

Um estudo de leprose dos citros com uso do modelo autológico

Luziane Franciscon¹, Paulo Justiniano Ribeiro Júnior², Elias Teixeira Krainski²

¹Embrapa Suínos e Aves, Caixa Postal 21, CEP 89700-000, Concórdia, SC. E-mail: luziane@cnpa.embrapa.br

²Universidade Federal do Paraná, Laboratório de Estatística e Geoinformação, Caixa Postal 19.081, CEP 81531-990, Curitiba, PR. E-mail: paulojus@ufpr.br, elias@leg.ufpr.br

Resumo – A leprose dos citros é considerada uma grave virose na citricultura nacional reduz a produção e o período de vida das plantas de citros. Causada por *Citrus leprosis vírus* (CiLV), trata-se de um ácaro que para seu controle químico são demandados altos custos. Conhecer a incidência da doença e do ácaro transmissor no tempo e espaço é de grande importância para descrever a dinâmica da doença e para controle da epidemia. Dados com estrutura espaço-temporal permitem agregar estudos de padrões, associações de dependência espacial e de tempo de contágio da doença entre plantas. Este trabalho apresenta uma aplicação do modelo autológico aos dados de incidência de leprose em citros coletados em diferentes pontos no tempo em um talhão de citros. O modelo autológico estende o modelo de regressão logística para acomodar efeitos espaciais. A essência desse modelo é a utilização de covariáveis obtidas a partir da variável resposta em posições vizinhas da localização considerada. Portanto, o elemento chave do modelo é a estrutura de vizinhança. É mostrado como o modelo autológico é apropriado para investigar doenças desse tipo, bem como Além disso, quantifica o efeito da presença da doença em diferentes estruturas de vizinhança.

Termos de indexação: Doenças em citros, estatística espacial, pseudo-verossimilhança.

Abstract –

Introdução

A citricultura brasileira apresenta números expressivos que traduzem a importância econômica e social que a atividade representa para a economia do país. A maior parte da produção é destinada para a indústria de suco, a qual responde por 53% do suco de laranja produzido no mundo e por 80% do suco concentrado que transita pelo mercado internacional. O maior estado produtor é São Paulo que participa com 79% da produção brasileira e emprega 400 mil trabalhadores na indústria de citros (Bassanezi & Laranjeira, 2007). Contudo, a produção é limitada devido

à presença de doenças nos pomares produtores, com destaque para as viroses, como a leprose dos citros (Marques et al., 2007). As doenças podem causar a erradicação de talhões inteiros de plantas, comprometendo a quantidade e a qualidade das frutas cítricas.

A leprose dos citros foi considerada, nas últimas décadas, a mais importante virose na citricultura nacional, pois reduz a produção e o período de vida das plantas de citros (Rodrigues, 2000). Causada por *Citrus leprosis vírus* (CiLV), é uma doença endêmica nas regiões produtoras do estado de São Paulo. O vírus é transmitido

exclusivamente pelo ácaro *Brevipalpus phoenicis* (Czermainski, 2006). A alimentação e circulação do ácaro na planta e entre as plantas afetadas gera o mecanismo de dispersão da doença (Bassanezi & Laranjeira, 2007).

A principal prática de controle são os acaricidas que geram de 10 a 15% dos custos da produção (Fundecitros, 2007). Porém a utilização em grande escala de produtos químicos, além de causar problemas induzindo o aparecimento de populações de ácaros resistentes, gera preocupações ecológicas e econômicas. Diante disso, há um aumento na busca de alternativas para controlar a leprose, entre elas, o uso comercial de diferentes tipos de agentes biológicos de controle e variedades resistentes ao vírus (Bastianel et al., 2006).

Além disso, conhecer a incidência da doença e do ácaro transmissor no tempo e espaço é de grande importância para descrever a dinâmica da doença e para controle da epidemia.

O movimento e a distribuição espacial de vetores que transmitem a virose afetam a disseminação nas plantas e o progresso da epidemia. Existem relativamente poucos estudos dos processos temporais e espaciais, relacionando a dinâmica da doença na população de plantas hospedeiras e a dinâmica populacional do vetor (Czermainski, 2006).

Os dados de plantas cítricas são, em geral, coletados ao longo do tempo, as plantas estão dispostas dentro de um talhão, em linhas e colunas com certa estrutura de vizinhança. A disposição das plantas permite estudar o padrão espacial de doenças cítricas. Em estudos para detectar o padrão espacial nos dados foram, inicialmente, propostas técnicas descritivas, como por exemplo: análise por quadrat, variograma, índices de associação. Mais recentemente além de detectar o padrão espacial dos dados, também foram introduzidos modelos estatísticos espaciais, tais como: o modelo autologístico, o modelo CAR (*Conditional Auto Regressive*) e modelo estocástico espacial. Enquanto que os dois

últimos modelos requerem técnicas de Monte Carlo com Cadeias de Markov (MCMC), o modelo autologístico trata o componente espacial como efeito fixo. O que permite ao modelo simples interpretabilidade, incorporando explicitamente a estrutura de dependência espacial. Diferentemente dos outros dois modelos que assumem efeitos aleatórios espaciais.

O modelo autologístico é flexível para modelar a incidência de doenças em plantas, considera o *status* da planta como resposta binária, e através de covariáveis de vizinhança, modela a dependência espacial ou autocorrelação espacial dos dados (Krainski et al., 2008). Os coeficientes de regressão dão à estimativa do acréscimo na probabilidade da presença, ou não, da doença. Na agricultura, esse modelo foi estudado para a análise da incidência de *Phytophthora* em pimentas de sino por Gumpertz et al. (1997), também foi estudado por Czermainski (2006) para análise de incidência de leprose dos citros, por Viola (2007) na análise de dados de um experimento com pimentões e para dados simulados e por Krainski et al. (2008), na análise de dados de incidência de morte súbita em citros.

O interesse do trabalho é propor estratégias de modelagem dos dados com questões que podem ser incorporadas no estudo quando ajustado o modelo autologístico: a estrutura de covariáveis de vizinhança, a dimensão temporal dos dados e o efeito defasado das covariáveis de vizinhança. Encontrar um modelo adequado com fatores que descrevam a dinâmica da doença e que de alguma forma auxiliem nas práticas de controle.

Material e Métodos

Os dados analisados nesse estudo foram cedidos pelo Fundecitrus (Fundo de Defesa da Citricultura). Trata-se de dados com incidência de leprose dos citros, em um

talhão da variedade valência sobre limoeiro cravo, plantado em 1996 e localizado no município de Santa Cruz do Rio Pardo/SP. O talhão possui 20 linhas de plantas com 58 plantas em cada linha. O espaçamento entre linhas é de 7,5 metros e entre plantas na linha é de 3,8 metros. Foram analisadas 23 avaliações entre os períodos de junho de 2003 e novembro de 2004. Nesse período não foram realizadas pulverizações com acaricidas no talhão, de modo a não afetar a população de ácaro da leprose.

As 23 avaliações foram feitas em intervalos variáveis de tempo, foram coletadas informações das incidências de doença e de ácaro em cada planta do talhão (0 para ausência do evento e 1 para presença). Foi coletado também o número de ácaro em cada planta.

O talhão de citros analisado no trabalho permite desenvolver uma metodologia de análise que poderá ser estendida para avaliar os demais talhões com incidência de doenças cítricas monitorados espacialmente.

Devido ao fato que, presença, ou não, da doença em determinada planta está altamente correlacionada com a presença, ou não, da doença em plantas vizinhas, espera-se que plantas próximas tenham características similares, particularmente, que a localização das plantas doentes apresente algum padrão espacial. Para os dados apresentados a proposta é de considerar a dependência espacial na modelagem estatística através do modelo autologístico.

Esse artigo estende o modelo logístico usual, pertencente a família dos modelos lineares generalizados (McCullagh & Nelder, 1989; Demétrio & Cordeiro, 2007). Construindo covariáveis de vizinhança uma vez que a configuração espacial dos dados está disponível.

O modelo autologístico calcula a probabilidade p_{ij} de determinada planta na i -ésima linha e j -ésima coluna estar doente como combinação linear do estado da doença (doente/sadia) nas plantas vizinhas e de

covariáveis usuais de um modelo de regressão. Tem-se, a equação do modelo dada por,

$$\text{logit}(p_{ij}) = \beta_0 + \sum_{k=1}^r \beta_k x_{ij} + \sum_{m=1}^s \lambda_m (y_{ij})^t$$

em que, x_{ij} representam as covariáveis usuais da planta na i -ésima linha e j -ésima coluna, $y_{(ij)}$ representam as covariáveis de vizinhança sem ij -ésima planta, particularmente o trabalho assume as seguintes covariáveis de vizinhança $L_{ij} = (y_{i-1,j} + y_{i+1,j})$; $C_{ij} = (y_{i,j-1} + y_{i,j+1})$; $S_{ij} = y_{i,j,t-1}$, que formam, respectivamente, a covariável dentro da linha, entre linhas e mesma planta no tempo anterior. Os β 's são parâmetros de regressão e λ 's são os parâmetros de autocorrelação espacial.

O método de estimação do modelo é por pseudo-verossimilhança (Besag, 1975). Esse método maximiza uma função de pseudo-verossimilhança para o modelo ajustado. Porém, os erros-padrão das estimativas de vizinhança calculados pelo método são inválidos é preciso recalculá-los de outra maneira. Gumpertz et al. (1997) usaram reamostragem com Amostrador de Gibbs para manter a configuração espacial dos dados. Krainski et al. (2008) descrevem o algoritmo de estimação para obter a reamostragem e as novas estimativas.

A essência do modelo autologístico é utilizar covariáveis obtidas a partir da variável resposta em posições vizinhas da localização considerada. Portanto, o elemento chave do modelo é a estrutura de vizinhança. No trabalho a estrutura de vizinhança considerada foi o efeito da planta nas linhas e colunas separadamente (He et al., 2003), além da mesma planta na avaliação anterior da analisada.

Krainski et al. (2008) ajustaram modelos autologísticos para incidência de morte súbita em citros de forma a analisar a presença da doença em plantas vizinhas no tempo contemporâneo à avaliação, no tempo

anterior e um modelo com ambos os tempos. Czermainski (2006) ajustou modelos para dados de leprose de citros observando a influência das covariáveis de vizinhança defasadas no tempo. Em ambos um modelo foi ajustado para cada avaliação coletada.

O fato de que as avaliações dos dados estão em diferentes períodos de tempo, permite que seja explorada a dimensão temporal dos dados. O modelo autologístico espaço-temporal considera os diferentes períodos de tempo entre as avaliações e ajusta apenas um modelo para todas as avaliações coletadas.

$$\text{logit}(p_{ijt}) = \beta_0 + \sum_{k=1}^r \beta_k x_{ij} + \Delta_t \sum_{m=1}^s \lambda_m (y_{ij})^t$$

em que, Δ_t é considerado *offset* no modelo, é o peso do valor da diferença entre os diferentes períodos de tempo (Besag, 1972). As covariáveis de vizinhança são indexadas pelo tempo

$L_{ijt} = (y_{i-1,j,t} + y_{i+1,j,t}); C_{ijt} = (y_{i,j-1,t} + y_{i,j+1,t});$
 $S_{ijt} = y_{i,j,t-1}$. O método de estimação é o mesmo método usado no modelo autologístico (atemporal). propostas diferentes formas de modelar a probabilidade de doença nas plantas, usando o modelo autologístico espaço-temporal.

As equações dos modelos estão na Tabela 1. A primeira proposta é de modelar os dados de uma avaliação t , utilizando a mesma planta na avaliação anterior $t-1$. Esse modelo (M1) permite verificar se o *status* da planta pode estar determinado pela condição dela mesma no tempo anterior. Ou seja, reflete o efeito da permanência da doença na planta.

Outra possibilidade é construir um modelo com covariáveis de vizinhança no período de tempo anterior $t-1$ (M2), que reflete o contágio da doença pela vizinhança do passado.

O terceiro modelo considera as covariáveis de vizinhança no período de tempo contemporâneo a avaliação, expresso pelo M3. Esse modelo reflete um contágio por fontes adicionais que não foram capturadas no

M2, talvez por outros aspectos que não só a influência das plantas vizinhas doentes.

A incidência de ácaro na planta é um importante fator no caso da planta apresentar a doença, essa covariável é acrescentada nos modelos anteriores e formam os três novos modelos sob análise M4, M5 e M6. Ao invés de incidência, talvez o número de ácaros na planta possa explicar melhor o fato da planta ter leprose. Os modelos M7, M8 e M9 avaliam esse efeito juntamente com as covariáveis de vizinhança. O estudo da relação espacial e temporal da associação entre a incidência de plantas infestadas por *B. phoenicis* e da incidência de leprose dos citros, pode trazer estimativas mais precisas sobre o crescimento da epidemia, no tempo e no espaço. O que viabiliza o controle do vetor e práticas de intervenção para a retirada de órgãos afetados, nas áreas de localização dos focos da doença ao invés da ação na área total (Czermainski, 2006).

Os modelos considerados sugerem diferentes formas de explicar a dispersão da doença, o interesse é selecionar um modelo adequado. Um critério de seleção de modelos comumente usado é o critério de informação de Akaike (AIC). O AIC é dado por $2 * \log(L(\hat{\lambda}, y)) + 2k$, em que k é o número de parâmetros incluídos no modelo.

Os procedimentos de análise foram feitos no *software* estatístico R (2008), usando os pacotes Rcitrus e strLattice.

Resultados e Discussão

A incidência de leprose iniciou com 0,1% até 32,76% na última avaliação realizada. Isso mostra que a doença foi aumentando com o passar do tempo. A estrutura espaço-temporal dos dados permite agregar um estudo de padrões, associações de dependência espacial e de tempo de contágio da doença entre plantas. Informações que ajudam na prática de controles da doença.

Foram ajustados aos dados diferentes modelos autologísticos espaço-temporais

descritos na Seção anterior e presentes na Tabela 1. O interesse em encontrar quais variáveis influenciam a presença da doença na planta é parte do trabalho. Através das covariáveis de vizinhança no tempo contemporâneo, uma avaliação anterior, e na mesma planta na avaliação anterior, além de medir os efeitos da incidência e do número de ácaros na planta.

As estimativas e a significância dos parâmetros de cada modelo estão na Tabela 2. As estimativas foram significativas em todos os modelos ajustados. A covariável de vizinhança com maior efeito foi a covariável entre plantas na linha, o espaçamento entre plantas é menor, se comparado ao espaçamento entre linhas de plantas. Indica que plantas mais próximas são mais susceptíveis ao contágio da doença por suas vizinhas doentes.

Não deixa de ser significativa (valor $p < 0,001$) a covariável de vizinhança entre as linhas de plantas em todos os modelos ajustados. Nos modelos em que foram incluídas as covariáveis de ácaro, o efeito também foi significativo tanto para incidência quanto para número de ácaro na planta (valor $p < 0,001$). Mostra que a presença do ácaro na planta resulta na doença, uma vez que o ácaro é o principal vetor da doença (Czermainski, 2006).

A escolha do modelo mais adequado aos dados de leprose de citros foi guiada pelo valor do AIC. Que selecionou como melhor, o modelo com menor valor. Para o caso, o modelo mais adequado aos dados, entre os ajustados, foi o modelo M1.

Uma vantagem do modelo autológico é a possibilidade de quantificar a probabilidade de doença em uma particular planta dado o *status* das plantas vizinhas ou da mesma planta no tempo anterior. No presente estudo a escolha do modelo M1 permite dizer que a doença na mesma planta em uma avaliação anterior é mais importante para explicar a doença que o *status* das plantas vizinhas. Através dos coeficientes do modelo pode-se calcular a probabilidade da planta estar doente

dado que ela estava ou não doente na avaliação anterior. Os resultados desses cálculos mostram que a probabilidade da planta apresentar doença dado que estava doente na avaliação anterior é de 0,517, caso não estivesse doente a probabilidade é de 0,102. Em outras palavras, uma planta que estava doente na avaliação anterior tem aproximadamente 5 vezes mais chance de estar doente que uma planta que estava sadia. O que mostra que em um período de tempo de aproximadamente 22 dias, intervalo aproximado entre as avaliações, uma planta com sintomas permanece com a doença. Evidencia que a planta após adquirir os sintomas da leprose não volta a ficar sadia.

Conclusões

O modelo autológico descreve de forma explícita o padrão espacial de doenças em plantas. Além disso, quantifica o efeito da presença da doença em diferentes estruturas de vizinhança.

Encontrar a probabilidade da planta estar doente dado o *status* da vizinhança auxilia no conhecimento da dinâmica da doença, e indiretamente auxilia nas práticas de controle.

Os dados de incidência de leprose dos citros apresentaram dependência espacial entre as plantas. Quando a planta está doente a evidência é que permaneça doente, além de existir o contágio entre plantas vizinhas, especialmente dentro das linhas onde o espaçamento é menor.

Agradecimentos

Ao Fundecitros, e ao pesquisador Renato Beozzo Bassanezi pela concessão dos dados. Esse trabalho é parte da dissertação de mestrado do primeiro autor no Departamento de Ciências Exatas, Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba.

Referências

- BASSANEZI, R.B.; LARANJEIRA, F.F. Spatial patterns of leprosis and its mite vector in commercial citrus groves in Brazil. **Plant Pathology**, v.56, p.97-106, 2007.
- BASTIANEL M.; FREITAS-ASTÚA J.; KITAJIMA E.W. ; MACHADO, M.A. The citrus leprosis pathosystem. **Summa Phytopathologica**, v.32, p.211-220, 2006.
- BESAG, J. Nearest-neighbour systems and the auto-logistic model for binary data. **Journal of the Royal Statistics Society, Series B**, v. 34, p.75-83, 1972.
- BESAG, J. Statistical analysis of non-lattice data. **The Statistician**, v.24, p.179-195, 1975.
- CZERMAISNKI, A.B.C. **Dinâmica espaço-temporal de populações do patossistema leprose dos citros em condições naturais de epidemia**. 2006. 91p. Tese (Doutorado) – Universidade de São Paulo, Piracicaba.
- DEMÉTRIO, C.G.B.; CORDEIRO, G.M. Modelos lineares genrelizados. In: SIMPÓSIO DE ESTATÍSTICA APLICADA À EXPERIMENTAÇÃO AGRONÔMICA, 12., 2007, Santa Maria. **Minicurso**. Santa Maria: UFSM, 2007. p.159.
- FUNDECITRUS. **Manual da leprose dos citros**. Araraquara: Fundecitrus, 2006. 12p. Disponível em : <http://www.fundecitrus.com.br/manuais.html>. Acesso em: 17 jan. 2008.
- 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. **Journal of Agricultural, Biological and Environmental Statistics**, v.2, p.131-156, 1997.
- HE, F.; ZHOU, J.; ZHU, H. Autologistic Regression Model for the Distribution of Vegetation. **Journal of Agricultural, Biological, and Environmental Statistics**, v.8, p.205-222, 2003.
- KRAINSKI, E.T.; RIBEIRO JR, P.J.; BASSANEZI, R.B.; FRANCISCON, L. Autologistic model with an application to the citrus sudden death disease. **Scientia Agricola**, (em prelo).
- MARQUES, J.P.R.; FREITAS-ASTÚA J.; KITAJIMA, E.W.; APPEZZATO-DA-GLÓRIA, B. Losões foliares e de ramos de laranja-doce causadas pela leprose-dos-citros. **Pesquisa Agropecuária Brasileira**, v.42, p.1531-1536, 2007.
- McCULLAGH, P.; NELDER, J.A. **Generalized linear models**. 2.ed. London: Chapman and Hall, 1989. 511p.
- R DEVELOPMENT CORE TEAM. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. 2008. Disponível em <<http://www.r-project.org>>. Acesso em: 13 jan. 2008.
- RODRIGUES, J.C.V.; **Relações patógeno-vetor-planta no sistema leprose dos citros**. 2000. 168p. Tese (Doutorado) – Universidade de São Paulo, Piracicaba.
- VIOLA, D.; **Deteção e modelagem de padrão espacial em dados binários e de contagem**. 2007. 119p. Tese (Doutorado) – Universidade de São Paulo, Piracicaba.
- ZHU, J; HUANG, H.; WU, J. Modeling spatial-temporal binary data using markov random fields. **Journal of Agricultural, Biological, and Environmental Statistics**, v.10, p. 212-225, 2005.

Tabela 1. Equações dos modelos autolísticos espaço-temporais propostos para o ajuste dos dados de leprose.

Modelos	Equações
M1	$\beta_0 + \lambda_1 S_{(i,j,t-1)}$
M2	$\beta_0 + \lambda_1 L_{(i,j,t-1)} + \lambda_2 C_{(i,j,t-1)}$
M3	$\beta_0 + \lambda_1 L_{(i,j,t)} + \lambda_2 C_{(i,j,t)}$
M4	$\beta_0 + \beta_1 x_{ij} + \lambda_1 S_{(i,j,t-1)}$
M5	$\beta_0 + \beta_1 x_{ij1} + \lambda_1 L_{(i,j,t-1)} + \lambda_2 C_{(i,j,t-1)}$
M6	$\beta_0 + \beta_1 x_{ij1} + \lambda_1 L_{(i,j,t)} + \lambda_2 C_{(i,j,t)}$
M7	$\beta_0 + \beta_1 x_{ij2} + \lambda_1 S_{(i,j,t-1)}$
M8	$\beta_0 + \beta_1 x_{ij2} + \lambda_1 L_{(i,j,t-1)} + \lambda_2 C_{(i,j,t-1)}$
M9	$\beta_0 + \beta_1 x_{ij2} + \lambda_1 L_{(i,j,t)} + \lambda_2 C_{(i,j,t)}$

Tabela 2. Estimativas dos parâmetros e valores p para os três modelos ajustados aos dados de leprose.

Modelos	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\lambda}_1$	$\hat{\lambda}_2$	AIC
M1	-2,283**	-	1,623**	-	18223,40
M2	-2,669**	-	0,747**	1,075**	25674,77
M3	-2,448**	-	0,835**	0,868**	26149,11
M4	-2,485**	1,500**	0,801**	-	26769,44
M5	-2,800**	0,605**	0,688**	1,024**	33820,69
M6	-2,584**	0,663**	0,776**	0,809**	34393,32
M7	-2,341**	1,519**	0,025**	-	27085,15
M8	-2,701**	0,018**	0,692**	1,023**	34064,10
M9	-2,478**	0,018**	0,808**	0,840**	34697,90

**significativos a 1% de probabilidade.