aleatoriamente espalhadas por toda a área. Existe uma fonte externa de infecção, pois $\alpha_1 > 0$, o que leva ao aparecimento de vários focos espalhados pela área toda. Como α_2 também é maior do que zero também existe a infecção de plantas sadias por plantas doentes dentro da própria área. A diferença entre esses dois padrões é que na Figura 1C, o parâmetro α_2 é maior, o que simula uma situação em que a transmissão da doença ocorre entre plantas bem mais próximas, levando à formação de pequenos grupos ou agregados mais bem definidos do que em 1B,

em que a infecção das plantas sadias pelas doentes ocorre em distâncias maiores, aproximando-se mais de um padrão espacial aleatório. Assim, tem-se nas Figura 1C e 1D, padrões espaciais moderado e fortemente agregado, respectivamente, ambos na presença de uma fonte externa de infecção. Discussões mais detalhadas sobre os parâmetros do modelo apresentado em (6) e sobre o processo de simulação, podem ser encontradas em Lima (2005).

Para as sete estatísticas de caracterização do padrão espacial usadas neste artigo, foram utilizados *quadrats* 4×4 e 8×8 . Portanto, tem-se um total de 32 situações diferentes de análise: 4 níveis de incidência da doença, 4 padrões espaciais diferentes e 2 tamanhos de *quadrats*. Para cada uma das estatísticas foram obtidos as médias e os erros-padrão das 1.000 estimativas de cada uma das 32 situações. Todo o procedimento de simulação e análise pelas diversas medidas de agregação foi implementado em código próprio, utilizando linguagem Fortran 90.

3 Resultados e discussão

Médias e erros-padrão das medidas de agregação espacial obtidas para as simulações são apresentadas nas Tabelas 1 e 2. A Tabela 3 apresenta as porcentagens de padrões espaciais indicados como significativamente diferentes de um padrão espacial aleatório, baseando-se nas distribuições χ^2 , para os índices de dispersão e normal, no caso do coeficiente de autocorrelação espacial.

De maneira geral, observa-se nas Tabelas 1 e 2 que os valores médios para os três índices de dispersão, dois coeficientes de correlação intraclasse e dois coeficientes de autocorrelação espacial se comportam da forma esperada, ou seja, à medida que as plantas doentes se apresentam mais agregadas, identifica-se um aumento nos valores médios das estimativas.

No padrão espacial aleatório, as médias das estimativas dos três índices de dispersão foram próximas de um e dos coeficientes de correlação intraclasse e autocorrelação espacial foram nulas ou bem próximas de zero para a maioria das situações simuladas, conforme esperado para esse tipo de padrão espacial (Upton e Fingleton, 1985; Campbell e Madden, 1990; e Bergamin Filho et al., 2002). Um pequeno viés negativo pode ser observado nos índices VM e I_δ quando se aumenta o número de plantas doentes, sendo mais acentuado para o primeiro. Porém, conforme pode ser observado na Tabela 3, essa diminuição não afeta a caracterização do padrão espacial como aleatório, pelo contrário, à medida que se aumenta o número de plantas doentes, tem-se uma melhor caracterização desse padrão espacial, quando se utilizam os índices de dispersão.

No padrão espacial fortemente agregado sem a presença de uma fonte externa de infecção (4), em que se têm os maiores valores estimados das sete estatísticas conforme Tabelas 1 e 2, garante-se que praticamente todos os látices simulados sejam classificados como padrão espacial não aleatório (Tabela 3). Uma exceção é o caso do coeficiente de autocorrelação espacial, estimado considerando *quadrats* 8×8 , com pior desempenho quando se tem uma baixa incidência de doença (5%). Entretanto, à medida que a incidência da doença aumenta, ocorre um aumento

Tabela 1 - Médias e erros-padrão das 1.000 estimativas dos índices de dispersão VM , I_δ de Morisita e D , obtidas a partir dos quatro padrões espaciais simulados

Padrão espacial	Quadrat	% plantas doentes	Média (erro-padrão)		
			VM	I_δ	D
(1) $\alpha_1 = 0,1$ $\alpha_2 = 0,0$	4×4	5	0,96 (0,16)	0,95 (0,21)	1,01 (0,17)
		10	0,90 (0,16)	0,93 (0,09)	0,99 (0,17)
		20	0,80 (0,13)	0,93 (0,04)	1,00 (0,17)
		30	0,71 (0,12)	0,94 (0,02)	1,00 (0,17)
	8×8	5	0,97 (0,35)	0,99 (0,11)	1,01 (0,37)
		10	0,91 (0,34)	0,98 (0,05)	1,01 (0,37)
		20	0,80 (0,28)	0,98 (0,02)	0,99 (0,35)
		30	0,71 (0,25)	0,98 (0,01)	1,00 (0,35)
(2) $\alpha_1 = 0,1$ $\alpha_2 = 1,0$	4×4	5	1,68 (0,39)	1,88 (0,50)	1,77 (0,41)
		10	1,81 (0,37)	1,51 (0,23)	2,01 (0,41)
		20	1,81 (0,33)	1,26 (0,10)	2,26 (0,41)
		30	1,66 (0,31)	1,14 (0,06)	2,35 (0,45)
	8×8	5	2,47 (0,98)	1,45 (0,30)	2,60 (1,03)
		10	3,00 (1,10)	1,30 (0,16)	3,32 (1,22)
		20	3,28 (1,23)	1,17 (0,09)	4,08 (1,53)
		30	3,12 (1,13)	1,11 (0,05)	4,41 (1,61)
(3) $\alpha_1 = 0,1$ $\alpha_2 = 2,0$	4×4	5	1,98 (0,41)	2,26 (0,52)	2,08 (0,43)
		10	2,43 (0,43)	1,91 (0,27)	2,69 (0,48)
		20	2,76 (0,43)	1,55 (0,13)	3,43 (0,54)
		30	2,72 (0,40)	1,36 (0,08)	3,85 (0,56)
	8×8	5	2,38 (0,94)	1,42 (0,28)	2,51 (0,99)
		10	3,31 (1,21)	1,35 (0,18)	3,66 (1,34)
		20	4,21 (1,47)	1,24 (0,11)	5,23 (1,83)
		30	4,36 (1,40)	1,17 (0,07)	6,17 (1,98)
(4) $\alpha_1 = 0,0$ $\alpha_2 = 2,0$	4×4	5	8,17 (1,41)	10,21 (1,81)	8,59 (1,48)
		10	8,96 (1,15)	6,06 (0,73)	9,93 (1,27)
		20	8,80 (0,92)	3,47 (0,29)	10,94 (1,15)
		30	8,05 (0,73)	2,48 (0,15)	11,39 (1,03)
	8×8	5	21,29 (7,59)	7,21 (2,32)	22,37 (7,98)
		10	27,44 (6,71)	5,01 (1,02)	30,41 (7,43)
		20	29,64 (5,32)	3,16 (0,40)	36,83 (6,61)
		30	27,83 (4,47)	2,35 (0,22)	39,36 (6,32)

Tabela 2 - Médias e erros-padrão das 1.000 estimativas dos coeficientes de correlação intraclasse (ρ_F e ρ_D) e dos I de Moran de primeira ordem (I_1) e de segunda ordem (I_2), obtidas a partir dos quatro padrões espaciais simulados

Padrão espacial	Quadrat	% plantas doentes	Média (erro-padrão)			
			ρ_F	ρ_D	I_1	I_2
(1) $\alpha_1 = 0,1$ $\alpha_2 = 0,0$	4×4	5	0,00 (0,01)	0,00 (0,01)	-0,01 (0,08)	-0,01 (0,06)
		10	0,00 (0,01)	0,00 (0,01)	-0,01 (0,09)	-0,01 (0,06)
		20	0,00 (0,01)	0,00 (0,01)	-0,01 (0,09)	-0,02 (0,06)
		30	0,00 (0,01)	0,00 (0,01)	-0,01 (0,09)	-0,01 (0,06)
	8×8	5	0,00 (0,01)	0,00 (0,01)	-0,07 (0,19)	-0,07 (0,12)
		10	0,00 (0,01)	0,00 (0,01)	-0,06 (0,17)	-0,07 (0,11)
		20	0,00 (0,01)	0,00 (0,01)	-0,07 (0,18)	-0,07 (0,11)
		30	0,00 (0,01)	0,00 (0,01)	-0,06 (0,17)	-0,07 (0,12)
(2) $\alpha_1 = 0,1$ $\alpha_2 = 1,0$	4×4	5	0,05 (0,02)	0,05 (0,03)	0,08 (0,09)	0,16 (0,11)
		10	0,07 (0,02)	0,07 (0,02)	0,13 (0,10)	0,08 (0,08)
		20	0,08 (0,02)	0,08 (0,02)	0,19 (0,09)	0,12 (0,09)
		30	0,09 (0,02)	0,09 (0,03)	0,22 (0,10)	0,14 (0,09)
	8×8	5	0,02 (0,01)	0,03 (0,01)	-0,01 (0,17)	-0,04 (0,12)
		10	0,03 (0,02)	0,04 (0,02)	-0,01 (0,18)	-0,03 (0,13)
		20	0,04 (0,02)	0,05 (0,02)	0,02 (0,17)	-0,02 (0,13)
		30	0,05 (0,02)	0,05 (0,02)	0,03 (0,17)	-0,01 (0,12)
(3) $\alpha_1 = 0,1$ $\alpha_2 = 2,0$	4×4	5	0,07 (0,02)	0,07 (0,02)	0,16 (0,11)	0,12 (0,10)
		10	0,11 (0,03)	0,11 (0,03)	0,21 (0,11)	0,16 (0,10)
		20	0,16 (0,04)	0,16 (0,04)	0,27 (0,11)	0,22 (0,11)
		30	0,19 (0,04)	0,19 (0,04)	0,31 (0,12)	0,25 (0,12)
	8×8	5	0,02 (0,01)	0,02 (0,01)	0,07 (0,18)	0,01 (0,13)
		10	0,04 (0,01)	0,04 (0,02)	0,12 (0,18)	0,05 (0,14)
		20	0,06 (0,02)	0,07 (0,03)	0,17 (0,18)	0,08 (0,15)
		30	0,07 (0,03)	0,08 (0,03)	0,20 (0,17)	0,10 (0,14)
(4) $\alpha_1 = 0,0$ $\alpha_2 = 2,0$	4×4	5	0,50 (0,10)	0,51 (0,10)	0,55 (0,06)	0,44 (0,06)
		10	0,58 (0,08)	0,59 (0,08)	0,69 (0,05)	0,60 (0,06)
		20	0,65 (0,07)	0,66 (0,08)	0,78 (0,05)	0,71 (0,06)
		30	0,68 (0,07)	0,69 (0,07)	0,80 (0,05)	0,74 (0,06)
	8×8	5	0,32 (0,12)	0,34 (0,13)	0,21 (0,12)	0,10 (0,10)
		10	0,44 (0,11)	0,47 (0,12)	0,33 (0,09)	0,19 (0,10)
		20	0,53 (0,10)	0,57 (0,10)	0,49 (0,08)	0,34 (0,11)
		30	0,57 (0,09)	0,61 (0,10)	0,55 (0,09)	0,41 (0,11)

na porcentagem de padrões espaciais caracterizados corretamente. Esse resultado evidencia os efeitos do tamanho do *quadrat* e do número de plantas doentes sobre o coeficiente de autocorrelação I de Moran.

Comparando-se os três índices de dispersão, pode-se observar na Tabela 3 a concordância entre os resultados de VM e I_δ , sendo suas porcentagens um pouco menores do que as obtidas com D , nos casos de padrão espacial aleatório e moderadamente agregado com a presença de uma fonte externa de infecção. Essa diferença de D para VM e I_δ pode ser explicada pelo fato de que esses dois últimos são baseados no mesmo conceito básico, da relação variância-média, enquanto D foi proposto considerando a relação variância observada com a variância esperada sob a condição de um padrão espacial aleatório. Uma conclusão interessante, uma vez que os índices foram estimados para os mesmos látices simulados, é que o índice D caracterizou uma maior porcentagem de padrões aleatórios e moderadamente agregados como agregados.

Na Tabela 3, pode-se ainda verificar que a porcentagem de classificação de padrões espaciais agregados e moderadamente agregados, com uma fonte externa de infecção (padrões 2 e 3), como padrões aleatórios é bem maior quando se utiliza o coeficiente de autocorrelação espacial I de Moran, tanto de primeira como de segunda ordem. Esse desempenho pior indica que esse coeficiente não detectou o pequeno grau de agregação das plantas doentes, principalmente quando se têm níveis menores de infecção ou menor incidência da doença. Porém, o teste utilizado considera que o coeficiente I de Moran tem distribuição assintoticamente normal, o que, de acordo com Assunção (1999), não é uma boa pressuposição. Esse autor sugere a utilização dos testes aleatorizados que, com os recursos computacionais atualmente disponíveis, seriam a melhor escolha para realizar testes de significância no caso de autocorrelação espacial evitando uso de distribuições aproximadas. Porém, como em outros testes utilizados neste trabalho foram consideradas distribuições assintóticas, optou-se por manter a aproximação para o coeficiente I de Moran.

Com relação às respostas obtidas considerando diferentes níveis de incidência de uma doença em plantas, observa-se na Tabela 3 que existe uma tendência de a caracterização equivocada diminuir com o aumento no número de plantas infectadas, pelo menos para as incidências consideradas aqui, entre 5% e 30%; com exceção do padrão espacial 4 ou fortemente agregado. Xu e Ridout (1998) verificaram que a correlação intraclasse e a autocorrelação espacial aumentam com a incidência da doença, até atingir um máximo e diminuem à medida que a incidência da doença se aproxima de 100%, comportamento que também foi verificado neste trabalho (Tabela 2). Este mesmo comportamento foi observado para as estimativas médias dos índices de dispersão (Tabela 1). Portanto, se o número de plantas doentes é muito pequeno ou muito grande, as estatísticas utilizadas neste estudo não são adequadas para se caracterizarem padrões espaciais, pois o número de plantas infectadas será insuficiente para a formação de grupos que caracterizem os agregados ou a área apresentará uma taxa de incidência da doença tão elevada, que não há sentido em se falar na descrição do padrão espacial da doença.

Tabela 3 - Porcentagens de resultados significativamente diferentes do padrão espacial aleatório, para VM , I_δ de Morisita, D , I de Moran de primeira ordem (I_1) e segunda ordem (I_2), considerando 1.000 simulações dos quatro padrões espaciais

Padrão espacial	<i>Quadrat</i>	% plantas doentes	% significativamente diferentes					
			VM	I_δ	D	I_1	I_2	
(1) $\alpha_1 = 0, 1$ $\alpha_2 = 0, 0$	4×4	5	3,3	3,3	4,5	4,3	5,6	
		10	1,0	1,0	4,4	5,7	6,1	
		20	0,3	0,3	3,8	4,7	4,7	
		30	0,0	0,0	3,6	4,3	5,8	
	8×8	5	4,1	4,1	6,3	5,8	7,2	
		10	2,7	2,7	6,8	3,1	6,1	
		20	0,5	0,5	4,1	3,8	5,6	
		30	0,2	0,2	4,0	4,9	7,0	
	(2) $\alpha_1 = 0, 1$ $\alpha_2 = 1, 0$	4×4	5	84,7	84,7	88,6	27,4	25,5
			10	94,3	94,3	97,7	48,0	40,9
			20	95,3	95,3	99,9	71,5	60,3
			30	88,8	88,8	99,9	80,8	70,4
8×8		5	80,1	80,1	83,5	8,3	7,9	
		10	90,9	90,9	94,0	10,1	11,6	
		20	92,9	92,9	97,3	11,1	13,1	
		30	92,4	92,4	98,9	11,2	13,4	
(3) $\alpha_1 = 0, 1$ $\alpha_2 = 2, 0$		4×4	5	97,8	97,8	98,8	55,8	58,1
			10	99,9	99,9	100	73,6	74,0
			20	100	100	100	87,6	87,5
			30	100	100	100	91,9	90,6
	8×8	5	77,5	77,5	81,7	19,5	21,6	
		10	95,3	95,3	97,3	29,6	28,6	
		20	99,1	99,1	99,6	37,5	37,2	
		30	99,1	99,1	100	47,8	42,7	
	(4) $\alpha_1 = 0, 0$ $\alpha_2 = 2, 0$	4×4	5	100	100	100	100	100
			10	100	100	100	100	100
			20	100	100	100	100	100
			30	100	100	100	100	100
8×8		5	100	100	100	44,5	35,0	
		10	100	100	100	90,5	79,8	
		20	100	100	100	99,0	92,4	
		30	100	100	100	99,4	97,0	

Os tamanhos de *quadrats* utilizados nas análises exercem influência sobre as estimativas médias obtidas (Tabelas 1 e 2), quando se tem algum grau de agregação das plantas doentes. Nota-se que, quando se considera um *quadrat* maior nas análises, as estimativas médias dos índices de dispersão VM e D aumentam e, ao contrário, I_δ , ρ_F , ρ_D e I de Moran, diminuem. Resultados semelhantes foram encontrados por Xu e Ridout (2000, 2001), que concluíram que, com o

aumento do tamanho do *quadrat*, tem-se uma diminuição no coeficiente de correlação intraclasse e um aumento no índice de dispersão (D). Conforme salientado por Xu e Ridout (2000), esse resultado é aparentemente contraditório, pois menores estimativas de ρ indicam menor agregação e maiores de D indicam maior agregação. No entanto, de acordo com esses mesmos autores, isso enfatiza a importância de se interpretarem corretamente as estatísticas espaciais. Apesar das estimativas de VM e D tenderem a ser maiores para *quadrats* 8×8 , enquanto o I_δ apresenta comportamento contrário, a partir dos resultados apresentados na Tabela 3, observa-se que as porcentagens para VM e I_δ são exatamente as mesmas.

Nos látices simulados com algum grau de agregação, na presença de uma fonte externa de infecção (padrões espaciais 2 e 3), classifica-se um número menor de casos como aleatório nas análises com *quadrats* 4×4 . Esse efeito é mais evidente no caso dos coeficientes de autocorrelação de Moran (I_1 e I_2) e pode ser devida à uma diminuição do poder do teste aplicado, uma vez que, nesse caso o tamanho da amostra é muito maior, ou seja, 64 *quadrats* contra 16 para o caso de *quadrats* 8×8 . Porém, com os resultados apresentados, não é possível identificar se a pior *performance*, quando se consideram *quadrats* maiores, é devida ao tamanho do *quadrat* ou a uma diminuição do tamanho da amostra. Segundo Xu e Ridout (2000, 2001), estatísticas espaço-temporais são muito influenciadas pelo tamanho e forma do *quadrat*.

Conclusões

Apesar de se ter uma caracterização correta do padrão espacial agregado na maioria das situações estudadas, utilizando-se os índices de dispersão, coeficientes de correlação intraclasse e autocorrelação espacial, problemas foram detectados quando o nível de agregação não é muito forte, ou seja, quando as plantas doentes não estão muito agregadas e formando grupos muito bem definidos. Esses resultados estão coerentes com trabalhos anteriores como Xu e Ridout (1998, 2000 e 2001) e confirmam os efeitos do tamanho do *quadrat* e níveis de incidência da doença sobre as estimativas de estatísticas utilizadas na caracterização do padrão espacial. Porém, deve-se ter muito cuidado na interpretação de tais estatísticas e dos resultados dos testes aplicados, para os quais são consideradas distribuições assintóticas. Trabalhos futuros devem ser conduzidos, buscando avaliar especificamente o efeito de se considerarem distribuições assintóticas.

Não foi testada a hipótese de $\rho = 0$, uma vez que não foi encontrado na literatura testes para o coeficiente de correlação intraclasse que pudessem ser aplicados de uma maneira rápida e direta como para as demais estatísticas consideradas. Neste caso, uma alternativa seria utilizar teste da razão de verossimilhança baseado na distribuição beta-binomial estendida, porém, este teste pode apresentar instabilidades numéricas e comprometer o estudo de simulação. Portanto, alternativas mais simples para se testar essa hipótese devem ser estudadas.

Este trabalho concentrou-se nas metodologias baseadas em *quadrats* mas outras abordagens são possíveis, tais como: técnicas baseadas em processos pontuais (Amorin et al., 2007) e ajustes de modelos, como os autológicos (Gumpertz,

Graham e Ristaino, 1997) ou de campos aleatórios gaussianos (Ortiz et. al., 2004). Além disso, com as estatísticas estudadas não é possível discriminar entre padrões espaciais agregados diferentes, como os obtidos nos padrões espaciais simulados 3 ($\alpha_1 = 0,1; \alpha_2 = 2,0$) e 4 ($\alpha_1 = 0,0; \alpha_2 = 2,0$). Como complemento podem ser ajustados modelos estocásticos espaço-temporais (Gibson, 1997a; Gibson, 1997b; Lima, 2005), que permitem essa identificação dos tipos diferentes de padrões espaciais agregados.

Agradecimentos

Este trabalho é parte da tese de doutorado do primeiro autor no Departamento de Ciências Exatas, Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, e foi realizado com o apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes), pelo Programa Institucional de Capacitação Docente e Técnica (PICDT) e Programa de Doutorando no País com Estágio no Exterior (PDEE). O segundo autor é Pesquisador 1C do CNPq.

LIMA, R. R. de; DEMÉTRIO, C. G. B.; RIBEIRO JUNIOR, P. J.; RIDOUT, M. S. Comparing quadrat-based methods to characterize spatial patterns of plant diseases. *Rev. Mat. Estat.*, São Paulo, v.24, n.4, p.7-26, 2006.

- **ABSTRACT:** *Information about the spatial-temporal dynamics is of fundamental importance in epidemiological studies for describing and understanding the development of diseases, for developing efficient sampling plans, for planning controlled experiments, for evaluating the effect of different treatments, and for determining crop losses. Several different methods have been used to characterize spatial aggregation, such as the computation of dispersion index, intraclass correlation, and the use of spatial autocorrelation techniques. These methods consider count data from quadrats, which are sample units. With these methods one obtains a direct measure of heterogeneity of disease incidence, which is a function of the spatial interaction of the plants that are inside the quadrats. Some problems were identified by using data simulations, and we can conclude that the characterization of spatial patterns using the indexes considered here must be interpreted with caution, in particular if only one is adopted for a particular data analysis. There are discrepancies between results and their behavior under effects of different quadrat sizes and levels of incidence of the diseases.*
- **KEYWORDS:** *Dispersion index; intraclass correlation; simulation; spatial autocorrelation; spatial pattern.*

Referências

AMORIM, L. et al. Spatial pattern of trees affected by black spot in citrus groves in Brazil. *Plant Dis.*, St. Paul, 2007. In Press

- ASSUNÇÃO, R. M. A New proposal to adjust Moran's I for population density. *Stat. Med.*, Chichester, v.18, p.2147-2162, 1999.
- BASSANEZI, R. B. et al. Spatial and temporal analyses of citrus sudden death as a tool to generate hypotheses concerning its etiology. *Phytopathology*, St. Paul, v.93, n.4, p.502-512, 2003.
- BERGAMIN FILHO, A. et al. Análise espacial de epidemias. *Rev. Anu. Patol. Plantas*, Passo Fundo, v.10, p.155-218, 2002.
- CAMPBELL, C. L.; MADDEN, L. V. *Introduction to plant disease epidemiology*. New York: John Wiley, 1990. 560p.
- CLIFF, A. D.; ORD, J. K. *Spatial processes: models and applications*. London: Pion, 1981. 266p.
- COCHRAN, W. G. The statistical analysis of field counts of diseased plants. *J. R. Stat. Soc.*, London, v.3, n.1, p.49-67, 1936. Supplement.
- FERRANDINO, F. J. A distribution-free method for estimating the effects of aggregated plant damage on crop yield. *Phytopathology*, St. Paul, v.79, n.11, p.1229-1232, 1989.
- FLEISS, J. L. Measuring nominal scale agreement among many raters. *Psychol. Bull.*, Washington, v.76, p.378-382, 1971.
- GIBSON, G. V. Investigating mechanisms of spatiotemporal epidemic spread using stochastic models. *Phytopathology*, St. Paul, v.87, n.2, p.139-146, 1997a.
- GIBSON, G. V. Markov Chain Monte Carlo methods for fitting spatiotemporal stochastic models in plant epidemiology. *J. R. Stat. Soc. Sr. C: Appl. Stat.*, London, v.46, n.2, p.215-233, 1997b.
- GIBSON, G. V.; AUSTIN, E. J. Fitting and testing spatiotemporal stochastic models with application in plant epidemiology. *Plant Pathol.*, Oxford, v.45, p.172-184, 1996.
- GOTTWALD, T. R. Spatio-temporal analysis and isopath dynamics of citrus scab in nursery plots. *Phytopathology*, St. Paul, v.85, n.10, p.1082-1092, 1995.
- GOTTWALD, T. R. et al. Examination of the effect of aphid vector population composition on the spatial dynamics of citrus tristeza virus spread by stochastic modeling. *Phytopathology*, St. Paul, v.89, n.7, p.603-608, 1999.
- GUMPERTZ, M. L.; GRAHAM, J. M.; RISTAINO, J. B. Autologistic model of spatial pattern of *Phytophthora* epidemic in bell pepper: effects of soil variables on disease presence. *J. Agric., Biol. Environ. Stat.*, Alexandria, v.2, n.2, p.131-156, 1997.
- HINDE, J. P.; DEMÉTRIO, C. G. B. *Overdispersion: models and estimation*. In: SINAPE, 13., 1998, Caxambu. *Anais...* Caxambu: ABE, 1998. 73p.
- HUGHES, G. Spatial heterogeneity in crop loss assessment models. *Phytopathology*, St. Paul, v.78, n.7, p.883-884, 1988.

- HUGHES, G.; MADDEN, L. V. Aggregation and incidence of disease: some implications for sampling. *Aspects Appl. Biol.*, Warwick, v.37, p.25-31, 1994.
- HUGHES, G.; MADDEN, L. V. Some methods allowing for aggregated patterns of disease incidence in the analysis of data from designed experiments. *Plant Pathol.*, Oxford, v.44, n.6, p.927-943, 1995.
- HUGHES, G. et al. Validating mathematical models of plant-disease progress in space and time. *IMA J. Math. Appl. Med. Biol.*, Oxford, v.14, p.85-112, 1997.
- LIMA, R. R. *Modelagem espaço-temporal para dados de incidência de doenças em plantas*. 2005. 149f. Tese (Doutorado em Estatística e Experimentação Agronômica) – Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, 2005.
- MADDEN, L. V. Aggregation of *Colletotrichum acutatum* in response to simulated rain episodes. *J. Phytopathol.*, Berlin, v.138, n.2, p.145-156, 1993.
- MADDEN, L. V.; NUTTER JÚNIOR, F. W. Modeling crop losses at field scale. *Can. J. Plant Pathol.*, Guelph, v.17, n.2, p.124-137, 1995.
- MADDEN, L. V.; HUGHES, G. Plant disease incidence: distributions, heterogeneity, and temporal analysis. *Annu. Rev. Phytopathol.*, Palo Alto, v.33, p.529-564, 1995.
- MADDEN, L. V.; HUGHES, G. Sampling for plant disease incidence. *Phytopathology*, St. Paul, v.89, n.11, p.1088-1103, 1999.
- MARION, G.; GIBSON, G.; RENSHAW, E. Estimating likelihoods for spatio-temporal models using importance sampling. *Stat. Comp.*, London, v.13, n.2, p.111-119, 2003.
- ORTIZ, G. C. et al. Padrões espaciais na epidemiologia da pinta preta dos citrus. Curitiba: LEG/UFPR, 2004. Relatório Técnico
- PETHYBRIDGE, S. J.; MADDEN, L. V. Analysis of spatiotemporal dynamics of virus spread in an Australian hop garden by stochastic modeling. *Plant Dis.*, St. Paul, v.87, n.1, p.56-62, 2003.
- PRENTICE, R. L. Binary regression using an extended beta-binomial distribution, with discussion of correlation induced by covariate measurement errors. *J. Am. Stat. Assoc.*, Alexandria, v.81, n.394, p.321-327, 1986.
- RIDOUT, M. S.; XU, X.-M. Relationships between several quadrat-based statistical measures used to characterize spatial aspects of disease incidence data. *Phytopathology*, St. Paul, v.90, n.6, p.568-575, 2000.
- RIDOUT, M. S.; DEMÉTRIO, C. G. B.; FIRTH, D. Estimating intraclass correlation for binary data. *Biometrics*, Washington, v.55, n.1, p.137-148, 1999.
- RIPLEY, B. D. *Spatial Statistics*. New York: John Wiley, 1981. 252p.
- TODD, H. A note on random associations in a square point lattice. *J. R. Stat. Soc.*, Cambridge, v.7, n.1, p.78-82, 1940.

- TURECHEK, W. W.; MADDEN, L. V. Analysis of the association between the incidence of two spatially aggregated foliar diseases of strawberry. *Phytopathology*, St. Paul, v.90, n.2, p.157-170, 2000.
- TURECHEK, W. W.; MADDEN, L. V. Effect of scale on plant disease incidence and heterogeneity in a spatial hierarchy. *Ecol. Model.*, Amsterdam, v.144, n.1, p.77-95, 2001.
- UPTON, G. J. G.; FINGLETON, B. *Spatial data analysis by example: point pattern and quantitative data*. New York: John Wiley, 1985. v.1, 410p.
- VAN MAANEN, A.; XU, X. -M. Modelling plant disease epidemics. *Eur. J. Plant Pathol.*, Dordrecht, v.109, n.7, p.669-682, 2003.
- XU, X.-M.; RIDOUT, M. S. Effects of initial epidemic conditions, sporulation rate, and spore dispersal gradient on the spatio-temporal dynamics of plant disease epidemics. *Phytopathology*, St. Paul, v.88, n.10, p.1000-1012, 1998.
- XU, X.-M.; RIDOUT, M. S. Effects of quadrat size and shape, initial epidemic conditions, and spore dispersal gradient on spatial statistics of plant disease epidemics. *Phytopathology*, St. Paul, v.90, n.7, p.738-750, 2000.
- XU, X.-M.; RIDOUT, M. S. Effects of prevailing wind direction on spatial statistics of plant disease epidemics. *J. Phytopathol.*, Berlin, v.149, n.2, p.155-166, 2001.
- ZADOKS, J. C.; VAN DEN BOSCH, F. Expansion and spatial spread of disease. *Annu. Rev. Phytopathol.*, Palo Alto, v.32, p.503-521, 1994.

Recebido em 22.02.2005.

Aprovado após revisão em 10.12.2006.