

O ESTUDO DE MEDIDAS REPETIDAS: UMA APLICACAO A DADOS DE FORRAGEIRAS

Alfredo Ribeiro de Freitas¹, Waldomiro Barioni Júnior², Reinaldo de Paula Ferreira²,
Camila Regina Destefani³, Aline Rodrigues dos Santos³, Adônis Moreira¹

Introdução

Na pesquisa quando se avalia apenas uma variável, geralmente três situações de análises estatística são consideradas: a) testes para efeitos fixos em regressão simples e múltipla e análise de variância (ANOVA) - quando a característica é avaliada apenas uma vez na unidade experimental; b) análises de medidas repetidas quando duas ou mais avaliações, geralmente até um máximo entre 10 a 12 no tempo, são feitas na unidade experimental; c) séries temporais - quando um conjunto de observações são feitas seqüencialmente ao longo do tempo (geralmente mais de 12).

As análises dos itens a) e b) são geralmente utilizadas quando os dados são provenientes de delineamentos experimentais e, uma situação que requer bastante cuidado é a dos blocos casualizados com parcelas divididas (*split-plot*), um dos mais utilizados na agricultura; o *split-plot* envolve os fatores experimentais A e B, sendo que os níveis destes são aleatoriamente atribuídos às parcelas principais e subparcelas, respectivamente. Geralmente, doses de adubação e variedades, níveis de irrigação e variedades, variedades e cortes das plantas no tempo, profundidades de aração e doses de adubação, entre outros, são utilizados como fatores A e B. Um exemplo comum é o experimento com forrageiras em que se avalia a produção em cortes das plantas no tempo; os acessos ou variedades, doses de adubo, etc, correspondem aos tratamentos A, que são alocados nas parcelas e os cortes de avaliação aos tratamentos B, que são alocados nas subparcelas. Experimentos dessa natureza tem sido erroneamente analisados como *split-plot*. Como o tempo de corte não pode ser alocado aleatoriamente às subparcelas, a análise apropriada para essa situação é a de medida repetidas-MR (Littell et al., 1996; Littell et al., 1998, Reiezigel, 1999; Freitas et al., 2001). Na análise

¹ Pesquisador da Embrapa Pecuária Sudeste, Rodovia Washington Luís, km 234, Caixa Postal 339, CEP 13560-970 - São Carlos-SP. Endereço eletrônico: ribeiro@cnpse.embrapa.br, adonis@cnpse.embrapa.br, Bolsista do CNPq.

² Pesquisador da Embrapa Pecuária Sudeste - barioni@cnpse.embrapa.br, reinaldo@cnpse.embrapa.br

³ Bacheranda em Estatística da Universidade Federal de São Carlos, Rodovia Washington Luís, km 235, São Carlos - São Paulo - Brasil, CEP 13565-905, bolsista de IC da FAPESP. Endereço eletrônico: alinear@gmail.com

proci-2007.00117
FRE
2007

SP-2007.00117

Estudo de medidas repetidas:
2007 SP-2007.00117



17117-1

salta a seguir do CD

de MR as unidades experimentais ou indivíduos são aleatoriamente alocados aos tratamentos (fator A) e são avaliados no tempo, podendo ser também no espaço, como por exemplo, profundidade. A ordem das observações realizadas no indivíduo ou na unidade experimental é fundamental, sendo as mais próximas mais correlacionadas.

Dentre alguns exemplos de MR na agricultura, pode-se citar o crescimento corporal dos animais a produção de leite, a contagem de ecto e endoparasitos, a produção em cortes em forrageiras, o tempo de armazenamento e a qualidade de produtos.

Quanto às séries temporais (item c), correspondem ao estudo de observações que se apresentam de forma ordenadas no tempo, geralmente obtidas de processos observacionais, e resultam da combinação de vários efeitos: tendência, componente sazonal e componente aleatório. Como comentado no item b) sobre medidas repetidas, a ordem das observações é fundamental e as mais próximas são mais estreitamente correlacionadas. O propósito principal no estudo de séries temporais é a construção de modelos matemáticos que possibilitem visão simplificada da realidade ou fenômeno contido nos dados. Várias são as finalidades: predição (fazer previsões de valores futuros com base em valores passados), controle de processos (controle estatístico de qualidade), descrição do fenômeno (estudo do comportamento da série através de gráficos). Dentre alguns exemplos da aplicação das séries temporais na agricultura, tem-se os dados climáticos (precipitação pluviométrica, temperatura diária, velocidade do vento), crescimento de colônia de bactérias, de insetos, sazonalidade de população de insetos, predição de cenários do ambiente, entre outros (EHLERS, 2005).

No presente estudo foi analisada a produção de matéria seca, kg/ha de um experimento de melhoramento de alfafa realizado em blocos casualizados; nas parcelas foram distribuídas aleatoriamente cinco cultivares de alfafa e nas subparcelas foram considerados 20 cortes consecutivos (meses). Foi comparada a análise considerando-se o efeito de cortes como subparcela e também como medidas repetidas.

MATERIAL E MÉTODOS

Foi analisada a produção de matéria seca, kg/ha, obtida de um experimento de melhoramento de alfafa realizado na Embrapa Pecuária Sudeste, São Carlos, SP, com início em junho de 2004. O delineamento experimental foi o de blocos casualizados, duas repetições, e parcelas (constituídas de cinco fileiras de cinco metros de comprimento e a bordadura uma fileira de cada lado e 0,50m de cada extremidade da parcela) divididas; os tratamentos aplicados aleatoriamente nas parcelas corresponderam

aos cultivares: Barbara, Crioula, P30, P5715 e LEN 4; nas subparcelas foram considerados 20 cortes consecutivos, realizados quando aproximadamente 10% das plantas estavam em florescimento. A análise considerando-se o efeito de cortes como subparcela e como medidas repetidas (LITTELL et al., 1996, 1998), foram realizadas, respectivamente, por meio dos procedimentos GLM (método 1) e MIXED (método 2) do SAS (SAS Institute, 2002-2003). Foi utilizado o modelo: $y_{ijk} = \mu + \alpha_i + \delta_{ij} + t_k + (\delta t)_{ik} + \varepsilon_{ijk}$, em que y_{ijk} é a resposta no corte k na unidade experimental j no tratamento aplicado às parcelas i; μ é a média global; α_i é o efeito fixo do tratamento i aplicado às parcelas; δ_{ij} é o efeito aleatório da unidade experimental j no grupo de tratamento i, t_k é o efeito fixo do corte k, $(\delta t)_{ik}$ é o efeito de interação e ε_{ijk} é o erro aleatório no corte k, na unidade experimental j e no grupo de tratamento i.

Para a análise considerando-se o efeito de cortes como subparcela, δ_{ij} e ε_{ijk} são assumidos independentes e $\delta_{ij} \sim N(0, \sigma^2_d)$ e $\varepsilon_{ijk} \sim N(0, \sigma^2)$. Para a análise de MR δ_{ij} são assumidos independentes, enquanto que os ε_{ijk} na mesma unidade experimental são correlacionados e refletem a variação dentro da unidade experimental, dada por $V(\varepsilon_{ijk}) = R$. Foram avaliadas as seguintes estruturas de R: Auto Regressiva de Primeira Ordem: AR(1), Auto Regressiva de Média Móvel: ARMA (1,1), Componentes de variância: VC e Simetria Composta: CS e selecionada aquela com o menor valor de AIC ((Wolfinger, 1993).

RESULTADOS E DISCUSSÃO

Os valores do critério AIC das estruturas de covariâncias avaliadas: AR(1), ARMA(1,1), VC e CS, foram, respectivamente, 1434,8; 1417,7; 1464,6 e 1421,9. Dessa forma, a mais adequada foi a ARMA(1,1). Quanto aos valores de $Pr > F$ para os efeitos de tratamentos (0,8855), cortes ($< 0,0001$) e interação tratamentos x cortes (0,9901), houve significância apenas para o efeito de cortes. Como os dados eram balanceados os dois métodos foram concordantes quanto às estimativas de efeitos fixos. As médias de tratamentos, obtidas por quadrados mínimos e respectivos erros-padrão, para as variedades: Barbara, Crioula, P30, P5715 e LEN 4, nessa ordem, foram $1653,1 \pm 34,8$; $1519,5 \pm 34,8$; $1526,2 \pm 36,5$; $1599,7 \pm 34,8$ e $1434,2 \pm 34,8$ para o método 1 e, $1653,1 \pm 158,2$; $1519,5 \pm 158,2$; $1526,2 \pm 158,5$; $1599,7 \pm 158,5$ e $1434,2 \pm 34,8$ para o método 2. Observa-se que os erros-padrão obtidos por máxima verossimilhança (método 2), são maiores do que os obtidos por quadrados mínimos (método 1). Esta superioridade pode ser também observada nos intervalos de confiança, com 95% de

probabilidade, para as médias de tratamentos e de cortes (Figura 1), principalmente para o efeito de tratamentos. Esta divergência observada entre os métodos quanto aos os erros-padrão e demais estatísticas derivadas destes, tais como, testes de hipóteses e intervalos de confiança, é explicada pela diferença no cálculo do erro-padrão. No método 1 e 2, o erro-padrão é calculado, respectivamente, pela raiz quadrada de $\sigma^2 L(X'X)L'$ e $L(X'V^{-1}X)L'$, em que L é a matriz de hipótese, σ^2 é o quadrado médio residual, L é a matriz de hipótese, V a matriz de variâncias e covariâncias.

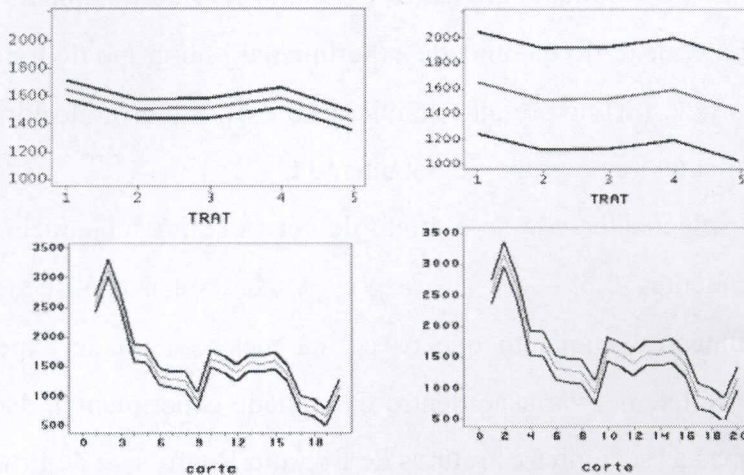


Figura 1- Intervalos de confiança com 95 % de probabilidade da estimativa de médias de produção de matéria seca: PMS, kg/ha, obtidos por meio do GLM (gráficos à esquerda) e MIXED do SAS (gráficos à direita): as três linhas em ordem decrescente de valor de PMS, indicam, o limite superior, média e o limite inferior.

A Figura 2 apresenta a tendência da correlação considerando uma combinação da estrutura da matriz ARMA(1,1) em função de: efeitos aleatórios σ_d^2 (blocos/tratamentos), σ^2 (residual), lag (comprimento do intervalo de tempo entre medidas), e dos parâmetros de ARMA(1,1): γ e ρ . A tendência é dada por: $[\sigma_d^2 + \sigma^2(\gamma\rho^{\text{lag}})]/(\sigma_d^2 + \sigma^2)$.

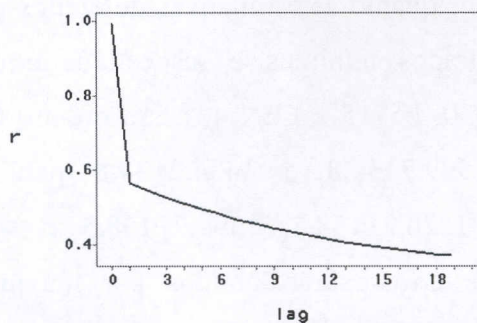


Figura 2- Tendência da correlação obtida da matriz Auto Regressiva de Média Móvel em função do comprimento do intervalo de tempo entre medidas (lag)

CONCLUSÕES

A análise de medidas repetidas efetuada por meio do procedimento MIXED do SAS em que $V(\varepsilon_{ijk}) = R$, foi mais eficiente do que a executada por meio do procedimento GLM, $V(\varepsilon_{ijk}) = I\sigma_e^2$ na análise de dados de produção de matéria seca de alfafa considerando-se o efeito de cortes como subparcela (método 1) e como medidas repetidas (método 2).

A estrutura de covariância R selecionada para modelar a correlação existente entre os dados de PMS nos 20 cortes foi a Auto Regressiva de Média Móvel.

Os resultados dos métodos foram diferentes; os erros-padrão obtidos por meio do MIXED foram maiores do que os obtidos por quadrados mínimos, superioridade também observada nos intervalos de confiança, com 95% de probabilidade para as estimativas de médias dos efeitos fixos.

REFERÊNCIAS BIBLIOGRÁFICAS

- FREITAS, A. R. de; PRIMAVESI, O.; CORREA, L. A.; PRIMAVESI, O.; POTT, E. B.; MASCIOLLI, A. S. Repeated measurement analyses of forages in cropping systems. In: INTERNATIONAL GRASSLAND CONGRESS, 19., 2001, Piracicaba. **Proceedings...** Piracicaba: FEALQ, 2001. p.1046-1047.
- LITTELL, R. C.; HENRY, P. R.; AMMERMAN, C. B. Statistical analysis of repeated measures data using SAS procedures. **Journal of Animal Science**, v.76, p.1216-1231, 1998.
- LITTELL, R. C.; MILLIKEN, G. A.; STROUP, W. W.; WOLFINGER, R. D. **SAS System for Mixed Models**. Cary: Statistical Analysis System Institute, 1996. 633p.
- REIEZIGEL, J. Analysis of experimental data with repeated measurement. **Biometrics**, v.55, p.1059-1063, 1999.
- SAS Institute. User's Guide. versão 9.1.3, versão para Windows. Cary, NC, USA, 2002-2003.
- WOLFINGER, R. Covariance structure selection in general mixed models. **Commun. Statistic. Simula.**, v. 2, n. 4, p.1079-1106, 1993.