

Identificação e modelagem da autocorrelação residual no ajuste do modelo de Wood às curvas de lactação de cabras

Identification and modeling of residual autocorrelation in the adjustments of Wood's model to lactation curves of goats

Ana Lúcia Puerro de Melo^{I*} Robledo de Almeida Torres^I José Ivo Ribeiro Júnior^{II}
Fabyano Fonseca e Silva^{II} Marcelo Teixeira Rodrigues^I Felipe Gomes da Silva^I Luiz Fernando Brito^I
Luiz Fernando Brito^I

RESUMO

Objetivou-se com este trabalho apresentar uma metodologia de identificação e modelagem da autocorrelação residual considerando ajustes individuais do modelo de Wood às lactações de cabras leiteiras e também avaliar a influência de tal modelagem na qualidade do ajuste. O modelo de Wood foi ajustado individualmente às lactações, considerando três estruturas residuais. Na primeira, assumiu-se independência dos erros (EI) para todas as lactações, na segunda, assumiu-se a estrutura de erros autoregressivos de primeira ordem (ARI) para todas as lactações e, na terceira, nomeada por EI-ARI, foi utilizada a estrutura de erros ARI somente para as lactações que apresentaram autocorrelação residual, segundo o teste de Durbin-Watson, e de EI para as demais. As três situações de ajuste foram comparadas pelos percentuais de convergência e pelas médias dos quadrados médios dos erros (QME) e dos coeficientes de determinação ajustados (R^2_{aj}). As médias dos QME e dos R^2_{aj} apresentaram valores semelhantes nas três situações de estrutura residual. No entanto, o modelo com estrutura EI-ARI apresentou maior convergência, o que consiste em uma vantagem, já que permite que um maior número de animais seja avaliado quanto à sua curva de lactação. Portanto, em função da maior convergência obtida, o ajuste do modelo de Wood com a estrutura EI-ARI consiste na opção mais indicada para grandes conjuntos de dados.

Palavras-chave: Durbin-Watson, erros independentes, resíduos autoregressivos.

ABSTRACT

The objective of this research was to present a methodology for identification and modeling of residual autocorrelation considering individual adjustments of the Wood's model to lactation dairy goats and evaluate the influence of

such modeling in the quality of adjustment. The Wood's model was adjusted individually for lactations in three different ways, the first have assumed independence of errors (IE) for all lactations, the second have assumed autoregressives first order errors (ARI) for all lactations and the third, named (IE-ARI), was used the ARI errors structure only for lactations that showed residual autocorrelation according to Durbin-Watson test, and the IE errors structure for the other lactations. The three ways of adjustment were compared by the percentage of convergence and the average of the mean square errors (MSE) and coefficients of determination adjusted (R^2_{adj}). The average of MSE and R^2_{aj} were very similar in the three cases of residual structure. However, the model with IE-ARI residual structure showed a higher rate of convergence, which is an advantage, as it allows a greater number of animals are evaluated for their lactation curve. Therefore, due to the increasing convergence obtained, the fit of the Wood's model with IE-ARI residual structure is the option most suitable for large data sets.

Key words: autoregressives errors, Durbin-Watson, independent errors.

INTRODUÇÃO

Apesar de a produção do leite de cabra no Brasil ser crescente, esse aumento não é necessariamente acompanhado da melhoria na produtividade dos animais. Uma forma eficiente de melhorar tal produtividade é por meio do estudo de suas curvas de lactação. A curva de lactação pode ser definida como a representação gráfica da produção de leite no decorrer da lactação, e seu estudo contribui

^IDepartamento de Zootecnia, Universidade Federal de Viçosa (UFV), 36570-000, Viçosa, MG, Brasil. E-mail: analupuerro@yahoo.com.br *Autor para correspondência.

^{II}Departamento de Informática, Setor de Estatística, UFV, Viçosa, MG, Brasil.

para o melhor entendimento do sistema de produção, auxiliando o produtor na previsão da produção de leite em determinado estágio de lactação e, também, na tomada de decisões quanto ao descarte ou manejo dos animais (COBUCI et al., 2001).

Em muitos estudos sobre a produção de leite, foram utilizados modelos estatísticos para descrever as curvas de lactação, sendo uma forma prática e consistente para a obtenção de informações sobre o pico de produção e a persistência da lactação (COBUCI et al., 2000; MCMANUS et al., 2003; MUNIZ et al., 2007; MENEZES et al., 2010).

Segundo GUIMARÃES et al. (2006), o modelo mais comumente utilizado para a estimativa de curvas de lactação é a função gama incompleta, proposta por WOOD (1967), que tem proporcionado bom ajuste das curvas de lactação. Além disso, modelos simples como o de WOOD (1967) têm a vantagem de apresentar uma interpretação biológica de seus parâmetros.

Em estudos envolvendo dados longitudinais como os de curvas de lactação e de crescimento, devido ao fato de um mesmo indivíduo ser repetidamente avaliado ao longo tempo, muitas vezes, a pressuposição de independência dos erros requerida pelo modelo de regressão adotado não é atendida, levando à ocorrência de autocorrelação residual.

Segundo MORETTIN & TOLOI (2006), se for detectada a presença de autocorrelação residual, ela deve ser modelada para que seu efeito seja eliminado e novos resíduos independentes, denominados de ruídos brancos, sejam obtidos. Uma alternativa para resolver esse problema é a modelagem da autocorrelação residual utilizando resíduos autoregressivos. Até o momento, não há referências sobre o efeito da autocorrelação residual no ajuste de modelos às curvas de lactação. No entanto, estudos com curvas de crescimento de bovinos vêm mostrando que os modelos com estrutura de erros autoregressivos apresentaram estimativas dos coeficientes mais próximas da realidade biológica e melhor qualidade de ajuste, sendo, portanto, mais eficientes para descrever a curva de crescimento do rebanho (MAZZINI et al., 2005; MENDES et al., 2008).

Assim, objetivou-se com este trabalho apresentar uma metodologia de identificação e modelagem da autocorrelação residual considerando ajustes individuais do modelo de Wood às lactações de cabras leiteiras. Além disso, avaliar a influência de tal modelagem na qualidade do ajuste.

MATERIAL E MÉTODOS

Os animais utilizados pertencem ao rebanho caprino do Departamento de Zootecnia da Universidade Federal de Viçosa (Viçosa-MG). Os controles foram registrados semanalmente, provenientes de ordenha mecânica, realizada duas vezes ao dia. Os animais foram mantidos em baias coletivas sob o sistema de estabulação livre.

Os dados foram coletados de 1997 a 2008. O banco de dados inicial foi composto por 84.377 registros de produção de leite no dia do controle (pldc), provenientes de 2.706 lactações, porém foram consideradas somente as lactações até a sexta ordem, com idades ao parto entre 11 e 91 meses, número de controles entre 19 e 51, duração entre 199 e 331 dias, de cabras com uma ou duas crias.

Após essas restrições, o modelo de Wood foi ajustado individualmente a 1.069 lactações de cabras das raças Alpina e Saanen e sem raça definida (SRD), de três formas diferentes, tendo em vista a presença de resíduos autocorrelacionados e sua posterior modelagem.

Na primeira forma, desconsideraram-se as possíveis autocorrelações, ou seja, foram assumidos erros independentes (EI) nos ajustes do modelo de Wood a todas as lactações. Na segunda, o modelo Wood com estrutura de erros autoregressivos de primeira ordem (AR1) foi ajustado a todas as lactações. Na terceira, nomeada de EI-AR1, foi ajustado o modelo com estrutura de erros AR1 somente para as lactações que apresentaram autocorrelação de primeira ordem significativa por meio do teste de Durbin-Watson, sendo que para as demais foi adotado o modelo com EI.

O teste de Durbin-Watson original (D) foi realizado por meio do PROC AUTOREG do SAS® (2002), sendo a significância da autocorrelação residual de primeira ordem constatada ao nível de 5% de probabilidade.

$$D = \frac{\sum_{t=2}^n [e_t - e_{(t-1)}]^2}{\sum_{t=1}^n e_t^2}, \text{ em que: } e_t \text{ é o resíduo no instante } t; e_{t-1}$$

é o resíduo no instante t-1; n = número de estágios de lactação.

Os modelos de Wood, assumindo EI e erros AR1 são apresentados, respectivamente, pelas expressões (1) e (2):

$y_t = at^b \exp^{-ct} + e_t$ (1), $y_t = at^b \exp^{-ct} + \phi_1 e_{t-1} + e_t$ (2), em que: y_t é a produção de leite no dia t; a, b e c são, respectivamente, os coeficientes associados com a produção média no início da lactação, a taxa média de

ascensão da produção até atingir a produção máxima e a taxa média de declínio da produção após atingir a produção máxima; t corresponde aos dias em lactação; $et-1$ é o resíduo no dia $t-1$; ϕ_1 é o coeficiente autoregressivo de primeira ordem e e é o resíduo no dia t , $e_t \sim \text{NID}(0, \sigma^2)$.

Para ajustar o modelo de Wood aos dados de produção de leite, foi utilizado o PROC MODEL do SAS® (2002), o qual possibilita estimar os coeficientes do modelo simultaneamente com o coeficiente autoregressivo pelo método dos quadrados mínimos, generalizado para modelos de regressão não linear via processo iterativo de Gauss-Newton. Para tanto, é necessário utilizar o comando %AR(y,1) quando se tem interesse na estimação do componente autoregressivo de primeira ordem. O método iterativo utilizado visa a minimizar a Soma do Quadrado do Erro, $\text{SQE}(\theta)$, que no modelo em questão é dada por:

$$\text{SQE}(\theta) = \sum_{i=1}^n [y_i - f(t, \hat{\theta})]^2, \text{ sendo } \hat{\theta} = [\hat{a}, \hat{b}, \hat{c}, \hat{\phi}_1].$$

Outros métodos tais como Yule-Walker e Máxima Verossimilhança (MORETTIN & TOLOI, 2006) também podem ser utilizados para tal fim.

Os valores iniciais dos coeficientes do modelo de Wood ($a = 1,5$; $b = 0,2$ e $c = 0,008$), necessários para a obtenção de suas estimativas mediante processo iterativo de Gauss-Newton, foram obtidos do trabalho de SILVA et al. (2005) com curvas de lactação de cabras Saanen.

A qualidade dos ajustes obtidos pelas três situações consideradas, EI, AR1 e EI-AR1 foram

comparadas pelos percentuais de convergência e pelos quadrados médios dos erros (QME) e coeficientes de determinação ajustados (R^2_{aj}), sendo que, para esses dois últimos, obtiveram-se as médias dos valores gerados nos ajustes individuais.

O percentual de convergência foi calculado pela relação entre o número de lactações para as quais o critério de convergência no ajuste foi alcançado em relação ao total de lactações para as quais se ajustou o modelo. No referido critério, também se considerou falha de convergência quando as estimativas dos coeficientes caracterizaram curvas atípicas, isto é quando os valores de a , b e c não atenderam às seguintes condições: $a > 0$, $0 < b < 1$ e $0 < c < 1$ e $0 < c < 1$.

Os coeficientes de determinação ajustados (R^2_{aj}) e os quadrados médios dos erros (QME) foram calculados, respectivamente, por meio das expressões

$$(3) \text{ e } (4): R^2_{aj} = \frac{(n-1)R^2 - p}{n-p-1} \quad (3), \quad \text{QME} = \frac{\text{SQE}}{\text{GLE}} \quad (4), \text{ em que:}$$

n é o número de observações utilizadas para ajustar a curva; p é o número de parâmetros do modelo incluindo o intercepto; R^2 é o coeficiente de determinação; SQE é a soma de quadrados do erro e GLE seu respectivo grau de liberdade.

RESULTADOS E DISCUSSÃO

A figura 1 apresenta as curvas estimadas para cada estrutura de erros considerada. Na confecção dessas curvas, os coeficientes representam as médias das estimativas dos coeficientes obtidos nos ajustes individuais, as quais se encontram na tabela 1.

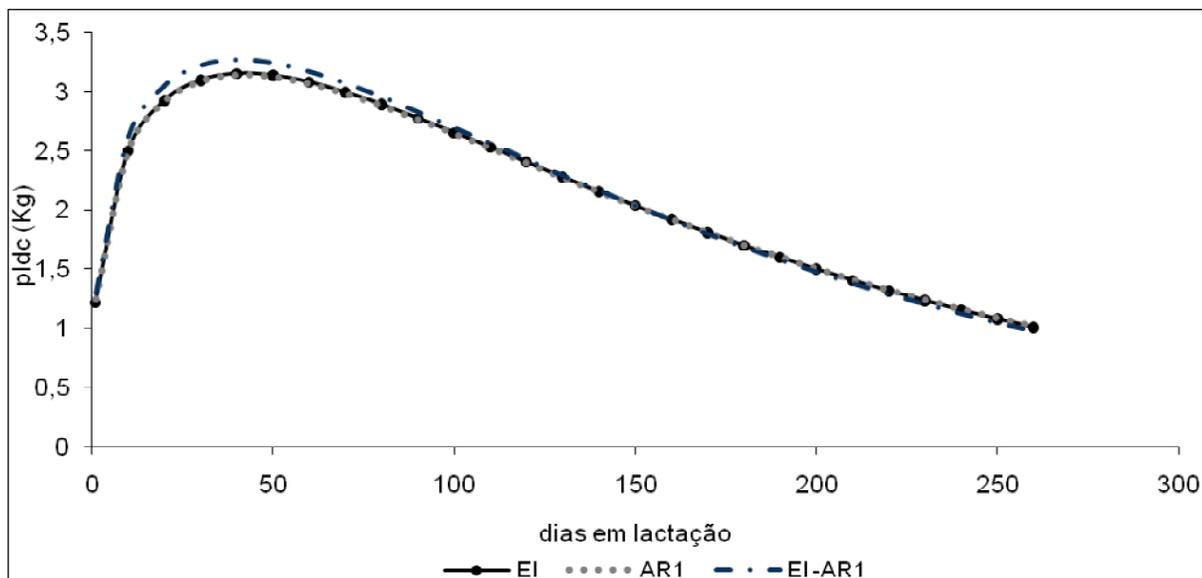


Figura 1 - Curvas de lactação estimadas pelo modelo de Wood nas três situações de estruturas residuais (EI, AR e EI-AR1), em que pldc é a produção de leite no dia do controle.

Tabela 1 - Médias, mínimos, máximos, desvios-padrão (DP) e coeficientes de variação (CV%) das estimativas dos coeficientes "a", "b" e "c" do modelo de Wood.

Estrutura	Coefficientes	Média	Mínimo	Máximo	DP	CV(%)
EI	a	1,2276	0,0418	4,1988	0,7762	63,23
	b	0,3435	0,0016	0,9958	0,2021	58,84
	c	0,0081	0,0001	0,0419	0,0038	46,68
AR1	a	1,2515	0,0427	4,337	0,7819	62,47
	b	0,3369	0,0051	0,9645	0,2029	60,22
	c	0,0080	0,0001	0,026	0,0036	44,57
	ϕ_1	0,1729	-0,5983	0,9852	0,2687	155,42
EI - AR1	a	1,3045	0,0418	4,0862	0,8287	63,53
	b	0,3406	0,0005	0,9626	0,2069	60,75
	c	0,0084	0,0006	0,0413	0,0039	46,64
	ϕ_1	0,1785	-0,5983	0,9642	0,2667	149,37

EI, AR1 e EI-AR1 correspondem, respectivamente, às estruturas de erros independentes, autoregressivos de primeira ordem para todas as lactações e autoregressivos de primeira ordem somente para as lactações nas quais se constatou autocorrelação significativa.

No presente estudo, as médias das estimativas dos coeficientes variaram de 1,2276 a 1,3045 para "a", de 0,3369 a 0,3435 para "b" e de 0,0080 a 0,0084 para "c" (Tabela 1). SILVA et al. (2005), em um estudo da curva de lactação de cabras Saanen por meio da análise Bayesiana do modelo de Wood, encontraram os valores de 0,9117 e 1,1255 para "a", de 0,1631 e 0,1006 para "b" e de 0,0095 e 0,0040 para "c", respectivamente, para primeira e segunda ordens de parto.

Na tabela 1, encontram-se, também, os valores das estimativas dos parâmetros autoregressivos de primeira ordem (f_1) para os casos em que foram utilizadas as estruturas residuais AR1 e IE-AR1. No caso do modelo com resíduos AR1, o parâmetro f_1 foi estimado para todas as lactações que convergiram (850). No caso do modelo EI-AR1, houve convergência para 945 lactações, sendo o parâmetro f_1 estimado para 446 lactações, as quais apresentaram autocorrelação significativa (96,41% de autocorrelações positivas), segundo o teste de Durbin-Watson.

Apesar das médias das estimativas dos coeficientes obtidos serem próximas àquelas relatadas na literatura (SILVA et al., 2005), observa-se que os coeficientes de variação (CVs) foram maiores que 30% (Tabela 1). Esses altos valores de CVs indicam uma grande variabilidade das estimativas dos coeficientes do modelo de Wood e, conseqüentemente, uma grande variabilidade na forma das curvas de lactação, a qual se deve a causas genéticas e/ou ambientais. Dessa forma, torna-se interessante um estudo futuro com o intuito de estimar parâmetros genéticos para os coeficientes do modelo de Wood, os quais poderiam indicar a proporção dessa variabilidade, devida a causas genéticas e conseqüentemente a viabilidade

de se alterar as curvas por meio de programas de seleção.

Nota-se também, na tabela 1, que os valores mínimos e máximos de algumas estimativas de "a", "b" e "c" foram muito próximos de zero e ainda que algumas outras de "b" foram muito próximas de um, ou seja, assumiram valores semelhantes àqueles utilizados como critério de não convergência. Este fato indica que para algumas lactações o modelo de Wood pode gerar curvas cuja forma se aproxima de um comportamento atípico, o qual deve ser melhor investigado com o intuito de descobrir suas causas, que podem ser devidas à inconsistência de dados, falta de habilidade do modelo ou influência de fatores genéticos e/ou ambientais.

Na tabela 2, são apresentados os resultados dos critérios de comparação dos ajustes provenientes das três diferentes estruturas residuais consideradas. Nela, se observa que os modelos com estrutura de erros independentes (EI) e autorregressivos de primeira ordem (AR1) apresentaram percentuais de convergência bastante semelhantes, próximos a 80%. Para a terceira forma de ajuste (EI-AR1), verificou-se uma maior porcentagem de convergência (88,40%), indicando que a utilização da estrutura AR1 proporcionou a convergência do ajuste do modelo de Wood para algumas lactações, nas quais esta condição não havia sido verificada com a estrutura EI.

Em relação aos critérios QME e R^2_{aj} , embora as diferenças tenham sido insignificantes, observou-se que, à medida que se incluiu a estrutura de erros autoregressivos, o QME diminuiu e o R^2_{aj} aumentou. Esses resultados concordam com aqueles apresentados por MAZZINI et al. (2005) e MENDES et al. (2008), os quais também relataram essa relação em estudos de curvas de crescimento de bovinos de corte.

Tabela 2 - Percentual de convergência e as médias e coeficientes de variação (CV) dos quadrados médios dos erros (QME) e dos coeficientes de determinação ajustado (R^2_{aj}), obtidos nos ajustes individuais do modelo de Wood.

Estrutura residual	Convergência (%)	-----QME-----		----- R^2_{aj} -----	
		-----Médias-----	-----CV (%)-----	-----Médias-----	-----CV(%)-----
EI	79,05	0,1794	77,27	0,6370	34,82
AR1	79,51	0,1679	73,85	0,6553	32,22
EI - AR1	88,40	0,1664	73,34	0,6593	31,27

EI, AR1 e EI-AR1 correspondem, respectivamente, às estruturas de erros independentes, autoregressivos de primeira ordem para todas as lactações e autoregressivos de primeira ordem somente para as lactações nas quais se constatou autocorrelação significativa.

GUIMARÃES et al. (2006) encontraram valores inferiores para R^2_{aj} ao comparar o ajuste de 12 modelos às curvas de lactação de caprinos, sendo 0,389 o valor referente ao modelo de Wood. De acordo com GUIMARÃES et al. (2006), esses resultados podem ser explicados pela utilização de dados individuais na estimativa dos parâmetros, visto que alguns pesquisadores utilizam a média de produção por período, o que tende a eliminar a variação dos dados.

Apesar de os resultados da comparação dos ajustes nas diferentes estruturas residuais terem sido semelhantes, o fato de a estrutura EI-AR1 ter apresentado uma maior porcentagem de convergência consiste em uma vantagem, já que permite que um maior número de animais seja avaliado quanto a sua curva de lactação.

CONCLUSÃO

O ajuste do modelo de Wood com erros autoregressivos de primeira ordem somente para as lactações que apresentaram erros autocorrelacionados significativamente consiste na opção mais indicada para grandes conjuntos de dados em função da maior convergência obtida.

REFERÊNCIAS

COBUCI, J.A. et al. Curva de lactação na raça Guzerá. **Revista Brasileira de Zootecnia**, v.29, n.5, p.1332-1339, 2000. Disponível em: <<http://www.scielo.br/pdf/rbz/v29n5/5654.pdf>>. Acesso em: 20 jul. 2011. doi: 10.1590/S1516-35982000000500011.

COBUCI, J.A. et al. Aspectos genéticos e ambientais da curva de lactação de vacas Guzerá. **Revista Brasileira de Zootecnia**, v.30, n.4, p.1204-1211, 2001. Disponível em: <<http://www.scielo.br/pdf/rbz/v30n4/6026.pdf>>. Acesso em: 20 jul. 2011. doi: 10.1590/S1516-35982001000500011.

GUIMARÃES, V.P. et al. Utilização de funções matemáticas no estudo da curva de lactação em caprinos. **Revista Brasileira de Zootecnia**, v.35, n.2, p.535-543, 2006. Disponível em: <<http://www.scielo.br/pdf/rbz/v35n2/a28v35n2.pdf>>. Acesso em: 20 jul. 2011. doi: 10.1590/S1516-35982006000200028.

MAZZINI, A.R.A. et al. Curva de crescimento de novilhos Hereford: heterocedasticidade e resíduos autoregressivos. **Ciência Rural**, v.35, n.2, p.422-427, 2005. Disponível em: <<http://www.scielo.br/pdf/cr/v35n2/a28v35n2.pdf>>. Acesso em: 20 jul. 2011. doi: 10.1590/S0103-84782005000200028.

MCMANUS, C. et al. Fatores que influenciam os parâmetros das curvas de lactação em cabras no distrito federal. **Revista Brasileira de Zootecnia**, v.32, n.6, p.1614-1623, 2003. Disponível em: <<http://www.scielo.br/pdf/rbz/v32n6s1/19680.pdf>>. Acesso em 20 jul. 2011. doi: 10.1590/S1516-35982003000700010.

MENDES, P.N. et al. Modelo logístico difásico no estudo do crescimento de fêmeas da raça Hereford. **Ciência Rural**, v.38, n.7, p.1984-1990, 2008. Disponível em: <<http://www.scielo.br/pdf/cr/v38n7/a29v38n7.pdf>>. Acesso em: 20 jul. 2011. doi: 10.1590/S0103-84782008000700029.

MENEZES, G.R.O. et al. Avaliação de medidas da persistência da lactação de cabras da raça Saanen sob modelo de regressão aleatória. **Revista Brasileira de Zootecnia**, v.39, n.8, p.1691-1698, 2010. Disponível em: <<http://www.scielo.br/pdf/rbz/v39n8/v39n8a10.pdf>>. Acesso em: 20 jul. 2011. doi:10.1590/s1516-35982010000800010.

MORETTIN, P.A.; TOLOI, C.M. **Análise de séries temporais**. 2.ed. São Paulo: Edgard Blucher, 2006. 538p.

MUNIZ, J.A. et al. Evaluation of lactation curve of low-yielding gir cows: a Bayesian approach. **Acta Scientiarum Animal Science**, v.29, n.1, p.79-83, 2007. Disponível em: <<http://periodicos.uem.br/ojs/index.php/ActaSciAnimSci/article/viewFile/262/162>>. Acesso em: 20 jul. 2011. doi: 10.4025/actascianimsci.v29i1.262.

SAS INSTITUTE. **SAS/ETS® User's guide**. Volumes 1 and 2. Cary, NC, 2002. 1161p.

SILVA, F.F. et al. Abordagem Bayesiana da curva de lactação de cabras Saanen de primeira e segunda ordem de parto. **Pesquisa Agropecuária Brasileira**, v.40, n.1, p.27-33, 2005. Disponível em: <<http://www.scielo.br/pdf/pab/v40n1/23238.pdf>>. Acesso em: 20 jul. 2011. doi: 10.1590/S0100-204X2005000100004.

WOOD, P.D.P. Algebraic model of the lactation curve in cattle. **Nature**, v.216, p.164-165, 1967. Disponível em: <<http://www.nature.com/nature/journal/v216/n5111/pdf/216164a0.pdf>>. Acesso em: 20 jul. 2011. doi: 10.1038/216164a0.