

# NÚMERO DE AMOSTRAGENS DA PRODUÇÃO DE LEITE PARA ESTIMATIVA DA PRODUÇÃO MÉDIA DIÁRIA DE LEITE EM VACAS CHAROLÊS

PAULO SANTANA PACHECO,<sup>1</sup> JOÃO RESTLE,<sup>2</sup> ALINE KELLERMANN DE FREITAS,<sup>3</sup> IVAN LUIZ BRONDANI,<sup>4</sup> JOSÉ HENRIQUE SOUZA DA SILVA<sup>4</sup> E DARI CELESTINO ALVES FILHO<sup>4</sup>

1. Pós-doutorando do PPGCA/UFG, bolsista PDJ-CNPq. E-mail: pspacheco@pop.com.br

2. Departamento de Produção Animal/EV/UFG

3. Doutoranda do PPGZ/UFRGS, bolsista CNPq

4. Departamento de Zootecnia/UFSM

## RESUMO

Objetivou-se com este experimento avaliar modelos preditos com diferentes números de amostragens da produção de leite visando estimar a produção média diária de leite de vacas Charolês até os 217 dias de lactação. Foram utilizadas 27 vacas, sendo as amostragens da produção de leite realizadas pelo método direto com ordenha, aos 14, 42, 70, 98, 126, 154, 182 dias e ao desmame (217 dias). Utilizou-se o procedimento Stepwise para seleção dos modelos de regressão linear múltipla, objetivando estimar a produção média diária de leite de acordo com o número de amostragens da produção de leite. Os dados foram submetidos à análise residual através dos testes de heterocedasticidade da variância (estatística  $\chi^2$ ), normalidade (estatística W de Shapiro-Wilk) e diagnóstico de observações influentes (três observações foram excluídas de um n inicial=30), além do diagnóstico de multicolinearidade. Para a validação dos modelos de regressão selecionados, utilizou-se a estatística

PRESS (predição da soma de quadrados do erro). Os dias das amostragens da produção de leite selecionados foram: 154 para 1 amostragem; 70 e 154 para 2 amostragens; 14, 70 e 154 para 3 amostragens; 14, 42, 70 e 154 para 4 amostragens; 14, 42, 70, 126 e 154 para 5 amostragens; 14, 42, 70, 126, 154 e 182 para 6 amostragens e 14, 42, 70, 98, 126, 154 e 182 para 7 amostragens. Os  $R^2$  ajustados para modelos com 1; 2; 3; 4; 5; 6 e 7 amostragens foram, respectivamente, 0,72; 0,87; 0,97; 0,97; 0,99; 0,99 e 0,99. Consideraram-se todos os modelos preditos com diferentes números de amostragens da produção de leite satisfatórios na estimativa da produção média diária de leite. Na possibilidade de executar diversas amostragens ao longo do período de lactação, três amostragens da produção de leite, em datas estrategicamente escolhidas ao longo do período de lactação, são suficientes para predizer com elevada precisão a produção média diária de leite.

**PALAVRAS-CHAVES:** Análise residual, lactação, regressão linear múltipla, seleção de variáveis independentes, técnicas de validação de modelos de regressão.

## ABSTRACT

### NUMBER OF COLLECTIONS OF MILK YIELD TO ESTIMATE AVERAGE DAILY MILK YIELD OF CHAROLAIS BEEF COWS

The objective of the experiment was to evaluate different number of collections of milk yield to prediction of models to estimate average daily milk yield of Charolais beef cows until 217 days of lactation. The eight evaluations of milk yield were taken from 27 cows, by direct method through milking, at 14, 42, 70, 98, 126, 154, 182 days and weaning (217 days). For the selection of multiple linear

regression models, with the objective to estimate the average daily milk yield according to number of milk collections, Stepwise procedure was used. The data were submitted to residual analysis by tests of heterocedasticity of variance ( $\chi^2$  statistic), normality (W statistic of Shapiro-Wilk) and diagnosis of outliers (three observations were excluded from original n=30), beyond multicollinearity diagnosis. For

validation of the selected regression models, the prediction error sum of squares statistic (PRESS), was used. The selected days of milk yield collections were: day 98 for 1 collection; days 70 and 98 for 2 collections; days 14, 70, and 98 for 3 collections; days 14, 70, 98 and 154 for 4 collections; days 14, 70, 98, 154 and 182 for 5 collections; days 14, 42, 70, 98, 154 and 182 for 6 collections and days 14, 42, 70, 98, 126, 154 and 182 for 7 collections. The adjusted  $R^2$  for models with 1; 2; 3; 4; 5; 6 and 7 collections were,

respectively, 0.72; 0.87; 0.97; 0.97; 0.99; 0.99 and 0.99. All the predicted models with different number of collections of milk yield were satisfactory to estimate average daily milk yield. In a possibility of to execute various collections of milk yield during lactation period, three milk collections chosen in strategic dates during lactation period were sufficient to estimate with high precision the average daily milk yield.

**KEY WORDS:** Lactation, multiple linear regression, selection of independent variables, residual analysis, techniques of validation of regression models.

## INTRODUÇÃO

Em bovinos de corte, a avaliação da produção de leite tem grande importância pela sua relação direta com o desempenho pré-desmame dos(as) bezerros(as). Estudos demonstraram correlações positivas e superiores a 0,40 entre produção de leite e peso ao desmame dos bezerros (ALENCAR, 1987; RIBEIRO & RESTLE, 1991, RESTLE et al., 2004).

No entanto, PACHECO et al. (2009) comentam sobre as várias dificuldades enfrentadas na coleta destas amostras e consequências desfavoráveis para os animais, como o aumento considerável do estresse, em virtude do pouco ou quase inexistente manejo em vacas de corte, se comparadas com vacas produtoras de leite, ocasionando dificuldade na realização da ordenha quando se utiliza o método direto; desgaste físico de vacas e bezerros quando são percorridas distâncias razoáveis dos poteiros de pastejo até o centro de manejo e vice-versa; aumento do estresse dos bezerros por causa da separação de suas mães conforme as metodologias empregadas; frequência acentuada das avaliações em rebanhos que utilizam o desmame convencional, necessitando considerável quantidade de tempo e mão de obra para a realização das amostragens.

Além disso, em vários estudos relacionados com a produção de leite de vacas de corte, desenvolvidos nas mais variadas situações ambientais e com diversos grupos genéticos, o número ou frequência de amostragens foi determinado previamente, considerando manejo, disponibi-

lidade de mão de obra e de tempo, quantidade de animais a serem avaliados etc., ocasionando diferentes intervalos entre as amostragens. Com isso, há trabalhos com intervalos de amostragens a cada 21 dias (CERDÓTES et al., 2004), a cada 28 dias (RIBEIRO et al., 1991; RESTLE et al., 2003), a cada 30 dias (HOLLOWAY et al., 1985; ALENCAR et al., 1988; ALBUQUERQUE et al., 1993; BROWN & BROWN, 2002), a cada 50 dias (FREETLY & CUNDIFF, 1998), a cada 90 dias (ALENCAR et al., 1985; ALENCAR, 1987; ALENCAR et al., 1992) ou a cada 100 dias (JOHNSON et al., 2003).

Portanto, objetivou-se, com este estudo, avaliar modelos preditos com diferentes números de amostragens da produção de leite visando estimar a produção média diária de leite de vacas Charolês.

## MATERIAL E MÉTODOS

O experimento foi desenvolvido no Setor de Bovinocultura de Corte do Departamento de Zootecnia da Universidade Federal de Santa Maria, no município de Santa Maria, RS, localizado na região fisiográfica Depressão Central, a 153 m de altitude, apresentando, segundo classificação de Köppen, clima subtropical úmido (cfa) (MORENO, 1961). A temperatura média anual varia de 14,3 a 25,2 °C, com média das mínimas de 9,7 °C no mês de agosto e média das máximas de 29,9 °C no mês de janeiro, sendo que geadas podem ocorrer de abril a outubro. A média anual da umidade relativa do ar é de 73%

e a precipitação é de 1650,9 mm (RIO GRANDE DO SUL, 1979).

Utilizaram-se trinta vacas da raça Charolês, com idades variando de três a nove anos, tomadas ao acaso do mesmo rebanho experimental e submetidas às mesmas condições de manejo no período anterior ao início do experimento.

No período pós-parto, vacas e bezerros(as) foram submetidos a dois tipos de pastagem: pastagem nativa do parto ao desmame dos bezerros, com lotação de 1,0 vaca mais bezerro (a)/ha ou pastagem cultivada de inverno composta de aveia branca (*Avena sativa*) mais azevém (*Lolium multiflorum*) do parto até o final do mês de novembro e após em pastagem de capim-papuã (*Brachiaria plantaginea*), até o desmame dos bezerros, com lotação de 1,76 vaca mais bezerro (a)/ha.

O período de parição foi de setembro a dezembro, sendo os bezerros(as) obtidos(as) de inseminação artificial (primeiros 45 dias do período reprodutivo) ou de monta natural (45 dias seguintes). Durante o período de inseminação, utilizou-se sêmen de seis touros da raça Charolesa e de seis da raça Nelore e para o repasse em monta natural, quatro touros Charolês e quatro touros Nelore. Assim, produziram-se bezerros (as) puros(as) e cruzados(as) F1 Charolês x Nelore.

Realizaram-se as avaliações da produção de leite aos 14, 42, 70, 98, 126, 154, 182 dias e ao desmame. O desmame ocorreu com idade média dos bezerros de 217 dias. O método utilizado para determinar a produção de leite foi o direto, com ordenha manual, segundo metodologia descrita por RIBEIRO et al. (1991) e RESTLE et al. (2003). As vacas eram separadas de seus bezerros pela manhã (um dia antes da ordenha) e soltas em piquete anexo. Por volta das 18h, eram recolhidas novamente até o centro de manejo, permanecendo com o bezerro por trinta minutos, tempo suficiente para esgotar o úbere. Após mamarem, os bezerros eram separados novamente das vacas e ficavam contidos no centro de manejo, sendo as vacas soltas em piquete com pasto e água. No dia seguinte, às 6h, iniciava-se a ordenha. Para realização da ordenha, a vaca era contida pelo pescoço e a pata traseira esquerda era amarrada com corda, recebendo, então, 3 U.I. de ocitocina via intramuscular,

a fim de estimular a liberação do leite. A ordenha foi realizada, após o úbere ser lavado e massageado, apenas nos dois quartos do lado direito, com duração aproximada de cinco minutos. A produção obtida era multiplicada por dois, para se obter a produção do úbere, sendo também corrigida para 24 horas, através da seguinte fórmula:

$$\text{Produção leite (l/dia)} = \frac{\text{produção [l] dos dois quartos do úbere} * 2 * 60 \text{min} * 24 \text{h}}{\text{tempo [min] entre a última mamada e a ordenha}}$$

Inicialmente, os dados foram ajustados mediante um modelo matemático em que se consideraram os efeitos de sexo do bezerro, a composição racial do(a) bezerro(a) (puros ou cruzados FL Nelore x Charolês), o tipo de pastagem, e as covariáveis: ordem de parição e idade da vaca (anos), utilizando-se o procedimento GLM (SAS, 1997).

Procedeu-se à análise da matriz de correlação de Pearson entre a produção média diária de leite (0 a 217 dias) e a produção de cada dia da amostragem da produção de leite (14, 42, ..., 217 dias), por meio do procedimento CORR (SAS, 1997).

Para seleção dos modelos de regressão objetivando estimar a produção média diária de leite de acordo com o número de amostragens da produção de leite, optou-se pelo procedimento STEPWISE (SAS, 1997), considerando  $P < 0,05$  a probabilidade da estatística F para entrada e permanência de cada variável preditora no modelo.

Procedeu-se à análise residual através dos testes de heterocedasticidade da variância (estatística  $\chi^2$ ) pela opção SPEC do procedimento REG (SAS, 1997) e normalidade (estatística W de Shapiro-Wilk) pela opção NORMAL do procedimento UNIVARIATE (SAS, 1997).

Foi realizado diagnóstico de observações influentes ou *outliers* nas estimativas dos parâmetros das variáveis preditoras, através da análise da estatística resíduo padronizado de Student (RStudent), obtida pela opção INFLUENCE do procedimento REG (SAS, 1997). Consideraram-se observações com RStudent superior a 2, em valores absolutos, como influentes, e três observações foram excluídas da análise. Assim, o  $n=30$  passou a ser igual a 27.

O diagnóstico de multicolinearidade entre os coeficientes de regressão utilizados para estimar a produção média diária de leite, representados pelos dias de amostragem da produção de leite, foi realizado conforme sugestões de FREUND & LITTEL (1991) e SOUZA (1998), através da análise da matriz de correlação de Pearson, e das medidas fator de inflação da variância (FIV), autovalor ( $\lambda$ ), índice de condição (IC) e proporção da variância associada com cada autovalor. As medidas de multicolinearidade foram obtidas pelas opções VIF e COLLINOINT do procedimento REG (SAS, 1997).

Para a validação dos modelos de regressão múltipla selecionados, utilizou-se a estatística predição da soma de quadrados do erro (PRESS) no cálculo de um  $R^2$  aproximado para validação ou predição do modelo proposto. Optou-se por esta alternativa por considerar observações do próprio conjunto de dados para seu cálculo. A estatística PRESS foi obtida pela opção PRESS do procedimento REG (SAS, 1997).

Informações sobre a teoria e estimação das estatísticas utilizadas na análise residual, nas

observações influentes, na multicolinearidade e na validação de modelos podem ser obtidas em MONTGOMERY & PECK (1982), FREUND & LITTEL (1991), SAS (1997), DRAPER & SMITH (1998) e SOUZA (1998).

## RESULTADOS E DISCUSSÃO

As médias para produção média diária de leite ao longo do período de lactação e a matriz contendo os coeficientes de correlação de Pearson entre as diferentes datas de amostragem da produção de leite constam na Tabela 1.

A produção média diária de leite foi de 4,5 L/dia para o período de lactação de 217 dias, valor superior ao de 3,40 L/dia relatado no estudo de RIBEIRO et al. (1991) em 182 dias de lactação e muito próximo aos 4,48 L/dia relatados por MELTON et al. (1967) em 175 dias lactação e MARSHALL et al. (1976) em 210 dias de lactação. Todos estes estudos comentados avaliaram vacas Charolês, primíparas e múltíparas, com idades entre 2 e 10 anos.

**TABELA 1.** Médias e desvios-padrão (dp), e coeficientes de correlação de Pearson entre as produções médias diárias de leite em vacas Charolês

		Produção média diária de leite, L (dp) <sup>1</sup>							
	0 a 217 dias	14	42	70	98	126	154	182	217
	4,5 (1,2)	5,9 (1,8)	5,1 (1,7)	5,1 (2,0)	5,0 (1,7)	4,4 (1,5)	4,0 (1,7)	2,8 (1,8)	1,7 (0,9)
		Coeficientes de correlação de Pearson							
14	0,69**								
42	0,70**	0,40*							
70	0,68**	0,29	0,66**						
98	0,82**	0,40*	0,49**	0,79**					
126	0,68**	0,45**	0,15	0,15**	0,51**				
154	0,85**	0,43*	0,45**	0,38*	0,66**	0,74**			
182	0,68**	0,31	0,34	0,17	0,39*	0,63**	0,86**		
217	0,31	0,09	0,37	0,32	0,27	0,20	0,20	-0,01	

<sup>1</sup> 14, 42, 70, 98, 126, 154, 182 e 217 são os dias da amostragem da produção de leite.

\*\* P<0,01; \* P<0,05.

De acordo com a matriz de correlação de Pearson, com exceção da avaliação aos 217 dias (desmame), a produção de todas as amostragens

de leite (14, 42, ..., 182 dias) correlacionaram-se significativamente (P<0,05) com a produção média diária de toda a lactação (0 a 217 dias),

com coeficientes de correlação variando de 0,68 a 0,85, ou seja, de elevada magnitude. Resultados semelhantes foram relatados por PACHECO et al. (2009), avaliando a lactação de vacas Nelore nas mesmas datas de amostragem do presente estudo, variando de 0,57 a 0,89, dos 14 aos 182 dias de amostragem. No presente estudo, a data de amostragem da produção de leite que apresentou maior coeficiente de correlação com a produção de toda a lactação (0 a 217 dias) foi aos 154 dias (0,85),

enquanto que, em vacas Nelore, PACHECO et al. (2009) verificaram maior coeficiente de correlação ( $r=0,89$ ) entre o 98º dia com a produção de toda a lactação (0 a 217 dias).

As estatísticas utilizadas na escolha dos modelos de regressão para estimar a produção média diária de leite em 217 dias de lactação, com diferentes números de amostragens da produção de leite, estão apresentadas Tabela 2.

**TABELA 2.** Resumo das estatísticas utilizadas na escolha dos modelos para estimar a produção média diária de leite em vacas Charolês, de acordo com o procedimento Stepwise

Passo (Step)	Dia da amostragem da produção de leite incluída no modelo	Número de amostragens da produção de leite incluídas no modelo	R <sup>2</sup> parcial	R <sup>2</sup>	R <sup>2</sup> ajustado <sup>1</sup>	F	Pr>F
1	154	1	0,7299	0,7299	0,7191	67,57	<,0001
2	70	2	0,1511	0,8810	0,8711	30,47	<,0001
3	14	3	0,0892	0,9702	0,9663	68,87	<,0001
4	42	4	0,0077	0,9779	0,9739	7,66	0,0112
5	126	5	0,0134	0,9914	0,9893	32,64	<,0001
6	182	6	0,0026	0,9939	0,9921	8,54	0,0084
7	98	7	0,0046	0,9986	0,9980	61,23	<,0001
8	217	8	0,0014	1,0000	1,0000	infity	<,0001

<sup>1</sup> R<sup>2</sup> ajustado =  $1 - \frac{(n-i) * (1 - R^2)}{n-p}$ , em que  $i$  é igual a 1, pois o intercepto está incluído no modelo;  $n$  é o número de observações utilizadas para estimar o modelo e  $p$  é o número de parâmetros do modelo, incluindo intercepto.

SAS (1997) esclarece que, neste procedimento, o valor da estatística F parcial é utilizado para inclusão e permanência das variáveis no modelo. As variáveis a serem utilizadas no modelo de regressão são adicionadas uma a uma, considerando o maior valor da estatística F parcial ( $P<0,05$ ). Para cada nova variável adicionada, o procedimento Stepwise verifica todas as variáveis já incluídas no modelo, e exclui a que não apresenta valor F parcial, suficientemente, significativo ( $P<0,05$ ), para permanecer no modelo.

O uso do R<sup>2</sup> ajustado é justificado por SAS (1997) em casos de regressão linear múltipla, visando corrigir o aumento do R<sup>2</sup> com a inclusão de novos parâmetros no modelo, mesmo que estes não tenham relação com a variável dependente.

Analisando os valores dos coeficientes de determinação (R<sup>2</sup>) parciais, nota-se que, a partir da quarta amostragem da produção de leite incluída no modelo, os incrementos foram menos expressivos. Considerando o R<sup>2</sup> ajustado, com uma (no 154º dia de lactação), duas (70º e 154º dias de lactação) e três (14º, 70º e 154º dias de lactação) amostragens da produção de leite, o modelo estimado pelo procedimento Stepwise explicou, respectivamente, 71,91; 87,11 e 96,63% da variação da produção média diária de leite até os 217 dias de lactação. Em vacas Nelore, PACHECO et al. (2009) relataram incrementos expressivos nos valores dos coeficientes de determinação (R<sup>2</sup>) parciais até a quarta coleta de leite incluída no modelo, sendo as datas selecionadas

para uma amostragem ao 98° dia ( $R^2$  ajustado = 0,7900), para duas amostragens aos 70° e 98° dias ( $R^2$  ajustado = 0,8718), para três amostragens aos 14°, 70° e 98° dias ( $R^2$  ajustado = 0,9064), e para quatro amostragens aos 14°, 70°, 98° e 154° dias ( $R^2$  ajustado = 0,9596).

Procedeu-se aos diagnósticos em regressão, através de técnicas formais e visuais, para detectar desvios das hipóteses clássicas do modelo linear, conforme SOUZA (1998). Esse diagnóstico passa por duas etapas: análise residual e análise de multicolinearidade. Conforme SOUZA (1998), esses desvios avaliados na análise residual seriam os

resíduos que: não apresentam variância constante; possuem erros de especificação na forma funcional da resposta; suas variáveis residuais não são independentes; seus extremos ocorrem nos valores das variáveis independentes (observações influentes) e suas variáveis residuais não têm distribuição comum normal.

Na Tabela 3, consta a análise residual formal, o que permite avaliar a inconstância da variância dos resíduos pelo teste de heterocedasticidade através da estatística Qui-quadrado ( $\chi^2$ ), e a normalidade dos resíduos pela estatística W de Shapiro-Wilk (quando  $n < 2000$ ), conforme SOUZA (1998).

**TABELA 3.** Resumo da análise de heterocedasticidade e de normalidade dos resíduos dos modelos selecionados para estimativa da produção média diária de leite em vacas Charolês, de acordo com o número de amostragens da produção de leite

Estatísticas <sup>1</sup>	Número de amostragens da produção de leite							
	1	2	3	4	5	6	7	8
$\chi^2$	0,44	4,24	8,52	13,61	16,38	23,71	26,13	-
Pr > $\chi^2$	0,8025	0,5148	0,4828	0,4790	0,6927	0,5925	0,5656	-
W	0,9746	0,9768	0,9545	0,9674	0,9490	0,9735	0,9874	0,9734
Pr < W	0,7275	0,7846	0,2754	0,5371	0,2027	0,6983	0,9800	0,6942

<sup>1</sup>  $\chi^2$  = valor da estatística Qui-quadrado (análise de heterocedasticidade dos resíduos); W = valor da estatística W de Shapiro-Wilk (análise de normalidade dos resíduos).

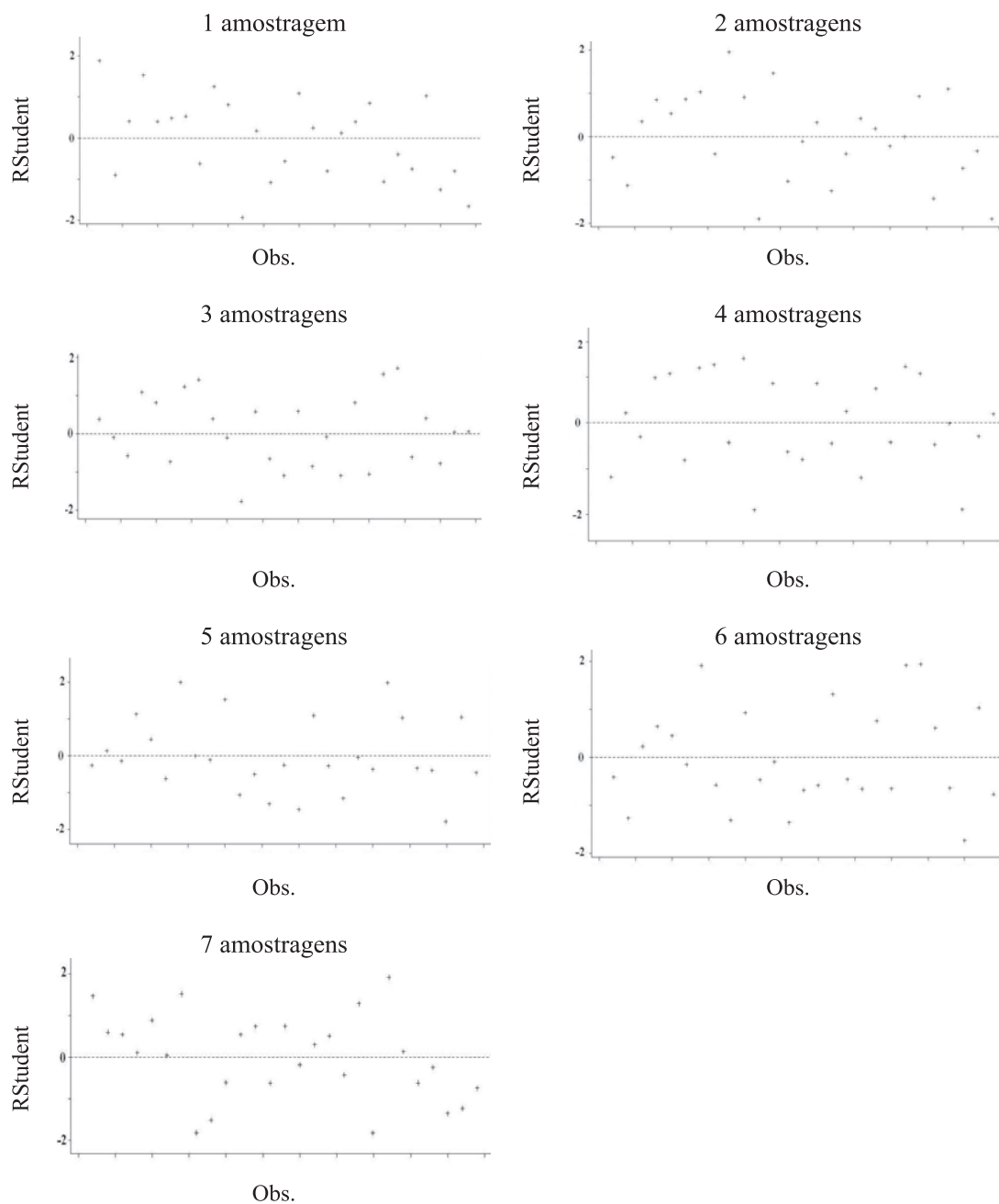
O fato de as probabilidades das respectivas estatísticas não terem sido significativas em ambas as análises indica que os modelos selecionados de acordo com o número de amostragens da produção de leite atendem às pressuposições de resíduos com variância constante e com distribuição normal.

Na Figura 1, consta a análise residual visual de observações influentes, para cada modelo selecionado de acordo com o número de amostragens da produção de leite, utilizando-se o resíduo padronizado de Student (RStudent) e os valores observados da produção média diária de leite.

Conforme MONTGOMERY & PECK (1982), uma observação influente é uma observação extrema, ou seja, observação atípica do restante dos dados, apresentando elevado desvio-padrão da média. Falhas na amostragem dos dados

e nos equipamentos de medida são exemplos de eventos que resultam em observações influentes. SOUZA (1998) comenta que, de um modo geral, para modelo linear, todas as estatísticas geradas pelo método de quadrados mínimos podem ser substancialmente afetadas ou influenciadas por umas poucas observações. Desse modo, é muito importante a identificação de tais observações influentes e a avaliação de seus impactos.

Dentre as várias medidas de avaliação de influência, utilizou-se o resíduo padronizado de Student (RStudent), considerando observações possivelmente influentes aquelas com valores absolutos maiores que 2 para esta estatística (SAS, 1997). Após remoção de três observações influentes, de uma amostragem inicial de trinta observações, verifica-se na Figura 1 que esta exclusão foi uma medida corretiva eficiente.



**FIGURA 1.** Análise gráfica de observações influentes dos modelos selecionados, considerando resíduo de Student (RStudent) e valores observados da produção média diária de leite (Obs.), de acordo com o número de amostragens da produção de leite em vacas Charolês.

Em relação à análise de multicolinearidade, em uma primeira etapa, verificam-se, na Tabela 1, elevados coeficientes de correlação entre as variáveis preditoras dos modelos. Isto pode ser indicativo de algum grau de confundimento entre estas variáveis preditoras, embora a severidade da

multicolinearidade não deve ser quantificada somente pela magnitude da correlação, devendo ser avaliadas outras estatísticas específicas, conforme apresentado na Tabela 4.

Para esta segunda etapa do diagnóstico de multicolinearidade, foi utilizado ajuste para o in-

tercepto, seguindo a recomendação de FREUND & LITTEL (1991), visando reduzir de modo substancial os valores dos índices de condição (IC) e desprezar os efeitos da multicolinearidade nas estatísticas associadas. Conforme SOUZA (1998), a multicolinearidade pode ser um problema e necessitar ação corretiva quando o interesse inferencial concentra-se nas avaliações dos impactos dos coeficientes individuais  $\beta_j$  na resposta, como é o caso dos modelos que utilizam regressão múltipla. FREUND & LITTEL (1991),

SAS (1997) e SOUZA (1998) sugerem problemas de multicolinearidade quando a medida fator de inflação da variância (FIV) for maior que 10, autovalor ( $\lambda$ ) quando menor que 0,01, índice de condição (IC) quando maior que 30 (entre 10 e 30 possível problema) e para proporção da variância associada com os autovalores quando maior que 0,80. Com isso, no presente estudo (Tabela 4), a multicolinearidade não influenciou a estimativa dos modelos de regressão linear múltipla.

**TABELA 4.** Diagnóstico de multicolinearidade (intercepto ajustado)<sup>1</sup>

Dia da amostragem da produção de leite	$\Lambda$	IC	Proporção da variação associada aos autovalores							
			Dia da amostragem da produção de leite							
			14	42	70	98	126	154	182	217
14	3,91	1,00	0,02	0,01	$6 \cdot 10^{-3}$	$7 \cdot 10^{-3}$	0,01	$5 \cdot 10^{-3}$	$5 \cdot 10^{-3}$	$5 \cdot 10^{-3}$
42	1,53	1,59	$2 \cdot 10^{-3}$	0,03	0,03	$3 \cdot 10^{-3}$	0,03	$5 \cdot 10^{-3}$	0,02	0,07
70	0,85	2,14	0,08	0,01	0,02	$3 \cdot 10^{-3}$	0,03	$9 \cdot 10^{-4}$	$2 \cdot 10^{-7}$	0,56
98	0,73	2,31	0,63	$5 \cdot 10^{-5}$	0,02	$9 \cdot 10^{-3}$	0,01	$5 \cdot 10^{-3}$	0,02	0,04
126	0,61	2,53	$4 \cdot 10^{-6}$	0,24	0,02	0,07	0,06	$1 \cdot 10^{-3}$	0,04	$8 \cdot 10^{-3}$
154	0,19	4,51	0,25	0,50	$4 \cdot 10^{-3}$	0,01	0,72	0,02	0,09	0,22
182	0,13	5,57	0,01	0,14	0,65	0,32	0,13	0,11	0,18	0,01
217	0,06	7,95	$2 \cdot 10^{-3}$	0,05	0,26	0,57	$7 \cdot 10^{-5}$	0,84	0,62	0,07
FIV			1,55	2,67	4,82	5,95	3,15	9,35	5,49	1,45

<sup>1</sup>  $\lambda$  = autovalor;

$$IC = \text{índice de condição} = IC_i = \sqrt{\frac{\lambda_{\text{máximo}}}{\lambda_i}};$$

$$FIV = \text{fator de inflação da variância} = FIV_i = \frac{1}{(1 - R_i^2)}.$$

Os coeficientes de regressão dos modelos selecionados para estimativa da produção média diária de leite em raças Charolês, de acordo com o número de amostras de leite, constam na Tabela 5.

Na Tabela 6 apresenta-se o resumo das estatísticas utilizadas na análise de validação dos modelos, para verificar se o modelo predito terá sucesso em sua aplicação final, ou seja, na predição de dados.

MONTGOMERY & PECK (1982) citam três procedimentos usados para validar um modelo de regressão: a análise dos coeficientes do modelo e valores preditos baseado em experiência própria

ou resultados de outros estudos; amostragem de novos dados para avaliar a capacidade de predição do modelo; e o uso dos próprios dados utilizados na predição do modelo (empregado no presente estudo). Neste procedimento, os dados originais são separados em duas partes: uma para construir o modelo de regressão e a outra para estudar a capacidade de predição do referido modelo. Trata-se de separação dos dados que é feita com o uso da estatística PRESS (soma de quadrados predito do erro) (ALLEN, 1974), sendo utilizada para computar um  $R^2$  aproximado para validação ou predição do modelo de regressão desenvolvido.



**TABELA 5.** Coeficientes de regressão dos modelos selecionados para estimativa da produção média diária de leite em vacas Charolês, de acordo com o número de amostragens da produção de leite

Número de amostragens da produção de leite	Dia(s) da(s) amostragem(s) da produção de leite	Coeficientes de regressão representando os dias de amostragens da produção de leite									
		Intercepto	14	42	70	98	126	154	182	217	
1	154	2,076						0,599			
2	70, 154	1,234			0,253			0,488			
3	14, 70, 154	0,424	0,225		0,222			0,401			
4	14, 42, 70, 154	0,339	0,211	0,087	0,181			0,385			
5	14, 42, 70, 126, 154	0,042	0,180	0,125	0,184		0,159	0,280			
6	14, 42, 70, 126, 154, 182	0,053	0,184	0,119	0,198		0,161	0,210	0,071		
7	14, 42, 70, 98, 126, 154, 182	0,022	0,181	0,138	0,130	0,117	0,140	0,141	0,103		
8	14, 42, 70, 98, 126, 154, 182, 217	-7,10*10 <sup>-15</sup>	0,188	0,125	0,125	0,125	0,125	0,125	0,125	0,063	

**TABELA 6.** Resumo das estatísticas utilizadas na análise de validação dos modelos selecionados para estimativa da produção média diária de leite em vacas Charolês, de acordo com o número de amostragens da produção de leite

Estatísticas	Número de amostragens da produção de leite							
	1	2	3	4	5	6	7	8
GL <sub>Erro</sub> <sup>1</sup>	25	24	23	22	21	20	19	18
SQ <sub>Total</sub> <sup>2</sup>	37,62	37,62	37,62	37,62	37,62	37,62	37,62	37,62
PRESS <sup>3</sup>	11,90	5,31	1,51	1,37	0,60	0,45	0,14	1,43*10 <sup>-26</sup>
S <sub>PRESS</sub> <sup>4</sup>	0,14	0,10	0,05	0,05	0,04	0,03	0,02	6,64*10 <sup>-15</sup>
R <sup>2</sup> <sub>Validação</sub> <sup>5</sup>	0,6837	0,8589	0,9599	0,9636	0,9841	0,9880	0,9963	1,0000

<sup>1</sup> GL<sub>Erro</sub> = graus de liberdade do erro.

<sup>2</sup> SQ<sub>Total</sub> = soma de quadrados total.

<sup>3</sup> PRESS = soma de quadrados predito do erro =  $\sum_{i=1}^n [y_i - \hat{y}_{(i)}]^2$

<sup>4</sup> S<sub>PRESS</sub> = desvio-padrão de PRESS =  $\frac{\sqrt{PRESS}}{GL_{Erro}}$

<sup>5</sup> R<sup>2</sup><sub>Validação</sub> =  $1 - \frac{PRESS}{SQ_{Total}}$ .

Verifica-se que os R<sup>2</sup> das validações foram muito próximos dos R<sup>2</sup> ajustados apresentados na Tabela 2, demonstrando boa precisão dos modelos de regressão selecionados para estimativa da produção de leite média diária em vacas Charolês. Resposta semelhante ao uso da estatística PRESS para validação de modelos de regressão para estimativa da produção de leite em vacas Nelore foi relatada por PACHECO et al. (2009).

## CONCLUSÕES

Todos os modelos preditos com diferentes números de amostragens da produção de leite foram considerados satisfatórios na estimativa da produção média diária de leite até os 217 dias de lactação em vacas Charolês.

Na possibilidade de executar diversas amostragens ao longo do período de lactação, três

amostragens da produção de leite, realizadas aos 14, 70 e 154 dias, são suficientes para predizer com elevada precisão a produção média diária de leite.

## REFERÊNCIAS

- ALBUQUERQUE, L.G.; ELER, J.P.; COSTA, M.J.R.P. Produção de leite e desempenho do bezerro na fase de aleitamento em três raças bovinas de corte. **Revista Brasileira de Zootecnia**, v. 22, n. 5, p. 745-754, 1993.
- ALENCAR, M.M. Efeitos da produção de leite sobre o desenvolvimento de bezerros Canchim. **Revista da Sociedade Brasileira de Zootecnia**, v.16, n.1, p. 1-13, 1987.
- ALENCAR, M.M.; JUNQUEIRA FILHO, A.A.; PARANHOS, N.E. Produção de leite em vacas da raça Canchim. **Revista da Sociedade Brasileira de Zootecnia**, v. 14, n. 3, p. 358-366, 1985.
- ALENCAR, M.M.; RIBEIRO, R.P.; VERÍSSIMO, C.J.; DURAM, J.T.; MORO, M.E.G. Efeitos da produção de leite das vacas sobre o desenvolvimento pós-desmama de bezerros da raça Canchim. **Pesquisa Agropecuária Brasileira**, v. 27, n.1, p.105-110, 1992.
- ALENCAR, M.M.; RUZZA, F.J.; PORTO, E.J.S. Desempenho produtivo de fêmeas das raças Canchim e Nelore. III. Produção de leite. **Revista Brasileira de Zootecnia**, v. 17, n. 4, p. 317-328, 1988.
- ALLEN, D.M. The relationship between variable selection and data augmentation and a method of prediction. **Technometrics**, v.16, p. 125-127, 1974.
- BROWN, M.A.; BROWN JR., A.H. Relationship of milk yield and quality to preweaning gain of calves from Angus, Brahman and reciprocal-cross cows on different forage systems. **Journal of Animal Science**, v. 80, p.2522-2527, 2002.
- CERDÓTES, L.; RESTLE, J.; ALVES FILHO, D.C.; NÖRNBERG, M.F.B.L.; NÖRNBERG, J.L.; HECK, I.; SILVEIRA, M.F. Produção e composição do leite de vacas de quatro grupos genéticos submetidas a dois manejos alimentares no período de lactação. **Revista Brasileira de Zootecnia**, v. 33, n. 3, p. 610-622, 2004.
- DRAPER, N.R.; SMITH, H. **Applied regression analysis**. 3. ed. New York: Wiley Series in Probability and Statistics, 1998. 705 p.
- FREETLY, H.C.; CUNDIFF, L.V. Reproductive performance, calf growth, and milk production of first-calf heifers sired by seven breeds and raised on different levels of nutrition. **Journal of Animal Science**, v. 76, p.1513-1522, 1998.
- FREUND, R.J.; LITTEL, R.C. **SAS system for regression**. 2<sup>nd</sup> ed. Cary, North Caroline: SAS Institute Incorporation, 1991.
- HOLLOWAY, J.W.; BUTTS, W.T.; McCURLEY, J.R.; BEAVER, E.E.; PEELER, H.L.; BACKUS, W.L. Breed x nutritional environment interactions for beef female weight and fatness, milk production and calf growth. **Journal of Animal Science**, v. 61, n. 6, p. 1354-1363, 1985.
- JOHNSON, C.R.; LALMAN, D.L.; BROWN, M.A.; APPEDDU, L.A.; BUCHANAN, D.S.; WETTEMANN, R.P. Influence of milk production potential on forage dry matter intake by multiparous and primiparous Brangus females. **Journal of Animal Science**, v. 81, p. 1837-1846, 2003.
- MARSHALL, D.A.; PARKER, W.R.; DINKEL, C.A. Factors affecting efficiency to weaning in Angus, Charolais and reciprocal cross cows. **Journal of Animal Science**, v. 43, p. 1176-1187, 1976.
- MELTON, A.A.; RIGGS, J.K.; NELSON, L.A.; CARTWRIGHT, T.C. Milk production, composition and calf gains of Angus, Charolais and Hereford cows. **Journal of Animal Science**, v. 26, p. 804-809, 1967.
- MONTGOMERY, D.C.; PECK, E.A. **Introduction to linear regression analysis**. USA: John Wiley & Sons, Inc./Wiley series in Probability and Mathematical Statistics, 1982. 504 p.
- MORENO, J.A. **Clima do Rio Grande do Sul**. Porto Alegre: Secretaria da Agricultura, 1961. 42 p.
- PACHECO, P.S.; RESTLE, J.; BRONDANI, I.L.; SILVA, J.H.S.; MIOTTO, F.R.C.; NEIVA, J.N.M. Número de coletas na predição de modelos para estimativa da produção média diária de leite em vacas Nelore. **Ciência Animal Brasileira**, v. 10, n. 1, p. 158-168, 2009.
- RESTLE, J.; PACHECO, P.S.; MOLETTA, J.L.; BRONDANI, I.L.; CERDÓTES, L. Grupo genético e nível nutricional pós-parto na produção e composição do leite de vacas de corte. **Revista Brasileira de Zootecnia**, v. 32, n.3, p.585-597, 2003.

RESTLE, J.; PACHECO, P.S.; PASCOAL, L.L.; PADUA, J.T.; MOLETTA, J.L.; FREITAS, A.K.; LEITE, D.T. Efeito da pastagem, da produção e composição do leite no desempenho de bezerros de diferentes grupos genéticos. **Revista Brasileira de Zootecnia**, v. 33, n. 3, p. 691-703, 2004.

RIBEIRO, E.L.A.; RESTLE, J. Desempenho de terneiros Charolês e Aberdeen Angus puros e seus mestiços com Nelore. **Pesquisa Agropecuária Brasileira**, v. 26, n. 8, p.1145-1151, 1991.

RIBEIRO, E.L.A.; RESTLE, J.; PIRES, C.C. Produção e composição do leite de vacas Charolês e Aberdeen Angus amamentando bezerros puros ou mestiços. **Pesquisa Agropecuária Brasileira**, v. 26, n. 8, p.1267-1273, 1991.

RIO GRANDE DO SUL. Secretaria da Agricultura. Departamento de Pesquisa. Instituto de Pesquisas Agronômicas. Observações meteorológicas do Estado do Rio Grande do Sul. **Boletim Técnico**, Porto Alegre, n. 3, 1979, 270 p.

SAS. STATISTICAL ANALYSIS SYSTEMS. **Sas Institute: User's Guide**. Version 6, Cary, NC: 1997. v. 2, 1052 p.

SOUZA, G.S. **Introdução aos modelos de regressão linear e não-linear**. Brasília: Embrapa-SPI/Embrapa-SEA, 1998. 505 p.

---

Protocolado em: 11 jul. 2008 Aceito em: 23 set. 2008.