J. Dairy Sci. 94:1601–1604
doi:10.3168/jds.2010-3865

Short communication: Best prediction of 305-day lactation yields with regional and seasonal effects

J. B. Cole,*[1] D. J. Null,* and A. De Vries†
*Animal Improvement Programs Laboratory, ARS, USDA, Beltsville, MD 20705-2350
†Department of Animal Sciences, University of Florida, Gainesville 32611

ABSTRACT

In the United States, lactation yields are calculated using best prediction (BP), a method in which test-day (TD) data are compared with breed- and parity-specific herd lactation curves that do not account for differences among regions of the country or seasons of calving. Complete data from 538,090 lactations of 348,123 Holstein cows with lactation lengths between 250 and 500 d, records made in a single herd, at least 5 reported TD, and twice-daily milking were extracted from the national dairy database and used to construct regional and seasonal lactation curves. Herds were assigned to 1 of 7 regions of the country, individual lactations were assigned to 3-mo seasons of calving, and lactation curves for milk, fat, and protein yields were estimated by parity group for regions, seasons, and seasons within regions. Multiplicative pre-adjustment factors (MF) also were computed. The resulting lactation curves and MF were tested on a validation data set of 891,806 lactations from 400,000 Holstein cows sampled at random from the national dairy database. Mature-equivalent milk, fat, and protein yields were calculated using the standard and adjusted curves and MF, and differences between 305-d mature-equivalent yields were tested for significance. Yields calculated using 50-d intervals from 50 to 250 d in milk (DIM) and using all TD to 500 DIM allowed comparisons of predictions for records in progress (RIP). Differences in mature-equivalent milk ranged from 0 to 51 kg and were slightly larger for first-parity than for later parity cows. Milk and components yields did not differ significantly in any case. Correlations of yields for 50-d intervals with those using all TD were similar across analyses. Yields for RIP were slightly more accurate when adjusted for regional and seasonal differences.
Key words: best prediction, milk yield, regional effect

Received September 27, 2010.
Accepted December 2, 2010.
[1]Corresponding author: john.cole@ars.usda.gov

Short Communication

In the United States, lactation yields are calculated using best prediction (Cole et al., 2009), a method in which test-day (**TD**) data are compared with breed- and parity-specific herd lactation curves that do not account for differences among regions of the country or seasons of calving. Test-day models allow animals to have lactation curves of different shapes based on factors such as parity, season of calving, and region of the country (Stanton et al., 1992). Multiplicative factors (**MF**) have been used for many years to adjust records for factors such as milking frequency (Shook et al., 1980; Wilmink, 1987) and to project records in progress (**RIP**; Wiggans, 1985).

The objectives of this study were to develop region-, season-, and region-by-season-specific lactation curves for the mean and SD of 305-d milk (**M**), fat (**F**), and protein (**P**) yields and compare them to MF and the standard lactation curves (**LC**) currently used by BP.

Lactation curves for M, F, and P yields were estimated by parity group (first versus later) for regions (**RLC**; n = 14), seasons (**SLC**; n = 8), and regions-by-seasons (**RSLC**; n = 56) with a data set constructed using the edits described in Cole et al. (2009). Herds were assigned to 1 of 7 regions of the country (Hare et al., 2004), and individual lactations were assigned to 3-mo seasons of calving (March–May, June–August, September–November, and December–February) and 1 of 2 parity groups (first or later). The TD data were grouped into either 15-d (1 to 300 DIM) or 30-d (>300 DIM) intervals based on TD DIM. Means and SD of M, F, and P yields of the groups were modeled using the Wood formula (Wood, 1967) fit using the NLIN procedure of SAS/STAT software (SAS Institute, 2007). The data were regressed toward the mean as a result of grouping them into intervals, but the 2 shape parameters of the curves are of greater interest than the scale parameter.

As an alternative to using many LC, the accurate calculation of which may be difficult for small breeds, MF were used to pre-adjust TD records for regional and seasonal effects. Regional MF were calculated as

Table 1. Correlations of 305-d mature-equivalent milk, fat, and protein yields calculated by best prediction from all available test-days and only test-days occurring on or before the indicated DIM using standard lactation curves (LC); curves calculated by region (RLC), season (SLC), and region-by-season (RSLC); and multiplicative pre-adjustment factors for region (RMF), season (SMF), and region-by-season (RSMF)

		DIM[1]					
Trait	Adjustment	1	50	100	150	200	250
Milk (kg)	LC	0.965	0.884	0.922	0.956	0.978	0.991
	RLC	0.965	0.884	0.922	0.956	0.978	0.991
	RSLC	0.965	0.887	0.925	0.958	0.980	0.993
	SLC	0.966	0.886	0.925	0.958	0.980	0.993
	RMF	0.942	0.880	0.922	0.957	0.980	0.993
	RSMF	0.939	0.880	0.922	0.957	0.980	0.993
	SMF	0.964	0.884	0.923	0.958	0.980	0.993
Fat (kg)	LC	0.967	0.873	0.912	0.949	0.975	0.990
	RLC	0.966	0.874	0.911	0.948	0.974	0.990
	RSLC	0.966	0.876	0.914	0.951	0.976	0.992
	SLC	0.967	0.877	0.914	0.951	0.976	0.992
	RMF	0.953	0.871	0.912	0.950	0.976	0.991
	RSMF	0.951	0.869	0.911	0.950	0.976	0.991
	SMF	0.966	0.873	0.913	0.950	0.976	0.992
Protein (kg)	LC	0.930	0.878	0.919	0.952	0.976	0.990
	RLC	0.929	0.877	0.917	0.952	0.975	0.990
	RSLC	0.929	0.879	0.920	0.954	0.978	0.992
	SLC	0.929	0.879	0.921	0.954	0.977	0.992
	RMF	0.902	0.872	0.918	0.953	0.977	0.992
	RSMF	0.901	0.871	0.917	0.953	0.977	0.992
	SMF	0.929	0.878	0.920	0.954	0.978	0.992

[1]Differences between complete and partial lactation yields using no test-day (1 DIM) or test-days in successive 50-d intervals (50, 100, 150, 200, and 250 DIM; n = 891,806). Yields did not differ between adjustment methods for any DIM group ($P > 0.05$).

$m_{tlr} = y_{tl}/\bar{y}_{tlr}$, where m_{tlr} = is the multiplicative factor for the tth time period (15- or 30-d), lth parity group (first versus later), and rth region; $\bar{y}_{tlr}$ = average TD yield for the tth time period, lth parity group, and rth region; and y_{tl} = average TD yield for the tth time period and lth parity group over all regions. Season and region-by-season curves were calculated similarly.

The RLC, RSLC, SLC, and MF were validated as in Cole et al. (2009) by using data from 891,806 lactations from 400,000 Holstein cows with lactation lengths between 250 and 500 d, records made in a single herd, at least 5 reported TD, and twice-daily milking extracted from the national dairy database. Herds were assigned to regions and lactations to seasons of calving as described above. Test-day data were grouped into 50-d periods (1 to 50 DIM, 1 to 100 DIM, . . . , 1 to 305 DIM), and 305-d yields were calculated using each adjustment method. The resulting lactation yields were correlated with 305-d yields computed using LC, and differences between them were calculated.

Seasonal MF varied only slightly, ranging from 0.97 (December–February) to 1.02 (June–August), with cows calving in the winter months producing more than cows calving in the summer. The regional MF ranged from 0.89 (Northwest) to 1.09 (Southeast), indicating that yields for animals in the Northwest region are overestimated and those in the Southeast region underestimated. Region-by-season factors were similar to regional MF, with values between 0.86 (Northwest, September–November) and 1.10 (Southeast, March–May). Factors <1 adjust yields down, whereas factors >1 adjust them up.

Correlations of 305-d mature-equivalent yields based on LC with projected 305-d mature-equivalent yields from 50-d intervals using RLS, RSLC, SLC, and MF for all lactations are shown in Table 1. Results were similar for all methods of adjustment, although early in lactation the lactation curve-based adjustments produced slightly better correlations than did the MF. However, as DIM increased, the correlations became similar for all methods. These results suggest that MF-based adjustments may be preferable to those based on LC because of the comparative ease of calculating the former, and also because the dairy industry understands and is comfortable with multiplicative adjustments.

Scale and shape parameters of lactation curves varied by season, region, and region-by-season, although it was difficult to directly interpret those values. The scale parameter of the Wood function changed the y-intercept of the curve, whereas the 2 scale parameters changed the shape of the curve. First-parity cows generally had smaller scale and larger shape parameters than did later parity cows, which is in agreement with

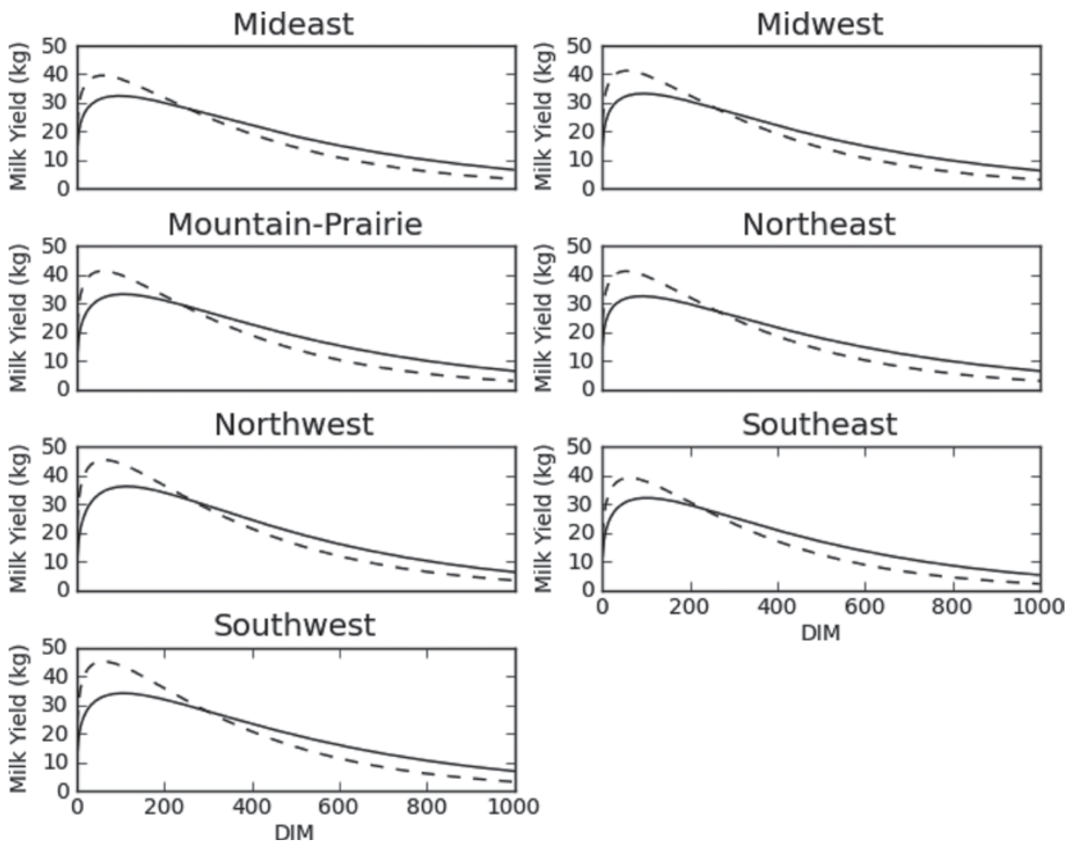

Figure 1. Lactation curves for mean milk yield (kg) of Holstein dairy cattle for first (solid line) and later (broken line) parity groups in 7 regions of the United States.

the results of Cole et al. (2009). Visual inspection of the regional curves for milk yield (Figure 1) shows little practical difference among the curves.

Little seasonal variation occurred among multiplicative factors, with cows freshening in the spring or summer producing slightly less 305-d milk than those calving in the fall or winter. Regional differences were larger, with cows in the Southeast having the lowest average mature-equivalent milk yields and those in the Northwest having the highest. Region-by-season factors appeared to be dominated by regional rather than seasonal effects. Differences among parity groups also were small.

Carabaño et al. (1990) compared cows in California, New York, and Wisconsin and, based on genetic correlations among the regions, concluded that there were no substantial genotype-by-environment interactions affecting milk, fat, or protein. Even if modest genotype-by-environment effects existed, the heterogeneous variance adjustments (Wiggans and VanRaden, 1991) used in the calculation of PTA would account for most of those differences. Regional and seasonal effects on yield may have little or no effect on genetic evaluations, and their effect on predictions of daily yield used to make management decisions probably are also small.

Decision support systems use inputs such as milk yield to determine the value of alternative management strategies, such as the decision to keep or cull a cow. Accurate estimates of yield are needed to forecast income from milk sales and predict costs associated with production; if seasonal or regional effects are large, then costly errors may be made by ignoring them. The largest difference in mature-equivalent milk yield between SLC and any alternative method tested was 51 kg, which has a value of approximately \$16 (Cole et al., 2010) over the course of a 305-d lactation; differences were similar for actual 305-d milk. For early RIP the maximum difference among methods was \$78 of mature-equivalent milk or \$94 of actual milk. These differences among methods of adjustment account for only \$0.05 to \$0.31 of income per day of lactation, and are dwarfed by costs associated with feed and reproduction.

Although this study focused on 305-d lactation yields used to compute genetic evaluations, methods to adjust daily yields for seasonal and regional effects also are needed. A general approach is to pre-adjust daily or TD data to remove regional and seasonal effects, use best prediction to calculate daily or lactation totals, and then post-adjust the data. Best prediction can be used to calculate daily from TD data, and post-adjustments can be used to adjust the predicted daily yield values to reflect seasonal effects associated with season of calving of individual cows and regional effects associated with individual herds. Further research is needed to determine the effectiveness of such an approach.

The adjustments presented in this study had small effects on mature-equivalent milk, fat, and protein, as well as actual production (data not shown), although effects of climate on production are well known. Several reasons for this exist, the most likely of which is that the regions used in this study did not accurately group areas with similar climatic conditions. This results in an averaging-out of effects such that less variation than expected occurs among regions. Grouping nearby states is convenient, but regions may need to be defined in a more appropriate way, such as by the Köppen-Geiger climate classification map (Kottek et al., 2006). For example, in this study Idaho and Utah were assigned to different regions, but most cattle in those states are located in areas with very similar climates (NOAA, 2010).

Effects of seasonal variation, most notably heat stress, can be ameliorated using a variety of technologies (Bucklin et al., 1991). Effective use of cooling systems can eliminate most effects of heat stress, but there is no way to easily identify the farms using those technologies. In their work on genetic improvement of resistance to heat stress, Ignacy Misztal's group at the University of Georgia identified bias in bull genetic evaluations resulting from the use of cooling systems (I. Misztal, University of Georgia, Athens, personal communication, 2010), and it is reasonable to assume that a similar bias exists in our results.

Finally, the 3-parameter Wood function may not have the ability to accurately describe seasonal or regional variations in lactation curves. Many different functions for modeling lactation curves have been described, several of which can describe a wider range of shapes, but Dematawewa et al. (2007) concluded that the Wood curve was adequate for calculating 305-d yields.

Adjustment of TD data for regional and seasonal effects had little effect on predictions of mature-equivalent milk, fat, or protein yield. This suggests that regional effects on yield are small, or that the region definitions used in this study may not properly reflect actual climatic differences among states. Actual yields were better predicted by RSLC, which may be useful for on-farm decision support. The MF are preferred over RSLC for ease of computation.

ACKNOWLEDGMENTS

The cooperation of the dairy records processing centers [AgriTech Analytics (Visalia, CA), AgSource Cooperative Services (Verona, WI), Dairy Records Management Systems (Raleigh, NC, and Ames, IA), and DHI Computing Services (Provo, UT)] in supplying lactation yield data is acknowledged. Two anonymous reviewers provided valuable feedback on the manuscript.

REFERENCES

Bucklin, R. A., L. W. Turner, D. K. Beede, D. R. Bray, and R. W. Hemken. 1991. Methods to relieve heat stress for dairy cows in hot, humid climates. Appl. Eng. Agric. 7:241–247.

Carabaño, M. J., K. M. Wade, and L. D. Van Vleck. 1990. Genotype by environment interactions for milk and fat production across regions of the United States. J. Dairy Sci. 73:173–180.

Cole, J. B., D. J. Null, and P. M. VanRaden. 2009. Best prediction of yields for long lactations. J. Dairy Sci. 92:1796–1810.

Cole, J. B., P. M. VanRaden, and Multi-State Project S-1040. 2010. AIPL Research Report NM$4: Net merit as a measure of lifetime profit: 2010 revision. Accessed Sep. 21, 2010. http://aipl.arsusda.gov/reference/nmcalc.htm.

Dematawewa, C. M. B., R. E. Pearson, and P. M. VanRaden. 2007. Modeling extended lactations of Holsteins. J. Dairy Sci. 90:3924–3936.

Hare, E., H. D. Norman, and J. R. Wright. 2004. Duration of herd participation in Dairy Herd Improvement milk recording in the United States. J. Dairy Sci. 87:2743–2747.

Kottek, M., J. Grieser, C. Beck, B. Rudolf, and F. Rubel. 2006. World map of the Köppen-Geiger climate classification updated. Meteorolog. Zeitsch. 15:259–263.

NOAA. 2010. United States Climate Page. Accessed Sep. 21, 2010. http://www.esrl.noaa.gov/psd/data/usclimate/states.fast.html.

SAS Institute. 2007. SAS OnlineDoc 9.1.3. SAS Institute Inc., Cary, NC. Accessed Dec. 11, 2007. http://support.sas.com/onlinedoc/913/docMainpage.jsp.

Shook, G. E., L. P. Johnson, and F. N. Dickinson. 1980. Factors for improving accuracy of estimates of test-interval yield. DHI Lett. 56(4):9–24.

Stanton, T. L., L. R. Jones, R. W. Everett, and S. D. Kachman. 1992. Estimating milk, fat, and protein lactation curves with a test day model. J. Dairy Sci. 75:1691–1700.

Wiggans, G. R. 1985. Calculating lactation records. National Cooperative Dairy Herd Improvement Program Handbook, Fact Sheet G-1. Extension Service, USDA, Washington, DC.

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

Wilmink, J. B. M. 1987. Adjustment of test-day milk, fat and protein yield for age, season and stage of lactation. Livest. Prod. Sci. 16:335–348.

Wood, P. D. P. 1967. Algebraic model of the lactation curve in cattle. Nature 216:164–165.