

Análise Espacial e Temporal de Epidemias de Patógenos de Solo

Armando Bergamin Filho; Bernhard Hau; Lílian Amorim; Francisco Ferraz Laranjeira

A epidemia é um sistema e como tal comporta-se como um todo em resposta a estímulos dirigidos a qualquer parte. Obviamente o patógeno, o hospedeiro e a doença constituem-se em subsistemas do sistema epidemia. Dependendo do interesse do observador, qualquer um dos três subsistemas pode ser elevado à categoria de sistema e estudado isoladamente: isto é fácil de ser visualizado no caso do hospedeiro, que não depende nem do patógeno nem da doença para sua existência, mas também é possível no caso do patógeno, seja quando se considera sua sobrevivência na ausência do hospedeiro, seja quando se considera seu crescimento saprofítico em meio outro que a planta. A doença é a interação entre patógeno e hospedeiro, sob a influência do ambiente e do homem. A análise epidemiológica temporal e espacial, apresentada a seguir, aplica-se indistintamente tanto a epidemias causadas por patógenos aéreos quanto àquelas causadas por patógenos veiculados pelo solo.

Análise temporal

A curva de progresso da doença, usualmente expressa pela plotagem da proporção de doença *versus* tempo, é a melhor representação de uma epidemia. Por intermédio dela, interações entre patógeno, hospedeiro e ambiente podem ser caracterizadas, estratégias de controle avaliadas, níveis futuros de doença previstos e simuladores verificados.

J.E. Vanderplank definiu, do ponto de vista epidemiológico, dois grupos de doenças: as **doenças de juros compostos** e as **doenças de juros simples**. No primeiro grupo, *plantas infectadas durante o ciclo da cultura servirão de fonte de inóculo para novas infecções durante o mesmo ciclo*. É o caso típico da ferrugem asiática da soja, por exemplo, cujo agente causal (*Phakopsora pachyrhizi*), em condições favoráveis, pode produzir uma geração a cada 7-10 dias. Esta situação é análoga ao crescimento de capital a juros compostos, onde os juros ganhos rendem novos juros; no caso de doenças de juros compostos, plantas doentes rendem novas plantas doentes durante o ciclo da cultura. Para que isto ocorra, está implícita uma movimentação do patógeno a partir de plantas doentes em direção a novos sítios de infecção. No segundo grupo, *plantas infectadas durante o ciclo da cultura não servirão de fonte de inóculo para novas infecções*.

ções durante o mesmo ciclo. É o caso típico de muitas doenças veiculadas pelo solo, como a murcha do algodoeiro, por exemplo, cujo agente causal (*Fusarium oxysporum* f. sp. *vasinfectum*) coloniza principalmente o interior do xilema das plantas infectadas. Esta situação é análoga ao crescimento de capital a juros simples, onde os juros ganhos não rendem novos juros; no caso de doenças de juros simples, o aumento gradativo do número de plantas doentes durante o ciclo da cultura não é devido, primariamente, à movimentação do patógeno a partir de plantas doentes em direção a novos sítios de infecção e, sim, ao inóculo original, neste caso clamidósporos, previamente existente no solo.

Como seria a curva de progresso típica de cada um desses grupos? Para o caso das doenças de juros compostos, considerando que plantas doentes (ou lesões) dão origem a novas plantas doentes (ou novas lesões) no mesmo ciclo da cultura, a velocidade de aumento da doença é proporcional à própria quantidade de doença em cada instante. Assim, se uma lesão der origem a 10 lesões, 10 lesões darão origem a 100, 100 a 1000, 1000 a 10.000 e assim por diante. Esta cinética de crescimento é expressa matematicamente pela equação diferencial

$$dx / dt = rx \quad (1)$$

onde dx/dt é a velocidade de aumento da doença, x , a quantidade de doença e r , a taxa de infecção. A integração de 1 leva a

$$x = x_0 \exp(rt) \quad (2)$$

onde x_0 é a quantidade de doença no tempo t_0 . A curva descrita pela equação 2 tem a forma típica de um J e é conhecida como curva exponencial.

Para o caso das doenças de juros simples, considerando que plantas doentes (ou lesões) não dão origem a novas plantas doentes (ou novas lesões) no mesmo ciclo da cultura, a velocidade de aumento da doença não tem qualquer relação com a quantidade de doença em cada instante. Como já discutido, o aumento gradativo do número de plantas doentes durante o ciclo da cultura é função do inóculo original previamente existente. A quantidade de inóculo existente é, na maioria dos casos, desconhecida, mas, por conveniência, considerada constante durante cada período de vegetação. A fração de plantas que se torna doente (x) depende da frequência de contatos efetivos entre hospedeiro e patógeno (inóculo original). Contato efetivo é definido como aquele contato que leva à doença. Assim,

$$dx / dt = QR \quad (3)$$

onde Q é a quantidade de inóculo previamente existente e R , a taxa de infecção. O produto QR representa o número de contatos efetivos. Tanto Q quanto R são considerados constantes. A integração de 3 resulta em

$$x = x_0 + QRt \quad (4)$$

onde x_0 é a quantidade de doença no tempo t_0 . A curva descrita pela equação 4 é uma linha reta.

Seriam os modelos exponencial (equação 2) e linear (equação 4) espelhos fiéis da realidade? Simulariam eles com razoável precisão o crescimento da doença em condições naturais? As epidemias reais, para vários patossistemas, mostram um acordo parcial entre estes modelos e a realidade: aparentemente, com quantidades pequenas de doença, os modelos ficam próximos da realidade; à medida que a quantidade de doença aumenta, aumenta também o divórcio entre realidade e modelo. E, pensando bem, as coisas não poderiam ser diferentes: tanto o modelo exponencial quanto o linear permitem à quantidade de doença crescer até o infinito. Nenhum processo biológico comporta-se dessa maneira: leveduras ou bactérias cultivadas em meio de cultura não crescem ao infinito, pois o meio, entre outras causas, esgota-se; a quantidade de doença, do mesmo modo, não pode tender ao infinito, pois seu crescimento é limitado, entre outras causas, pela crescente indisponibilidade de tecido sadio. Um fator de correção obviamente faz-se necessário, fator este que reduza a velocidade de crescimento da doença proporcionalmente à diminuição da oferta de tecido sadio. A equação 1 (juros compostos) pode, assim, ser alterada para

$$dx / dt = rx(1 - x) \quad (5)$$

onde $(1 - x)$ representa a quantidade de tecido sadio (x , neste contexto, é sempre expresso em proporção de doença). A integração de 5 produz

$$\ln(x / (1 - x)) = \ln(x_0 / (1 - x_0)) + rt \quad (6)$$

Em conseqüência, o valor da taxa r (chamada de **taxa aparente de infecção**) é calculado por

$$r = (1 / t)(\ln(x / (1 - x)) - \ln(x_0 / (1 - x_0))) \quad (7)$$

A curva descrita pela equação 6 tem a forma de S, é conhecida pelo nome de **curva logística** e pode ser linearizada plotando-se na ordenada $\ln(x / (1 - x))$ ao invés de x . O valor de $\ln(x / (1 - x))$ é conhecido pelo nome de **logito** de x . É digno de nota que este modelo, o modelo logístico, confunde-se com o modelo exponencial para baixas quantidades de doença (aproximadamente 5% ou 0,05 de proporção de doença). As diferenças são crescentes à medida que x aproxima-se de 1.

Pelo mesmo raciocínio, a equação 3 (juros simples) pode ser alterada para

$$dx / dt = QR(1 - x) \quad (8)$$

onde $(1 - x)$ representa a quantidade de tecido sadio. A integração de 8 produz

$$\ln(1 / (1 - x)) = \ln(1 / (1 - x_0)) + QRt \quad (9)$$

O produto QR (quantidade de inóculo inicial e taxa de infecção) é calculado por

$$QR = (1/t)(\ln(1/(1-x)) - \ln(1/(1-x_0))) \quad (10)$$

A curva descrita pela equação 9 é conhecida pelo nome de **curva monomolecular** (inicialmente usada para descrever reações químicas monomoleculares de primeira ordem) e pode ser linearizada plotando-se na ordenada $\ln(1/(1-x))$ ao invés de x . O valor $\ln(1/(1-x))$ é conhecido pelo nome de **monito** de x . A exemplo dos modelos exponencial e logístico, aqui também, para baixos valores de x (até aproximadamente 5% ou 0,05), os modelos linear e monomolecular confundem-se. As diferenças, porém, acentuam-se à medida que x aproxima-se de 1.

Análise espacial

Padrões espaciais ao acaso e agregado

O padrão espacial de uma doença numa linha de plantio ou numa parcela pode ser ao acaso ou agregado. Padrão ao acaso em doenças cujos patógenos são veiculados pelo ar significa que a probabilidade de um esporo cair sobre uma planta hospedeira é igual para todas as plantas hospedeiras. Assim, a ocorrência da doença não é influenciada pela distância até a fonte de inóculo, seja a fonte localizada na mesma planta ou na vizinhança próxima. Padrão espacial ao acaso está intimamente relacionado a iguais oportunidades de infecção. Se, no entanto, em condições naturais, o patógeno se dispersa apenas a curtas distâncias (por respingos de chuva, por exemplo), a probabilidade de uma planta ser infectada é maior para aquelas plantas situadas próximas à fonte de inóculo do que para aquelas situadas longe da fonte de inóculo. Neste caso, o padrão espacial da doença não será ao acaso e, sim, agregado, com focos visíveis ao redor das fontes primárias de inóculo. Agregação é apenas um dos possíveis desvios para o padrão espacial ao acaso ou aleatório. O outro desvio possível é a regularidade ou uniformidade. Padrão espacial regular de doença existe numa linha de plantio, por exemplo, caso plantas doentes e sadias ocorram de forma alternada. Esse padrão regular não ocorre de forma natural, mas pode ser provocado, por exemplo, com o plantio alternado de variedade suscetível e variedade resistente.

Padrões espaciais em linhas de plantio

Dois tipos de análise serão apresentados para investigar o padrão espacial de plantas doentes numa linha de plantio. Em ambos os casos assume-se que o 'status' de uma planta pode ser caracterizado apenas como sadio ou doente; severidade de doença não pode ser considerada neste contexto. Se as categorias binárias forem designadas por 0 (sadia) e 1 (doente), o padrão da doença na linha é dado por determinada seqüên-

cia de 0 e 1, por exemplo, 0 0 1 0 1 0 0 1 1 0 1 0 1 0 1 1 0 1 para uma linha contendo 20 plantas. A questão pertinente é se essa seqüência reflete um padrão ao acaso ou agregado de plantas doentes.

Teste 'run' - Neste teste, o número de 'runs' é considerado como um critério de padrão aleatório. Um 'run' é definido como uma seqüência de um ou mais símbolos idênticos, os quais são seguidos ou precedidos por um símbolo diferente ou por símbolo nenhum (no começo ou fim de uma linha). A seqüência seguinte de 10 plantas tem seis 'runs':

0 0 1 1 0 1 0 1 1 1
1 2 3 4 5 6

Se uma doença está se disseminando de planta para planta, plantas doentes devem estar agregadas, o que leva a poucos 'runs'. Por outro lado, se plantas doentes estiverem distribuídas ao acaso na linha, o número de 'runs' deve ser grande.

Para que se possa concluir estatisticamente a respeito de agregação ou distribuição espacial ao acaso, é necessário saber o número de 'runs' e sua variação esperada teórica. Assim, caso existam N plantas numa linha e m dentre elas estejam doentes, as fórmulas seguintes podem ser usadas para calcular o valor esperado $E(R)$ de 'runs' e sua variância $\sigma^2(R)$ assumindo-se padrão espacial ao acaso:

$$(Eq. 11) E(R) = 1 + 2m(N-m)/N$$

$$(Eq. 12) \sigma^2(R) = 2m(N-m)[2m(N-m)-N]/[N^2(N-1)]$$

Para realizar o teste, os 'runs' observados (R) e os esperados $E(R)$ são comparados. Para um número razoavelmente alto de plantas (usualmente $N > 20$), é útil proceder-se a uma standardização por meio do cálculo de Z_R , o qual tem distribuição normal:

$$(Eq. 13) Z_R = [R + 0,5 - E(R)] / \sigma(R)$$

A constante 0,5 é chamada de correção para continuidade e é introduzida para minimizar o efeito da natureza discreta do número de 'runs'. O valor de Z_R será um grande número negativo se houver agregação porque R será, nesse caso, muito menor que $E(R)$. Como o teste é usualmente aplicado somente para diferenciar distribuição espacial agregada de distribuição espacial ao acaso, rejeita-se a hipótese de distribuição espacial ao acaso se $Z_R < -1,64$.

Alguns exemplos teóricos serão apresentados a seguir, sempre considerando 20 plantas numa linha, 12 das quais doentes: portanto, $N = 20$; $m = 12$.

Exemplo I. $\frac{0\ 1\ 1\ 0\ 0\ 1\ 0\ 1\ 1\ 1\ 0\ 1\ 0\ 1\ 1\ 0\ 1\ 0\ 1\ 1}{1\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 9\ 10\ (\dots)\ 14}$ $R = 14$

Número de 'runs' esperados para distribuição espacial ao acaso de acordo com a Eq. 11:

$$E(R) = 1+2 \cdot 12 \cdot 8/20 = 10,6$$

A variância correspondente é (Eq. 12):

$$\sigma^2(R) = 2 \cdot 12 \cdot 8 [2 \cdot 12 \cdot 8 - 20] / [20 \cdot 20 - 19] = 4,3453$$

Para o valor standardizado (Eq. 13):

$$Z_R = [14+0,5-10,6]/4,3453^{1/2} = 1,87$$

Como o número observado de 'runs' é maior que o número esperado, a seqüência examinada não exibe um padrão agregado, o que também é mostrado pelo valor positivo de Z_R .

Exemplo II. $\frac{0\ 1\ 1\ 1\ 0\ 0\ 0\ 1\ 1\ 0\ 1\ 1\ 0\ 0\ 1\ 1\ 1\ 1\ 1\ 0}{1\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 9}$ $R = 9$

Como N e m são os mesmos do exemplo anterior, $E(R)$ e $\sigma^2(R)$ também têm o mesmo valor. Z_R pode ser determinado por:

$$Z_R = [9+0,5-10,6]/4,3453^{1/2} = -0,527$$

Neste caso, o número observado de 'runs' é menor que aquele esperado no caso de padrão espacial ao acaso, mas a hipótese da nulidade não pode ser rejeitada e a seqüência de plantas ainda é considerada aleatória.

Exemplo III. $\frac{0\ 1\ 1\ 1\ 1\ 1\ 1\ 0\ 0\ 0\ 1\ 1\ 1\ 0\ 0\ 0\ 0\ 1\ 1\ 1}{1\ 2\ 3\ 4\ 5\ 6}$ $R = 6$

Com esses valores, Z_R é determinado como:

$$Z_R = [6+0,5-10,6]/4,3453^{1/2} = -1,967$$

Aqui, como Z_R é menor que -1,64, a hipótese da nulidade é rejeitada e o padrão espacial na linha é considerado agregado.

Os três exemplos discutidos mostram que é necessário haver um considerável desvio da aleatoriedade para que a hipótese alternativa de agregação seja aceita. Um exame superficial do exemplo II levaria a maioria das pessoas a optar pelo padrão agregado, fato desmentido pelo teste objetivo.

Teste de ‘doublet’ - Na análise de ‘doublet’, o número de ‘doublets’, isto é, duas plantas doentes adjacentes, é usado como o critério de decisão. As plantas dentro da linha são outra vez caracterizadas como 0 (planta sadia) e 1 (planta doente). Se existirem 10 plantas numa linha exibindo o padrão 0 0 1 1 0 1 0 1 1 1, três ‘doublets’ podem ser identificados. Se o padrão espacial da doença for agregado, o número de ‘doublets’ será grande; o contrário é verdadeiro no caso de um padrão espacial ao acaso.

Também neste tipo de análise, o número observado de ‘doublets’ será comparado com o número esperado de acordo com a hipótese da aleatoriedade. Considerando N plantas numa linha, com m plantas doentes, as fórmulas seguintes são válidas para o cálculo do número esperado de ‘doublets’ $E(D)$ e sua variância $\sigma^2(D)$:

$$(Eq. 14) E(D) = m(m-1)/N$$

$$(Eq. 15) \sigma^2(D) = m(m-1)[N(N-1)+2N(m-2)+N(m-2)(m-3)-(N-1)m(m-1)]/N^2(N-1)$$

Para conduzir o teste de ‘doublet’, o número observado de ‘doublets’ (D) é comparado com o número esperado $E(D)$. Para um grande número de plantas ($N > 20$), é novamente possível calcular um valor estandardizado de Z_D com base na distribuição normal:

$$(Eq. 16) Z_D = (D - E(D)) / \sigma(D)$$

No caso de agregação, o valor observado D será maior que o esperado $E(D)$ e, assim, Z_D terá um grande valor positivo. Uma linha de plantas exibirá padrão agregado se $Z_D > 1,64$ ($P = 0,05$).

As mesmas seqüências dos três exemplos anteriores (teste ‘run’) serão apresentadas a seguir para o teste de ‘doublet’ (com 20 plantas numa linha, 12 das quais doentes: $N = 20$; $m = 12$).

Exemplo I. 0 1 1 0 0 1 0 1 1 1 0 1 0 1 1 0 1 0 1 1 $D = 5$

Note que uma seqüência 111 define dois ‘doublets’; uma seqüência 11111 define quatro ‘doublets’, etc. O esperado número de ‘doublets’, segundo a hipótese de aleatoriedade, é dada pela Eq. 14:

$$E(D) = 12 \cdot 11 / 20 = 6,6$$

A variância dos ‘doublets’, de acordo com a Eq. 15 é:

$$\sigma^2(D) = 12 \cdot 11 [20 \cdot 19 + 2 \cdot 20 \cdot 10 + 20 \cdot 10 \cdot 9 - 19 \cdot 12 \cdot 11] / [20 \cdot 20 \cdot 19] = 1,25$$

Com a Eq. 16, calcula-se o valor estandardizado (Z_D):

$$Z_D = [5 - 6,6] / 1,25^{1/2} = -0,98$$

Como o número observado de ‘doublets’ é menor que o número esperado, o padrão da doença certamente não é agregado, o que também se reflete no valor de Z_D .

Exemplo II. 0 1 1 1 0 0 0 1 1 0 1 1 0 0 1 1 1 1 1 0 $D = 8$

Valores de $E(D)$ e $\sigma^2(D)$ são idênticos ao exemplo anterior (mesmos valores de N e m). O valor de Z_D é calculado de acordo com a Eq. 16:

$$Z_D = [8+0,5-6,6]/1,25^{1/2} = 1,70$$

Neste caso, o número observado de ‘doublets’ é maior que o esperado sob a hipótese de aleatoriedade. O valor de Z_D mostra que a hipótese da nulidade pode ser rejeitada em favor da hipótese alternativa, o que significa que as plantas doentes exibem padrão agregado na linha de plantio.

Exemplo III. 0 1 1 1 1 1 1 0 0 0 1 1 1 0 0 0 0 1 1 1 $D = 9$

$$Z_D = [9+0,5-6,6]/1,25^{1/2} = 2,59$$

Neste caso, Z_D é muito maior que 1,64 (o limite a 5% de probabilidade) e, assim, o padrão de plantas doentes é considerado agregado.

Para o exemplo teórico II, como demonstrado, os testes ‘run’ e de ‘doublet’ apresentaram resultados diferentes. No teste ‘run’, concluiu-se que o padrão espacial de plantas doentes era ao acaso enquanto que no teste de ‘doublet’ a aleatoriedade foi rejeitada e o padrão agregado, aceito. Qual desses testes é mais confiável? Após comparação dos dois testes com dados simulados, considera-se superior a análise de ‘runs’.

Para parcelas experimentais com muitas linhas de plantio, sugere-se que as diversas linhas sejam combinadas para formar uma linha maior. Esta é, então, analisada com os métodos apresentados anteriormente. Deve-se levar em conta, no entanto, a transição entre diferentes linhas: assim, um par de plantas doentes não deve ser contado como um ‘doublet’ se as plantas doentes estiverem em linhas adjacentes.

Padrões espaciais em parcelas ou campos experimentais

Diversos métodos estão disponíveis na literatura para a análise de padrões espaciais em parcelas ou campos experimentais como, por exemplo, comparação com distribuições estatísticas, índices de dispersão, autocorrelação, áreas isópatas, lei de Taylor, etc. Há excelentes artigos e livros que tratam da interpretação de dados espaciais nos contextos da epidemiologia e da ecologia. Uma breve descrição desses métodos será dada a seguir.

Comparação com distribuições estatísticas

Para analisar o padrão espacial da doença em parcelas ou campos experimentais, uma moldura pode ser colocada sobre a parcela de tal forma a dividi-la em 'quadrats'. 'Quadrats' são definidos como uma unidade de amostragem e têm usualmente a forma retangular, com dimensões a critério do pesquisador. Em cada 'quadrat', o número de pontos (ou plantas doentes) é determinado e uma distribuição de frequência é construída. A distribuição de frequência, obviamente, dependerá do número de 'quadrats' ou do tamanho do 'quadrat'. Como o tamanho ideal do 'quadrat' para cada situação não é conhecido, recomenda-se a análise utilizando-se diferentes tamanhos.

Para analisar o padrão espacial da Fig. 1, cada parcela foi dividida em 36 'quadrats' (Fig. 2). A próxima etapa da análise é contar os 'quadrats' que contêm nenhum ponto (ausência de plantas doentes), um ponto, dois pontos, etc. A distribuição de frequência conseguida para os três casos está apresentada na Tabela 1. As três distribuições de frequência são obviamente diferentes: a Fig. 2B tem alta incidência de 'quadrats' com nenhum ponto, conseqüência da concentração de plantas doentes em determinada área da parcela; a Fig. 2C tem predominância de 'quadrats' com um ou dois pontos (com uma exceção com três pontos), o que é uma clara indicação de padrão regular; a Fig. 2A tem alta incidência de 'quadrats' com 0, 1 e 2 pontos.

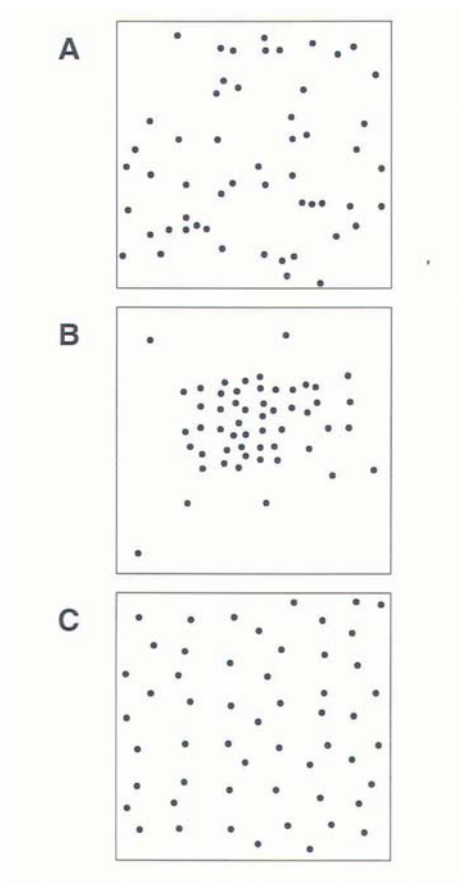


Fig. 1. Padrões espaciais ao acesso (A), agregado (B) e regular (C) de pontos.

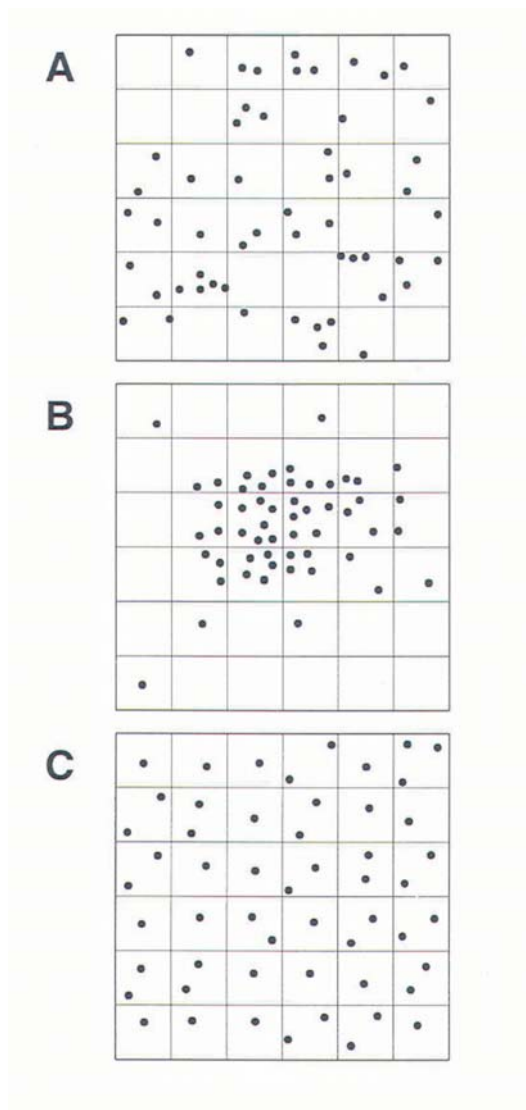


Fig. 2. “Quadrats” (6 x 6) para determinar as distribuições de freqüência dos três padrões espaciais da Fig. 1.

De modo semelhante aos outros testes discutidos neste capítulo, a pergunta que precisa ser respondida é qual a distribuição de freqüência esperada no caso dos pontos apresentarem um padrão espacial ao acaso. A resposta pode ser dada com base na distribuição binomial, que é caracterizada pelos parâmetros p e n . O parâmetro p é a probabilidade de um ponto se localizar num determinado ‘quadrat’. Como há um total de 36 ‘quadrats’, esta probabilidade é dada por $p = 1/36 = 0,027$. O parâmetro n é o número de todos os pontos (plantas doentes) existentes, neste caso $n = 54$. A probabilidade que um ‘quadrat’ contenha r pontos pode ser calculada usando-se a distribuição binomial:

$$(Eq. 17) \quad P(r) = \binom{n}{r} p^r (1-p)^{n-r} \quad \text{com} \quad \binom{n}{r} = n!/[r!(n-r)!]$$

Tabela 1. Distribuição de frequência para os padrões espaciais da Fig. 2.

Pontos por 'quadrat'	Número de 'quadrats' em		
	Fig. 2A	Fig. 2B	Fig. 2C
0	9	16	0
1	11	7	19
2	9	4	16
3	4	3	1
4	2	3	0
5	1	1	0
6	0	1	0
7	0	1	0

O número de 'quadrats' com r pontos esperados é obtido multiplicando-se $P(r)$ pelo número total de 'quadrats' na parcela. As frequências observadas e esperadas para a Fig. 2A estão apresentadas na Tabela 2. Há alto grau de semelhança entre elas, o que indica padrão espacial aleatório para a Fig. 2A. Um método objetivo para comparar as duas distribuições de frequência é o teste Chi-quadrado.

Tabela 2. Frequências observadas e esperadas para as distribuições binomial e de Poisson na Fig. 2A.

Distribuição binomial

	Número de pontos por 'quadrat'						
	0	1	2	3	4	5	6
Frequência observada	9	11	9	4	2	1	0
Frequência esperada	7,86	12,13	9,19	4,55	1,65	0,47	0,11

Distribuição de Poisson

	Número de pontos por 'quadrat'						
	0	1	2	3	4	5	6
Frequência observada	9	11	9	4	2	1	0
Frequência esperada	8,03	12,05	9,04	4,51	1,69	0,51	0,131

O cálculo das probabilidades $P(r)$ com a Eq. 17 pode ser muito trabalhosa para altos valores de n . Se n for grande e p , pequeno, a distribuição binomial pode ser substituída pela mais conveniente distribuição de Poisson. Esta distribuição estatística depende somente de um parâmetro (λ), que determina o número médio de pontos por 'quadrat'. A probabilidade que r pontos ocorram num determinado 'quadrat' é dada por:

$$(Eq. 18) \quad P(r) = \lambda^r \exp(-\lambda)/r!$$

Para a determinação do número esperado de 'quadrats', a probabilidade $P(r)$ é novamente multiplicada pelo número total de 'quadrats', neste caso 36 (Fig. 2A). O parâmetro λ é o número total de pontos dividido pelo número de 'quadrats': $\lambda = 54/36 = 1,5$. As distribuições de freqüência observada e esperada mostram grande concordância entre si (Tabela 2). Há também alto grau de similaridade quando se compara a distribuição binomial com a distribuição de Poisson (Tabela 2): ambas descrevem situação de aleatoriedade. Uma diferença entre as duas distribuições é que para a binomial o número máximo de pontos num 'quadrat' é restrito (valor máximo = n); para a distribuição de Poisson, ao contrário, probabilidades positivas podem ser calculadas mesmo para valores mais altos.

A distribuição de Poisson só deve ser aplicada para dados binários caso a freqüência de doença seja menor que 20%. Além das distribuições binomial e de Poisson, ambas específicas para padrões espaciais ao acaso, outras distribuições estatísticas podem ser ajustadas às freqüências observadas no caso de ocorrer agregação, como as distribuições binomial negativa (para dados de contagem) e beta-binomial (para dados binários).

Mapas

O mapeamento (bi ou tridimensional) é uma técnica útil na visualização e discernimento de possíveis relações entre dados e variáveis, tais como fatores edáficos, hídricos e, também, fontes de inóculo. Neste último caso, atenção especial é dada àquelas fontes localizadas fora da área avaliada, que podem ajudar a melhor compreender o patossistema. Em adição ao mapeamento, áreas ou linhas isópatas podem ser construídas, por meio de 'softwares' apropriados, evidenciando locais de mesmo nível de doença.

Índices de dispersão

O cálculo da maioria dos índices de dispersão D (relação variância/média, índices de Lloyd, índice de Morisita, etc.) descritos na literatura é baseado na relação entre a variância e a média do conjunto de dados da área experimental considerada. Esses índices podem indicar padrões espaciais regulares ($D < 1$), aleatórios ($D = 1$) ou agregados ($D > 1$) e são válidos para dados de contagem (como o número de lesões por planta ou o número de insetos por folha).

O índice de Morisita (IM), por exemplo, é calculado pela fórmula:

$$(Eq. 19) \quad IM = Q \left[\sum_{i=1}^Q X_i (X_i - 1) \right] / [N^2]$$

onde Q é o total de 'quadrats', X_i , o total de pontos no 'quadrat' i , e N , o total de todos os pontos ($=\sum X_i$). Para o exemplo da Fig. 2 ($N = 54$, $Q = 36$), os resultados são os seguintes:

- **Fig. 2A:** $IM = 36 \cdot 86 / (54 \cdot 53) = 1,082$
- **Fig. 2B:** $IM = 36 \cdot 154 / (54 \cdot 53) = 1,937$
- **Fig. 2C:** $IM = 36 \cdot 38 / (54 \cdot 53) = 0,478$

Esses resultados indicam que o padrão da Fig. 2A é ao acaso, o da Fig. 2B, agregado, e o da Fig. 2C, regular. O teste de Chi-quadrado pode ser empregado para confirmar objetivamente essas indicações.

Pode-se demonstrar que índices de dispersão de modo geral (como o índice de Morisita), baseados que são nas premissas da distribuição de Poisson, não devem ser usados para dados de incidência (especialmente para incidências maiores que 20%). Para esse tipo de dado (binário), a melhor aproximação de uma condição de aleatoriedade é dada pela distribuição binomial e, assim, o índice de dispersão apropriado deve ser outro. De fato, índice de dispersão, a rigor, é a relação entre a variância observada e a variância teórica do processo em estudo ($D = \text{variância observada} / \text{variância teórica}$). Deriva desta definição a sempre utilizada relação variância observada e média ($D = \text{variância observada} / \text{média}$), uma vez que a variância da distribuição de Poisson é sua própria média. Já para a distribuição binomial, a variância é igual a $V_{bin} = p(1-p)/n$, onde p é a incidência na parcela e n , o número de plantas por 'quadrat'. Já a variância observada é calculada por $V_{obs} = \sum (X_i - np)^2 / n^2 (N-1)$, onde $\sum X_i$ é o somatório do número de plantas sintomáticas em cada 'quadrat' i e N é o número total de 'quadrats' em cada área.

A utilização de índices de dispersão apropriados para dados de incidência vem, de forma gradual, ocupando espaço na literatura fitopatológica. Apesar de muito usados, os índices de dispersão têm a grande limitação (assim como a comparação com distribuições estatísticas) de não levarem em conta a posição relativa de cada medida.

Lei de Taylor

Populações de diferentes espécies apresentam graus de agregação característicos. A partir de dados de distribuição espacial de diversas populações de uma espécie, pode-se relacionar linearmente o logaritmo da variância observada (V_{obs}) e o logaritmo da média:

$$(Eq. 20) \quad \log(V_{obs}) = \log(A) + b \log(\text{média})$$

Esta é a chamada lei de Taylor. Os parâmetros da equação, assim, representam a característica espacial de uma dada espécie. Mais especificamente, Taylor enfatizou o emprego do parâmetro b como um índice de agregação: quando $b > 1$, o padrão espacial é agregado (tanto mais agregado quanto maior for o valor de b); quando $b = 1$, ao acaso; quando $b < 1$, regular.

A lei de Tylor é inadequada para dados de incidência e pode ser modificada para a forma:

$$(Eq. 21) \quad \log(V_{obs}) = \log(A) + b \log(\text{variância teórica})$$

Para o caso de dados binários (como incidência de doença) a variância teórica é a variância binomial (V_{bin}), já referida anteriormente. Assim,

$$(Eq. 22) \quad \log(V_{obs}) = \log(A) + b \log(V_{bin})$$

Uma distribuição espacial ao acaso é indicada por $b = 1$ e $A = 1$ [$\log(A) = 0$], isto é, $\log(V_{obs}) = \log(V_{bin})$, ou, variância observada = variância binomial. Isto é o mesmo que dizer que $D = 1$ para todo o conjunto de dados. Se $b = 1$ e $A > 1$, então D é fixo e igual a A para todo o conjunto de dados, isto é, D não varia com p . Se $b > 1$, segue-se que $\log(V_{obs})$ aumenta com $\log(V_{bin})$ numa taxa maior que V_{bin} , isto é, D varia com mudanças em p . Neste contexto, b é considerado um índice de agregação, embora a interpretação do valor de b nunca deva ser separada da interpretação do valor de A .

A grande vantagem da lei de Taylor sobre outros métodos de estudo do padrão espacial de doenças é que a agregação de todo o conjunto de dados é descrito pelo menor número de parâmetros (A e b). Outras abordagens (como os índices de dispersão, por exemplo) requerem um parâmetro (D) para cada subconjunto de dados. Adicionalmente, os parâmetros da lei de Taylor permitem uma visão dinâmica de como muda a agregação da doença em função do tempo (e de maiores incidências). Quando a lei de Taylor for usada para comparar diferentes doenças ou uma mesma doença em diferentes condições, cuidado deve ser tomado para que as amostragens sejam feitas considerando-se gama semelhante de intensidades de doença, caso contrário resultados errôneos podem ser obtidos.

Estudos de caso

Padrão espacial da clorose variegada dos citros e de outras doenças transmitidas por cigarrinhas

A clorose variegada dos citros (CVC), causada por *Xylella fastidiosa*, tem sua epidemiologia temporal e espacial bem estudada no Brasil. Três aspectos serão destacados neste item com respeito à análise espacial: teste 'run', lei de Taylor e áreas isópatas.

Aplicando o teste 'run' (Eq. 11-13) a diversos pomares do Estado de São Paulo, Laranjeira (1997) não encontrou evidência de maior agregação nas linhas de plantio (que apresentam distância entre plantas menor que entre linhas). Resultados idênticos foram também relatados por Gottwald et al. (1993) para pomar localizado em Bebedouro-SP. Neste caso, os autores encontraram, inclusive, maior agregação entre linhas do que na linha de plantio. Assim, tratos culturais (roçagem, gradagem, aplicação de defensivos e adubos, etc.), sempre realizados ao longo da linha de plantio, não parecem contribuir para a disseminação da bactéria. A resposta de escape de cigarrinhas, ou seja, o vôo induzido por perturbações mecânicas no seu ambiente, é variável conforme o gênero do vetor. Assim, é possível que as cigarrinhas capazes de transmitir *X. fastidiosa* em citros não tenham uma boa resposta de escape ou, nesse movimento, não tenham preferência pela planta mais próxima. Dessa maneira, a passagem de máquinas no talhão não induz os insetos ao vôo ou, caso induza, os vetores não tendem a pousar nas plantas mais próximas (dentro das linhas de plantio, em função do menor espaçamento). Esta última hipótese parece ser a verdadeira, conforme demonstrado a seguir, com a aplicação da lei de Taylor para dados binários.

Levantamentos bimensais realizados em seis talhões durante 20 meses resultaram em $b = 1,13$ (Eq. 22), com $R^2 = 0,983$, para a lei de Taylor (LARANJEIRA, 1997). O valor de 1,13 diferiu de 1 pelo teste Chi-quadrado, o que implica em padrão agregado de plantas sintomáticas. No entanto, a diferença da unidade é bastante pequena, sugerindo plantas doentes apenas levemente agregadas nos pomares avaliados. Leve agregação indica que as cigarrinhas vetoras realizam vôos curtos de uma árvore para outra sem, porém, necessariamente, pousarem em plantas adjacentes àquelas já colonizadas. Digno de nota é que outros patossistemas envolvendo cigarrinhas vetoras e diferentes tipos de patógeno apresentam valores de b bastante semelhantes: 'aster yellows' em alface/fitoplasma, $b=1,18$; enfezamento do milho/fitoplasma e espiroplasma, $b=1,12$; 'rayado fino' do milho/vírus, $b=1,19$. Esses resultados indicam que o padrão espacial de doenças com vetor é muito mais função do comportamento do vetor do que de fatores característicos do patógeno ou do hospedeiro.

A análise de áreas isópatas permitiu a Laranjeira (1997) postular que talhões mais velhos com alta incidência de CVC são as principais fontes de inóculo da bactéria para pomares recém-implantados. Ao contrário do que acontece com o mal de Pierce da videira, também causado por *X. fastidiosa*, outras hospedeiras do patógeno parecem desempenhar papel secundário na epidemiologia da CVC. Adicionalmente, o estudo de áreas isópatas mostrou, em dois casos, o papel importante desempenhado por mudas contaminadas no início da epidemia.

Padrão espacial e etiologia do amarelecimento fatal do dendezeiro em Benevides, PA

Um aspecto raramente discutido na literatura fitopatológica, em contraste com a literatura médica, é o emprego de padrões espaciais de plantas doentes para inferir a respeito da natureza (biótica ou abiótica) do agente causal. Um dos motivos para tal é que a obtenção de dados consistentes sobre o padrão espacial de plantas doentes requer, via de regra, vários anos de coleta, tempo geralmente suficiente para que o agente causal (biótico ou abiótico) seja identificado inequivocamente. Uma exceção é o declínio dos citros, ainda de etiologia desconhecida, mas cada vez mais considerado de causa abiótica em virtude de suas características epidemiológicas.

O trabalho de Laranjeira et al. (1998), aqui discutido, analisa a distribuição espacial de plantas de dendê (*Elaeis guineensis*) com sintomas de amarelecimento fatal (AF) como contribuição ao esclarecimento de sua etiologia. Mesmo tendo ciência que um determinado padrão não indica necessariamente o processo que o originou, a premissa básica do trabalho é que doenças bióticas de causa desconhecida devem apresentar padrão espacial de plantas doentes semelhante ao padrão espacial encontrado em doenças de causa biótica já devidamente caracterizadas. Apesar de ser assunto recente e ainda controvertido, as doenças de causa biótica, especialmente quando o hospedeiro é perene, exibem padrões definidos de aparecimento e de crescimento de focos, além de mostrarem agregação estatisticamente significativa de plantas doentes a partir de incidências relativamente elevadas, usualmente entre 10% e 15%.

Duas propriedades marcantes caracterizaram as epidemias de AF em Benevides, Pará, sob o enfoque espacial, conforme quantificado pelo índice de dispersão ($D = V_{obs} / V_{bin}$) (LARANJEIRA et al. 1998): primeiro, a forte e freqüente agregação de plantas afetadas (plantas sintomáticas encontravam-se agregadas em 93,75% das situações analisadas) e, segundo, as incidências p extremamente baixas nas quais a agregação foi estatisticamente detectada. Via-de-regra, doenças de causa infecciosa (ou biótica), por sua própria infecciosidade, necessitam de limiares de incidência maiores para que agregação significativa seja detectada, quando comparadas com doenças de causa não-infecciosa (ou abiótica). Para as epidemias de AF analisadas por Laranjeira et al. (1998), valores de incidência tão baixos como 3,2% ou 3,8% foram suficientes para que agregação significativa fosse detectada. Discorrendo sobre o mesmo assunto, van de Lande (1993) vai além e relata agregação significativa para o AF em Victoria, Suriname, para incidência de 0,7% ($P < 0,05$). Em outros blocos, incidências de 1,5%, 8,3% e 9,3% também foram suficientes para que agregação significativa fosse detectada. No mesmo estudo (VAN DE LANDE, 1993), ainda com o dendezeiro, uma doença sabidamente de causa biótica ('marchitez sorpresiva', causada por *Phytomonas* e transmitida por insetos pentatomídeos) só mostrou agregação a incidências consideravelmente mais elevadas ($19,1\% < p < 34,4\%$). A julgar por esses indícios, os processos que dão origem aos padrões espaciais de uma doença de causa biótica ('marchitez sorpresiva') e os processos que dão origem aos padrões espaciais do AF são bastante diferentes.

Houve relação significativa entre $\log(V_{obs})$ e $\log(V_{bin})$ (Eq. 22) para cada uma das situações analisadas, com valores de F^2 entre 0,970 e 0,997. Os valores de b e de A foram estatisticamente maiores que 1, para todas as análises, mostrando agregação de plantas sintomáticas em todos os blocos. A amplitude de variação do parâmetro b entre blocos foi pequeno ($1,321 < b < 1,597$). Análise conjunta dos dados, considerando todos os blocos e todas as datas de avaliação, resultou em $b = 1,282$, $\log(A) = 0,891$ e $F^2 = 0,901$. Neste caso, também, b e A diferiram de 1, confirmando o padrão espacial agregado do AF. Os resultados obtidos com a lei de Taylor confirmam e ampliam aqueles obtidos com os índices de dispersão, ou seja, plantas com sintomas de AF exibem acentuada agregação desde incidências extremamente baixas, tendência que se acentua em função do tempo e do conseqüente aumento da incidência (LARANJEIRA et al. 1998).

A análise de áreas isópatas não permitiu a identificação de nenhum padrão espacial definido para o AF. Assim, áreas de maior incidência de plantas sintomáticas (focos) podem ser encontradas nas bordas dos blocos, mas podem, também, ser encontradas no centro dos blocos. Em alguns casos, a situação é ainda mais confusa, pois focos de plantas sintomáticas são encontrados tanto nas bordas como no centro dos blocos (LARANJEIRA et al. 1998). Com relação ao crescimento dos focos, os mapas de áreas isópatas mostram progressão radial, às vezes com velocidades iguais em todas as direções, dando origem a formas circulares, às vezes com velocidades maiores na direção Norte-Sul (ou Sul-Norte), dando origem a formas alongadas nesse sentido, às vezes com velocidades maiores no sentido Leste-Oeste (ou Oeste-Leste), dando origem a formas alongadas nessa direção (LARANJEIRA et al. 1998).

A ausência de padrão que caracteriza o aparecimento e o crescimento de focos de AF também sugere o envolvimento de processos formadores não-bióticos (LARANJEIRA et al. 1998). Em sua análise de áreas isópatas para a clorose variegada dos citros, Laranjeira (1997) encontrou, de forma consistente, os primeiros focos de plantas doentes nas bordas dos blocos, indicação clara que o patógeno, e seu vetor, sobrevivem nos pomares mais velhos infectados por *X. fastidiosa*. Nenhum padrão pôde ser definido para o AF, uma vez que focos foram encontrados indistintamente nas bordas e no centro dos talhões.

Deve-se mencionar, finalmente, a associação entre proximidade de riachos ou áreas alagadas e maiores incidências de AF. Essa associação também sugere causa abiótica para o AF e está de acordo com as conclusões de análise epidemiológica temporal feita na mesma área e na mesma época (BERGAMIN FILHO et al. 1998).

Influência da espécie do vetor na epidemiologia da tristeza dos citros

A tristeza dos citros foi observada pela primeira vez no Brasil em 1937, na região do vale do rio Paraíba, São Paulo. Seu agente causal, o vírus da tristeza (*Closterovirus*), é um dos patógenos de maior importância econômica da cultura, afetando

do-a em todos os países produtores. No Brasil, o vírus é transmitido de maneira semipersistente por diversas espécies de afídeos, principalmente *Toxoptera citricida* e *Aphis gossypii*. A primeira coloniza o citros, que é seu principal hospedeiro, e transmite o vírus com alta eficiência; a segunda migra apenas acidentalmente para o citros, não o coloniza, e transmite o vírus com baixa eficiência. *T. citricida* não ocorria nos Estados Unidos e nos países da América Central e Caribe até o início da década de 90. Ainda não ocorre na região mediterrânea.

O progresso temporal da tristeza quando *T. citricida* está presente no pomar é tão mais rápido que a causa para tal não deve ser apenas sua maior eficiência de transmissão. Estudos experimentais conduzidos na Espanha, Flórida e Califórnia (na ausência de *T. citricida* e presença de *A. gossypii*) mostraram que, em média, 95% das plantas de um pomar tornam-se sintomáticas após 8 a 15 anos; na Costa Rica e República Dominicana (na presença de *T. citricida* e *A. gossypii*), no entanto, a mesma incidência de tristeza é alcançada em apenas 2 a 6 anos.

Uma explicação adicional (além da maior eficiência de transmissão) para essa dramática mudança no progresso temporal da tristeza deve ser procurada no comportamento espacial de seus vetores. É sugerido que *A. gossypii* tem o hábito de, ao deixar uma planta cítrica, voar para plantas distantes (entre 100 a 200 m), muitas vezes deixando o pomar e interrompendo a cadeia de transmissão do patógeno. Esse comportamento do vetor produz um padrão espacial ao acaso de plantas com sintomas de tristeza, fato comprovado recentemente, por meio da lei de Taylor (Eq. 22), para pomares da Califórnia e da Espanha: os parâmetros A e b não diferiram da unidade. Em pomares onde *T. citricida* estava presente, ao contrário, a lei de Taylor mostrou padrões espaciais agregados de plantas sintomáticas, tanto na Costa Rica (onde b foi maior que 1) quanto na República Dominicana (onde A foi maior que 1), fato explicado pelo hábito de *T. citricida*, ao se locomover, voar para plantas próximas daquela previamente colonizada.

Dispersão a longa ou a curta distância: qual traz maior vantagem para o patógeno?

O primeiro cientista a se preocupar com essa questão foi Vanderplank (1967), algumas décadas atrás. Ele baseou sua argumentação para respondê-la na dispersão de *Phytophthora infestans* em batata e sua conclusão foi surpreendente para a época: “Patógenos com apenas um mecanismo de disseminação – seja para curtas distâncias, seja para longas distâncias – estão mal servidos”. Mecanismos de dispersão para curtas distâncias, diz Vanderplank (1967), confinam o patógeno a focos já ocupados por lesões ou plantas doentes, limitando o livre progresso da epidemia às áreas limítrofes do foco, cada vez menores em relação à área total do foco. Mecanismos de dispersão para longas distâncias levam propágulos do patógeno para bem longe da fonte primária de inóculo, muitas vezes para fora da plantação, impedindo o rápido progresso da epidemia.

Uma mistura dos dois mecanismos, portanto, parece ser indispensável para que patógenos possam ter sucesso na arte de causar epidemias (e, conseqüentemente, na arte de sobreviver). Dispersão a curta distância é necessária para que o patógeno colonize e reproduza-se abundantemente numa área recém conquistada; dispersão a longa distância é necessária para que o patógeno escape do local já conquistado e, assim, amplie sua distribuição geográfica. Vanderplank (1967) postula, ainda, que o mecanismo para curtas distâncias deve se repetir com maior freqüência, já que esta é uma atividade recorrente durante o ciclo da doença, ao contrário do mecanismo para longas distâncias, que deve se constituir num evento raro (mas não menos importante). Vanderplank (1967) também deixa claro que dividir os mecanismos de dispersão em apenas dois grupos (curtas e longas distâncias) é uma simplificação excessiva que não espelha a realidade: na verdade, cada patógeno bem sucedido certamente desenvolveu durante sua evolução inúmeros mecanismos de dispersão, cada um mais apropriado para determinadas distâncias, o que permite sua sobrevivência sob várias condições de ambiente.

Foram necessários 25 anos e o desenvolvimento do computador para que essas previsões de J.E. Vanderplank pudessem ser verificadas por meio de modelo de simulação. Num trabalho de grande repercussão, Zawolek e Zadoks (1992) não só confirmaram a necessidade de mais de um mecanismo de dispersão para que patógenos possam ser bem sucedidos, como chegaram a quantificar, ainda que de forma preliminar, a freqüência de ocorrência de cada um deles numa situação ideal para o patógeno: 80% para eventos de disseminação a curta distância e 20% para eventos de disseminação a longa distância.

O reexame do patossistema tristeza dos citros, descrito nesta mesma seção, à luz das hipóteses de Vanderplank (1967), e Zawolek e Zadoks (1992), ilustra de forma inequívoca a importância de diferentes mecanismos de dispersão para a competência epidêmica e a capacidade de sobrevivência de patógenos em geral: o vírus da tristeza na América do Norte e na América Central, sem *Toxoptera citricida*, só tinha mecanismo de dispersão a longa distância e, por esse motivo, apresentava importância secundária; a introdução de *T. citricida* mudou completamente o cenário; *T. citricida* (com seu hábito de colonizar plantas cítricas vizinhas) representa o mecanismo de curta distância para o patógeno ocupar áreas próximas das fontes primárias de inóculo.

Referências

BERGAMIN FILHO, A.; AMORIM, L.; LARANJEIRA, F.F.; BERGER, R.D.; HAU, B. Análise temporal do amarelecimento fatal do dendezeiro como ferramenta para elucidar sua etiologia. **Fitopatologia Brasileira**, v. 23, p. 391-396, 1998.

GOTTWALD, T.R.; GIDTTI, F.B.; SANTOS, J.M. ; CARVALHO, A.C. Preliminary spatial and temporal analysis of Citrus Variegated Chlorosis (CVC) in São Paulo, Brazil. In: IOCV CONFERENCE, 12., 1993, Riverside. **Proceedings...** Riverside:International Organization of Citrus Virologists, 1993. p. 327-335

LARANJEIRA, F.F. **Dinâmica espacial e temporal da clorose variegada dos citros**. 1997. Dissertação (Mestrado) – Escola Superior de Agricultura Luiz de Queiroz, Piracicaba.

LARANJEIRA, F.F.; AMORIM, L.; BERGAMIN FILHO, A.; BERGER, R.D. ; HAU, B. Análise espacial do amarelecimento fatal do dendezeiro como ferramenta para elucidar sua etiologia. **Fitopatologia Brasileira**, v. 23, p. 397-403, 1998.

VAN DE LANDE, H.L. Spatio-temporal analysis of spear rot and ‘marchitez sorpresiva’ in African oil palm in Surinam. **Netherlands Journal of Plant Pathology**, v. 99, p. 129-138, 1993.

VANDERPLANK, J.E. 1967. Spread of plant pathogens in space and time. In: GREGORY, P.H. ; MONTEITH, J.L. (Ed.). **Airborne microbes**. Cambridge: Cambridge University, 1967. p. 227-247.

ZAWOLEK, M.W. ; ZADOKS, J.C. Studies on focus development: An optimum for the dual dispersal of plant pathogens. **Phytopathology**, v. 82, p. 1288-1297, 1992.