



Comparação entre ensaios longitudinais no desenvolvimento do cancro cítrico.

João Pedro Serenini¹, Terezinha A. Guedes¹, Ana Beatriz T. Martins¹, Everton da Costa¹, Bruna Rafaela Barbieri¹ e William Mário de Carvalho Nunes¹.

¹Universidade Estadual de Maringá

RESUMO

O cancro cítrico causado pela bactéria *Xanthomonas citri* subsp. *citri* traz vários prejuízos às lavouras e indústrias de citros. A identificação de variedades de laranjas resistentes a tal patógeno pode ajudar a contornar tal problema. Para isso foi realizado um estudo envolvendo dois ensaios, os quais foram avaliados a partir de modelos generalizados mistos.

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

1 INTRODUÇÃO

O cancro cítrico, causado pela bactéria *Xanthomonas citri* subsp. *citri* tem trazido grandes prejuízos às indústrias ao longo do tempo. As lesões causadas pela infecção por tal bactéria aparecem nas folhas, ramos e frutos, iniciando com lesões necróticas, geralmente circundadas por um halo amarelo. Com o passar do tempo os sintomas tendem a aumentar podendo gerar rachaduras nos frutos, o que contribui na infecção de outros patógenos, além de ocasionar queda dos frutos conforme aumenta a gravidade dos sintomas [2].



Figura 1: Lesão de cancro cítrico.

Para medir o quanto o tecido foliar é suscetível ao patógeno, uma ferramenta utilizada é a avaliação por meio do diâmetros de lesão causada pela doença, que verifica o desenvolvimento da doença nas diversas variedades avaliadas [3].

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. O banco de dados possui dois ensaios e além do diâmetro apresenta mais três variáveis: a variedade das plantas, os dias após a inoculação e as folhas de cada variedade. 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 em cada variedade, buscando-se também diferenças significativas em cada ensaio realizado nesse experimento, através de modelos mistos generalizados.

Entre os diversos modelos estatísticos existentes para analisar dados longitudinais, os modelos mistos se destacam por serem capazes de incorporar a dependência e a estrutura de correlação dos erros, [5]. Segundo [4], observações feitas no mesmo indivíduo não possuem correlação nula e os modelos mistos são capazes de incorporar tal correlação na modelagem.

2 METODOLOGIA

As amostras foram obtidas a partir da inoculação da bactéria em 16 plantas, 5 folhas por planta, num total de 6 perfurações por folha. O diâmetro das lesões provocadas foi avaliado aos 19, 25, 33, 39, 46, 61, 68 e 76 dias após a inoculação (DAI).

O conjunto de dados é do tipo longitudinal, formado por dois ensaios num total de 3840 observações em cada, contendo 4 variáveis, sendo elas, as variedades de laranja utilizadas no experimento (Trat), as folhas das plantas em cada variedade (Folhas), 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.

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 à 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 [1]:

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

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

Tipicamente se deseja relacionar os parâmetros da distribuição com os preditores do modelo. Para isso é feita 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.

Neste trabalho, a distribuição inversa gaussiana com função de ligação logarítmica $IG(\mu, \sigma^2)$ foi utilizada para analisar o banco de dados com o objetivo de avaliar a resistência das 16 variedades de laranja.

3 RESULTADOS E DISCUSSÕES

O conjunto de dados analisado, contém dois ensaios, os quais se deseja comparar, cada qual com 3840 observações do diâmetro médio das lesões (16 Variedades \times 8 Dias \times 80 Folhas \times 6 Repetições), considerando todas as covariáveis e o efeito do tempo. Para se obter as estimativas dos parâmetros foram utilizado o *software* SAS com a proc GLIMMIX.

Os modelos foram ajustados na variável resposta y_{ijk} , $i = 1, 2, \dots, 16$ (número de variedades), $j = 1, 2, \dots, 8$ (dias após a inoculação), $k = 1, 2, \dots, 80$ (número de folhas). A seleção do modelo foi baseada no critério de informação de Akaike (AIC), na log-verossimilhança penalizada, 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 a distribuição inversa gaussiana

ID	Modelo
1	$IG_1 \quad \eta = \beta_0$
2	$IG_2 \quad \eta_i = \beta_0 + \alpha_i \text{Variedade}$
3	$IG_3 \quad \eta_{ij} = \beta_0 + \alpha_i \text{Variedade} + \theta_j \text{DAI}$
4	$IG_4 \quad \eta_{ijk} = \beta_0 + \alpha_i \text{Variedade} + \theta_j \text{DAI} + u_{0k}$
5	$IG_5 \quad \eta_{ijk} = \beta_0 + \alpha_i \text{Variedade} + \theta_j \text{DAI} + u_{1k} \text{Folha}$
6	$IG_5 \quad \eta_{ijk} = \beta_0 + \alpha_i \text{Variedade} + \theta_j \text{DAI} + u_{0k} + u_{1k} \text{Folha}$

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 a k -ésima folha.

Levando em consideração os testes de razão de verossimilhança e as medidas de decisão o modelo IG_5 foi escolhido como o melhor modelo.

As Tabelas (2) e (3) apresentam as estimativas da parte aleatória para o modelo selecionado referente ao primeiro e segundo ensaio respectivamente.

Tabela 2: Estimativas da parte aleatória para o primeiro ensaio.

Estimativas dos parâmetros de covariância			
Param. cov.	Sujeito	Estimativa	Erro padrão
Intercepto	Folha	0.003746	0.001043
DAI	Folha	0.01301	0.000969
Residual		0.01396	0.000348

Tabela 3: Estimativas da parte aleatória para o segundo ensaio.

Estimativas dos parâmetros de covariância			
Param. cov.	Sujeito	Estimativa	Erro padrão
Intercepto	Folha	0.000350	0.000160
DAI	Folha	0.002506	0.000259
Residual		0.007474	0.000187

As estimativas em tais tabelas representam a variância dos efeitos aleatórios (b_i) e a variância dos erros, respectivamente. Em tais tabelas estão sendo representadas o quanto da variabilidade dos dados está sendo explicada pelo intercepto aleatório em cada folha (Intercepto), pelo coeficiente angular aleatório em cada folha com relação aos dias (DAI) e o quanto da variabilidade não pode ser explicada pelo modelo (Residual).

As Tabelas (4) e (5) apresentam as estimativas da parte fixa para o modelo selecionado referente ao primeiro e segundo ensaio.

Tabela 4: Estimativas da parte fixa para o primeiro ensaio.

Variável	Estimativa	Erro padrão	Valor t	Pr > t
Intercept	0.5037	0.03676	13.70	< .0001
MORCOTT	-0.1993	0.04826	-4.13	< .0001
NATAL 24	-0.2804	0.04820	-5.82	< .0001
NATAL 26	-0.3574	0.04814	-7.42	< .0001
NATAL 30	-0.1849	0.04828	-3.83	0.0001
NATAL M9-350	-0.3366	0.04816	-6.99	< .0001
NATAL M9-324	-0.2284	0.04824	-4.73	< .0001
PERA 329	-0.2241	0.04824	-4.64	< .0001
PERA 331	-0.2119	0.04825	-4.39	< .0001
PERA 436	-0.3247	0.04816	-6.74	< .0001
PERA 460	-0.05277	0.04840	-1.09	0.2757
RUBI 251	-0.2290	0.04824	-4.75	< .0001
RUBI 353	-0.01164	0.04844	-0.24	0.8101
VALENCIA	-0.2257	0.04824	-4.68	< .0001
WESTIN 1	0.03671	0.04850	0.76	0.4491
WESTIN 3	-0.05163	0.04840	-1.07	0.2861
BAHIA 25	0	.	.	.
25	-0.1096	0.02006	-5.47	< .0001
33	0.004516	0.02017	0.22	0.8229
39	0.06031	0.02022	2.98	0.0030
46	-0.03372	0.02013	-1.67	0.0946
61	0.03090	0.02020	1.53	0.1266
68	0.1102	0.02028	5.44	< .0001
76	0.1394	0.02031	6.86	< .0001
19	0	.	.	.

Tabela 5: Estimativas da parte fixa para o segundo ensaio.

Variável	Estimativa	Erro padrão	Valor t	Pr > t
Intercept	0.3599	0.01500	23.99	< .0001
MORCOTT	-0.1497	0.01880	-7.96	< .0001
NATAL 24	-0.1250	0.01882	-6.64	< .0001
NATAL 26	-0.08825	0.01887	-4.68	< .0001
NATAL 30	-0.09901	0.01885	-5.25	< .0001
NATAL M9-350	-0.08592	0.01887	-4.55	< .0001
NATAL M9-324	-0.1134	0.01884	-6.02	< .0001
PERA 329	-0.1141	0.01884	-6.06	< .0001
PERA 331	-0.08577	0.01887	-4.55	< .0001
PERA 436	-0.1172	0.01883	-6.22	< .0001
PERA 460	-0.01445	0.01895	-0.76	0.4459
RUBI 251	-0.02985	0.01893	-1.58	0.1150
RUBI 353	-0.01139	0.01896	-0.60	0.5479
VALENCIA	-0.04497	0.01892	-2.38	0.0175
WESTIN 1	-0.00664	0.01896	-0.35	0.7263
WESTIN 3	-0.05964	0.01890	-3.16	0.0016
BAHIA 25	0	.	.	.
25	-0.02059	0.01019	-2.02	0.0437
33	-0.01641	0.01019	-1.61	0.1079
39	0.01434	0.01022	1.40	0.1612
46	0.07208	0.01028	7.01	< .0001
61	0.08655	0.01030	8.40	< .0001
68	0.1254	0.01034	12.12	< .0001
76	0.1852	0.01041	17.78	< .0001
19	0	.	.	.

A Figura 2 apresenta os gráficos de resíduos para o modelo selecionado referente ao primeiro ensaio, onde não se notam evidências de que os pressupostos do mesmo não foram atendidos.

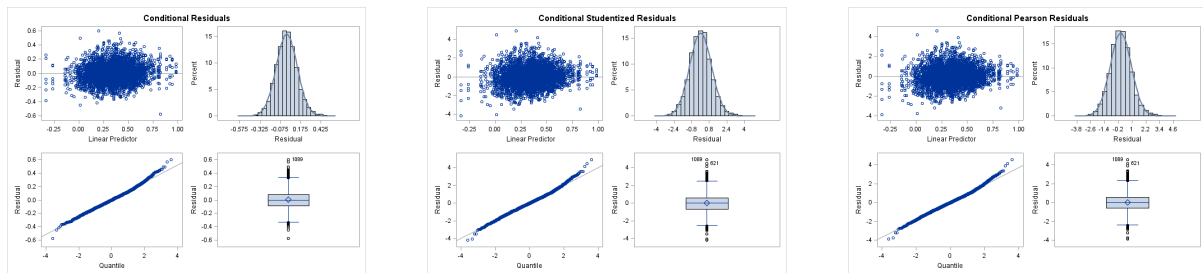


Figura 2: Resíduos condicionais, padronizados e de Pearson para o modelo IG_5 .

Por outro lado, a Figura 3 mostra os gráficos de resíduos para o modelo IG_5 referente ao segundo ensaio onde, também, não se notam evidências de que os pressupostos deste modelo não foram atendidos.

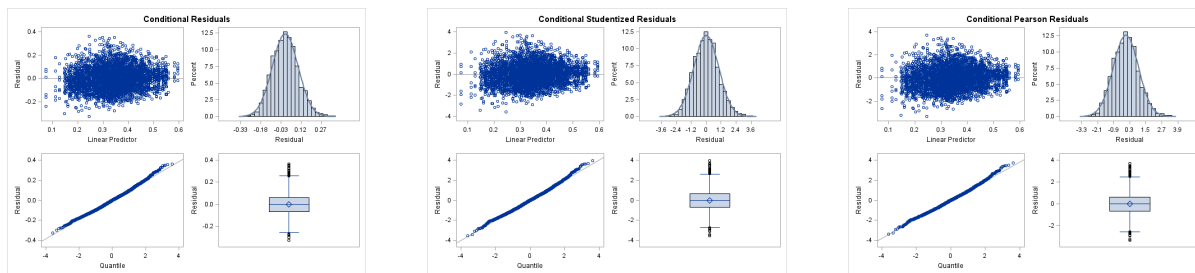


Figura 3: Resíduos condicionais, padronizados e de Pearson para o modelo IG_5 .

4 CONCLUSÃO

De forma geral, o modelo IG_5 se mostrou adequado para descrever o comportamento do cancro cítrico nas variedades de laranja avaliadas em ambos os ensaios. A variedade BAHIA 25, assim como o DAI 19 foram tomados como *baseline*, ou seja, as estimativas de todas as variedades foram comparadas com a da BAHIA, assim como as estimativas de todos os DAI foram comparados com a do dia 19, logo observa-se que no primeiro ensaio a variedade Natal 26 foi a que se mostrou mais resistente ao cancro cítrico, pois é a que apresentou a menor estimativa em relação as outras variedade o que indica que em comparação com a variedade BAHIA ela foi a que teve o menor crescimento médio do diâmetro no decorrer do tempo, seguido pelas variedades Natal M9-350 e Pera 436. Já a variedade Westin 1 foi a que se mostrou menos resistente à doença, seguida pelas variedades Rubi 353 e Westin 3. Para o segundo ensaio observou-se que a variedade Morcot foi a que se mostrou mais resistente, seguida por Natal 24 e Pera 436. Já a variedade Westin 1, novamente, foi a que se mostrou menos resistente à doença, seguida por Rubi 357 e Pera 460. 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 à modelagem e fornecendo resultados confiáveis.

Referências

- [1] DOBSON, A. J. Exponential family and generalized linear models. An Introduction to Generalized Linear Models. Boca Raton, FL: Chapman and Hall/CRC, 2002.
- [2] LARANJEIRA, F. F.; AMORIM, L.; FILHO, A. B.; AGUILAR-VILDOSO, C. I.; FILHO, H. D. C.; JUNIOR, D. M.; NEGRI, J. D.; PIO, R.; JUNIOR, J. P. Fungos, procaríotos e doenças abióticas. MATTOS JUNIOR, D.; DE NEGRI, JD; PIO, RM, p. 511-566, 2005.
- [3] NOCITI, L. A.; CAMARGO, M.; NETO, J. R.; FRANCISCHINI, F. J.; JR, J. B. Agressividade de linhagens de *xanthomonas axonopodis* pv. *aurantifolii* tipo c em lima ácida 'galego'. Fitopatologia Brasileira, SciELO Brasil, v. 31, p. 140-146, 2006.
- [4] PINHEIRO, J. C. *Topics in mixed effects models*. Tese (Doutorado) - UNIVERSITY OF WISCONSIN-MADISON, 1994.
- [5] WALLACE, D.; GREEN, S. B. Analysis of repeated measures designs with linear mixed models. Lawrence Erlbaum Associates Publishers, 2002.