

REGRESSÃO ALEATÓRIA: NOVA TECNOLOGIA PODE MELHORAR A QUALIDADE DAS AVALIAÇÕES GENÉTICAS

LUCIA GALVÃO DE ALBUQUERQUE^{1,2}

¹ Pesquisadora do CNPq

² Professora Adjunta - Departamento de Zootecnia, FCAV/UNESP, Jaboticabal/SP

INTRODUÇÃO

Em gado de corte, normalmente, para características de crescimento, os valores genéticos são preditos e os componentes de variâncias são estimados, considerando-se os pesos às idades padrão – como, por exemplo: nascimento, desmama, ano e final - ou ganhos em pesos entre estas idades. Da mesma forma, em gado de leite, faz-se avaliação genética para a produção de leite ajustada para 305 dias (ou para a média de duração da lactação do rebanho), ou para a produção no dia do controle leiteiro. Em qualquer destes casos, são formadas classes arbitrárias e as produções (leite ou peso) são ajustadas a uma determinada idade ou dias em lactação e os componentes de variância e os valores genéticos dos animais são obtidos por meio de análises uni ou multi-características. Com estas últimas, as correlações entre as produções nas diferentes idades ou dias, são consideradas diferentes de zero, entretanto, não se assume qualquer estrutura de covariâncias entre as medidas.

Mais recentemente, os modelos de regressão aleatória (MRA) têm sido propostos como uma alternativa para modelar características, como produção de leite ou crescimento, que são medidas repetidamente na vida dos animais, denominadas características repetidas ou dados longitudinais (Schaeffer e Dekkers, 1994). Em contraste com os modelos multi-características, que fornecem predições para determinados pontos ou idades, os modelos de regressão aleatória permitem a predição de valores genéticos para a curva de crescimento ou lactação como um todo, para qualquer ponto desejado na escala de tempo utilizada e para funções da curva. Com estes modelos não há necessidade de se criar classes de desempenho arbitrárias ou de utilizar fatores de ajuste para uma determinada idade (dias em lactação) uma vez que esta está incluída na matriz de delineamento. Além disto, permitem uma melhor utilização dos dados, já que todas as medidas do animal e de seus parentes são utilizadas para avaliação do mesmo, com potencial aumento da acurácia de seleção.

Com modelos de regressão aleatória, ajusta-se uma curva de regressão fixa para levar em conta a tendência média da população e, no mínimo, mais duas equações de regressão aleatória, para o efeito genético direto e para o ambiente permanente de animal, uma vez que existem medidas repetidas. A curva genética de cada animal é predita como um desvio da curva média, fixa. Desta forma, em lugar de valores genéticos para determinados pontos da curva (idades padrão), são obtidos valores genéticos para os coeficientes de regressão genéticos. A partir dos coeficientes de regressão aleatórios, pode-se descrever a curva de crescimento (lactação) genética do animal; encontrar DEP's para qualquer idade desejada, mesmo para idades em que o animal não tenha sido medido (desde que esta esteja dentro da amplitude de idades existente nos dados) e examinar funções da curva de crescimento ou lactação (ex. taxa de crescimento entre duas idades ou persistência da lactação). Os modelos de regressão aleatória têm sido utilizados para modelar uma grande variedade de características em diversas espécies, tais como suínos, bovinos de leite e de corte, com destaque, principalmente, para a modelagem de dados de produção de leite no dia do controle em bovinos de leite (Jamrozik et al., 1997; Van der Werf et al., 1998).

O objetivo do presente capítulo é apresentar uma revisão sucinta sobre modelos de regressão aleatória e sua aplicação para estimação de componentes de variância e avaliação genética dos animais.

MEDIDAS REPETIDAS

Medidas que são feitas seqüencialmente no mesmo animal, ao longo do tempo, as chamadas “medidas repetidas”, estão recebendo, nos últimos anos, uma maior atenção dos pesquisadores em melhoramento animal. Segundo Van der Werf e Schaeffer (1997) este tipo de característica merece um tratamento estatístico especial, uma vez que, o padrão de covariâncias entre medidas repetidas é bem estruturado e, para que se possam fazer inferências sobre este tipo de dados, é importante modelar esta estrutura de covariâncias. Normalmente, para características como produção de leite ou crescimento, a correlação entre as medidas tomadas ao longo do tempo diminui quando o espaço de tempo entre elas aumenta.

Características medidas diversas vezes durante a vida do animal têm sido estudadas de várias formas. A maneira mais simples para analisá-las é assumir que todas as medidas representam a mesma característica e usar um modelo de repetibilidade. Entretanto, neste modelo, pressupõe-se que a correlação entre as medidas repetidas seja igual à unidade, portanto, todas as (co) variâncias genéticas e fenotípicas entre as diferentes medidas são de mesma magnitude. Isto é, os mesmos genes controlam o desempenho ao longo do tempo.

A produção de leite nos diversos controles é um exemplo típico de medida repetida e o mais trabalhado no contexto do melhoramento animal. Em gado de leite, tem-se usado a produção no dia do controle para estimar a produção de leite em 305 dias, ignorando desta forma os efeitos de meio específicos a cada controle, bem como a forma da curva de lactação (Ptak e Schaeffer, 1993, Jamrozik e Schaeffer, 1997).

Uma alternativa tem sido os modelos multi-características, denominados “test-day models”, com covariâncias não estruturadas (Meyer et al., 1989; Pander e Hill, 1993, Gadini, 1997), assumindo que a produção em cada controle é uma característica diferente. Nestes modelos as covariâncias entre as produções nos diversos controles variam e as correlações podem ser menores que a unidade, mas não é feita qualquer pressuposição sobre a estrutura de covariâncias. Em análises genéticas, o número de parâmetros a ser estimado é equivalente a $[t(t+1)n]/2$, onde t é o número de características ou medidas repetidas por indivíduo e n é o número de componentes de variância estimados por característica (Meyer e Hill, 1997; Olori, 1997). Isto torna as análises computacionalmente lentas e exige muita memória. Segundo Thompson e Hill (1990) e Meyer e Hill (1997), este tipo de modelo de análise não estruturado só será viável computacionalmente e produzirá estimativas acuradas quando o número de medidas repetidas for pequeno e o de indivíduos medidos for grande. Ptak e Schaeffer (1993) apresentaram um modelo alternativo, de forma a considerar as produções nos diversos controles como a mesma característica, usando um modelo de repetibilidade, mas empregando uma função de regressão fixa para ajustar para a forma da curva. Segundo Meyer e Hill (1997), neste caso, é necessário que se conheça a forma da curva com antecedência, o que nem sempre é possível.

Os modelos de repetibilidade são usados, freqüentemente, por serem mais simples, já que requerem menor esforço computacional e menos parâmetros que os modelos multi-características. Embora este último, em geral, pareça mais correto por permitir correlações diferentes entre as medidas, apresenta a desvantagem de ser super parametrizado quando se têm medidas em várias idades. Além disto, neste modelo, a matriz de covariâncias não é estruturada, podendo não ser a mais desejável quando as medidas são tomadas ao longo de uma trajetória.

Uma outra possibilidade para trabalhar com características que se repetem é descrever a trajetória por meio de um modelo matemático (regressão sobre o tempo), estimar os parâmetros e funções da curva e selecionar para os mesmos. Diversos trabalhos de pesquisa têm sido realizados para estimar fatores genéticos e de meio que afetam os parâmetros de curvas de crescimento e lactação (e.g: Oliveira, 1995; Gonçalves, 1994), e os mesmos podem ser utilizados como critérios de seleção. Entretanto, esta é uma alternativa que raramente tem sido implementada na prática por grandes programas de melhoramento de gado de corte ou leite.

Outras opções que têm merecido a atenção dos melhoristas nos últimos anos são: as funções de covariância e os modelos de regressão aleatória (RA). Segundo Kirkpatrick et al. (1990 e 1994), características que se alteram com a idade podem ser representadas por uma trajetória, em função do tempo. Elas apresentam um determinado valor em cada uma das infinitas (possíveis) idades e cada valor pode ser considerado como uma característica diferente. Assim, tais trajetórias são denominadas de características de “dimensão infinita”.

FUNÇÕES DE COVARIÂNCIA E MODELOS DE REGRESSÃO ALEATÓRIA

Uma função de covariância (FC) é o equivalente de “dimensão infinita” da matriz de covariâncias em uma análise multivariada de “dimensão finita”. Ou, como definido por Van der Werf e Schaeffer (1997), é uma função contínua que fornece as (co) variâncias de características medidas em diferentes pontos de uma trajetória, descrevendo a covariância entre medidas tomadas em determinadas idades como uma função destas idades.

Kirkpatrick e Heckman (1989) apresentam três vantagens dos modelos de FC sobre os modelos tradicionais, multi-características: 1- as FC produzem uma descrição para todo ponto ao longo de uma escala contínua de medidas (tempo), permitindo que as covariâncias entre idades não mensuradas sejam obtidas facilmente por interpolação. Não há necessidade de se fazer qualquer pressuposição sobre a forma da curva (lactação, crescimento); 2- cada FC tem um conjunto de

autovalores e autofunções (análogos aos auto-vetores) que fornecem informações sobre a direção na qual a curva média (crescimento, lactação, etc.) tem maior chance de ser modificada por seleção, pelo fato de apresentarem maior variância genética. Estes autores mostraram que as FC permitem a estimação de uma função de seleção gradiente contínua, isto é, levando em conta os efeitos de seleção em todos os pontos. Assim, as FC permitem uma predição da resposta à seleção mais acurada; 3- modelos de FC tornam mais eficientes a utilização dos dados e estimação de parâmetros de interesse. Ajustando-se polinômios da menor ordem possível (necessária), evita-se a superparametrização, minimizando a variância do erro e reduzindo problemas de correlações amostrais negativas (Meyer e Hill, 1997). Kirkpatrick et al. (1994) acrescentam que, com este método, pode-se levar em conta a data correta em que os dados são coletados, mesmo quando os dados são tomados de forma irregular.

Van der Werf e Schaeffer (1997) resumiram as vantagens das FC relatando que as mesmas permitem: modelar de forma acurada a estrutura de variâncias - covariâncias das características; prever as estruturas de covariâncias em qualquer ponto de uma escala contínua de tempo; dar maior flexibilidade na utilização de medidas tomadas em qualquer momento ao longo da trajetória, sem a necessidade de corrigi-las para idades padrão; analisar padrões de covariância associados a determinadas alterações da característica ao longo da trajetória.

Uma função de covariâncias, que dá a covariância entre as idades m e l , pode ser representada como:

$$f(a_m^*, a_l^*) = \sum_{i=0}^{k-1} \sum_{j=0}^{k-1} \phi_i(a_m^*) \phi_j(a_l^*) K_{ij}$$

onde:

ϕ_i é o $i^{\text{ésimo}}$ polinômio; k é a ordem do polinômio que está sendo ajustado; a^* são as idades padronizadas para a amplitude em que o polinômio é definido e K_{ij} são os elementos da matriz \mathbf{K} , que é a matriz de coeficientes da função de covariâncias.

Em notação matricial, assumindo que $k = t$ ($t =$ número de idades), a matriz de covariâncias observada (Σ) pode ser escrita como:

$$\Sigma = \Phi \mathbf{K} \Phi' \text{ e } \mathbf{K} = \Phi^{-1} \Sigma (\Phi^{-1})'$$

onde Φ é uma matriz $t \times k$ de polinômios ortogonais (Meyer e Hill, 1997).

Kirkpatrick et al. (1990 e 1994) modelaram as FC usando polinômios ortogonais, chamados polinômios de Legendre (ϕ_k), que apresentam a seguinte fórmula:

$$\phi_k(a^*) = \frac{1}{2} \sqrt{\frac{2k+1}{2}} \sum_{m=0}^{[k/2]} (-1)^m \binom{k}{m} \binom{2k-2m}{k} (a^*)^{k-2m},$$

onde: a^* são as idades padronizadas para a amplitude em que o polinômio é definido $[-1 \text{ a } +1]$ e $[.]$ indica que os valores da fração são arredondados para baixo, assumindo o valor do inteiro mais próximo. As idades são padronizadas utilizando-se a fórmula:

$$a_i^* = \frac{2(a_i - a_{\min})}{a_{\max} - a_{\min}} - 1$$

sendo a_{\max} e a_{\min} as idades máxima e mínima, respectivamente, existentes nos dados (Abramowitz e Stegun, 1965). O exemplo utilizado por Kirkpatrick et al. (1990) considerou o peso de camundongos em três idades, 2, 3 e 4 semanas e uma função de ordem 3 (mesmo número de idades), o que daria:

$$\phi = \begin{bmatrix} \phi_0(-1) & \phi_1(-1) & \phi_2(-1) \\ \phi_0(0) & \phi_1(0) & \phi_2(0) \\ \phi_0(1) & \phi_1(1) & \phi_2(1) \end{bmatrix} = \begin{bmatrix} 0,707 & -1,225 & 1,581 \\ 0,707 & 0 & -0,791 \\ 0,707 & 1,225 & 1,581 \end{bmatrix}$$

Neste exemplo:

$$\Sigma = \begin{bmatrix} 436,0 & 522,3 & 424,2 \\ 522,3 & 808,0 & 664,7 \\ 424,2 & 664,7 & 558,0 \end{bmatrix} \text{ e } \mathbf{K} = \begin{bmatrix} 1348,0 & 66,5 & -112,0 \\ 66,5 & 24,3 & -14,0 \\ -112,0 & -14,0 & 14,5 \end{bmatrix}$$

A função de covariâncias neste caso é:

$$f(a_i^*, a_j^*) = 808,0 + 71,2(a_i^* + a_j^*) + 36,4 a_i^* a_j^* - 40,79((a_i^*)^2 a_j^* + a_i^* (a_j^*)^2) - 215,0((a_i^*)^2 + (a_j^*)^2) + 81,6(a_i^*)^2(a_j^*)^2.$$

No exemplo acima, a ordem do polinômio é igual ao número de idades existente nos dados, entretanto, o que se busca na prática é o ajuste de polinômios de menor ordem ($k < t$). Para a situação em que a ordem da função (k) é diferente do número de idades (t), Kirkpatrick et al. (1994) descreveram um procedimento de quadrados mínimos generalizados para determinar os coeficientes da FC a partir de uma matriz de covariâncias pré-estimada.

Meyer e Hill (1997) mostraram que os coeficientes da FC podem ser estimados diretamente dos dados, usando-se máxima verossimilhança restrita (REML), por meio de uma re-parametrização do algoritmo de REML, em modelos multi-variados (dimensão finita), já em uso para estimar componentes de variância (Meyer, 1991). A principal vantagem de usar REML, em lugar de quadrados mínimos generalizados, é que com REML, a matriz de coeficientes vai ser positiva definida (Van der Werf e Schaeffer, 1997).

O procedimento desenvolvido por Meyer e Hill (1997), pode ser resumido da forma como segue abaixo. Assumindo-se um modelo multi-característica, com animais tendo medidas em todas as "t" características, representado na forma mais simples: $\mathbf{y} = \mathbf{Xb} + \mathbf{Za} + \mathbf{e}$, onde: \mathbf{y} = vetor de observações; \mathbf{b} = vetor de efeitos fixos; \mathbf{a} = vetor de efeitos aleatórios aditivos de animal; \mathbf{e} = vetor de efeito aleatório residual; \mathbf{X} e \mathbf{Z} são matrizes de incidência para os efeitos fixos e aleatórios, respectivamente, e $\text{var}(\mathbf{y}) = \mathbf{ZGZ}' + \mathbf{R}$, $\text{var}(\mathbf{a}) = \mathbf{G}$, com $\mathbf{G} = \mathbf{A} \otimes \mathbf{G}_0$. Se as matrizes forem iguais para todas as características, $\text{var}(\mathbf{e}) = \mathbf{R} = \mathbf{I} \otimes \mathbf{R}_0$. Sendo, \mathbf{A} o numerador da matriz de parentesco entre os animais; \mathbf{G}_0 a matriz de variâncias e covariâncias entre as t características e \otimes o produto Kronecker. Assim, o logaritmo da função de máxima verossimilhança restrita é

$$\ln \varepsilon = -1/2 [\text{const} + N \ln |\mathbf{R}_0| + N_a \ln |\mathbf{G}_0| + t \ln |\mathbf{A}| + \ln |\mathbf{C}| + \mathbf{y}'\mathbf{P}\mathbf{y}]$$

onde: N_a é o número de animais nas análises, \mathbf{C} é a matriz de coeficientes para as equações de modelos mistos e $\mathbf{y}'\mathbf{P}\mathbf{y}$ é a soma de quadrados dos resíduos. Meyer e Hill (1997) re-escreveram a equação acima como uma função das matrizes de (co) variâncias entre os coeficientes da FC, \mathbf{K}_a e \mathbf{K}_e , para as FC genética aditiva e residual, respectivamente, em lugar de \mathbf{G}_0 e \mathbf{R}_0 . Desta forma,

$$\ln \varepsilon = -1/2 [\text{const} + N \ln |\mathbf{K}_e| + N_a \ln |\mathbf{K}_a| + (N + N_a) \ln |\Phi\Phi'| + t \ln |\mathbf{A}| + \ln |\mathbf{C}| + \mathbf{y}'\mathbf{P}\mathbf{y}]$$

onde: Φ é a matriz t x k com os polinômios ortogonais.

Neste contexto, é possível separar os efeitos de ambiente em permanentes e temporários, com $\mathbf{e} = \mathbf{r} + \boldsymbol{\varepsilon}$, e $\boldsymbol{\Sigma}_e = \boldsymbol{\Sigma}_r + \boldsymbol{\Sigma}_\varepsilon$, onde \mathbf{r} é o vetor de efeitos de ambiente permanente; $\boldsymbol{\varepsilon}$ o vetor de efeitos de ambiente temporário; $\boldsymbol{\Sigma}_r = \{\sigma_{r_{ij}}\}$ a matriz de (co)variâncias de ambiente permanente e $\boldsymbol{\Sigma}_\varepsilon = \mathbf{Diag}\{\sigma_{\varepsilon_i}^2\}$ a matriz de variâncias de ambiente temporário. Assim, a função de máxima verossimilhança restrita pode ser reescrita como:

$$\ln \varepsilon = -1/2 [\text{const} + N \ln |\boldsymbol{\Phi}\mathbf{K}\boldsymbol{\Phi}' + \mathbf{Diag}\{\sigma_{\varepsilon_i}^2\}| + N_a \ln |\mathbf{K}_a| + N_a \ln |\boldsymbol{\Phi}\boldsymbol{\Phi}'| + t \ln |\mathbf{A}| + \ln |\mathbf{C}| + \mathbf{y}'\mathbf{P}\mathbf{y}],$$

onde: \mathbf{K}_r é a matriz de coeficientes da função de covariâncias de ambiente permanente.

Este procedimento exige um menor esforço computacional que os modelos multi-variados com t características, uma vez que o número de parâmetros genéticos a serem estimados passa de $t(t+1)/2$ para $k_a(k_a+1)/2$, onde k_a é a ordem do polinômio empregado para o efeito genético. Segundo Van der Werf e Schaeffer (1997), o tempo para calcular cada função de verossimilhança não é reduzido, uma vez que o tamanho das equações de modelo misto é o mesmo daquele de uma análise com t características. Os autores ressaltam que o procedimento descrito por Meyer e Hill (1997) não é muito prático quando os dados são medidos de forma irregular, uma vez que o mesmo requer uma matriz de equações de modelo misto proporcional ao número de idades e não à ordem da função de covariâncias.

Meyer (1998a) mostrou que um modelo de FC é equivalente a um modelo de regressões aleatórias com as funções da idade (tempo) como covariáveis, e que as estimativas de REML dos coeficientes das funções de covariância podem ser obtidos como as covariâncias entre os coeficientes de regressão aleatórios. Um modelo de regressão aleatória, usando polinômios de Legendre da idade (tempo) como variável independente, pode ser escrito como:

$$y_{ij} = F + \sum_{m=0}^{k_a-1} \alpha_{im} \phi_m(a_{ij}^*) + \sum_{m=0}^{k_q-1} \delta_{im} \phi_m(a_{ij}^*) + \varepsilon_{ij},$$

onde y_{ij} é a medida do animal i no tempo j , F é um conjunto de efeitos fixos, α_{im} e δ_{im} são os $m^{\text{ésimos}}$ coeficientes de regressão aleatórios genético aditivo e de ambiente permanente para o animal i , respectivamente, e k_a e k_q são as ordens de ajuste dos polinômios correspondentes.

Neste contexto, a covariância entre duas medidas de um mesmo animal é dada por:

$$\begin{aligned} \text{Cov}(y_{ij}, y_{ij'}) &= \sum_{m=0}^{k_a-1} \sum_{l=0}^{k_a-1} \phi_m(a_{ij}^*) \phi_l(a_{ij'}^*) \text{Cov}(\alpha_{im}, \alpha_{il}) \\ &+ \sum_{m=0}^{k_q-1} \sum_{l=0}^{k_q-1} \phi_m(a_{ij}^*) \phi_l(a_{ij'}^*) \text{Cov}(\delta_{im}, \delta_{il}) + \text{Cov}(\varepsilon_{ij}, \varepsilon_{ij'}) \end{aligned}$$

onde: os dois primeiros termos são as funções de covariâncias genética aditiva e de ambiente permanente, respectivamente; e as covariâncias entre os coeficientes de regressão são iguais aos coeficientes das funções de covariâncias, isto é, matriz \mathbf{K} . Assumindo-se que os resíduos são distribuídos independentemente, com média zero e com variância (σ_ε^2), a $\text{Cov}(\varepsilon_{ij}, \varepsilon_{ij'}) = \sigma_\varepsilon^2$ para $j=j'$, e zero para $j \neq j'$ (Meyer, 1998a).

Ajustando-se um modelo de RA, assume-se, implicitamente, uma determinada estrutura de covariâncias entre os coeficientes de regressão e pode ser caracterizado por uma função de covariâncias (Meyer, 1998 a e c). Desta forma, estas duas propostas convergem e o modelo de FC descrito por Kirkpatrick et al. (1990) é equivalente a um caso especial de RA, como demonstrado por Meyer (1998a). Os coeficientes das FC podem ser estimados diretamente do modelo de RA, usando máxima verossimilhança restrita ou amostragem de Gibbs.

Em notação matricial o modelo acima pode ser descrito como:

$$y = \mathbf{XB} + \mathbf{Z}_1\boldsymbol{\alpha} + \mathbf{Z}_2\boldsymbol{\delta} + \boldsymbol{\varepsilon}, e$$

$$V \begin{bmatrix} \alpha \\ \delta \\ \varepsilon \end{bmatrix} = \begin{bmatrix} K_a \otimes A & 0 & 0 \\ 0 & K_p \otimes I & 0 \\ 0 & 0 & R \end{bmatrix}$$

onde: y é o vetor de observações, B o vetor de efeitos fixos (incluindo a curva de regressão fixa para modelar a trajetória populacional), α o vetor aleatório dos coeficientes genéticos aditivos diretos, δ o vetor de coeficientes de ambiente permanente de animal, X , Z_1 e Z_2 , são as matrizes de incidência correspondentes e ε o vetor de resíduos. K_a , e K_p são as matrizes de variâncias e covariâncias entre os coeficientes de regressão aleatórios para os efeitos genéticos aditivos e de ambiente permanente de animal, respectivamente. A é a matriz de parentesco aditivo e I uma matriz identidade de ordem igual ao número de animais com dados. Cabe lembrar que, utilizando-se modelos multi-características, os efeitos de ambiente permanente não são separados do resíduo.

PREDIÇÃO DE VALORES GENÉTICOS

As equações de modelo misto, em um modelo de regressão aleatória, podem ser representadas como:

$$\begin{bmatrix} X'R^{-1}X & X'R^{-1}Z_1 & X'R^{-1}Z_2 \\ Z_1'R^{-1}X & Z_1'R^{-1}Z_1 + A^{-1} \otimes K_a^{-1} & Z_1'R^{-1}Z_2 \\ Z_2'R^{-1}X & Z_2'R^{-1}Z_1 & Z_2'R^{-1}Z_2 + I \otimes K_p^{-1} \end{bmatrix} \begin{bmatrix} B \\ \alpha \\ \delta \end{bmatrix} = \begin{bmatrix} X'R^{-1}y \\ Z_1'R^{-1}y \\ Z_2'R^{-1}y \end{bmatrix}$$

Quando as variâncias são conhecidas, as soluções para α e δ são BLUP (Henderson, 1975). Cabe lembrar que as matrizes Z_1 e Z_2 , que são matrizes de incidência para os efeitos genético e de ambiente permanente de animal, não contém apenas zeros e uns, como em uma análise multi-características, incluindo as covariáveis (funções do tempo). Em conseqüência, as equações de modelo misto são mais densas que em uma análise multi-características padrão. As soluções para as equações de modelo misto são obtidas por processos iterativos, utilizando algoritmos já desenvolvidos para análises padrão.

Com o modelo acima, o vetor de valores genéticos normalmente obtido em uma análise multi-características, é substituído por um vetor com k_a coeficientes de regressão aleatórios genéticos aditivos (α_{im}). Colocando de outra forma, em uma análise multi-características, o número de efeitos aleatórios preditos para cada animal é igual ao número de características, ou ao número de idades nas quais os animais foram medidos, enquanto que, em uma análise de RA, o número de efeitos genéticos aditivos para cada animal é igual à ordem do polinômio utilizado. O valor genético de um animal (i) em um determinado tempo (t) é predito como:

$$VG(a_t)_i = \sum_{m=0}^{k_a-1} \alpha_{im} \phi_m(a_t) \quad (\text{Jamrozik, et al., 1997})$$

Valor genético para funções da curva, como por exemplo, a taxa de crescimento entre determinadas idades; persistência da lactação; mérito genético total na lactação que pode ser obtido pela integral da curva de lactação; entre outras podem ser estimadas. Assim, é possível a obtenção de mérito genético dos animais para funções de interesse específicas para cada programa de melhoramento. No caso de crescimento, por exemplo, pode ser importante identificar animais de crescimento rápido mas que não atinjam altos pesos à maturidade. Para produção de leite, no caso das raças Zebus especificamente, animais com maior persistência de lactação podem ser interessantes. Ainda são necessários trabalhos nesta área para definição das funções da curva a serem utilizadas em cada caso.

Nobre (2001) e Nobre et al. (2002) utilizando dados de crescimento de animais da raça Nelore, estimaram valores genéticos para pesos do nascimento a 550 dias de idade utilizando modelos multicaracterísticas e de regressão aleatória. As correlações entre os valores genéticos dos animais, obtidos com os dois tipos de modelo, variaram de 0,86 a 0,89 para o efeito genético direto e

de 0,84 a 0,88 para o efeito genético materno. Os autores sugeriram que as diferenças obtidas ocorreram por diferenças nos componentes de variância estimados nos dois tipos de análise e, em menor grau, ao maior número de dados e ao melhor ajuste para os efeitos de meio permitidos pela análise com modelos de regressão aleatória. Entretanto, ressaltaram que os valores genéticos obtidos por RA podem ser piores devido a problemas numéricos e na estimação dos componentes de variância. Schenkel et al. (2002b) compararam os valores genéticos para ganhos em peso, medidos a cada 28 dias em um período de 140 dias em teste, utilizando: 1) um procedimento em dois passos, no qual os pesos de cada animal foram ajustados por meio de um modelo fixo para então calcular o ganho de peso no período do teste, o qual foi então analisado utilizando uma análise padrão; 2) um modelo de regressão aleatório. Comparando as classificações dos animais obtidas a partir dos dois procedimentos, com base no valor genético para o ganho em peso em 140 dias de teste, descreveram correlações de *rank* variando de 0,98 (todos os animais) a 0,73 (1% melhores animais), indicando uma possível modificação de *rank*, principalmente, entre os melhores animais. Os autores sugeriram que, como os coeficientes de herdabilidade estimados pelo dois procedimentos foram muito semelhantes, estas diferenças em valores genéticos estimados devem ter ocorrido por diferenças nos modelos, uma vez que com RA foi possível ajustar os efeitos de ambiente permanente individualmente e efeitos de raça e ambiente temporário para cada medida, o que não foi possível no procedimento em dois passos. Meyer (2004), trabalhando com simulação de dados de crescimento de bovinos de corte, verificou que os modelos de regressão aleatória apresentaram acurácias, medidas pela correlação dos valores estimados com os simulados, de 5,7% a 8,3% maiores que os modelos multi-características. O autor atribuiu esta vantagem à melhor modelagem das variâncias e dos parâmetros genéticos fornecida pelos modelos de regressão aleatória.

CONSIDERAÇÕES FINAIS

Regressões polinomiais têm sido amplamente utilizadas para descrever o crescimento e a curva de lactação dos animais. Em geral, regressões sobre o tempo (idade) são consideradas como efeitos fixos para levar em conta tendências gerais ou dentro de classes de efeitos fixos. O uso de regressões aleatórias para modelar dados longitudinais não é recente e no contexto de modelos lineares foi apresentado por Henderson Jr (1982). Mais recentemente, as regressões aleatórias têm sido empregadas para estimação de parâmetros genéticos para uma variedade de características tais como: produção de leite e/ou seus constituintes (Schaeffer e Dekkers, 1994; Jamrozik e Schaeffer, 1997; Jamrozik et al., 1997; Olori, 1997; Olori et al., 1999; Van der Verf et al., 1998; Brotherstone et al., 2000; Jamrozik e Schaeffer, 2002; Togashi e Lin, 2002; Lin e Togashi, 2002), para descrição de curva de crescimento de suínos (Anderson e Pedersen, 1996), de gado de corte (Varona et al., 1997; Meyer, 2001a, 2002a) e de peixes (McKay et al. 2002); para estimação dos efeitos de interação genótipo-ambiente (Berry et al., 2002), de tendência genética (Tsuruta et al., 2002 e Lawlor, et al., 2002), de parâmetros genéticos de ingestão de alimentos em gado de corte (Schenkel et al., 2002a), entre outros. No Brasil, regressões aleatórias têm sido aplicadas a dados de produção de leite (El Faro, 2002; Costa et al., 2002, entre outros), crescimento de bovinos de corte (Sakaguti, 2000; Albuquerque e Meyer, 2001; Nobre et al., 2002; Cyrillo, 2003; Dias, 2004; Scarpel, 2004), reprodução de bovinos de corte (Mercadante, 2001).

Na maioria dos trabalhos, tem se utilizado regressões sobre polinômios de Legendre para modelar dados longitudinais. Estes modelos não requerem qualquer pressuposição sobre a forma da curva ou trajetória da característica com o tempo. Entretanto, em geral, requerem o ajuste de polinômios de altos graus, especialmente para modelar os efeitos de ambiente permanente de animal (Brotherstone et al., 2000; Meyer, 1999, 2000, 2001a) e, em conseqüência, apresentam um grande número de parâmetros a ser estimado. Polinômios de alto grau são associados a problemas numéricos (Kirkpatrick et al., 1994; Meyer, 1998a) e a grandes exigências computacionais.

Um aspecto importante a ser considerado para avaliação genética de um grande número de animais, como nos programas de avaliação nacionais, são os requerimentos computacionais. A utilização de modelos de regressão aleatória em lugar de modelos multi-características poderá aumentar o número de equações de modelos mistos (no caso de características de crescimento em bovinos de corte, por exemplo), a matriz de coeficientes dos efeitos aleatórios será mais densa (não irá incluir apenas zeros e uns) e, conseqüentemente, os requerimentos computacionais irão aumentar. Assim, alternativas para encontrar modelos mais parcimoniosos tem sido apresentadas. Pool e Meuwissen (2000) encontraram que a ordem do polinômio necessário para modelar as (co)variâncias de dados de produção de leite diminuiu de 5 para 4, quando apenas dados de lactações completas foram consideradas nas análises, e de 4 para 3, quando a heterogeneidade de

variâncias residuais foi considerada. Curvas paramétricas, geralmente com um pequeno número de parâmetros, podem ser utilizadas, com procedimentos Bayesianos como apresentado por Varona et al. (1997, 1999).

Alternativamente, covariâncias dentro de animal podem ser modeladas por uma função de variâncias combinada a uma função paramétrica (Pletcher e Geyer, 1999; Foulley et al., 2000). Mais comumente, são utilizadas funções de correlação estacionárias, isto é, as correlações dependem apenas da diferença em tempo entre as medidas (lag) e têm somente um ou dois parâmetros a serem estimados (Nuñez-Anton e Zimmerman, 2000). Pletcher e Geyer (1999) denominaram este modelo de "character process model" e uma extensão para correlações não estacionárias e comparações com modelos de regressão aleatória são apresentados por Jaffrézic e Pletcher (2000). Programas computacionais como SAS (1997), ASREML (Gilmour et al., 1999) e DXMRR (Meyer, 1998b) podem ser usados para ajustar funções de correlação paramétricas (Little, et al., 2000; Jaffrézic et al., 2002a; Freitas et al., 2002, Albuquerque e Meyer, 2002).

Com regressões aleatórias, são modelados os desvios de cada animal em relação à curva média, que pode ser paramétrica ou não, e a estrutura de covariâncias é derivada do modelo de regressão escolhido. Já com as funções de correlação paramétricas, as estruturas das covariâncias são modeladas diretamente (Jaffrézic et al., 2002a e b). Neste caso, a função de covariâncias entre observações tomadas em dois tempos, t_i e t_j , pode ser representada como:

$C(t_i, t_j) = \sigma_{t_i} \sigma_{t_j} \rho(t_i, t_j)$ onde $\sigma_{t_i}^2$ é a função de variâncias, normalmente uma função polinomial de tempo, e $\rho(t_i, t_j)$ é a função de correlação (Jaffrézic et al., 2002a).

Embora funções de correlação paramétricas sejam amplamente utilizadas em diversas áreas (Diggle et al., 1994), no contexto do melhoramento genético animal o seu uso ainda é limitado. Modelos de auto-correlação de primeira ordem têm sido propostos para modelar efeitos aleatórios em gado de leite, tais como efeitos de ambiente permanente (Harville, 1979; Mansour et al., 1985; Carvalheira et al., 2002) e grupos contemporâneos (Wade et al., 1993). Modelos com correlações estacionárias, mas que permitem mudanças nas variâncias com a idade, também foram desenvolvidos (Mansour et al. 1985; Pletcher e Geyer, 1999; Jaffrézic et al., 2002a e b). Meyer (2001b) apresentou um conjunto de modelos combinando regressões aleatórias sobre polinômios ortogonais de Legendre da idade para modelar as variâncias entre animais e uma função de variâncias associada a uma função de correlação paramétrica, estacionária e não estacionária, para modelar as covariâncias dentro de animais. Aplicando estes modelos a dados de peso de vacas adultas, o autor apontou que os mesmos foram mais parcimoniosos que os modelos de regressão aleatória, considerando polinômios de Legendre para os efeitos de ambiente permanente. Ressalta ainda que estes modelos podem eliminar estimativas de correlação erráticas e não explicáveis entre os pesos tomados em idades extremas. Resultado semelhante foi observado por Albuquerque e Meyer (2002), utilizando dados de peso do nascimento aos 630 dias de idade de animais da raça Nelore. As estimativas de parâmetros genéticos obtidas com um modelo de regressão aleatória, utilizando regressão sobre polinômio de Legendre da idade para os efeitos genéticos de animal e da mãe e ambiente permanente de animal e da mãe, foram semelhantes aos obtidos quando as correlações de ambiente permanente de animal foram modeladas utilizando um modelo de correlação paramétrica não estacionário, diminuindo em 12 a 19 o número de parâmetros a serem estimados.

Modelos de regressão aleatória estão sendo empregados para avaliações genéticas nacionais de bovinos leiteiros em vários países como, Holanda, Alemanha, Áustria, Canadá, entre outros. Em geral, as características avaliadas são produção de leite, gordura, proteína e células somáticas e, na maioria dos casos, polinômios de Legendre têm sido empregados (Roos, et al., 2001; Liu, et al., 2001; Emmerling, et al., 2002; Reinhardt, et al., 2002; Strabel, et al., 2003). Para bovinos de corte não encontramos referências sobre aplicações de modelos de RA para avaliações genéticas de rotina, possivelmente pela dificuldade de modelar não apenas os efeitos de animal, mas também, os efeitos maternos.

REFERÊNCIAS BIBLIOGRÁFICAS

- ABRAMOWITZ, M.; STEGUN, I. A. **Handbook of mathematical functions**. New York: Dover, 1965.
- ALBUQUERQUE, L.G.; MEYER, K. Estimates of covariance functions for growth from birth to 630 days of age in Nelore cattle. **Journal of Animal Science**, v.79, p.2776-2789, 2001.



- ALBUQUERQUE, L.G.; MEYER, K. Estimates of genetic covariance functions for growth assuming a parametric correlation structure for permanent environmental effects. In: WORLD CONGRESS ON GENETICS APPLIED TO LIVESTOCK PRODUCTION, 7., 2002, Montpellier, France. **Proceedings...** Montpellier, France, 2002. 1 CD-ROM.
- ANDERSON, S.; PEDERSEN, B. Growth and food intake curves for group-housed gilts and castrated male pigs. **Journal of Animal Science**, v.63, p.457-464, 1996.
- BERRY, D.P.; BUCKEY, F.; DILLON, P. et al. Genotype X environment interactions for body condition score, body weight and milk yield using random regression models. In: WORLD CONGRESS ON GENETICS APPLIED TO LIVESTOCK PRODUCTION, 7., 2002, Montpellier, France. **Proceedings...** Montpellier, France, 2002. 1 CD-ROM.
- BROTHERSTONE, S.; WHITE, I.M.S.; MEYER, K. Genetic modelling of daily milk yield using orthogonal polynomials and parametric curves. **Journal of Animal Science**, v., 70, p.417-416, 2000.
- CARVALHEIRA, J.; LICITRA, G.; AZZARO, G. et al. Genetic parameters of test-day milk yield for Holstein, Bruna and Modicana breeds in southeastern Sicily using an autoregressive multiple lactation model. In: WORLD CONGRESS ON GENETICS APPLIED TO LIVESTOCK PRODUCTION, 7., 2002, Montpellier, France. **Proceedings...** Montpellier, France, 2002. 1 CD-ROM.
- COSTA, C.N.; MELO, C.M.R.; MARTINEZ, M.L. et al. Estimation of genetic parameters for test day milk records of first lactation Gir cows in Brazil, using random regression. In: WORLD CONGRESS ON GENETICS APPLIED TO LIVESTOCK PRODUCTION, 7., 2002, Montpellier, France. **Proceedings...** Montpellier, France, 2002. 1 CD-ROM.
- CYRILLO, J.N.S.G. **Estimativas de funções de covariância para crescimento de machos Nelore utilizando modelos de regressão aleatória.** Jaboticabal: Universidade Estadual Paulista, 2003. 72p. Tese (Doutorado) – Universidade Estadual Paulista, 2003.
- DIAS, L.T. **Modelos de regressão aleatória para características de crescimento de bovinos da raça Tabapuã.** Jaboticabal: Universidade Estadual Paulista, 2004. 59p. Tese (Doutorado) – Universidade Estadual Paulista, 2004.
- DIGGLE, P.J.; LIANG, K.Y.; ZEGER, S.L. **Analysis of longitudinal data.** Oxford: Oxford Science Publications, Clarendon Press, 1994.
- EL FARO, L. **Estimação de componentes de (co)variância para produção de leite no dia do controle de primeiras lactações de vacas Caracu, aplicando-se “test-day models” de dimensão finita e modelos de regressão aleatória.** Jaboticabal: Universidade Estadual Paulista, 2002. 102p. Tese (Doutorado em Zootecnia) – Universidade Estadual Paulista, 2002.
- EMMERLING, R.; LIDAUER, M.; MANTYSAARI, E.A. Multiple lactation random regression test-day model for Simmental and Brown Swiss in Germany and Austria. In: INTERBULL MEETING, 2002, Interlaken, Switzerland. **Proceedings..** Interlaken, Switzerland, 2002. p.111-117.
- FOULLEY, J. –L.; JAFFRÉZIC, F.; ROBERT-GRANIÉ, C. EM-REML estimation of covariance parameters in Gaussian mixed models for longitudinal data analysis. **Genetics Selection Evolution**, v.32, p.129-141, 2000.
- FREITAS, A.R.; SILVA, L.O.C.; FILHO, K.E.; et al. Covariance structures in repeated measures of body weight of *Boa indicus* beef cattle in Brazil. In: WORLD CONGRESS ON GENETICS APPLIED TO LIVESTOCK PRODUCTION, 7., 2002, Montpellier, France. **Proceedings...** Montpellier, France, 2002. 1 CD-ROM.
- GADINI, C.H. **Genetic evaluation of test day production traits and somatic cell score.** Nebraska: University of Nebraska, 1997. 91p. Thesis (PHD), University of Nebraska, 1997.
- GERSTMAYR, S. Impact of the data structure on the reliability of the estimated genetic parameters in an animal model with maternal effects. **Journal of Animal Breeding and Genetics**, v. 109, p.321-336, 1992.
- GONÇALVES, T.M. **Estudo da curva de lactação de vacas da raça Gir.** Viçosa : Universidade



- Federal de Viçosa, 1994. 66p. Dissertação (Mestrado em Genética e Melhoramento) - Universidade Federal de Viçosa, 1994.
- GILMOUR, A. R.; CULLIS, B.R.; WELHAM, S.J. et al. **ASREML Reference Manual**. Australia: NSW Agriculture, 1999. 120p. (NSW Agriculture Biometric Bulletin, n.3).
- HARVILLE, D.A. Recursive estimation using mixed linear models with autoregressive random effects. In: Van Vleck, L.D., Searle, S.R. (Eds.) Variance Components and Animal Breeding. In: CONFERENCE IN HONOR OF C. R. HENDERSON, 1979, Ithaca. **Proceedings...** Ithaca, 1979. p.157-179.
- HENDERSON, C.R. Best linear unbiased estimation and prediction under a selection model. **Biometrics**, v.31, p.423-447, 1975.
- HENDERSON Jr.; C.R. Analysis of covariance in the mixed model: higher-level, nonhomogeneous, and random regressions. **Biometrics**, v.38, p.623-640, 1982.
- JAFFRÉZIC, F.; PLETCHER, S.D. Statistical models for estimating the genetic basis of repeated measures and other function-valued traits. **Genetics**, v. 156, p.913-922, 2000.
- JAFFRÉZIC, F.; WHITE, I.M.S.; THOMPSON, R.; et al. Contrasting models for lactation curve analysis. **Journal of Dairy Science**, v.85, p.968-975, 2002a.
- JAFFRÉZIC, F.; WHITE, I.M.S.; THOMPSON, R.; et al. Statistical models for the genetic analysis of longitudinal data. In: WORLD CONGRESS ON GENETICS APPLIED TO LIVESTOCK PRODUCTION, 7., 2002, Montpellier, France. **Proceedings...** Montpellier, France, 2002b. 1 CD-ROM.
- JAMROZIK, J.; SCHAEFFER, L. Estimates of genetic parameters for a test day model with random regressions for yield traits of first lactation Holsteins. **Journal of Dairy Science**, v.80, p.762-770, 1997.
- JAMROZIK, J.; SCHAEFFER, L. Bayesian comparison of random regression models for test-day yields in dairy cattle. In: WORLD CONGRESS ON GENETIC APPLIED TO LIVESTOCK PRODUCTION, 7., 2002, Montpellier, France. **Proceedings...** Montpellier, France, 2002. 1 CD-ROM.
- JAMROZIK, J.; SCHAEFFER, L.R.; DEKKERS, J.C.M. Genetic evaluation of dairy cattle using test day yields and random regression model. **Journal of Dairy Science**, v.80, p.1217-1226, 1997.
- KIRKPATRICK, M.; HECKMAN, N. A quantitative genetic model for growth, shape and other infinite-dimensional characters. **Journal of Mathematical Biology**, v.27, p.429-450, 1989.
- KIRKPATRICK, M.; LOFSVOLD, D.; BULMER, M. Analysis of the inheritance, selection and evolution of growth trajectories. **Genetics**, v.124, p.979-993, 1990.
- KIRKPATRICK, M.; HILL, W.G.; THOMPSON, R. Estimating the covariance structure of traits during growth and aging, illustrated with lactations in dairy cattle. **Genetical Research**, v.64, p.57-69, 1994.
- LAWLOR, T.J.; TSURUTA, S.; KLET, I.; et al. Use of random regression model to investigate changes in genetic parameters over time. In: WORLD CONGRESS ON GENETICS APPLIED TO LIVESTOCK PRODUCTION, 7., 2002, Montpellier, France. **Proceedings...** Montpellier, France, 2002. 1 CD-ROM.
- LIN, C.Y.; TOGASHI, K. Simultaneous improvement of lactation milk and persistency. In: WORLD CONGRESS ON GENETICS APPLIED TO LIVESTOCK PRODUCTION, 7. 2002, Montpellier. **Proceedings...** Montpellier, 2002. 1 CD-ROM.
- LITTLE, R.C.; PENDERGAST, J.; NATARAJAN, R. Modelling covariance structure in the analysis of repeated measures data. **Statist. Med.**, v. 19, p. 1793-1919, 2000.
- LIU, Z.; REINHARDT, F.; BUNGER, A.; et al. Application of a random regression model to genetic evaluations of test day yields and somatic cell scores in dairy cattle. In: INTERBULL MEETING, 2001, Budapeste, Hungary. **Proceedings...** Budapeste, Hungary, 2001. p.159-166.



- MANSOUR, H.; NORDHEIM, E.V.; RUTLEDGE, J.J. Maximum likelihood estimation of variance components in repeated measures designs assuming autoregressive errors. **Biometrics**, v. 41, p. 287-294, 1985.
- McKAY, L.T.; SCHAEFFER, L.R.; McMILLAN, I. Analysis of growth curves in Rainbow trout using random regression. In: WORLD CONGRESS ON GENETICS APPLIED TO LIVESTOCK PRODUCTION, 7., 2002, Montpellier, France. **Proceedings...** Montpellier, France, 2002. 1 CD-ROM.
- MERCADANTE, M.E.Z. **Análise de um experimento de seleção para crescimento em bovinos Nelore: respostas direta no peso ao sobreano e correlacionadas no tamanho e reprodução das matrizes.** Piracicaba: Escola Superior de Agricultura Luiz de Queiroz, 2001. 106p. Tese (Doutorado em Agronomia) – Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, 2001.
- MEYER, K. Estimating variances and covariances for multivariate animal models by Restricted Maximum Likelihood. **Genetics Selection Evolution**, v.23, p.67-83, 1991.
- MEYER, K. Estimating covariance functions for longitudinal data using a random regression model. **Genetics Selection Evolution**, v. 30, p.221-240, 1998a.
- MEYER, K. "DXMRR"- a program to estimate covariance functions for longitudinal data by restricted maximum likelihood. In: WORLD CONGRESS ON GENETICS APPLIED TO LIVESTOCK PRODUCTION, 6., 1998, Armidale, Australia. **Proceedings...** Armidale, Australia, 1998b, v.27, p. 465-466.
- MEYER, K. Modeling 'repeated' records: covariance functions and random regression models to analyze animal breeding data. In: WORLD CONGRESS ON GENETICS APPLIED TO LIVESTOCK PRODUCTION, 6. 1998, Armidale, Australia. **Proceedings...** Armidale, Australia, 1998c. v.25, p.517-520.
- MEYER, K. Estimates of genetic and phenotypic covariance functions for postweaning growth and mature weight of beef cows. **Journal of Animal Breeding and Genetics**, v.116, p.181-205, 1999.
- MEYER, K. Random regressions to model phenotypic variation in monthly weights of Australian beef cows. **Livestock Production Science**, v.65, p.19-38, 2000.
- MEYER, K. Estimates of direct and maternal covariance functions for growth of Australian beef calves from birth to weaning. **Genetics Selection Evolution**, v.33, p.1-28, 2001a.
- MEYER, K. Estimating genetic covariance functions assuming a parametric correlation structure for environmental effects. **Genetics Selection Evolution**, v. 33, p. 557-585, 2001b.
- MEYER, K. Estimates of covariance functions for growth of Australian beef cattle from a large set of field data. In: WORLD CONGRESS ON GENETICS APPLIED TO LIVESTOCK PRODUCTION, 7., 2002, Montpellier, France. **Proceedings...** Montpellier, France, 2002a. 1 CD-ROM.
- MEYER, K. Scope for a random regression model in genetic evaluation of beef cattle for growth. **Livestock Production Science**, v.88, p.69-83, 2004.
- MEYER, K.; HILL, W.G. Estimation of genetic and phenotypic covariance functions for longitudinal or 'repeated' records by restricted maximum likelihood. **Livestock Production Science**, v.47, p.185-200, 1997.
- MEYER, K.; GRASER, H.-U.; HAMMOND, K. Estimates of genetic parameters for first lactation test day production of Australian Black and White cows. **Livestock Production Science**, v.21, p.177-199, 1989.
- NOBRE, P.R.C. **Analyses of sequential weights of Nelore cattle using multiple trait and random regression models.** Viçosa: Universidade Federal de Viçosa, 2001. 137p. Tese (Doutorado em genética e melhoramento) – Universidade Federal de Viçosa, 2001.
- NOBRE, P.R.C.; MISZTAL, I.; TSURUTA, S.; et al. Genetic evaluation of growth in beef cattle with random regression model. . In: WORLD CONGRESS ON GENETICS APPLIED TO LIVESTOCK PRODUCTION, 7., 2002, Montpellier, France. **Proceedings...** Montpellier, France. 2002. 1 CD-ROM.



- NUÑEZ-ANTON, V.; ZIMMERMAN, D.L.; Modeling nonstationary longitudinal data. **Biometrics**, v.56, p. 699-705, 2000.
- OLIVEIRA, H.N. **Análise genético-quantitativa da curva de crescimento de fêmeas da raça Guzerá**. Ribeirão Preto: Faculdade de Medicina de Ribeirão Preto, 1995. 73p. Tese (Doutorado em Ciências) - Faculdade de Medicina de Ribeirão Preto/Universidade de São Paulo, 1995.
- OLORI, V.H. **Utilisation of daily milk records in genetic evaluation of dairy cattle**. University of Edinburgh. Edinburg: University of Edinburg, 1997. 230p Thesis (PHD) - University of Edinburg, 1997.
- OLORI, V.E.; HILL, W.G.; MCGUIRK, B.J.; et al. Estimating variance components for test day milk records by restricted maximum likelihood with random regression animal model. **Livestock Production Science**, v. 61, p. 53-63, 1999.
- PANDER, B.L.; HILL, W.G. Genetic evaluation of lactation yield from test day records on incomplete lactation. **Livestock Production Science**, v.37, p.23-36, 1993.
- PLETCHER, S. D.; GEYER, C.J. The genetic analysis of age-dependente traits: modeling the character process. **Genetics**, v.151, p.825-835, 1999.
- POOL, M.H.; MEUWISSEN, T.H.E. Reduction of the number of parameters needed for a polynomial random regression test day model. **Livestock Production Science**, v.64, p.133-145, 2000.
- PTAK, E.; SCHAEFFER, L.R. Use of test day yields for genetic evaluation of dairy sires and cows. **Livestock Production Science**, v.34, p.23-34, 1993.
- REINHARDT, F.; LIU, Z.; BUNGER, A.; et al. Impact of application of a random regression test day model production trait genetic evaluations in dairy cattle. In: INTERBULL MEETING, 2002, Interlaken, Switzerland. **Proceedings...** Interlaken, Switzerland, 2002. p.103-108.
- ROOS, A.P.W.; HARBERS, A.G.F.; JONG, G. Random regression test-day model in the Netherlands. In: INTERBULL MEETING, 2001, Budapeste, Hungary. **Proceedings..** Budapeste, Switzerland, 2001. p.155-158.
- SAKAGUTI, E.S. **Funções de covariâncias e modelos de regressão aleatória na avaliação genética do crescimento de bovinos jovens da raça Tabapuã**. Viçosa: Universidade Federal de Viçosa, 2000. 81p. Tese (Doutorado em Genética e Melhoramento) - Universidade Federal de Viçosa, 2000.
- SAS INSTITUTE. **SAS/STAT Software: changes and enhancements through release 6.12**. Cary, 1997.
- SCARPEL, L.C.P. **Estimativas de parâmetros e valores genéticos para peso corporal de bovinos da raça Guzerá usando-se regressão aleatória**. 2004, 71p. Tese (Doutorado) – Universidade Estadual Paulista, 2004.
- SCHAEFFER, L.R.; DEKKERS, J.C.M. Random regressions in animal models for test-day production in dairy cattle. In: WORLD CONGRESS ON GENETICS APPLIED TO LIVESTOCK PRODUCTION, 5., 1994, Guelph. **Proceedings...**Guelph, 1994, v.17, p.443-446.
- SCHENKEL, F.S.; DEVITT, C.J.B.; WILTON, J.W.; et al. Random regression analyses of feed intake of individually tested beef steers. In: WORLD CONGRESS ON GENETICS APPLIED TO LIVESTOCK PRODUCTION, 7., 2002, Montpellier, France. **Proceedings...** Montpellier, France 2002. 1 CD-ROM.
- SCHENKEL, F.S.; MILLER, S.P; JAMROZIK, J.; et al. Two-step and random regression analyses of weight gain of station-tested beef bulls. **Journal of Animal Science**, v.80, p.1497-1507, 2002.
- STRABEL, T.; SZYDA, J.; PTAK, E.; et al. Comparison of random regression test-day models for production traits of dairy cattle in Poland. In: INTERBULL MEETING, 2003, Rome, Italy. **Proceedings..** Rome, Italy, 2003. p.197-201.
- TOGASHI, K.; LIN, C.Y. Optimal strategy to modify the shape of the lactation curves. In: WORLD CONGRESS ON GENETICS APPLIED TO LIVESTOCK PRODUCTION, 7., 2002, Montpellier, France. **Proceedings...** Montpellier, France, 2002. 1 CD-ROM.



- TSURUTA, S.; MISZTAL, I.; LAWLOR, T.J.; et al. Estimation of changes of genetic parameters over time for type traits in Holsteins using random regression models. In: WORLD CONGRESS ON GENETICS APPLIED TO LIVESTOCK PRODUCTION, 7., 2002, Montpellier, France. **Proceedings...** Montpellier, France, 2002. 1 CD-ROM.
- VAN der WERF, J.; SCHAEFFER, L. **Random Regression in Animal Breeding**. Course Notes, Ontario: University of Guelph, 1997, 70p.
- VAN der WERF, J.H.J.; GODDARD, M.E.; MEYER, K. The use of covariance functions and random regressions for genetic evaluation of milk production based on test day records. **Journal of Dairy Science**, v.81, p.3300-3308, 1998.
- VARONA, L.; MORENO, C.; GARCÍA CORTÉS, L. A.; et al. Multiple trait genetic analysis of underlying biological variables of production functions. **Livestock Production Science**. v.47, p.201-209, 1997.
- VARONA, L.; MORENO, C.; GARCIA CÓRTEZ, L.A.; et al. Two-step versus joint analysis of Von Bertalanffy function. **Journal of Animal Breeding and Genetics**, v.116, p.331-338, 1999.
- WADE, K.M.; QUAAS, R.L.; VAN VLECK, L.D. Estimation of the parameters involved in a first order autoregressive process for contemporary groups. **Journal of Dairy Science**, v.76, p.3033-3040, 1993.