J. Dairy Sci. 85:(Jan.)

Deriving Lactation Yields from Test-Day Yields Adjusted for Lactation Stage, Age, Pregnancy, and Herd Test Date

G. R. Wiggans,[*] P. M. VanRaden,[*] J. Bormann,[†] J. C. Philpot,[*] T. Druet,[†,‡] and N. Gengler[†,‡]
*Animal Improvement Programs Laboratory, Agricultural Research Service, USDA, Beltsville, MD 20705-2350
[†]Animal Science Unit, Gembloux Agricultural University, B-5030 Gembloux, Belgium
[‡]National Fund for Scientific Research, B-1000 Brussels, Belgium

---

Received September 4, 2001.
Accepted November 21, 2001.
Corresponding author: G. R. Wiggans, e-mail: wiggans@aipl.arsusda.gov.

ABSTRACT

Lactation records for milk, fat, and protein yields were calculated from test-day data adjusted for the effects of lactation stage, age, previous days open, days pregnant, and test-day class (herd, test date, and milking frequency). Those lactation records reflect the improved accounting of environmental effects from a test-day model and can be combined with historical lactation records. Test-day data were adjusted with existing lactation multiplicative adjustments to maintain variance characteristics. Then additive adjustments for lactation stage, age, previous days open, and days pregnant were applied. The current multiplicative adjustments for previous days open were not applied because its effect was expected to differ by lactation stage. To remove genetic differences, the estimated breeding value from the previous evaluation divided by 305 was subtracted. Effects of test-day class, and permanent environment within and across parities were estimated within herd. The effect of test-day class was subtracted from adjusted test-day yield, and the breeding value restored. Those deviations then were combined with the best prediction procedure into a lactation measure. Heritabilities and repeatabilities of lactation records that were adjusted for test-day class were higher than for current lactation records. The adjusted records should improve the accuracy of evaluations and allow the use of test-day data as well as provide for the continued use of historical data when test-day data are not available.
(**Key words:** test-day model, genetic evaluation, yield traits)
Abbreviation key: TDALR = test-day adjusted lactation records, **HLR** = historical lactation records, **CLR** = current lactation records.

INTRODUCTION

Efforts are under way worldwide to implement test-day models so that the accuracy of genetic evaluations can be improved by accounting for environmental effects at the test-day level and for genetic differences in persistency and maturity rate (Swalve, 2000). The first requirement of a test-day model is accurate test-day data. Traditionally, milk yield data have been summarized on a 305-d basis. A test-day system must be able to acquire, validate, and store the approximate 10-fold increase in data required for a test-day model. Secondly, a method of integrating previous lactation records and adjusted test-day data must be utilized, or lactations without test-day data must be excluded from the evaluations.

Full test-day models using random regression (Schaeffer et al., 2000; Swalve, 2000) are currently being used or studied by several countries. Those models are computationally demanding and are not capable of timely processing of the large US dataset. Wiggans and Goddard (1997) proposed a two-step approach for a test-day model. The first step included analysis within herd to estimate the effects of herd test day. In New Zealand (Johnson, 1996), Australia (Jones and Goddard, 1990), and the northeastern US (Animal Breeding Group, 2000), a similar within-herd analysis is followed by conversion of the adjusted test-day values into a lactation value. These data can then be analyzed across herds. A similar two-step approach was shown to be theoretically equivalent to a class of random regression models by Gengler et al. (2000). In that approach, step 1 was the estimation of regressions, which can be done within herd or region, and step 2 was the estimation of effects on regressions, which combines information across herds or regions. Iteration across steps may be necessary.

The size of the US dairy population makes implementation of new procedures a complex and time-consuming process. The worldwide interest in test-day model evaluations and the unpredictable time required to complete development of a full test-day model suggest an approach based on incremental implementation. A first step is to account for environmental effects on a test-day basis, which should improve the accuracy of current evaluations. The reported procedure calculates lactation records from adjusted test-day yields. These test-day adjusted lactation records could be combined with previous lactation records in an existing lactation model.

MATERIALS AND METHODS

Data

Test-day data for calvings since January 1, 1990, were extracted from records that are provided routinely for USDA-DHIA genetic evaluations and from archive files that were provided by dairy records processing centers and universities. Test-day information is stored in a database that includes cow identification, herd code, calving date, DIM, milking frequency, milk yield, fat and protein percentages, and SCS; this file currently contains over 218 million test-day records.

In addition to a file of cow test days, a file of herd test days is maintained. The herd file is used to prevent invalid test dates from being accepted into the cow test-day file. Because cows sometimes change herds and the data are stored at the processing centers on a cow

or lactation basis, an incorrect herd code is sometimes associated with a test-day record. Validation against the herd test-day file allows detection of such errors. Only data from one herd per lactation is included in evaluations. The herd test-day file also stores herd-wide characteristics of milk recording, such as the number of milkings on test day included in the milk and components observations and if the milk recording was supervised. Test-day data were acquired for most herds; however, some gaps remained, particularly prior to 1996. Milking frequency is stored as a characteristic of the cow test day because cows in the same herd may be milked at different frequencies.
Data from all breeds were processed. The results for Brown Swiss, Jerseys, and Holsteins are reported. Similar procedures also have been developed for SCS, although no results are included. For consistency with the current evaluation procedure (Wiggans, 1997), only test-day data for lactations currently included in the USDA-DHIA evaluation system were selected. As a result, for cows that changed herd during a lactation, only data from the herd with the most data were used.

Model

The within-herd model applied separately to each trait was

$$\mathbf{y} = \mathbf{Xh} + \mathbf{Wc} + \mathbf{Zp} + \mathbf{e}$$

where **y** = vector of test-day yields adjusted additively (Bormann et al., 2002) for lactation stage, age, previous days open, days pregnant, and EBV and multiplicatively for calving age, calving season, and milking frequency; **h** = vector of fixed effects of test-day class (interaction of herd, test date, and milking frequency); **c** = vector of random effects of permanent environment across parities; **p** = vector of random effects of permanent environment within parity; and **e** = vector of random residual effects; and **X**, **W** and **Z** = incidence matrices linking **y** with **h**, **c**, and **p** respectively. The (co)variance matrices of **c**, **p**, and **e** were diagonal. A common residual variance was assumed across stage of lactation although it would be more realistic for this variance to change with DIM.

Preparation of Adjusted Test-Day Yields

A goal of the within-herd analysis was to prepare data that are compatible with historical lactation records (**HLR**) calculated using the test-interval method. For combined analysis, the variances obtained must be appropriate for the heterogeneous variance adjustment (Wiggans and VanRaden, 1991) included in the lactation analysis. To achieve this, the same multiplicative adjustments applied to lactation records were applied to test-day data, except for previous days open and milking frequency. Adjustment factors for milking frequency were derived from those of Karaca (1997). These factors, which attribute a smaller benefit to three times-a-day milking than previous factors, have been used routinely for lactations since being phased in for calvings during 1997 and 1998. For test-day adjusted lactation records (**TDALR**), the new factors were used for all lactations. Previous factors for lactation yield depended on the number of days milked three times, whereas the new factors adjust each test day separately.

The next step was to apply the additive adjustments of Bormann et al. (2002) for lactation stage, age, previous days open, and days pregnant. Those effects were not included in the analysis within herd to minimize the time required for calculation. Further research could partition those effects into global and herd influences. The days pregnant in the current lactation was from the reported last breeding date. When there was a subsequent calving, this could be confirmed or estimated when missing. When no breeding date was reported, and no subsequent calving date was available, 120 d open was assumed.

The EBV from the previous evaluation was divided by 305 and subtracted from the adjusted records to estimate the test-day class (herd, test day, milking frequency) effect. The purpose of this step was to avoid confounding between test-day class solutions and cow genetic merit. This step could be repeated as part of an overall iteration to more accurately account for genetic effects and could mimic an additional benefit of test-day models allowing joint estimation of test-day class and genetic effects. Because yields differ with stage of lactation, the true breeding value is likely to differ similarly. This approximation does not consider that, and so could be improved by using stage specific EBV. The previous breeding value does not account for permanent environmental effects within and across parities, so the within-herd analysis included a block iteration for within and across parities effects. Because a maximum of five parities were included for a cow, blocks had a maximum order of six: five within parity and one across parities effect. Variance ratios were derived as part of the work reported in Bormann et al. (2002). However, the ratios presented in Table 1 differ by trait, but not breed or region. The test-day class solutions were calculated by averaging the test-day deviations adjusted by the solutions for within and across parities. At least three cows with data for a test-day class were required for that test-day class to be used.

Table 1. Variance ratios used in estimation of test-day class effect.

	Milk	Fat	Protein
Across parities	4.6	11.6	6.6
Within parity	1.7	3.5	2.2

Best Prediction

The within and across parities solutions were calculated only to improve the accuracy of the test-day class solutions. The test-day class solution was subtracted from the adjusted test-day yields and the breeding value restored. The best prediction procedure (VanRaden, 1997) was applied to these deviations to calculate TDALR. This procedure also requires a herd average. Because the test-day yields are deviations around zero, with all environmental trend removed, but genetic trend remaining, the average breeding value within herd by year of calving was used as the herd average. To provide a seamless transition from the HLR to the TDALR, the mean of the first year of the test-day based records was calculated for both the HLR and TDALR, then all TDALR were adjusted by the difference in the means.

Some gaps were detected in the data, so the date to switch to TDALR was determined for each herd. Moving backward in time, by quarter, the switch was at the start of a quarter where the two previous quarters had TDALR with less than 80% of the average lactation length weight of the current lactation records (**CLR**). This criterion was designed so that once the TDALR were used, they continued to be used, but limited the loss of lactation records due to incomplete test-day data. The TDALR could fall below the 80% limit both for having fewer tests included in the record, and for lactation records missing completely. The example in Table 2 illustrates the steps to calculate TDALR.

Table 2. Example of steps in preparing test-day adjusted lactation records (TDALR), for a cow, born 1997-08-27, calving 1999-06-13, in parity 1, in Wisconsin.

Test-Day Yield Adjustment					
Step		**Explanation**	**Test Date**		
			1999-07-28	1999-08-27	1999-09-24
			46 DIM	76 DIM	104 DIM
				kg	
A	Extract daily milk yield	Test day milk	30.0	33.1	35.4
B	Multiply by calving age, calving season, and milking frequency factors	**A** × 1.3835(mature equivalent factor = 1.3835)	41.5	45.9	49.0
C	Subtract adjustments	Stage of lactation (**C1**)	3.0	2.7	2.3
		Age at test day effect (**C2**)	0.5	0.7	0.9
		Previous days open effect (**C3**)(none for first parity)	0	0	0
		Days pregnant effect (**C4**)(bred at 80 DIM)	0	0	0.5
		B - (**C1** + **C2** + **C3** + **C4**)	38.0	42.5	45.3
D	Subtract previous run estimated breeding value, divided by 305	**C** - 3.47(previous PTAm = 1165; value converted to kg and divided by 305 = 3.47)	34.5	39.0	41.8
E	Solve within herd	Test-day class effect	32	29	34
F	Subtract test-day class effect	**D** - **E**	2.5	10.0	7.8
G	Restore the breeding value by combining it with residual	**F** + 3.47	6.0	13.5	11.3

Using Adjusted Test-day Yields to Calculate TDALR

	Step	Explanation	Lactation
H	Calculate lactation yield with best prediction using average breeding value by year of calving as herd mean	Best prediction lactation yield(herd mean = 1210)	2575
I	Adjust mean of lactation yields so mean of first year equals mean of lactation for that year	Herd average adjusted lactation	900
J	Calculate herd average lactations (unadjusted)	Unadjusted herd average lactation	9728
K	Calculate TDALR; from unadjusted herd mean subtract adjusted herd mean, add best prediction lactation yield	**J - I + H**	11403

Comparison of Lactation Records

The TDALR were matched with the CLR, and the herd-year mean was removed. The CLR are a mixture of test-interval method before 1997 and best prediction without test-day adjustments for lactations starting 1997 and later. Correlations were then calculated. Correlations without removing the herd-year mean would be affected by the absence of the environmental trend in TDALR. As a measure of the benefit of the test-day adjustment, heritabilities, and repeatabilities were estimated from these matching records for Brown Swiss, Jersey, and samples of Holstein data in California, Pennsylvania, Texas, and Wisconsin using the system described and used by Bormann et al. (2002) based on Method R and the preconditioned conjugate gradient solver. Method R was chosen so large datasets could be used to estimate these parameters. Values were calculated from six computations using different random samples representing 50% of the records of each data set.

RESULTS AND DISCUSSION

Considerable computer processing time was required for the calculation of TDALR. Because processing can be done independently by herd, the Holstein file was divided into six groups that were run simultaneously on an 8-processor machine. This reduced the running time to about 10 h. The calculations involving matrix inversion for best prediction were the primary consumer of computer resources. Table 3 contains the counts of lactation records calculated by breed and the percentage of total lactations since 1990 that they represent.

Table 3. Counts of lactations with calvings since 1990 and those not replaced by test-day adjusted lactation records (TDALR).

Breed	Lactations since 1990	Not replaced	Percent not replaced
Brown Swiss	176,148	51,061	29
Jersey	1,030,680	183,189	18
Holstein	17,124,977	2,803,207	16

Correlation within herd year of TDALR and CLR are shown in Table 4. Values were all between 0.93 and 0.95 with lower correlations for protein and fat than for milk.

Table 4. Correlations within herd year of current lactation records (CLR) and test-day adjusted lactation records (TDALR).

Breed	Region	Milk and Fat Lactations	Milk	Fat	Protein Lactations	Protein
Brown Swiss	US	121,303	0.95	0.94	120,998	0.94
Jersey	US	839,968	0.95	0.94	835,615	0.94
Holstein	California	2,611,907	0.95	0.95	1,925,729	0.94
	Pennsylvania	1,581,885	0.95	0.94	1,581,282	0.93
	Texas	230,876	0.95	0.94	228,695	0.93
	Wisconsin	2,117,801	0.94	0.93	2,116,656	0.93

Mean heritabilities (Table 5), obtained from six random samples, showed slight increases of up to 0.02 points for those from TDALR compared with those from CLR. Standard deviations of samples were low with values around 0.01 or lower. Mean repeatabilities (Table 6) also were obtained from these six samples. Increases for repeatability were greater than for heritability with values up to 0.05 points. Standard-deviations of sample estimates tended to be lower for repeatabilities with all values below 0.01. In the absence of an exact statistical test, reported standard deviations suggested that differences between heritabilities and repeatabilities were in most cases significant. These results suggest that the new method to create lactation records was superior to the old method.

Table 5. Mean of heritabilities from six samples with sample SD for current lactation records (CLR) and test-day adjusted lactation records (TDALR).

Trait	Breed	Region	CLR		TDALR	
			Mean	SD	Mean	SD
Milk	Brown Swiss	US	0.290	0.018	0.314	0.012
	Jersey	US	0.313	0.004	0.325	0.002
	Holstein	California	0.247	0.005	0.270	0.007
		Pennsylvania	0.246	0.015	0.259	0.016
		Texas	0.227	0.005	0.246	0.005
		Wisconsin	0.279	0.013	0.299	0.013
Fat	Brown Swiss	US	0.261	0.011	0.277	0.014
	Jersey	US	0.243	0.001	0.252	0.003
	Holstein	California	0.257	0.006	0.270	0.006
		Pennsylvania	0.289	0.014	0.304	0.014
		Texas	0.227	0.007	0.244	0.008
		Wisconsin	0.300	0.013	0.309	0.012
Protein	Brown Swiss	US	0.282	0.016	0.293	0.016
	Jersey	US	0.271	0.004	0.278	0.003
	Holstein	California	0.227	0.004	0.240	0.004
		Pennsylvania	0.228	0.013	0.238	0.013
		Texas	0.211	0.007	0.232	0.005
		Wisconsin	0.250	0.009	0.271	0.008

Table 6. Mean of repeatabilities from six samples with sample SD for current lactation records (CLR) and test-day adjusted lactation records (TDALR).

Trait	Breed	Region	CLR		TDALR	
			Mean	SD	Mean	SD
Milk	Brown Swiss	US	0.470	0.005	0.485	0.006
	Jersey	US	0.495	0.003	0.523	0.002
	Holstein	California	0.427	0.003	0.478	0.002
		Pennsylvania	0.423	0.006	0.465	0.005
		Texas	0.409	0.005	0.443	0.004
		Wisconsin	0.440	0.005	0.489	0.005
Fat	Brown Swiss	US	0.429	0.006	0.432	0.004
	Jersey	US	0.433	0.002	0.457	0.001
	Holstein	California	0.438	0.004	0.471	0.003
		Pennsylvania	0.458	0.006	0.493	0.004
		Texas	0.394	0.004	0.423	0.004
		Wisconsin	0.472	0.004	0.507	0.004
Protein	Brown Swiss	US	0.472	0.002	0.484	0.003
	Jersey	US	0.468	0.003	0.492	0.003
	Holstein	California	0.419	0.003	0.459	0.004
		Pennsylvania	0.420	0.006	0.461	0.006
		Texas	0.407	0.003	0.440	0.002
		Wisconsin	0.428	0.004	0.477	0.004

CONCLUSIONS

More than 218 million test-day records were available for use in genetic evaluation, but gaps in the data from 1990 to 1996 required novel methods to use both TDALR and HLR. The TDALR were adjusted for the environmental effects of each test-day class. These records had slightly higher heritabilities and repeatabilities than lactation records without this adjustment. The effects of the current pregnancy were considered in addition to the adjustments currently applied to lactation records. Milking frequency was used to define the test-day effect so that separate effects were estimated for herds with cows milked at different frequencies. These records can be combined with HLR for the period before the availability of test-day data and used in the current evaluation system without alteration. This achieves the benefit of the more accurate accounting of environmental effects in the recent data without loss of the historical data. Estimation of genetic differences in the shape of the lactation curve and effect of parity will be considered in future research.

ACKNOWLEDGMENTS

This research was conducted with original or modified computer programs originally developed by I. Misztal, Animal and Dairy Science Department, University of Georgia, Athens. Nicolas Gengler, who is Chercheur Qualifié, and Tom Druet, who is Aspirant, acknowledge their financial support from the National Fund for Scientific Research, Brussels, Belgium.

REFERENCES

Animal Breeding Group, Animal Science Department, Cornell University. 2000. AI Bull Census. http://www.ansci.cornell.edu/abc/siretabs/about.html. Accessed Aug. 21, 2001.

Bormann, J., G. R. Wiggans, T. Druet, and N. Gengler. 2001. Estimating effects of permanent environment, lactation stage, age, and pregnancy on test-day yield. J. Dairy Sci. . J. Dairy Sci. 85(Jan.). Online. Available: http://www.adsa.org.

Gengler, N., A. Tijani, and G. R. Wiggans. 2000. Use of sequential estimation of regressions and effects on regressions to solve large multitrait test-day models. J. Dairy Sci. 83(Feb.). Online. Available: http://www.adsa.org.

Johnson, D. L. 1996. Estimation of lactation yield from repeated measures of test day yields. Proc. New Zealand Soc. Anim. Prod. 56:16-18.

Jones, L. P., and M. E. Goddard. 1990. Five years experience with the animal model for dairy evaluations in Australia. Proc. 4th World Congr. Genet. Appl. Livest. Prod., Edinburgh, Scotland, XIII:382-385.

Karaca, M. 1997. Determination of the sources of variation to adjust production records for three times a day milking. M.S. Thesis. Iowa State Univ. Ames, Iowa.

Schaeffer, L. R., J. Jamrozik, G. J. Kistemaker, and B. J. Van Doormaal. 2000. Experience with a test-day model. J. Dairy Sci. 83:1135-1144.

Swalve, H. H. 2000. Theoretical basis and computational methods for different test-day genetic evaluation methods. J. Dairy Sci. 83:1115-1124.

VanRaden, P.M. 1997. Lactation yields and accuracies computed from test-day yields and (co)variances by best prediction. J. Dairy Sci. 80:3015-3022.

Wiggans, G. R. 1999. Genetic evaluation systems in the United States: present method and proposed changes. Pages 19-31 *in* Proc. 32nd Inter. Symp. Anim. Prod.: Advances in Technology, Accuracy and Management, Milan, Italy.

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.

Wiggans G. R., and P. M. VanRaden. 1991. Method and effect of adjustment for heterogeneous variance. J. Dairy Sci. 74:4350-4357.