

Visualização e análise de dados em estudos de patologia pós-colheita

Aline de Holanda N. Maia¹; Daniel Terao¹; Alfredo José Barreto Luiz¹

¹ Embrapa Meio Ambiente; Jaguriúna, São Paulo, Caixa Postal 69, CEP: 13.820-000,
e-mail: ahmaia@cnpma.embrapa.br

Modern data graphics can do much more than simply substitute for statistical tables. At their best, graphics are instruments for reasoning about quantitative information. Often, the most way to describe, explore and summarize a set of numbers – even a very large data set – is to look at pictures of those numbers. Furthermore, of all methods for analyzing and communicating statistical information, well designed data graphics are usually the simplest and at the same time the most powerful (Edward R. Tuft, *The Visual Display of Quantitative Information*, 2001).

A necessidade de estudos para desenvolvimento de tecnologias de redução de resíduos tóxicos em frutas e vegetais, decorrentes do controle fitossanitário pós-colheita, é ainda hoje um imperativo no Brasil. A utilização de métodos adequados para análise de ensaios de dados provenientes de estudos de sanidade pós-colheita é importante componente no processo de geração dessas tecnologias, entre elas o uso de bioagentes de controle. Além do uso de métodos adequados de análises descritivas e inferenciais, o uso de métodos de visualização dos resultados é de fundamental importância para a comunicação das ferramentas estatísticas geradas (TUFTE, 2001).

A eficiência de métodos de controle de doenças pós-colheita em frutos é geralmente avaliada em testes onde são mensuradas variáveis de diferentes naturezas ao longo do tempo, como por exemplo, a porcentagem de frutos infectados, o tamanho médio das lesões por fruto ou ainda escores indicadores da severidade da doença. A escolha dos métodos adequados para análise de tais variáveis depende tanto do delineamento experimental utilizado, quanto da natureza das variáveis de interesse. Variáveis com distribuição não Normal, como contagens de esporos, status do fruto (infectado ou não infectado) ou ainda escores de severidade requerem métodos especialmente delineados para esses tipos de resposta, como os métodos para análise de tabelas de contingência baseados em quadrados mínimos ponderados (QMP; STOKES et al., 2000) ou a classe de modelos denominada “Modelos Lineares Generalizados” (MLG; McCULLAGH & NELDER, 1989). Quando variáveis são observadas ao longo do tempo numa mesma unidade experimental, caracterizando um estudo longitudinal ou de medidas repetidas é necessária especial atenção para a modelagem das prováveis correções temporais entre as medidas, o que pode ser considerado por modelos que constituem uma generalização dos MLG, os “Modelos Lineares Generalizados Mistos” (MLGM; BRESLOW & CLAYTON, 1983).

Apresentamos para discussão, alguns exemplos de análises inferenciais para testes de hipóteses científicas pertinentes e métodos descritivos gráficos para visualização e análise dos resultados de ensaios para avaliação da eficiência de bioagentes de controle da podridão de *Fusarium pallidoroseum* em melão (*Cucumis melo*).

Inicialmente, tratamentos com perfis similares são identificados, considerando as avaliações longitudinais para uma determinada resposta, via análise de agrupamentos, cujo resultado é apresentado numa árvore de classificação ou dendograma. Gráficos descritivos dos perfis são construídos para cada grupo; para visualizar a evolução dos escores de severidade da doença, são construídos gráficos de barras empilhadas. O estudo da evolução temporal da doença nos frutos é uma informação crítica para o planejamento do armazenamento e transporte. Um

exemplo de análise inferencial de uma variável não Normal, utilizando o método QMP, é introduzido para comparação de porcentagens de infecção observadas em frutos tratados com bioagentes e frutos tratados com um fungicida tradicional e com o controle absoluto.

Nos exemplos discutidos a seguir, foram utilizados dados de dois ensaios, ambos com delineamento inteiramente casualizado, para avaliar a eficiência de bioagentes de controle sobre a podridão de *Fusarium* (*Fusarium pallidoroseum*) em melões da variedade Gália. Foram inoculados dez frutos para cada tratamento e a porcentagem de frutos infectados foi quantificada em diferentes tempos após a infecção.

No experimento I, foram avaliados onze bioagentes potenciais, um fungicida convencional e um controle absoluto (Tabela 1). As avaliações foram feitas aos 8, 10, 12 e 14 dias após a infecção. No experimento II, foi excluído um bioagente do experimento anterior. As avaliações correram aos 7, 9 e 11 dias após a inoculação.

Nas análises descritivas, tratamentos com perfis similares são identificados, considerando as avaliações longitudinais para uma determinada resposta, via análise de agrupamentos. Como exemplo, utilizamos os dados de porcentagem de frutos infectados em cada tratamento nas avaliações realizadas aos 8, 10, 12 e 14 dias após a inoculação. Na análise de agrupamentos, foi utilizado o PROC CLUSTER do software SAS[®]/STAT (SAS[®], 2008), considerando como métrica a distância euclidiana e como método de agrupamento o método do centróide. Gráficos descritivos dos perfis são construídos para cada grupo, de modo a representar a evolução temporal da doença nos frutos. Os frutos foram classificados com relação à severidade da doença em quatro classes de controle (total, alto, médio e baixo) de acordo com a magnitude dos escores de severidade (1 a 10). Foram construídos gráficos de barras empilhadas para visualizar a evolução da porcentagem de frutos em cada classe ao longo do tempo.

Como exemplo do uso de métodos inferenciais, apresentamos a análise das porcentagens de frutos infectados no Experimento I e do diâmetro médio de lesão no Experimento II. O método QMP foi usado para comparar os bioagentes testados com o controle absoluto e com o fungicida tradicional, aos dez dias após a infecção (análise de contrastes usando o CONTRAST Statement do PROC CATMOD, do software SAS[®]/STAT). Foi utilizado um modelo saturado para obtenção de estimativas das porcentagens de infecção em cada tratamento. Para contrastes entre tratamentos foi usado o teste de Wald. O efeito dos tratamentos sobre o diâmetro médio de lesão (média dos diâmetros transversal e longitudinal) foi avaliado via MLG, com distribuição Normal e função de ligação identidade, utilizando o PROC MIXED do SAS[®]/STAT. Os diâmetros médios de lesão para os bioagentes foram comparados com os respectivos diâmetros dos tratamentos com fungicida tradicional e controle absoluto por meio de testes F para contrastes.

Análise descritiva dos perfis temporais de infecção por Fusarium

Foram identificados quatro grupos de tratamentos em relação à similaridade dos perfis de evolução temporal da doença, mensurada pela porcentagem de frutos infectados (PFI) (Figura 1). No grupo I (SB, BL, TEST), a PFI já era elevada aos 8 dias, evoluindo para valores superiores a 90% até o final do período de avaliação (14 dias). Nos grupos 2 (IMZ, L-21, L-62) e 3 (BS, BS+BL, L-17, L-41) a PFI era inicialmente baixa mas atingiu valores elevados ao final da avaliação. A rapidez de evolução foi mais acentuada no grupo 3. No grupo 4, a PFI permaneceu estável ao longo do tempo, com valores em torno de 50% (Figura 2). A análise gráfica preliminar indica que o grupo 2 é mais eficiente para controle da doença aos 10 dias; nesse grupo, os bioagentes *Pichia membranifaciens* e *Debaryomyces hansenii* apresentaram eficiência similar ao fungicida convencional (Imazalil).

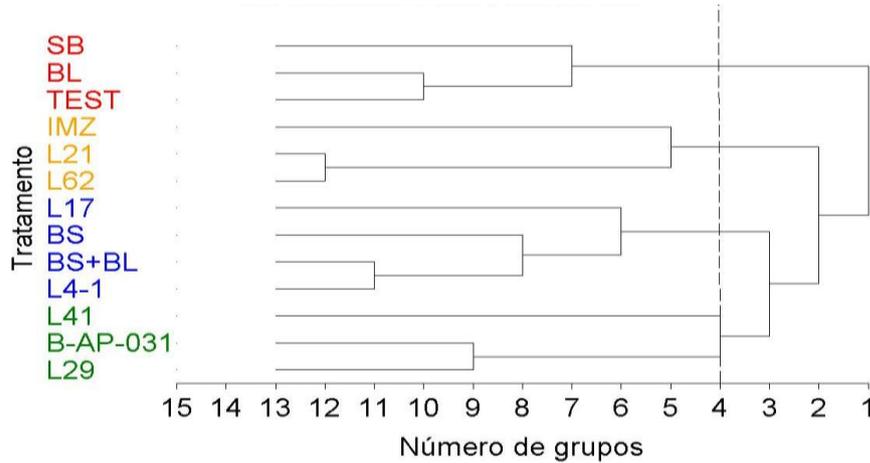


Figura 1. Árvore de classificação ou dendograma representando o agrupamento dos tratamentos com relação à similaridade entre perfis de evolução temporal de infecção de 8 a 14 dias após inoculação do agente causal da podridão de Fusarium, (*Fusarium pallidoroseum*), no Experimento I.

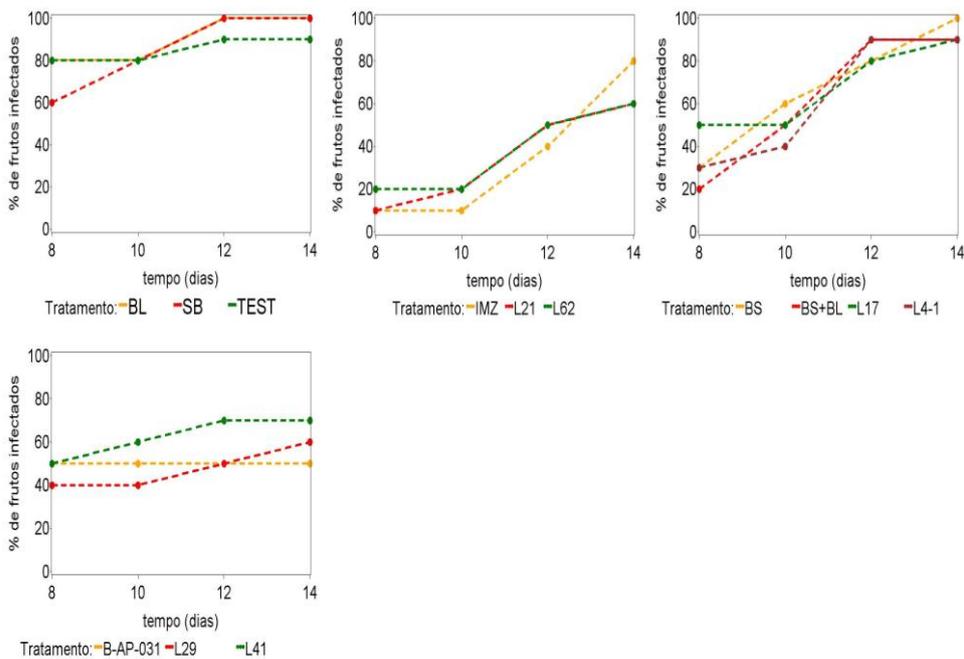


Figura 2. Evolução temporal da infecção período de 8 a 14 dias após inoculação do agente causal da podridão de Fusarium, no Experimento I (*Fusarium pallidoroseum*).

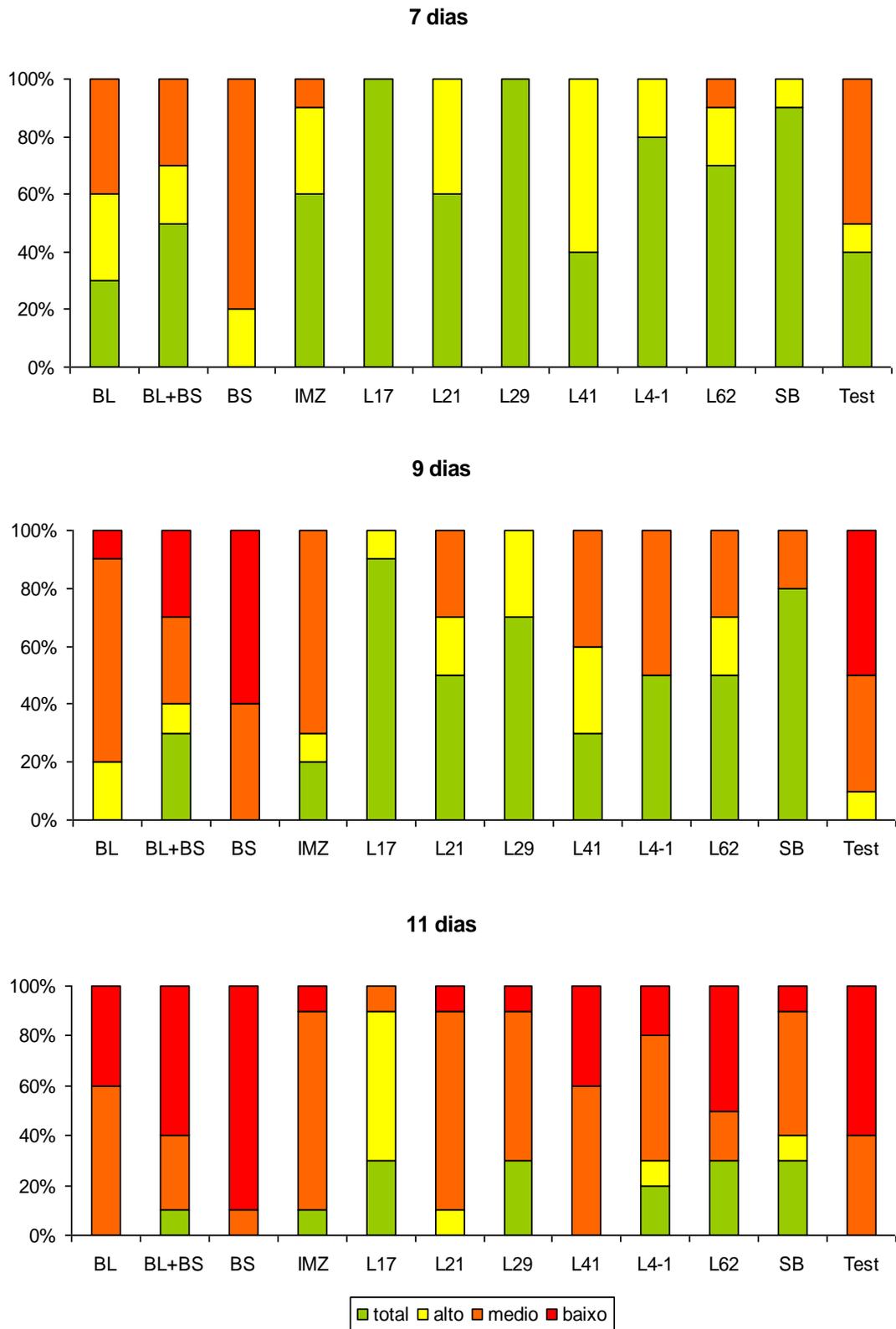


Figura 3. Porcentagem de frutos não infectados em cada classe de controle (total, alto, médio e baixo) da podridão de Fusarium, para os diversos bioagentes avaliados, fungicida tradicional e controle absoluto, aos 7, 9 e 11 dias após a inoculação (Experimento II).

Tabela 1. Comparação da porcentagem de frutos infectados para os diversos bioagentes em contraste com o fungicida tradicional e o controle absoluto, aos 10 dias após a inoculação, no Experimento I.

Tratamento	% de frutos infectados	Erro-padrão	P ₁ *	P ₂ *
<i>Bacillus subtilis</i> , isolado 2 – B-AP-031	50	16	0,1384	0,0301
<i>Bacillus licheniformis</i> (BL)	80	13	1,0000	<0,0001
<i>Bacillus subtilis</i> , isolado 1 - BS	60	15	0,3173	0,0059
<i>B. subtilis</i> + <i>B. licheniformis</i> (BS+BL)	50	16	0,1384	0,0301
<i>Rhodotorula mucilagenosa</i> (L17)	50	16	0,1384	0,0301
<i>Pichia membranifaciens</i> (L21)	20	13	0,0008	0,5271
<i>Pichia guilliermondii</i> (L29)	40	15	0,0455	0,0986
<i>Pichia sp</i> (L4-1)	40	15	0,0455	0,0986
<i>Sporobolomyces roseus</i> (L41)	60	15	0,3173	0,0059
<i>Debaryomyces hansenii</i> (L62)	20	13	0,0008	0,5271
<i>Sporodibolus pararoseus</i> (SB)	80	13	1,0000	<0,0001
Fungicida convencional, Imazalil	10	09	<0,0001	-
Controle absoluto, água (TEST)	80	13	-	<0,0001

*** Nível de significância nominal do teste Qui-quadrado de Wald para contraste dos bioagentes com a testemunha (*) e com o fungicida tradicional (**).

A análise de contrastes confirma a maior eficiência dos bioagentes *Pichia membranifaciens* e *Debaryomyces hansenii*, resultado consistente com o obtido na análise de agrupamentos. Três dos bioagentes avaliados não diferiram do controle absoluto (água); os demais apresentaram algum nível de eficiência na redução do tamanho das lesões. Apenas o agente *R. mucilagenosa* (L17) causou redução no tamanho das lesões; a estimativa do diâmetro médio de lesão foi cerca de metade ($p=0,0053$) da obtida para o fungicida tradicional, o que indica um alto potencial de controle da podridão de Fusarium. A análise visual do gráfico de barras empilhadas confirma esse potencial: mesmo aos 11 dias após a inoculação, o controle da doença é alto ou total para 90% dos frutos. Esse bioagente não havia se destacado no bioensaio anterior onde foi mensurada apenas a porcentagem de frutos infectados sem considerar o tamanho das lesões (**Tabela 1**). Os demais bioagentes não diferiram do fungicida tradicional ($p > 0,11$), exceto *B. subtilis* que apresentou menor eficiência no controle da doença, com lesões superiores a 24mm de diâmetro ($p=0,0028$).

Apesar da disponibilidade de métodos mais adequados para a análise de variáveis não Normais, seu uso é ainda restrito na análise de dados biológicos onde prevalecem abordagens mais tradicionais que utilizam transformações de dados. A divulgação de novas abordagens é importante para o uso mais eficiente das informações geradas em ensaios biológicos.

Novos ensaios serão realizados utilizando um grupo de frutos como unidade experimental. Dados dessa natureza serão modelados via MLGM, considerando a distribuição Binomial, com função de ligação logito e o fruto considerado efeito aleatório. Quando é utilizado apenas um fruto como unidade experimental, não é possível investigar a presença de sub ou sobredispersão.

Neste trabalho, as variáveis porcentagem de frutos infectados e diâmetro médio de lesão foram obtidas em diferentes ensaios. A sua mensuração simultânea permite uma melhor exploração da natureza multivariada das avaliações para uso na seleção de bioagentes promissores.

Optamos por apresentar análises inferenciais para apenas uma data de avaliação, com o objetivo de introduzir o uso de modelos mais adequados para modelagem de variáveis não Normais (porcentagem de frutos infectados) e a introdução de efeitos aleatórios (para diâmetro de lesão). Em trabalhos subsequentes, será considerada a natureza longitudinal dos dados, de fundamental importância nesse tipo de problema, onde a evolução temporal da doença é informação crítica para o planejamento de processos de transporte e armazenamento de frutos.

REFERÊNCIAS

BRESLOW, N.E.; CLAYTON, D.G. Approximate inference in generalized linear mixed models". **JASA**, n. 88, v.421, p. 9–25, 1983.

McCULLAGH, P.; NELDER, J. A. **Generalized linear models**, 2nd. ed. London: Chapman & Hall, 511 p., 1989.

SAS INSTITUTE INC. **SAS/STAT user's guide/** Version 9.2. Cary, NC: SAS Institute Inc., 2008.

STOKES, M.E.; DAVIS, C.S.; KOCH, G.G. **Categorical data analysis using the SAS system**. 2.ed. Cary, NC: SAS Institute Inc, 2000. 629p.

TUFTE, E. F. **The Visual Display of Quantitative Information**, 2nd. ed. Cheshire, Graphics Press, Cheshire, 2001.