

6-6-2005

# Genetic Evaluation of Dairy Cattle with Test-Day Models with Autoregressive Covariance Structures and with a 305-d Model

R. M. Sawalha

*University of Nebraska - Lincoln*

Jeffrey F. Keown

*University of Nebraska - Lincoln, jkeown1@unl.edu*

Stephen D. Kachman

*University of Nebraska-Lincoln, steve.kachman@unl.edu*

L. Dale Van Vleck

*University of Nebraska - Lincoln, dvan-vleck1@unl.edu*

Follow this and additional works at: <http://digitalcommons.unl.edu/animalscifacpub>

 Part of the [Animal Sciences Commons](#)

---

Sawalha, R. M.; Keown, Jeffrey F.; Kachman, Stephen D.; and Van Vleck, L. Dale, "Genetic Evaluation of Dairy Cattle with Test-Day Models with Autoregressive Covariance Structures and with a 305-d Model" (2005). *Faculty Papers and Publications in Animal Science*. Paper 153.

<http://digitalcommons.unl.edu/animalscifacpub/153>

This Article is brought to you for free and open access by the Animal Science Department at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Faculty Papers and Publications in Animal Science by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

## Genetic Evaluation of Dairy Cattle with Test-Day Models with Autoregressive Covariance Structures and with a 305-d Model

R. M. Sawalha,<sup>1</sup> J. F. Keown,<sup>1</sup> S. D. Kachman,<sup>2</sup> and L. D. Van Vleck<sup>1,3</sup>

<sup>1</sup>Department of Animal Science, and

<sup>2</sup>Department of Statistics, University of Nebraska, Lincoln 68583-0908

<sup>3</sup>Roman L. Hruska U.S. Meat Animal Research Center, Agricultural Research Service, USDA, Lincoln, NE 68583-0908

### ABSTRACT

This study compared genetic evaluations from 3 test-day (TD) models with different assumptions about the environmental covariance structure for TD records and genetic evaluations from 305-d lactation records for dairy cows. Estimates of genetic values of 12,071 first-lactation Holstein cows were obtained with the 3 TD models using 106,472 TD records. The compound symmetry (CS) model was a simple test-day repeatability animal model with compound symmetry covariance structure for TD environmental effects. The AR<sub>s</sub> and AR<sub>e</sub> models also used TD records but with a first-order autoregressive covariance structure among short-term environmental effects or residuals, respectively. Estimates of genetic values with the TD models were also compared with those from a model using 305-d lactation records. Animals were genetically evaluated for milk, fat, and protein yields, and somatic cell score (SCS). The largest average estimates of accuracy of predicted breeding values were obtained with the AR<sub>s</sub> model and the smallest were with the 305-d model. The 305-d model resulted in smaller estimates of correlations between average predicted breeding values of the parents and lactation records of their daughters for milk and protein yields and SCS than did the CS and AR<sub>e</sub> models. Predicted breeding values with the 3 TD models were highly correlated (0.98 to 1.00). Predicted breeding values with 305-d lactation records were moderately correlated with those with TD models (0.71 to 0.87 for sires and 0.80 to 0.87 for cows). More genetic improvement can be achieved by using TD models to select for animals for higher milk, fat, and protein yields, and lower SCS than by using models with 305-d lactation records.

**(Key words:** autoregressive covariance, genetic evaluation, milk yield, test day)

**Abbreviation key:** AR(1) = first-order autoregressive, CS = compound symmetry, PBV = predicted breeding values, TD = test day.

### INTRODUCTION

The current genetic evaluation system of dairy cattle in the United States utilizes 305-d lactation records. The length of an actual lactation may be considerably less than the arbitrarily defined length of 305 d, which requires extension and adjustment factors to generate the appropriate records. An alternative approach is the analysis of test-day (TD) records, which is more computationally demanding and may require estimation of many parameters (Jensen, 2001). The simplest, but perhaps least efficient, TD model, however, assumes that monthly TD records are separate traits, each of which could be analyzed with single-trait models (Swalve, 1995). Alternatively, monthly TD records can be viewed as different but correlated traits to be analyzed with multivariate analyses, which would require the estimation of a large number of parameters if no particular covariance structure is assumed (Trus and Buttazzoni, 1990; Meyer, 1991). For estimation of parameters, the use of a series of 2-trait models could be considered to limit the computational requirements per analysis (Gadini, 1997). Another way to reduce the computational burden would be to use canonical decomposition of covariance matrices, which reduces the number of parameters to be estimated compared with an unstructured multivariate approach (Ducrocq and Besbes, 1993; Wiggans and Goddard, 1997).

Test-day records from a cow may also be viewed as repeated measures of a single trait within a lactation. With the simple repeatability model, genetic and environmental correlations and variances are assumed constant among all TD records, which may not be a realistic assumption (Henderson, 1984). The first-order autoregressive covariance structure [AR(1)] allows for non-constant covariances among repeated records but requires only one more parameter than the simple repeatability model, which assumes a compound symmetry (CS) covariance structure.

Received December 14, 2004.

Accepted June 6, 2005.

Corresponding author: Rami M. Sawalha; e-mail: rami.sawalha@sac.ac.uk.

Quaas (1984) suggested the use of an AR(1) process to model the residual covariance structure when animals have repeated measures. An AR(1) covariance structure has been used to model covariances among contemporary groups (dates of calving within fixed effect of herd; Wade et al., 1990, 1993). Carneiro et al. (2002) tested the power of the AR(1) model to retrieve variance components from simulated multiple-lactation TD records. They compared the fit of the AR(1) model for datasets that had or did not have an AR(1) (co)variance structure for environmental effects. They concluded that the analysis with an AR(1) covariance structure was effective in detecting the presence or absence of autocorrelated effects in the simulated data.

Fixed regression models assume the shape of the lactation curve to be a fixed effect that can be modeled with multiple regression with covariates with different complexity, as in Ptak and Schaeffer (1993), or mathematical functions based on the biological nature of milk secretion, as in Wood (1967). Random regression models, although considered extensions of the fixed regression models (Jensen, 2001), may model the genetic and environmental components of the shape of the lactation curve better than some fixed regression models. However, random regression models often perform poorly in estimating dispersion parameters at early and late stages of lactation. In an early study of applying random regression models, Kettunen et al. (1998) reported negative genetic correlations between early and late TD records within lactation. Another drawback of polynomial random regression models (e.g., Legendre polynomials) is the large number of dispersion parameters to be estimated [ $0.5q(q+1)$  for each random effect in the model, where  $q$  is the order of the Legendre polynomial]. The fit of random regression models is usually improved by increasing of the order of the polynomial, but requires a large number of parameters to be estimated. First-order autoregression models, on the other hand, require estimation of only one additional parameter (the correlation coefficient between pairs of sequential TD) besides the variance component for the random effect. In a study comparing autoregression and random regression models for prediction of missing TD records, Meuwissen and Pool (2001) found similar predictions with both types of models. However, the number of estimated dispersion parameters and prediction equations was 4 times more and 2 times less with the random regression model than with the autoregression models, respectively. An AR(1) covariance structure among the environmental effects of TD records, however, has not been evaluated, nor have genetic evaluations with an AR(1) model been compared with those from a 305-d model or with those from simple TD repeatability models.

The goal of this study was to compare TD models with CS and first-order autoregressive covariance structures and the traditional 305-d model for genetic evaluations of first-lactation Holstein cows. Two different autoregressive covariance structures were compared: one model assumed an AR(1) covariance structure among TD short-term environmental effects and the other model assumed an AR(1) covariance structure among residual effects of TD records. The models with the autoregressive structure were chosen to model the environmental covariance structure among TD records within lactation with fewer parameters to be estimated compared with multivariate analyses. Traits studied were milk, fat, and protein yields, and SCS.

## MATERIALS AND METHODS

### Data and Editing

Test-day and 305-d lactation records for first lactations of US Holstein cows were obtained from Dairy Records Management Systems (DRMS) in Raleigh, NC. Test-day intervals were approximately 30 d from 6 to 305 DIM. Records were assigned to TD (1 through 10) based on the DIM at the time of recording rather than the ordinal sequence of the records. Only lactations with at least 5 TD records were included in the final data set. Records with yields and SCS greater than 3 SD from means of the records were deleted to eliminate possible outliers. To ensure data connectedness and to avoid contemporary groups with single records, each sire was required to have at least 10 daughters, and all herds were required to have at least 15 cows in the data set. The final data set included first-lactation records of 12,071 Holstein cows (with 106,472 TD milk records measured on 16,637 different calendar days within herds) calving from 1996 through 2001. A total of 410 sires were represented with daughters distributed among 214 herds.

### Models and Methodology

Test-day models with different definitions of environmental covariance structures were chosen to try to account for correlations among environmental effects in a computationally simple way and to compare predictions of breeding values with those from a 305-d model similar to that currently used in the United States. Predicted breeding values (PBV) were obtained with 4 single-trait animal models. The first 3 models used TD records with different assumptions about the covariance structure for environmental or residual effects. The TD models included additive genetic, permanent or short-term environmental, and residual effects as random effects. The CS model assumed a CS covariance

structure for environmental effects. Two TD models included a first-order autoregressive covariance structure, fitted for short-term environmental ( $AR_s$ ) or residual effects ( $AR_e$ ). Another model that attempted to partition the nongenetic components into permanent environmental, short-term environmental with autoregressive covariance structure, and residual effects was tested but excluded due to lack of fit to the data.

The equations for the linear mixed model in matrix notation for the TD models were:

$$y = X\beta + Z_1 a + Z_2 c + e \text{ for CS and } AR_e \text{ models, and}$$

$$y = X\beta + Z_1 a + Z_3 s + e \text{ for the } AR_s \text{ model}$$

where  $y$  is the vector of TD records for a trait,  $\beta$  is the vector of fixed effects,  $a$  is the vector of random animal additive genetic effects,  $c$  is the vector of random cow permanent environmental effects,  $s$  is the vector of random short-term environmental effects,  $e$  is the vector of random TD residual effects, and  $X$ ,  $Z_1$ ,  $Z_2$  and  $Z_3$  are incidence matrices relating TD observations to fixed, random animal additive genetic, random permanent environmental, and short-term environmental effects, respectively.

The vector of first moments for TD models was assumed to be  $E[y] = X\beta$  and the second moments about the means were assumed to be:

$$\text{Var} \begin{bmatrix} a \\ c \\ e \end{bmatrix} = \begin{bmatrix} A_N \sigma_a^2 & 0 & 0 \\ 0 & I_n \sigma_c^2 & 0 \\ 0 & 0 & I_n \otimes I_t \sigma_e^2 \end{bmatrix} \text{ for the CS model,}$$

$$\text{Var} \begin{bmatrix} a \\ s \\ e \end{bmatrix} = \begin{bmatrix} A_N \sigma_a^2 & 0 & 0 \\ 0 & I_n \otimes AR_s \sigma_s^2 & 0 \\ 0 & 0 & I_n \otimes I_t \sigma_e^2 \end{bmatrix}$$

for the  $AR_s$  model, and

$$\text{Var} \begin{bmatrix} a \\ c \\ e \end{bmatrix} = \begin{bmatrix} A_N \sigma_a^2 & 0 & 0 \\ 0 & I_n \sigma_c^2 & 0 \\ 0 & 0 & I_n \otimes AR_e \sigma_e^2 \end{bmatrix}$$

for the  $AR_e$  model,

where  $\sigma_a^2$ ,  $\sigma_c^2$ ,  $\sigma_s^2$ , and  $\sigma_e^2$  are variances of additive genetic, permanent environmental, short-term environmental, and residual effects, respectively,  $A_N$  is the numerator relationship matrix of order  $N$  (the number of animals in the pedigree file),  $I_n$  is an identity matrix of order  $n$

(the number of cows with records),  $I_t$  is an identity matrix with order  $t$  (the number of TD records of a cow),  $AR_s$  and  $AR_e$  are first-order autoregressive correlation matrices of order  $t$  among TD short-term environmental and residual effects, respectively, and  $\otimes$  is the direct product operator.

The 305-d model was also an animal model, but with 305-d lactation records instead of the TD records and with only a single record per animal. The linear mixed model equation for the 305-d model was:

$$y = X\beta + Za + e,$$

where  $y$  is the vector of 305-d records for a trait,  $\beta$  is the vector of fixed effects,  $a$  is the vector of random animal additive genetic effects,  $e$  is the vector of random residual effects, and  $X$  and  $Z$  are incidence matrices relating observations to fixed and random animal additive genetic effects, respectively. The assumption regarding the first moment for the 305-d model was that  $E[y] = X\beta$ . The random additive genetic and residual effects were assumed to be independent with respect to each other and to have a covariance matrix of:

$$\text{Var} \begin{bmatrix} a \\ e \end{bmatrix} = \begin{bmatrix} A_N \sigma_a^2 & 0 \\ 0 & I_n \sigma_e^2 \end{bmatrix},$$

where  $A_N$  is the numerator relationship matrix of order  $N$  (the number of animals in the pedigree),  $I_n$  is an identity matrix of order  $n$  (the number of cows with records), and  $\sigma_a^2$  and  $\sigma_e^2$  are the variances of additive genetic and residual effects, respectively.

The 4 models included age at calving and bST treatment as fixed effects. The 3 TD models also included fixed effects of herd test date (HTD), whereas the 305-d model included fixed effects of herd-year-season of calving (HYS). To adjust for the nonlinear shape of the lactation curve, DIM (DIM/30.5) for each TD record was included as a quadratic polynomial for the TD models. Data were analyzed using the ASREML program, release 1.0 (Gilmour et al., 2002).

Average estimates of accuracy of PBV with TD and 305-d models were obtained for milk, fat, and protein yields and SCS using estimates of prediction error variance (PEV) of PBV and additive genetic variance ( $\sigma_a^2$ ) (accuracy =  $\sqrt{1 - \text{PEV}/\sigma_a^2}$ ). The prediction error variances were calculated from the diagonal elements of the inverse of the mixed model coefficient matrix corresponding to the PBV. Estimates of variance components and autoregression correlations with the 4 models using the same data set were presented in a previous study (Sawalha et al., 2005). The different models were also



**Table 1.** Average estimates of accuracy and prediction error variance (PEV) of predicted breeding values of milk, fat, and protein yields and SCS of sires and cows.<sup>1,2</sup>

Trait	Model <sup>3</sup>	Sires		Cows	
		Accuracy	PEV	Accuracy	PEV
Milk	CS	0.65	1.24	0.51	1.63
	AR <sub>s</sub>	0.67	1.19	0.53	1.60
	AR <sub>e</sub>	0.63	1.30	0.50	1.67
	305-d	0.60	1.40	0.47	1.73
Fat	CS	0.65	0.23	0.52	0.30
	AR <sub>s</sub>	0.67	0.22	0.54	0.29
	AR <sub>e</sub>	0.64	0.24	0.51	0.31
	305-d	0.60	0.26	0.48	0.32
Protein	CS	0.63	0.12	0.50	0.15
	AR <sub>s</sub>	0.65	0.11	0.52	0.15
	AR <sub>e</sub>	0.61	0.12	0.49	0.16
	305-d	0.56	0.14	0.45	0.16
SCS	CS	0.54	0.12	0.43	0.14
	AR <sub>s</sub>	0.56	0.12	0.44	0.14
	AR <sub>e</sub>	0.53	0.12	0.42	0.15
	305-d	0.49	0.13	0.39	0.15

<sup>1</sup>Estimates of accuracy were obtained as:  $\text{accuracy} = \sqrt{1 - \text{PEV}/\sigma_a^2}$ , where PEV is prediction error variance of predicted breeding values and  $\sigma_a^2$  is the additive genetic variance.

<sup>2</sup>PEV are presented on daily basis for all models except for fat and protein yields where they were multiplied by 100.

<sup>3</sup>The CS model is a simple repeatability model with a compound symmetry covariance structure for test-day (TD) environmental effects; AR<sub>s</sub> model assumes a first-order autoregressive covariance structure for TD short-term environmental effects; AR<sub>e</sub> model assumes a first-order autoregressive covariance structure for TD residual effects; and 305-d model is an animal model using 305-d lactation records.

compared for their ability to predict lactation records of progeny based on PBV of the parents with progeny records removed from the data used to calculate PBV of the parents. Correlations were calculated between the average PBV of the parents and the lactation records of their daughters with different models. Correlations were also calculated among PBV with the different models for both cows and sires.

Approximate average standard errors of prediction of average PBV by year of birth with AR<sub>e</sub> model were

obtained as:  $\sqrt{\sum_{i=1}^n \text{PEV}/n^2}$  where n is the number of

evaluated animals per year. The PBV and their standard errors for yield traits were converted to the 305-d basis unless otherwise specified.

## RESULTS AND DISCUSSION

Average estimates of accuracy and prediction error variances of PBV with all models for milk, fat, and protein yields, and SCS are presented in Table 1. Among TD models, the AR<sub>s</sub> model resulted in the largest average estimates of accuracy and, in most cases, the smallest estimates of prediction error variance. The CS model resulted in greater average estimates of accuracy than the AR<sub>e</sub> model. The 305-d model

resulted in less reliable PBV, as the average estimates of accuracy of PBV were greater with any of the TD models than with the 305-d model for all traits. These results suggest that TD models account better for the variations in milk, fat, protein, and SCS records than the 305-d model. Consequently, the use of TD models for genetic evaluation may increase the rate of genetic gain by reducing the chance of selecting genetically poor sires and cows. As expected, sires had larger averages of accuracy and smaller averages of prediction error variance than did cows for all traits and with all models.

Prediction of future records is often the most important application of genetic evaluation. "Most applications of genetic evaluation are essentially problems in prediction of future records" (Henderson, 1984). The 4 models were compared by how well they predicted lactation records of future progeny. Table 2 shows correlations between the average PBV of the parents and the lactation records of their daughters with the CS, AR<sub>s</sub>, and AR<sub>e</sub> models and the 305-d model. With the 305-d model, correlations between the averages PBV of the parents and actual lactation records of their daughters for milk and protein yields and SCS were only slightly smaller compared with CS and AR<sub>e</sub> models, which may indicate somewhat less reliable predictions of records of future progeny using 305-d model for some milk traits. The differences in correlations with differ-

**Table 2.** Correlations between average predicted breeding values of the parents<sup>1</sup> and actual first-lactation records of their progeny with test-day (TD) and 305-d models for milk, fat, and protein yields, and SCS.

Model <sup>2</sup>	Milk	Fat	Protein	SCS
CS model	0.13	0.14	0.15	0.11
AR <sub>s</sub> model	0.12	0.14	0.13	0.12
AR <sub>e</sub> model	0.13	0.14	0.14	0.11
305-d model	0.12	0.14	0.13	0.09

<sup>1</sup>Breeding values of parents were predicted after removing the records of their daughters that had actual lactation records.

<sup>2</sup>CS, AR<sub>s</sub>, and AR<sub>e</sub> models are TD models; CS model is a repeatability model with a compound symmetry covariance structure for TD environmental effects; AR<sub>s</sub> model assumes a first-order autoregressive covariance structure for TD short-term environmental effects; AR<sub>e</sub> model assumes a first-order autoregressive covariance structure for TD residual effects; and 305-d model is an animal model using 305-d lactation records.

ent TD models were not large and did not follow a definite trend for the 4 traits in this study. Correlations between records of progeny and the average PBV of the parents tended to be smaller for SCS than for yield traits.

Strabel and Szwaczkowski (1999) concluded the same when comparing the correlation between the records of daughters and the index of their parents obtained with TD models with different definitions of contemporary groups and a 305-d lactation model. However, the correlations they reported (0.25 to 0.31) were twice as large as those obtained in the current study (0.12 to 0.13). The difference in magnitudes of the correlations between the 2 studies may be due to differences in the datasets and the connectedness between records by the pedigree files. Differences in models between the 2 studies and especially accounting for fixed effects and the shape of the lactation curve may also have had an effect on correlations between records of daughters and the average PBV of their parents.

Correlations between PBV with different models were estimated for milk, fat, and protein yields, and SCS. Correlations for milk yield are in Table 3 and for

**Table 3.** Correlations among predicted breeding values with test-day (TD) models<sup>1</sup> (CS, AR<sub>s</sub>, and AR<sub>e</sub>) and with 305-d model for sires (above diagonal) and cows (below diagonal) for milk yield.

	CS model	AR <sub>s</sub> model	AR <sub>e</sub> model	305-d model
CS model		0.98	0.99	0.81
AR <sub>s</sub> model	0.98		1.00	0.86
AR <sub>e</sub> model	0.99	0.99		0.85
305-d model	0.83	0.86	0.86	

<sup>1</sup>CS model is a repeatability model with a compound symmetry covariance structure for TD environmental effects; AR<sub>s</sub> model assumes a first-order autoregressive covariance structure for TD short-term environmental effects; AR<sub>e</sub> model assumes a first-order autoregressive covariance structure for TD residual effects; and 305-d model is an animal model using 305-d lactation records.

**Table 4.** Correlations among predicted breeding values with test-day (TD) models<sup>1</sup> (CS, AR<sub>s</sub>, and AR<sub>e</sub> models) and with 305-d model for sires (above diagonal) and cows (below diagonal) for SCS.

	CS model	AR <sub>s</sub> model	AR <sub>e</sub> model	305-d model
CS model		0.99	1.00	0.87
AR <sub>s</sub> model	0.99		1.00	0.87
AR <sub>e</sub> model	1.00	1.00		0.87
305-d model	0.86	0.86	0.87	

<sup>1</sup>CS model is a repeatability model with a compound symmetry covariance structure for TD environmental effects; AR<sub>s</sub> model assumes a first-order autoregressive covariance structure for TD short-term environmental effects; AR<sub>e</sub> model assumes a first-order autoregressive covariance structure for TD residual effects; and 305-d model is an animal model using 305-d lactation records.

SCS are in Table 4. Correlations for fat and protein yields were similar to those for milk yield and will be discussed but not shown in tables. Predicted breeding values with different TD models were highly correlated for both sires and cows with correlations in the range of 0.98 to 1.00 for the 4 traits. This result may be partially due to TD models sharing identical fixed effects and only differing in one random component. Carvalheira et al. (1998) reported slightly smaller correlations between PBV for TD milk yields compared with this study with models with or without the AR(1) covariance structure among environmental effects within lactations.

Predicted breeding values with the 305-d model were only moderately correlated with those of TD models for milk, fat, and protein yields, and SCS (0.71 to 0.87 for sires and 0.80 to 0.87 for cows). Among the TD models, PBV with the AR<sub>s</sub> model had slightly larger correlations with PBV with 305-d model for all traits for sires and for all traits except SCS for cows. The moderate correlations between evaluations with TD and 305-d models show a possibility of substantial change in the rank of elite sires and cows when the 305-d model is substituted with a TD model.

Ptak and Schaeffer (1993) reported a similar correlation (0.88) between PBV for sires for milk yield with a 305-d lactation model and with a TD model with CS. Swalve (1995) reported slightly greater correlations (0.89 to 0.92) between PBV with a 305-d model and PBV with TD model assuming CS for milk, fat, and protein yields for sires and cows. Reents et al. (1995) reported similar correlations (0.75 to 0.86) between PBV for SCS using TD and 305-d lactation models for both sires and cows.

Tables 5 and 6 present changes in ranks of top sires and potential elite bull-dams for milk yield and SCS, respectively, for the 4 different models. Similar results were found for fat and protein yields. At least 86%, but in most cases more than 90%, of the 100 sires with the

**Table 5.** Numbers of elite sires (above diagonal) and cows (below diagonal) with the largest 100 predicted breeding values in common between test-day (TD) and 305-d models<sup>1</sup> for milk yield.

	CS model	AR <sub>s</sub> model	AR <sub>e</sub> model	305-d model
CS model		86	91	68
AR <sub>s</sub> model	81		94	72
AR <sub>e</sub> model	89	88		71
305-d model	45	53	51	

<sup>1</sup>CS, AR<sub>s</sub>, and AR<sub>e</sub> models are TD models; CS model is a repeatability model with a compound symmetry covariance structure for TD environmental effects; AR<sub>s</sub> model assumes a first-order autoregressive covariance structure for TD short-term environmental effects and AR<sub>e</sub> model assumes a first-order autoregressive covariance structure for TD residual effects; and 305-d model is an animal model using 305-d lactation records.

highest PBV for milk, fat, and protein yields and with the lowest PBV for SCS would have been selected with all TD models. These results are consistent with the large correlations found among PBV with different TD models for sires. Nevertheless, some changes in ranking occurred from one evaluation model to another. Fewer elite cows were ranked the same with different TD models (75 to 92%) than sires.

The ranking of top 100 sires and cows was quite different between TD and 305-d models. Fewer cows than sires were in common among the top 100 animals from evaluations with the 305-d model and one of the TD models. Similar results were reported by Reents et al. (1995) and Swalve (1995). The reduced number of similarly ranked elite cows with the different models is due to differences in accuracy of PBV for sires and cows. The larger number of available cows than sires to select from may have also contributed to the differences in numbers ranked the same for sires and cows.

The correlations between PBV for sires with different models were not consistently different from those for cows, even though average estimates of accuracy of PBV

**Table 6.** Numbers of elite sires (above diagonal) and cows (below diagonal) with the smallest 100 predicted breeding values in common between test-day (TD) and 305-d models<sup>1</sup> for SCS.

	CS model	AR <sub>s</sub> model	AR <sub>e</sub> model	305-d model
CS model		94	96	74
AR <sub>s</sub> model	75		98	78
AR <sub>e</sub> model	90	81		77
305-d model	29	27	31	

<sup>1</sup>CS, AR<sub>s</sub>, and AR<sub>e</sub> models are TD models; CS model is a repeatability model with a compound symmetry covariance structure for TD environmental effects; AR<sub>s</sub> model assumes a first-order autoregressive covariance structure for TD short-term environmental effects; AR<sub>e</sub> model assumes a first-order autoregressive covariance structure for TD residual effects; and 305-d model is an animal model using 305-d lactation records.

**Table 7.** Averages of predicted breeding values for yield traits<sup>1</sup> (kg/lactation) and SCS predicted with the AR<sub>e</sub><sup>2</sup> model for the top 100 sires ranked with the 4 different models.

Model	Milk	Fat	Protein	SCS
CS model	348	15.9	10.7	-0.28
AR <sub>s</sub> model	350	15.8	10.7	-0.28
AR <sub>e</sub> model	352	15.9	10.7	-0.28
305-d model	303	13.7	8.7	-0.25

<sup>1</sup>Averages of predicted breeding values for yield traits with test-day (TD) models were converted to the lactation basis by multiplying by 305.

<sup>2</sup>CS, AR<sub>s</sub>, and AR<sub>e</sub> models are TD models; CS model is a repeatability model with a compound symmetry covariance structure for TD environmental effects; AR<sub>s</sub> model assumes a first-order autoregressive covariance structure for TD short-term environmental effects; AR<sub>e</sub> model assumes a first-order autoregressive covariance structure for TD residual effects; and 305-d model is an animal model using 305-d lactation records.

for cows were always smaller than for sires (Table 2). The differences in accuracy of PBV did not have a clear effect on the correlation between the PBV for sires compared with those for cows. The effect of accuracy may be more apparent when comparing the rankings of the top 100 cows and sires with different models, as in Tables 5 and 6. Strabel and Szwaczkowski (1999) reported smaller correlations between PBV for sires than for cows in 6 of 28 comparisons of TD models with different definitions of contemporary groups and with a 305-d model for milk yield.

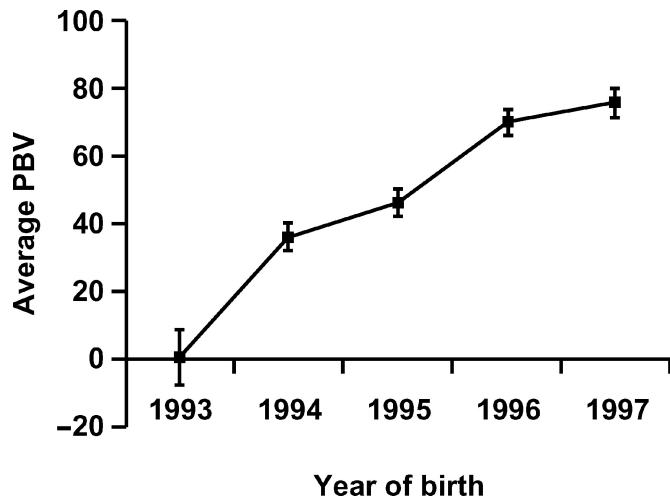
Animals were ranked with each of the 4 models to select the best 100 sires and cows for each trait. Then averages of PBV with AR<sub>e</sub> model were calculated for the 100 elite sires (Table 7) and cows (Table 8) selected with each of the 4 different models. The averages of PBV of the selected 100 elite sires and cows with CS and AR<sub>s</sub> models were very similar to the averages of PBV of those selected with the AR<sub>e</sub> model. The largest

**Table 8.** Averages of predicted breeding values for yield traits<sup>1</sup> (kg/lactation) and SCS predicted with AR<sub>e</sub> model<sup>2</sup> for the top 100 cows ranked with the 4 different models.

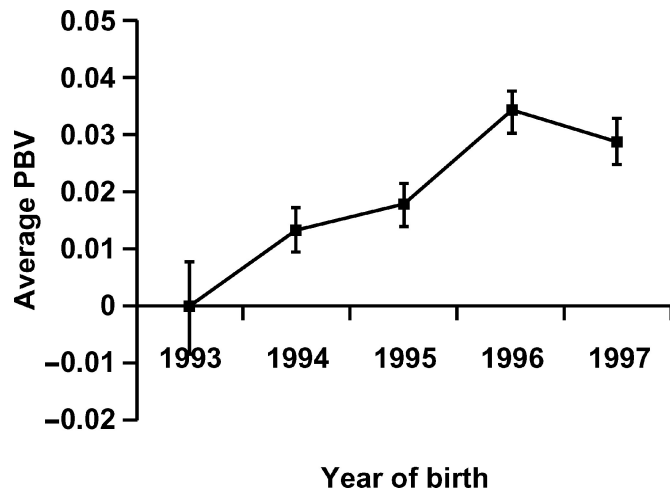
Model	Milk	Fat	Protein	SCS
CS model	615	26.6	18.8	-0.38
AR <sub>s</sub> model	616	26.6	18.8	-0.38
AR <sub>e</sub> model	618	26.6	18.9	-0.38
305-d model	540	20.9	15.7	-0.29

<sup>1</sup>Averages of predicted breeding values for yield traits with test-day (TD) models were converted to the lactation basis by multiplying by 305.

<sup>2</sup>CS, AR<sub>s</sub>, and AR<sub>e</sub> models are TD models; CS model is a repeatability model with a compound symmetry covariance structure for TD environmental effects; AR<sub>s</sub> model assumes a first-order autoregressive covariance structure for TD short-term environmental effects; AR<sub>e</sub> model assumes a first-order autoregressive covariance structure for TD residual effects; and 305-d model is an animal model using 305-d lactation records.



**Figure 1.** Average predicted breeding values (PBV) and approximate standard errors of prediction for milk yield (kg/305-d lactation) of cows by year of birth with the  $AR_e$  model. The  $AR_e$  model assumes a first-order autoregressive covariance structure for test-day residual effects.



**Figure 2.** Average predicted breeding values (PBV) and approximate standard errors of prediction for SCS of cows by year of birth with the  $AR_e$  model. The  $AR_e$  model assumes a first-order autoregressive covariance structure for test-day residual effects.

differences were between the averages of PBV of sires and cows selected with 305-d model and any TD model. This result agrees with the proportions of the 100 elite sires and cows in common among different models as shown in Tables 5 and 6. Averages of PBV of the 100 elite cows were always greater than those of the 100 elite sires for yield traits and smaller for SCS, which is due to the differences in the proportion of animals chosen when selecting the top 100 cows compared with the top 100 sires (i.e., the pool of cows from which to select was much larger).

Figures 1 and 2 show averages of PBV obtained with the  $AR_e$  model by year of birth for milk yield and SCS, respectively. The trends in averages of PBV by year of birth for fat and protein yields were similar to milk yield and are not presented. Average PBV by year of birth is a measure of genetic trend, and differences for consecutive years may be considered as realized genetic gain assuming the  $AR_e$  model is the "true" model. The estimated average annual realized genetic gains per lactation were 18.9 kg for milk yield, 0.25 kg for fat yield, and 0.45 kg for protein yield. Carvalho et al. (1998) reported 38 kg of annual realized genetic gain in milk yield using a multiple-lactation TD model with  $AR(1)$  covariance structures for environmental effects within and among lactations.

Average annual realized genetic change of SCS was unfavorable. Measures of SCS are often used as indicators for genetic selection against mastitis (Reents et al., 1995). The increase in the average of PBV of SCS may be due to positive (unfavorable) genetic correlations between SCS and yield traits, for which the population

has been selected. Zhang et al. (1994) reported no significant change for averages of PBV of first-lactation SCS records by year of birth of cows.

## CONCLUSIONS

The use of autoregressive covariance structure for short-term environmental effects for TD milk, fat, protein, and SCS records can result in more accurate predictions of breeding values than the CS covariance structure. The ranks of the top 100 sires and top 100 cows based on PBV were similar with different TD models. The correlations between PBV with TD models and the 305-d lactation model were generally smaller than the large correlations among PBV from TD models. Test-day and 305-d models can generate different genetic evaluations and rankings of both sires and cows.

## REFERENCES

- Carvalho, J. G. V., R. W. Blake, E. J. Pollak, R. L. Quaas, and C. V. Duran-Castro. 1998. Application of an autoregressive process to estimate genetic parameters and breeding values for daily milk yield in a tropical herd of Lucerna cattle and in United States Holstein herds. *J. Dairy Sci.* 81:2738-2751.
- Carvalho, J., E. J. Pollak, R. L. Quaas, and R. W. Blake. 2002. An autoregressive repeatability animal model for test-day records in multiple lactations. *J. Dairy Sci.* 85:2040-2045.
- Ducrocq, V., and B. Besbes. 1993. Solution of multiple trait animal models with missing data on some traits. *J. Anim. Breed. Genet.* 110:81-92.
- Gadini, C. H. 1997. Genetic evaluation of test day production traits and somatic cell scores. Ph.D. Diss., Univ. Nebraska, Lincoln.
- Gilmour, A. R., B. J. Gogel, B. R. Cullis, S. J. Welham, and R. Thompson. 2002. ASREML User Guide. Release 1.0. VSN International Ltd., Hemel Hempstead, UK.
- Henderson, C. R. 1984. Application of Linear Models in Animal Breeding. Univ. Guelph, Guelph, ON, Canada.



- Jensen, J. 2001. Genetic evaluation of dairy cattle using test-day models. *J. Dairy Sci.* 84:2803–2812.
- Kettunen, A., E. A. Mäntysaari, I. Strandén, J. Pöso, and M. Lidauer. 1998. Estimation of genetic parameters for first-lactation test-day milk production using random regression models. *Proc. 6th World Congr. Genet. Appl. Livest. Prod., Armidale, Australia.* XXIII:307–310.
- Meuwissen, T. H. E., and M. H. Pool. 2001. Auto-regressive versus random regression test-day models for the prediction of milk yields. *Proc. 2001 Interbull Meeting, Budapest, Hungary. Interbull Bull.* 27:172–178.
- Meyer, K. 1991. Estimating variances and covariances for multivariate animal models by restricted maximum likelihood. *Genet. Sel. Evol.* 23:67–83.
- Ptak, E., and L. R. Schaeffer. 1993. Use of test day yields for genetic evaluations of dairy sires and cows. *Livest. Prod. Sci.* 34:23–34.
- Quaas, R. L. 1984. Linear Prediction. Pages 1–76 in *BLUP School Handbook. Animal Genetics and Breeding Unit, Univ. of New England, N.S.W., Australia.*
- Reents, R., J. C. Dekkers, and L. R. Schaeffer. 1995. Genetic evaluation for somatic cell score with a test day model for multiple lactations. *J. Dairy Sci.* 78:2858–2870.
- Sawalha, R. M., J. F. Keown, S. D. Kachman, and L. D. Van Vleck. 2005. Evaluation of autoregressive covariance structures for test-day records of Holstein cows: Estimates of parameters. *J. Dairy Sci.* 88:2632–2642.
- Strabel, T., and T. Szwaczkowski. 1999. The use of test day models with small size of contemporary groups. *J. Anim. Breed. Genet.* 116:379–386.
- Swalve, H. H. 1995. The effect of test day models on the estimation of genetic parameters and breeding values for dairy yield traits. *J. Dairy Sci.* 78:929–938.
- Trus, D., and L. G. Buttazzoni. 1990. A multiple trait approach to modeling the lactation curve. *Proc. 4th World Congr. Genet. Appl. Livest. Prod., Edinburgh, UK.* XIII:492–495.
- Wade, K. M., R. L. Quaas, and L. D. Van Vleck. 1990. Mixed linear models with an autoregressive error structure. *Proc. 4th World Congr. Genet. Appl. Livest. Prod., Edinburgh, UK.* XIII:508–511.
- Wade, K. M., R. L. Quaas, and L. D. Van Vleck. 1993. Estimation of the parameters involved in a first-order autoregressive process for contemporary groups. *J. Dairy Sci.* 76:3033–3040.
- Wiggans, G. R., and M. E. Goddard. 1997. A computationally feasible test day model for genetic evaluation of yield traits in the United States. *J. Dairy Sci.* 80:1795–1800.
- Wood, P. D. P. 1967. Algebraic model of the lactation curve in cattle. *Nature* 216:164–165.
- Zhang, W. C., J. C. M. Dekkers, G. Banos, and E. B. Burnside. 1994. Adjustment factors and genetic evaluation for somatic cell score and relationships with other traits of Canadian Holsteins. *J. Dairy Sci.* 77:659–665.