Effects of Errors in the Age Adjustment of First Lactations

R. H. MILLER, B. T. McDANIEL, and R. D. PLOWMAN
U. S. Department of Agriculture, Beltsville, Maryland

Abstract

The Dairy Herd Improvement Association (DHIA) age factors of Kendrick and McDaniel were compared using 10,620 first lactations made in Michigan Holstein herds during the period 1961-63. The effectiveness of the alternative age adjustments was judged by the impact on sire ranking and on the statistical properties of the standardized records. Bulls were found to rank almost identically, regardless of the manner of age adjustment. Rank correlations between proofs based on actual and age-corrected records were about .95. There were significant regressions of age-adjusted production on age, for both sets of factors. It was concluded that standard age adjustments may not be satisfactory in research applications requiring a high degree of accuracy.

Searle and Henderson (9) discussed alternative ways of adjusting production records for variation in age at calving. They compared the Statistical properties of lactation records adjusted for age by three different procedures: (1) multiplicative, (2) herd level, and (3) additive corrections. Among the criteria considered were repeatability, coefficients of variation, and residual variance due to age. Searle and Henderson concluded that there is no single criterion for choosing a best method of age correction. They did not directly consider age adjustments from the standpoint of their effects on sire evaluation.

Recently, many workers have suggested the use of separate age adjustments according to season of calving [Gravir and Hickman (2), McDaniel et al. (7), Syrstad (11), and Miller et al. (8)]. Van Vleck and Henderson (12) previously advocated simultaneous adjustment for age and season effects. Butcher (1), Gravir and Hickman (2), McDaniel et al. (7), and Miller et al. (8) also reported that milk and fat age adjustment factors should be different, since fat per cent shows a slight decline from first lactation to maturity.

The purpose of the present study was to compare the properties of first-lactation records adjusted for age by two different sets of multiplicative factors: a) the original Dairy Herd Improvement Association (DHIA) factors developed by Kendrick (5), and b) the new DHIA factors recently reported by McDaniel et al. (7). The latter factors employ different percentage corrections for two seasons of freshening, milk and fat, and various geographical regions of the United States.

Data

The data were 10,620 first lactations (ages at calving less than 36 months) from Michigan Holstein herds using artificial service. These data have been described in detail by Miller et al. (9). There were 1,003 herd years represented and the calving dates were restricted to 1961-63. Approximately 25% of the records were by naturally sired cows, although all herds used artificial service to some extent. Only records of registered cows were used. All records terminated by culling were extended to a 305-day basis. There were 1,095 different sires represented, but only 40 artificial insemination (AI) bulls were studied individually. The remaining sires were allotted to a group comparison of non-AI bulls versus the remaining AI sires.

Results and Discussion

Several aspects of age-adjusted records were considered. The primary objective was to determine whether the use of different sets of age factors would alter the ranking of bulls evaluated on the basis of first lactation daughters and contemporaries. The age range employed (up to 36 months) represents only about one-fourth of all lactations made in DHIA herds. However, major portions of the sire selection decisions in artificial insemination are based largely on progeny records in this age range.

In addition, certain statistical properties of age-adjusted first lactations were investigated. These included the effects of age adjustments on 1) coefficients of variation, 2) magnitude of estimated seasonal effects, 3) reduction in variation associated with age, and 4) year-to-year homogeneity of the regression of production on age.

Effects on sire evaluation. In another report (9), results of a study of three different meth-
LACTATION ADJUSTMENTS

Several methods of sire evaluation are presented. This discussion considered the relative rank correlations among the three evaluation procedures. Using the same data, sires were also evaluated on the basis of mature equivalent (ME) production of their progeny, as estimated by the Kendrick age factors and by the new DHIA factors reported by McDaniel et al. (7). At the same time sires were also evaluated on the actual first lactation production of progeny, using the following least-squares model:

\[ Y_{ijklmn} = \mu + h_i + a_j + b_k + (ab)_{jk} + s_{im} + e_{ijklmn} \]

where \( h_i \) is an effect common to observations in the \( i^{th} \) herd year, \( a_j \) is an effect present for all observations in the \( j^{th} \) season of calving, \( b_k \) is an element peculiar to the \( k^{th} \) year of birth, \( (ab)_{jk} \) is an effect possessed by all observations in the \( j^{th} \) season and \( k^{th} \) year of birth, \( s_{im} \) is an effect common to all daughters of the \( m^{th} \) sire in the \( l^{th} \) group (the index \( l \) refers to two groups: Group 1 represents all non-AI and residual-AI bulls, Group 2 represents the 40 bulls that were fitted uniquely), and \( e_{ijklmn} \) is a random error peculiar to the \( n^{th} \) daughter of the \( m^{th} \) sire.

Table 1 contains the rank correlations among herdmate comparison and least-squares sire proofs, for both sets of age factors and for milk and fat. In addition, rank correlations between all ME measures of breeding value and least-squares proofs for actual milk and fat yields are presented.

The comparisons of primary interest in Table 1 are those between proofs based on the different age factors for the same method of computation. For example, \( r_2 \) is the rank correlation for proofs based on ME milk-yield and computed by herdmate comparisons for the two alternative age-adjustment procedures. This correlation is 0.99. In fact, all similar correlations (yield trait held constant) between sire evaluations based on the different age factors are 0.99. Apparently, the use of different age adjustment factors had essentially no impact upon the relative standing of the 40 sires.

The new DHIA factors apply different percentage corrections to milk and fat and to the two different seasons of calving (November-June, July-October). The Kendrick factors are almost the same as the new DHIA factors for milk in the November-June season in the Midwest. That the use of different ME factors would alter ranking very little can be inferred from the high rank correlations between ME proofs and those based on actual production, regardless of method of computation. For example, the rank correlations between least-squares proofs based on actual milk yield and those computed from ME milk yield by the herdmate comparison are about .95. The correlations of proofs based on actual yield with ME measures of breeding value computed by least squares are around .99. This indicates that bulls which enter service during the same general period of time will rank about the same on actual first lactation production as on an ME basis. This supports the generally held view that true contemporay comparisons (based on only first lactations) will minimize ranking errors arising from inaccuracies of age factors.

It must be remembered that these results do not apply to bulls entering service at wider intervals of time, or to proofs compiled for all available lactations for bulls that have been in service for widely different periods of time.

Effects on analysis of variance. Table 2 presents the means, standard deviations, and coefficients of variation for the various production measures. The means for ME production based

### Table 1

<table>
<thead>
<tr>
<th></th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.79</td>
<td>.99</td>
<td>.79</td>
<td>.94</td>
<td>.78</td>
<td>.93</td>
<td>.76</td>
<td>.94</td>
</tr>
<tr>
<td>2</td>
<td>.80</td>
<td>.99</td>
<td>.77</td>
<td>.96</td>
<td>.76</td>
<td>.96</td>
<td>.78</td>
<td>.97</td>
</tr>
<tr>
<td>3</td>
<td>.81</td>
<td>.95</td>
<td>.80</td>
<td>.94</td>
<td>.79</td>
<td>.95</td>
<td>.81</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>.78</td>
<td>.96</td>
<td>.77</td>
<td>.97</td>
<td>.78</td>
<td>.96</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>.83</td>
<td>.99</td>
<td>.82</td>
<td>.99</td>
<td>.84</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>.83</td>
<td>.99</td>
<td>.83</td>
<td>.99</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>.82</td>
<td>.99</td>
<td>.84</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>.99</td>
<td>.99</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*1 = Herdmate milk (new DHIA factors).*
*2 = Herdmate fat (new DHIA factors).*
*3 = Herdmate milk (Kendrick factors).*
*4 = Herdmate fat (Kendrick factors).*
*5 = Least-squares milk (new DHIA factors).*
*6 = Least-squares fat (new DHIA factors).*
*7 = Least-squares milk (Kendrick factors).*
*8 = Least-squares fat (Kendrick factors).*
*9 = Least-squares milk (actual production).*
*10 = Least-squares fat (actual production).*

J. DAIRY SCIENCE VOL. 51, NO. 3
TABLE 2
Means, standard deviations, and coefficients of variation*

<table>
<thead>
<tr>
<th>Trait</th>
<th>Mean</th>
<th>SD</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>ME milk—</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>new DHIA factors</td>
<td>6,310</td>
<td>1,034</td>
<td>16.38</td>
</tr>
<tr>
<td>Kendrick factors</td>
<td>6,405</td>
<td>1,047</td>
<td>16.34</td>
</tr>
<tr>
<td>ME milk—</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kendrick factors</td>
<td>224</td>
<td>35.6</td>
<td>15.87</td>
</tr>
<tr>
<td>ME fat—</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kendrick factors</td>
<td>233</td>
<td>36.9</td>
<td>15.82</td>
</tr>
</tbody>
</table>

* Production in kilogram units.

on the new DHIA factors are lower than those computed for mature yield estimated by the Kendrick factors. This is expected because the new DHIA factors are lower, on the average, for ages under 36 months in the Midwest region. The standard deviations were lower for ME yield estimated using the new DHIA factors, but the coefficients of variation were virtually identical, indicating that variability among records changed only as a result of changes in the mean.

In the computation of least-squares estimates of sire effects, the model also contained season and birth-year effects. The model was fitted on a within herd and year of calving basis. Two seasons (November-April and May-October) and four birth-year categories (1958-61) were employed. The least-squares estimates of the effects are shown in Table 3 for the two different sets of age factors and for actual production. All seasonal effects were statistically significant except for actual fat production. The magnitude of the estimates of seasonal effects varied greatly according to the manner of age adjustment. Seasonal effects were small for actual production, but were larger when the new DHIA factors were used to compute ME yield. Production adjusted by the Kendrick factors showed somewhat smaller seasonal effects compared to ME production based on the new DHIA factors. The estimates of effects for the different sets of age factors are not strictly comparable because of the difference in the standard deviations.

The differences in the estimates of seasonal effects may be due in part to the presence of an age × season interaction. Van Vleck and Henderson (12), Wunder and McGilliard (13), and Syrstad (11) have reported such an interaction for the entire age range. It may be postulated that this interaction also exists within the first lactation age range. The new DHIA
age factors are different for the two seasonal
classifications and records age-adjusted by these
factors exhibited the greatest amount of sea-
sonal variation. If an age X season interaction
exists, the new DHIA factors may attribute a
portion of the age X season interaction to sea-
sons. The Kendrick factors are constant across
seasons and the ME records exhibit less sea-
sonal variance. It might be supposed that ap-
parent seasonal variation found in the analysis
of the actual records would be increased be-
cause age was ignored in the analysis, but this
apparently did not occur.

The size of the birth-year effects also varied
greatly, depending upon the mode of age ad-
justment. Birth-year differences were signifi-
cant for actual production (P < .01), for ME
yield based on the new DHIA factors (P < .05), and nonsignificant for ME production
computed with the Kendrick factors. These
relationships may be influenced by the con-
 founding of year of birth and age in the data.
Late-calving heifers would be present only in
1958 or 1959 birth years, due to the way in
which the data were selected. If the age factors
are inaccurate, birth-year comparisons may not
be an appropriate measure of genetic changes,
due to the presence of a bias arising from age
differences. The trends for both actual pro-
duction and ME yield based on the new DHIA
factors were distinctly negative. Estimates for
actual production indicate a decrease of about
48 kg of milk from 1958 to 1961, a conclusion
which is somewhat suspect. For example, the
comparison of the 1958 birth year to the 1961
birth year may be biased for actual production,
because the mean age of calving would be
higher for 1958. This bias may also be present
to a lesser extent for ME yield if age differences
remain. That age effects biased the birth-year
comparisons for actual yield was shown by re-
peating the analysis with the inclusion of a
quadratic regression on age. When this was
done there was no consistent pattern for birth-
year effects and the F-value was nonsignificant.

Effects on residual age variance and age X
year variance. The following model was fitted
to the data:

$$Y = \mu + h_{ij} + \beta_1(x_{ij} - \bar{x}) + \beta_2(x_{ijk} - \bar{x})
+ \beta_3(x_{ijkl} - \bar{x}) + \beta_4(x_{ij} - \bar{x})^2 + \beta_5(x_{ijk} - \bar{x})^2 + \beta_6(x_{ijkl} - \bar{x})^2 + \epsilon_{ijk}$$

where $h_{ij}$ refers to the joint effects of the $i$th
herd and the $j$th year of calving and $X_{ijk}$ refers
to the age at calving of the $k$th cow in the $j$th
year and the $i$th herd. According to this model
a separate quadratic regression is fitted for
each year of calving. The procedure for fitting

<table>
<thead>
<tr>
<th>Source</th>
<th>Within herds and years</th>
<th>Differences among year regressions (L)</th>
<th>Differences among year regressions (Q)</th>
<th>Average regression (L)</th>
<th>Average regression (Q)</th>
<th>Derivations from subclass regression</th>
</tr>
</thead>
</table>

- Mean squares in (kg$^2$) units.
- Indicates $P < .01$.  

J. DAIRY SCIENCE VOL. 51, No. 3
The results are presented in Table 4 for the different production criteria. In all cases the average linear and quadratic regressions on age were significant, but the individual year regressions were not significantly different from the average coefficient.

Table 4 indicates that neither set of age factors removed all of the variance associated with age at calving. The residual variance was greater for the new DHIA factors than for production based on the Kendrick factors. However, the coefficient of variation based on the error mean square was only slightly larger for the new DHIA factors (16.7 to 16.6).

The results in Table 4 indicate that there was little yearly variation in the relationship of age and first-lactation production, confirming the results of Hillers and Freeman (4). A stepwise procedure was employed to discard non-significant effects from the model. The analysis of variance and regression coefficients for the final model are given in Tables 5 and 6, respectively. In interpreting these results it must be remembered that the ranges of ages and years studied were very limited. If data characterized by wider variations in age and time are studied, interactions between ages and years may be observed.

The results in Table 5 and 6 indicate that neither set of age factors was very effective in reducing variation in first lactation records associated with age. The regressions of actual milk yield and the ME yields on age are shown graphically in Figure 1. Searle and Henderson (10) pointed out that it may not be desirable for age factors to remove all the apparent variance due to age, especially if other variables are confounded with age. Nevertheless, the present results indicate that the factors used removed very little of the age variation in production. The regression coefficients for ME yield in Table 6 are similar to those for actual production. These

<p>| TABLE 6 |
| Regression coefficients for quadratic regression of production on age* |</p>
<table>
<thead>
<tr>
<th>Production measure</th>
<th>Linear</th>
<th>Quadratic</th>
</tr>
</thead>
<tbody>
<tr>
<td>New DHIA factors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Milk</td>
<td>216 ± 47.5</td>
<td>-15.5 ± 3.72</td>
</tr>
<tr>
<td>Fat</td>
<td>7.7 ± 1.6</td>
<td>-0.1 ± 0.01</td>
</tr>
<tr>
<td>Kendrick factors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Milk</td>
<td>159 ± 48.0</td>
<td>-13.2 ± 3.76</td>
</tr>
<tr>
<td>Fat</td>
<td>5.4 ± 1.7</td>
<td>-0.1 ± 0.01</td>
</tr>
<tr>
<td>Actual production</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Milk</td>
<td>189 ± 38.1</td>
<td>-11.8 ± 3.06</td>
</tr>
<tr>
<td>Fat</td>
<td>6.6 ± 1.4</td>
<td>-0.1 ± 0.01</td>
</tr>
</tbody>
</table>

* Milk and fat coded in kilogram units.
Birth-year comparisons were estimated by least squares as a criterion of genetic changes. The results varied greatly, depending upon the mode of age adjustment. It was concluded that the comparisons may have been biased by age differences.

Quadratic regressions on age were fitted on an intra-herd-year basis. The regressions were significant for both ME measures and for actual yield. The regressions were homogeneous from one calving year to another. The residual variation was of comparable size for both ME measures, as reflected by the coefficients of variation.

These results indicate that age factors computed for an over-all population may not fit individual subsets of the population very closely. Research applications requiring precise age adjustment, such as the estimation of genetic trends, should consider other methods to account for age variation, such as regression.

References


