



Análise sobre a evolução do cancro cítrico via modelos mistos

João Pedro Serenini¹, Terezinha A. Guedes¹, Jean Carlos Cardoso¹, Juliana Glória Franco Roberto¹ e William Mário de Carvalho Nunes¹.

¹Universidade Estadual de Maringá

RESUMO

O cancro cítrico causado pela bactéria *Xanthomonas citri* subsp. *citri* traz vários prejuízos as lavouras e indústrias de citros. O uso de variedades cítricas resistentes a tal bactéria é capaz de ajudar no controle da doença, diminuindo o prejuízo. O presente trabalho avalia a resistência de 12 variedades de laranja doce (*Citrus sinensis*) a bactéria *Xanthomonas citri* subsp. *citri* através de modelos generalizados mistos.

Palavras chave: cancro cítrico, *Xanthomonas citri*, modelos mistos.

1 INTRODUÇÃO

O presente estudo, procura encontrar associações biológicas entre espécies de laranja doce no desenvolvimento do cancro cítrico, uma doença que causa grandes prejuízos a indústria.

As lesões causadas por tal doença aparecem nas folhas, ramos e frutos, iniciando com lesões necróticas, geralmente circundadas por um halo amarelo. Com o aumento dos sintomas pode-se gerar rachaduras nos frutos, servindo de porta de entrada para outros patógenos, além de ocasionar queda do mesmo com o passar do tempo.



Figura 1: Lesão de cancro cítrico.

Para medir a suscetibilidade do tecido foliar da planta, uma ferramenta que pode ser utilizada é a avaliação por meio do diâmetro de lesão, que é capaz de verificar a interação entre a *Xanthomonas citri* \times variedades.

Um dos principais modelos estatísticos existentes para analisar dados longitudinais capaz de incorporar a dependência e a estrutura de correlação dos erros são os modelos mistos, [3]. Segundo [2], observações feitas no mesmo indivíduo não possuem correlação nula e os modelos mistos são capazes de incorporar tal correlação na modelagem.

Como os modelos mistos permitem uma melhor descrição da correlação entre os indivíduos inerentemente associadas com a repetição das medidas, pode-se obter melhores inferências estatísticas a respeito das performances médias dos indivíduos, [3].

A avaliação da resistência genética pela inoculação da folha com agulha e a avaliação do sintoma pela medida do diâmetro da lesão, permite realizar estudos mais rápidos, com um menor custo e maior número de plantas ou repetições, além de se ocupar um menor espaço, quando comparado com experimentos realizados no campo.

A variável resposta abordada no presente estudo é o diâmetro da lesão nas folhas da planta causado pelo cancro cítrico, que foi avaliado em centímetros. Além disso o banco de dados apresenta mais três

variáveis: a variedade das plantas, os dias após a inoculação da bactéria nas folhas da planta e o sujeito, onde esse último é dado pelo tipo de variedade em cada dia após a inoculação. Tendo isso em vista, se deseja avaliar o comportamento da variável resposta com relação a esses outros três fatores, desejando-se compreender a influência de cada um deles no comportamento do diâmetro da lesão.

2 METODOLOGIA

Para avaliar a resistência de 12 genótipos de laranja doce (*Citrus sinensis*) à bactéria *Xanthomonas citri* subsp. *citri*, em casa de vegetação, pela medida do diâmetro das lesões [1] realizou um experimento envolvendo 12 variedades de laranja doce.

As amostras foram obtidas a partir da inoculação da bactéria em 6 folhas por planta, num total de 5 plantas por variedade, as quais receberam adubações periódicas de N,P,K. O diâmetro das lesões provocadas foi avaliado aos 18, 24, 33, 39, 46, 61, 68 e 75 dias após a inoculação (DAI), sendo que a presença de halo amarelo ao redor da necrose foi desconsiderado. O experimento foi realizado de Fevereiro a Maio de 2016. Para a mensuração dos diâmetros foi utilizado um paquímetro, sendo utilizados três porta enxertos diferentes, Swingle, Sunki e Limão cravo.

O conjunto de dados é do tipo longitudinal, formado por um total de 2880 observações contendo 4 variáveis, sendo elas, as variedades de laranja doce utilizadas no experimento (Geno), o número de plantas em cada variedade (sujeitos), os dias após a inoculação (DAI) e o diâmetro da lesão (Diam) sendo essa última a variável resposta, que é do tipo contínua, medida em centímetros.

O estudo tem por objetivo avaliar a resistência de 12 variedades de laranja doce (*Citrus sinensis*) a bactéria *X. citri* subsp. *citri*, em condições parcialmente controladas (casa de vegetação), por inoculação da folha com agulha e avaliação do diâmetro da lesão através de modelos generalizados mistos.

2.0.1 Modelo Linear Generalizado Misto (GLMM)

Um modelo generalizado misto é um modelo estatístico que busca descrever o comportamento de uma variável aleatória com relação a uma ou mais variáveis explicativas, onde estas podem ser tanto de efeitos fixos como aleatórios. A grande diferença desses modelos com relação a um modelo linear misto, é que nos generalizados a variável resposta não precisa estar necessariamente associada a distribuição normal, o que amplia a abrangência de tais modelos.

Para se especificar um modelo generalizado misto, a variável resposta condicionada aos efeitos aleatórios ($y_i|b$) deve seguir uma distribuição pertencente a família exponencial:

$$y_i|b \sim f_{Y_i|b}(y_i|\theta) \quad (1)$$

$$f_{Y_i|b}(y_i|\theta) = b(y_i) \exp\{\eta(\theta)^T T(y_i) - a(\theta)\} \quad (2)$$

Tipicamente se deseja relacionar os parâmetros da distribuição com os preditores do modelo, para isso é feito uma transformação da média μ_i , para que a mesma se relacione com tais preditores:

$$E(y_i|b) = \mu_i \quad (3)$$

$$g(E(y_i|b)) = g(\mu_i) = \eta_i = X_i\beta + Z_i b, \quad (4)$$

em que $g(\cdot)$ é uma função conhecida, chamada função de ligação, responsável por relacionar a média de y_i com os preditores lineares, X_i é a i -ésima linha da matriz modelo dos efeitos fixos, β é o vetor de parâmetros dos efeitos fixos, Z_i é a i -ésima linha da matriz modelo dos efeitos aleatórios e b é o vetor dos efeitos aleatórios.

A variância de $y_i|b$ é dada por $Var(y_i|b) = \phi^2 v(\mu_i)$, em que v é uma função de variância conhecida dada pela distribuição de y_i e ϕ é o parâmetro de escala da distribuição, o qual precisa ser estimado.

3 RESULTADOS E DISCUSSÕES

As distribuições log-normal $LN(\mu, \sigma)$ e gama $GA(\mu, \sigma)$ com função de ligação logarítmica foram utilizadas para analisar o banco de dados obtido por [1] com o objetivo de avaliar a resistência de 12 variedades de laranja doce. Vale notar que a distribuição gama foi reparametrizada de forma que suas estimativas possam ter melhores fins práticos em suas interpretações, em que $\mathbb{E}(Y) = \mu$, $Var(Y) = \sigma^2 \mu^2$, $\sigma^2 = 1/\alpha$ e $\mu = \alpha\beta$.

O conjunto de dados contém 2880 observações do diâmetro médio das lesões (12 Genótipos \times 8 Dias \times 60 Sujeitos \times 6 Repetições), considerando todas as covariáveis e o efeito do tempo. Para se obter as estimativas dos parâmetros foram utilizado o *software* R com o pacote *gamlss*, em que este permite que a parte sistemática do modelo seja expandida permitindo modelar não apenas o parâmetro de média, mas também outros parâmetros da distribuição. A função de verossimilhança penalizada l_p foi utilizada para estimar os parâmetros dos modelos.

Os modelos foram ajustados na variável resposta y_{ijk} , $i = 1, 2, \dots, 12$ (número de variedades), $j = 1, 2, \dots, 8$ (dias após a inoculação), $k = 1, 2, \dots, 60$ (número de sujeitos). A seleção do modelo foi baseada no Critério de Informação de Akaike (AIC), na Deviance Global (GD), no Critério bayesiano de Schwarz (BIC) e no teste de razão de Verossimilhança (LRT). Os modelos ajustados estão descritos na Tabela 1.

Tabela 1: Modelos propostos para as distribuições Log-normal e Gama

ID	Modelo
1	$M_0 \quad \eta = \beta_0$
2	$M_1 \quad \eta_i = \beta_0 + \alpha_i \text{Variedade}$
3	$M_2 \quad \eta_{ij} = \beta_0 + \alpha_i \text{Variedade} + \theta_j \text{DAI}$
4	$M_3 \quad \eta_{ijk} = \beta_0 + \alpha_i \text{Variedade} + \theta_j \text{DAI} + u_{0k}$
5	$M_4 \quad \eta_{ijk} = \beta_0 + \alpha_i \text{Variedade} + \theta_j \text{DAI} + u_{1k} \text{Sujeito}$
6	$M_5 \quad \eta_{ijk} = \beta_0 + \alpha_i \text{Variedade} + \theta_j \text{DAI} + u_{0k} + u_{1k} \text{Sujeito}$

Em tais modelos tem-se que α_i representa os efeitos devido a i -ésima variedade, θ_j os efeitos devido ao j -ésimo dia após a inoculação, u_k é o efeito da parte aleatória para o k -ésimo sujeito.

Levando em consideração os testes de razão de verossimilhança e as medidas de decisão foram escolhidos o modelo M_3 como o melhor modelo para ambas as distribuições.

A validade de tais modelos podem ser confirmadas verificando os pressupostos dos mesmos, os quais foram avaliados através da análise de resíduos. Tal análise foi baseada nos resíduos quantílicos aleatorizados e nos gráficos de diagnósticos apresentados nas Figuras (2) e (3) para os modelos Log-Normal e Gama, respectivamente.

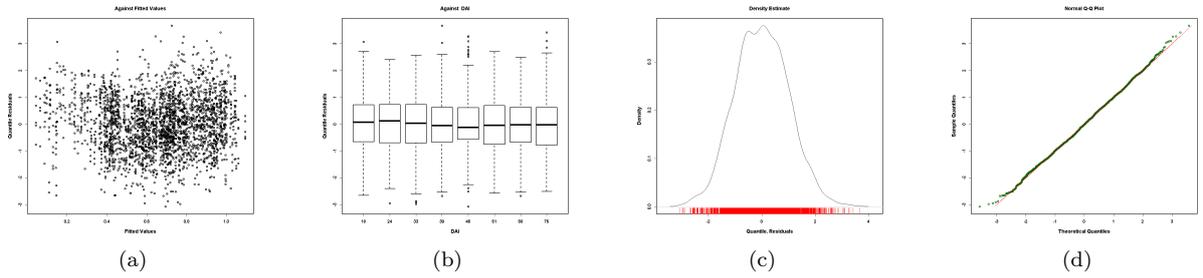


Figura 2: (a) Valores Ajustados vs Quantis Residuais; (b) Dias Após a Inoculação vs Quantis Residuais; (c) Quantis Residuais vs Densidade Estimada; (d) Quantis Teóricos vs Quantis Amostrais.

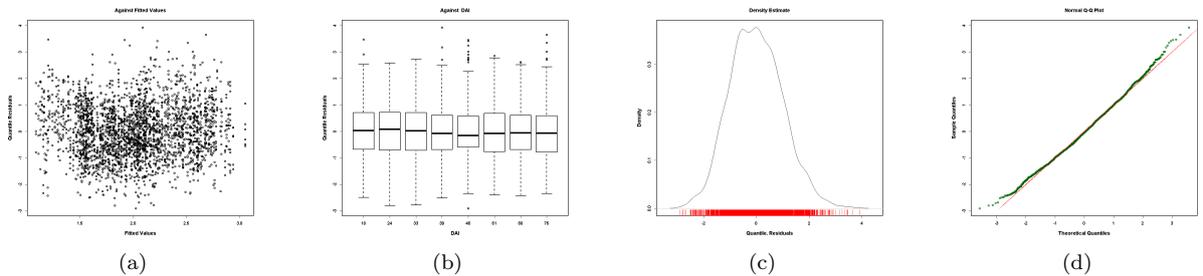


Figura 3: (a) Valores Ajustados vs Quantis Residuais; (b) Dias Após a Inoculação vs Quantis Residuais; (c) Quantis Residuais vs Densidade Estimada; (d) Quantis Teóricos vs Quantis Amostrais.

Nas Figuras (2), (3) (a) e (b) dos valores ajustados vs quantis residuais e do DAI vs quantis residuais, mostram que os resíduos estão distribuídos de forma aleatória em torno do zero, o que não revela nenhuma violação quanto a homoscedasticidade. Já os dois últimos, quantis residuais vs densidade estimada e

quantis teóricos *vs* quantis amostrais, indicam que os quantis residuais seguem uma distribuição Normal Padrão, que é requisito para um ajuste adequado.

Na Tabela (2), são apresentadas as estimativas dos parâmetros fixos do modelo Log-normal, na Tabela (3) são apresentadas as estimativas dos parâmetros fixos do modelo gama, para o parâmetro μ .

Tabela 2: Estimativas dos parâmetros do modelo Log-Normal para μ .

Tabela 3: Estimativas dos parâmetros do modelo Gama para μ .

Variável	Parâmetro	Estimativa	Erro Padrão	valor t	Pr(> t)	Variável	Parâmetro	Estimativa	Erro Padrão	valor t	Pr(> t)
Intercepto	β_0	0.45219	0.01658	27.278	$< 2e - 16$	Intercepto	β_0	0.47854	0.01652	28.973	$< 2e - 16$
Irradiada	α_1	-0.30186	0.01769	-17.061	$< 2e - 16$	Irradiada	α_1	-0.29417	0.01781	-16.516	$< 2e - 16$
Itapetinig	α_2	0.02746	0.01567	1.752	0.079875	Itapetinig	α_2	0.02708	0.01558	1.738	0.0823
Paloma	α_3	-0.02626	0.01564	-1.679	0.093186	Paloma	α_3	-0.02697	0.01565	-1.723	0.0849
PeraIAC	α_4	-0.03066	0.01491	-2.057	0.039787	PeraIAC	α_4	-0.03412	0.01486	-2.296	0.0217
PeraMga	α_5	-0.02102	0.01557	-1.350	0.177073	PeraMga	α_5	-0.02178	0.01544	-1.411	0.1585
PeraOri	α_6	0.06182	0.01633	3.787	0.000156	PeraOri	α_6	0.06406	0.01638	3.909	$9.47e - 05$
PrecOri	α_7	0.01207	0.01519	0.794	0.427121	PrecOri	α_7	0.01035	0.01521	0.680	0.4964
Puka	α_8	-0.06376	0.01512	-4.216	$2.56e - 05$	Puka	α_8	-0.06651	0.01510	-4.404	$1.10e - 05$
Sanguine	α_9	-0.14319	0.01593	-8.987	$< 2e - 16$	Sanguine	α_9	-0.14227	0.01588	-8.957	$< 2e - 16$
Valen	α_{10}	-0.34847	0.01779	-19.586	$< 2e - 16$	Valen	α_{10}	-0.34077	0.01757	-19.395	$< 2e - 16$
VALEN2	α_{11}	-0.35253	0.01783	-19.775	$< 2e - 16$	VALEN2	α_{11}	-0.34491	0.01767	-19.521	$< 2e - 16$
DAI 24	θ_1	0.12558	0.01502	8.361	$< 2e - 16$	DAI 24	θ_1	0.11186	0.01486	7.528	$6.93e - 14$
DAI 33	θ_2	0.30327	0.01608	18.855	$< 2e - 16$	DAI 33	θ_2	0.29303	0.01599	18.328	$< 2e - 16$
DAI 39	θ_3	0.27364	0.01506	18.173	$< 2e - 16$	DAI 39	θ_3	0.25854	0.01500	17.233	$< 2e - 16$
DAI 46	θ_4	0.27809	0.01469	18.926	$< 2e - 16$	DAI 46	θ_4	0.26140	0.01464	17.856	$< 2e - 16$
DAI 61	θ_5	0.43113	0.01600	26.953	$< 2e - 16$	DAI 61	θ_5	0.42012	0.01597	26.313	$< 2e - 16$
DAI 68	θ_6	0.53634	0.01681	31.900	$< 2e - 16$	DAI 68	θ_6	0.53002	0.01671	31.719	$< 2e - 16$
DAI 75	θ_7	0.48684	0.01616	30.135	$< 2e - 16$	DAI 75	θ_7	0.47733	0.01618	29.501	$< 2e - 16$

Os resultados apresentados nas Tabelas (2) e (3) mostram o quanto o diâmetro das lesões causados pela inoculação da bactéria *Xanthomonas citri* subsp. *citri* variam em média em relação a variedade Hamlin e o quanto o diâmetro das lesões variam em média ao tempo comparado ao dia 18. Nota-se que as variedades Valência e a Irradiada foram as que apresentaram o menor crescimento médio do diâmetro no decorrer do tempo, enquanto que a Perâ Oriçanga foi a que apresentou o maior, seguido pela Perâ Itapetininga e Precoce Oriçanga. Quanto aos dias após a inoculação nota-se que conforme o tempo passa o diâmetro tem uma tendência a aumentar de tamanho, havendo uma queda entre os dias 33 a 46, em que o dia 68 apresentou as maiores medidas do diâmetro. Tais resultados confirmaram o que a análise descritiva feita anteriormente havia mostrado.

4 CONCLUSÃO

De forma geral, tanto os modelos log-normal e gama se mostraram adequado para descrever o comportamento do cancro cítrico nas variedades de laranja doce avaliados. Observou-se que as duas variedades de Valência foram os que se mostraram mais resistentes ao cancro cítrico, seguido pela variedade Perâ Irradiada. Já a variedade Perâ Oriçanga foi a que se mostrou menos resistente a doença, seguida pelas variedades Perâ Itapetininga e Precoce Oriçanga. Sendo assim, pode-se dizer que esta metodologia mostrou-se adequada para a modelagem dos dados, pois possibilitou detectar as variedades mais suscetíveis ao cancro cítrico, atendendo os pressupostos necessários a modelagem e fornecendo resultados confiáveis.

Referências

- [1] FRANCO, J. G. Avaliação da resistência de genótipos de laranja doce (*citrus sinensis*) ao cancro cítrico em condições de casa de vegetação, 2016
- [2] PINHEIRO, J. C. *Topics in mixed effects models*. Tese (Doutorado) - UNIVERSITY OF WISCONSIN-MADISON, 1994.
- [3] WALLACE, D.; GREEN, S. B. Analysis of repeated measures designs with linear mixed models. Lawrence Erlbaum Associates Publishers, 2002.