July 1965

Change in Variance Components Associated with Milk Records with Time and Increase in Mean Production

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


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.
Change in Variance Components Associated with Milk
Records with Time and Increase in Mean Production

L. D. VAN VLECK
Department of Animal Husbandry, Cornell University, Ithaca, New York

Abstract
First-lactation milk records of artificially sired Holstein cows in New York were analyzed by a sire-by-herd variance component analysis separately for each year from 1954 to 1962. The two variables analyzed were 305-day, 2×, mature equivalent records and the mature equivalent records expressed as deviations from the average of their herd-mates. The total variance steadily increased with change in time. Most of the increase (r = .97) was accounted for by increase in mean production. The relative increase in the sire component of variance was slightly faster than the increase in residual variance. Thus, heritability estimated from the paternal half-sib correlation tended to increase slightly. The difference in variance with time could bias the heritability estimate from daughter-dam regression upward by about 10% in these data. This bias does not account for all the difference in heritability estimates from paternal half-sib correlation (.25) and daughter-dam regression (.40) for the records analyzed as deviations. For mature equivalent records the estimates from both methods were .36.

The difference in heritability estimates of first-lactation deviations from herd-mate averages between daughter-dam regression (.44) and paternal half-sib correlation (.25) reported by Bradford and Van Vleck (1) led to the present study. Studies with records of various relatives and with records of relatives in the same and different herds have not satisfactorily explained the difference in estimates, although genetic variance due to additive by additive and higher-order additive interaction effects and environmental correlations appear to be relatively unimportant (5, 7).

The results reported here were obtained from separate analyses of the data for each year, in an attempt to eliminate confounding of sire effects with year-season effects due to date of entry into service as a source of bias in estimating the paternal half-sib correlation. Another reason for estimating variance components for each year was to determine the trends in variance. Data were available to test this only for a few years.

Data and Methods
The data are the same as those used by Van Vleck and Bradford (6), except that only the first-lactation records were used. Briefly, these data consisted of first-lactation (305-day, 2×, M.E.) milk records of artificially sired Holstein daughters and their dams made in New York herds between 1950 and 1962, available in the files of the New York Dairy Records Processing Center. Variance and covariance components were estimated for a random effects model consisting of sire, herd, sire-by-herd interaction, and residual effects for both mature equivalent records and mature equivalent records expressed as deviations from their herd-mate averages (3). These analyses were made for each year of freshening of the daughters extending from 1954 through 1962. The records of cows freshening in the summer season (April-July) of freshening were excluded, since the average summer production is lower than the two other seasons. For example, the data for 1954 included records of cows freshening from August, 1954, to March, 1955, inclusive.

Daughter-dam regressions were computed from the within herd-sire variance and covariance components. Inspection of the results suggested that the variance component associated with the sire-by-herd interaction term either should be added to the residual component or interaction term eliminated from the model. Both procedures were followed in estimating the daughter-dam regressions, although the components for sire-by-herd interaction are also given in the results.

The numbers of records, herds, sires, and sire-by-herd subclasses, as well as the average records, both mature equivalent and deviations, for the analysis for the data of each year are given in Table 1.

Results and Discussion
The components of variance for each year are given in Table 2, for both the full model
TABLE 1

Number of records and daughter and dam means for each year

<table>
<thead>
<tr>
<th>Year</th>
<th>Records</th>
<th>Sub-classes</th>
<th>Herds</th>
<th>Sires</th>
<th>Dev.</th>
<th>M.E.</th>
<th>Dev.</th>
<th>M.E.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>(kg milk)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1954</td>
<td>624</td>
<td>578</td>
<td>372</td>
<td>55</td>
<td>-49</td>
<td>5635</td>
<td>61</td>
<td>5709</td>
</tr>
<tr>
<td>1955</td>
<td>979</td>
<td>900</td>
<td>541</td>
<td>81</td>
<td>-15</td>
<td>5767</td>
<td>130</td>
<td>5776</td>
</tr>
<tr>
<td>1956</td>
<td>1,386</td>
<td>1,261</td>
<td>666</td>
<td>92</td>
<td>-73</td>
<td>5654</td>
<td>145</td>
<td>5784</td>
</tr>
<tr>
<td>1957</td>
<td>1,445</td>
<td>1,276</td>
<td>687</td>
<td>97</td>
<td>-58</td>
<td>5846</td>
<td>173</td>
<td>5890</td>
</tr>
<tr>
<td>1958</td>
<td>1,673</td>
<td>1,457</td>
<td>754</td>
<td>123</td>
<td>64</td>
<td>6010</td>
<td>176</td>
<td>5857</td>
</tr>
<tr>
<td>1959</td>
<td>2,991</td>
<td>1,790</td>
<td>850</td>
<td>139</td>
<td>127</td>
<td>6148</td>
<td>152</td>
<td>6148</td>
</tr>
<tr>
<td>1960</td>
<td>2,509</td>
<td>2,319</td>
<td>1,074</td>
<td>155</td>
<td>-17</td>
<td>6228</td>
<td>157</td>
<td>5928</td>
</tr>
<tr>
<td>1961</td>
<td>3,454</td>
<td>2,786</td>
<td>1,273</td>
<td>160</td>
<td>-24</td>
<td>6279</td>
<td>178</td>
<td>5993</td>
</tr>
<tr>
<td>1962</td>
<td>2,915</td>
<td>2,381</td>
<td>1,201</td>
<td>150</td>
<td>16</td>
<td>6660</td>
<td>178</td>
<td>6071</td>
</tr>
</tbody>
</table>

TABLE 2

Components of variance for two statistical models for milk yield expressed in 10's of kilograms

<table>
<thead>
<tr>
<th>Deviations</th>
<th>Mature Equivalent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year</td>
<td>Sire</td>
</tr>
<tr>
<td>1954</td>
<td>495</td>
</tr>
<tr>
<td>1955</td>
<td>511</td>
</tr>
<tr>
<td>1956</td>
<td>484</td>
</tr>
<tr>
<td>1957</td>
<td>794</td>
</tr>
<tr>
<td>1958</td>
<td>618</td>
</tr>
<tr>
<td>1959</td>
<td>776</td>
</tr>
<tr>
<td>1960</td>
<td>692</td>
</tr>
<tr>
<td>1961</td>
<td>608</td>
</tr>
<tr>
<td>1962</td>
<td>969</td>
</tr>
</tbody>
</table>

and the model which ignores the interaction term. The estimates are not given as percentages, since that would obscure any trends to increase or decrease in the components of variance.

The most noticeable trends are the steady increases in residual and total variance for all analyses. The question then is, What is causing the increase or decrease in the components of variance? The regressions of the variance components on the yearly mean milk yield and on time are given in Table 3. Both mean yield and time ignoring the other account for most of the increase in total variance.

The close correlation between mean yield and time as shown in Table 1 led to a consideration of the regression of the variance components on mean yield and time independent of the other by multiple regression. The multiple regression did not give any better fit than the linear regression on mean yield. The conclusion, then, is that there is a very high correlation between mean and variance in these data. The variances of records expressed as deviations have a slightly higher relationship to mean yield than do mature equivalent records.

The close correlation between mean yield and time is shown in Table 1 led to a consideration of the regression of the variance components on mean yield and time independent of the other by multiple regression. The multiple regression did not give any better fit than the linear regression on mean yield. The conclusion, then, is that there is a very high correlation between mean and variance in these data. The variances of records expressed as deviations have a slightly higher relationship to mean yield than do mature equivalent records.

The close correlation between mean yield and time is shown in Table 1 led to a consideration of the regression of the variance components on mean yield and time independent of the other by multiple regression. The multiple regression did not give any better fit than the linear regression on mean yield. The conclusion, then, is that there is a very high correlation between mean and variance in these data. The variances of records expressed as deviations have a slightly higher relationship to mean yield than do mature equivalent records.

The close correlation between mean yield and time as shown in Table 1 led to a consideration of the regression of the variance components on mean yield and time independent of the other by multiple regression. The multiple regression did not give any better fit than the linear regression on mean yield. The conclusion, then, is that there is a very high correlation between mean and variance in these data. The variances of records expressed as deviations have a slightly higher relationship to mean yield than do mature equivalent records.

The close correlation between mean yield and time as shown in Table 1 led to a consideration of the regression of the variance components on mean yield and time independent of the other by multiple regression. The multiple regression did not give any better fit than the linear regression on mean yield. The conclusion, then, is that there is a very high correlation between mean and variance in these data. The variances of records expressed as deviations have a slightly higher relationship to mean yield than do mature equivalent records.

The close correlation between mean yield and time as shown in Table 1 led to a consideration of the regression of the variance components on mean yield and time independent of the other by multiple regression. The multiple regression did not give any better fit than the linear regression on mean yield. The conclusion, then, is that there is a very high correlation between mean and variance in these data. The variances of records expressed as deviations have a slightly higher relationship to mean yield than do mature equivalent records.

The close correlation between mean yield and time as shown in Table 1 led to a consideration of the regression of the variance components on mean yield and time independent of the other by multiple regression. The multiple regression did not give any better fit than the linear regression on mean yield. The conclusion, then, is that there is a very high correlation between mean and variance in these data. The variances of records expressed as deviations have a slightly higher relationship to mean yield than do mature equivalent records.

The close correlation between mean yield and time as shown in Table 1 led to a consideration of the regression of the variance components on mean yield and time independent of the other by multiple regression. The multiple regression did not give any better fit than the linear regression on mean yield. The conclusion, then, is that there is a very high correlation between mean and variance in these data. The variances of records expressed as deviations have a slightly higher relationship to mean yield than do mature equivalent records.

The close correlation between mean yield and time as shown in Table 1 led to a consideration of the regression of the variance components on mean yield and time independent of the other by multiple regression. The multiple regression did not give any better fit than the linear regression on mean yield. The conclusion, then, is that there is a very high correlation between mean and variance in these data. The variances of records expressed as deviations have a slightly higher relationship to mean yield than do mature equivalent records.

The close correlation between mean yield and time as shown in Table 1 led to a consideration of the regression of the variance components on mean yield and time independent of the other by multiple regression. The multiple regression did not give any better fit than the linear regression on mean yield. The conclusion, then, is that there is a very high correlation between mean and variance in these data. The variances of records expressed as deviations have a slightly higher relationship to mean yield than do mature equivalent records.

The close correlation between mean yield and time as shown in Table 1 led to a consideration of the regression of the variance components on mean yield and time independent of the other by multiple regression. The multiple regression did not give any better fit than the linear regression on mean yield. The conclusion, then, is that there is a very high correlation between mean and variance in these data. The variances of records expressed as deviations have a slightly higher relationship to mean yield than do mature equivalent records.

The close correlation between mean yield and time as shown in Table 1 led to a consideration of the regression of the variance components on mean yield and time independent of the other by multiple regression. The multiple regression did not give any better fit than the linear regression on mean yield. The conclusion, then, is that there is a very high correlation between mean and variance in these data. The variances of records expressed as deviations have a slightly higher relationship to mean yield than do mature equivalent records.

The close correlation between mean yield and time as shown in Table 1 led to a consideration of the regression of the variance components on mean yield and time independent of the other by multiple regression. The multiple regression did not give any better fit than the linear regression on mean yield. The conclusion, then, is that there is a very high correlation between mean and variance in these data. The variances of records expressed as deviations have a slightly higher relationship to mean yield than do mature equivalent records.
TABLE 3

Results of regression of variance (10 kg) on yearly mean (kg) and time (years)

<table>
<thead>
<tr>
<th>Component of variance</th>
<th>On mean</th>
<th>On time</th>
<th>On mean and time</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(b_{\text{VM}})</td>
<td>(s_{\text{VM}})</td>
<td>(r_{\text{VM}})</td>
</tr>
<tr>
<td>Deviations:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sire</td>
<td>.84</td>
<td>.25</td>
<td>.79</td>
</tr>
<tr>
<td>Herd</td>
<td>-.16</td>
<td>.86</td>
<td>-.07</td>
</tr>
<tr>
<td>(S \times H)</td>
<td>2.21</td>
<td>3.48</td>
<td>.23</td>
</tr>
<tr>
<td>Error</td>
<td>8.65</td>
<td>3.39</td>
<td>.69</td>
</tr>
<tr>
<td>Total</td>
<td>11.52</td>
<td>1.05</td>
<td>.97</td>
</tr>
<tr>
<td>Mature Equivalent:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sire</td>
<td>1.30</td>
<td>.45</td>
<td>.73</td>
</tr>
<tr>
<td>Herd</td>
<td>.25</td>
<td>.91</td>
<td>.10</td>
</tr>
<tr>
<td>(S \times H)</td>
<td>-1.71</td>
<td>3.05</td>
<td>-.21</td>
</tr>
<tr>
<td>Error</td>
<td>10.34</td>
<td>2.96</td>
<td>.80</td>
</tr>
<tr>
<td>Total</td>
<td>10.18</td>
<td>1.90</td>
<td>.90</td>
</tr>
<tr>
<td>Models:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model includes sire, herd, sire-by-herd, and residual effects.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Why the sire component is generally larger for mature equivalent records than for deviations is somewhat puzzling. A possibility is that there are some herd-by-season or month-of-freshening interaction effects, since the year for these analyses consisted of eight continuous months. Confounding of entry date into service for sires with season effects may lead to part of the interaction variance being forced to the sire component for the mature equivalent analyses. Expressing records as deviations from herd-mate averages should remove most of the season effects and also most of the herd-by-season interaction effects.

The within herd estimates of heritability from paternal half-sib correlations and daughter-dam regression are shown in Table 4. The regressions of these estimates on time were all

TABLE 4

Heritability estimates within herd from paternal half-sib correlation and within herd-sire daughter-dam regression

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Analysis: Model</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deviations:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>((S, H, S \times H, E))</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pat. sib</td>
<td>.24</td>
<td>.20</td>
<td>.21</td>
<td>.34</td>
<td>.27</td>
<td>.28</td>
<td>.22</td>
<td>.19</td>
<td>.28</td>
<td>.25</td>
<td>.0015</td>
</tr>
<tr>
<td>Dau.-dam</td>
<td>.24</td>
<td>.44</td>
<td>.44</td>
<td>.55</td>
<td>.40</td>
<td>.34</td>
<td>.40</td>
<td>.32</td>
<td>.49</td>
<td>.40</td>
<td>.0058</td>
</tr>
<tr>
<td>((S, H, E))</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pat. sib</td>
<td>.26</td>
<td>.19</td>
<td>.25</td>
<td>.32</td>
<td>.27</td>
<td>.30</td>
<td>.24</td>
<td>.20</td>
<td>.30</td>
<td>.26</td>
<td>.0025</td>
</tr>
<tr>
<td>Dau.-dam</td>
<td>.20</td>
<td>.37</td>
<td>.43</td>
<td>.51</td>
<td>.39</td>
<td>.32</td>
<td>.40</td>
<td>.34</td>
<td>.47</td>
<td>.38</td>
<td>.0123</td>
</tr>
<tr>
<td>Mature equivalent:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>((S, H, S \times H, E))</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pat. sib</td>
<td>.19</td>
<td>.28</td>
<td>.36</td>
<td>.55</td>
<td>.40</td>
<td>.44</td>
<td>.33</td>
<td>.33</td>
<td>.40</td>
<td>.36</td>
<td>.0137</td>
</tr>
<tr>
<td>Dau.-dam</td>
<td>.22</td>
<td>.49</td>
<td>.41</td>
<td>.37</td>
<td>.40</td>
<td>.33</td>
<td>.33</td>
<td>.23</td>
<td>.47</td>
<td>.36</td>
<td>.0003</td>
</tr>
<tr>
<td>((S, H, E))</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pat. sib</td>
<td>.21</td>
<td>.27</td>
<td>.41</td>
<td>.52</td>
<td>.39</td>
<td>.44</td>
<td>.34</td>
<td>.33</td>
<td>.41</td>
<td>.37</td>
<td>.0127</td>
</tr>
<tr>
<td>Dau.-dam</td>
<td>.20</td>
<td>.42</td>
<td>.39</td>
<td>.37</td>
<td>.39</td>
<td>.32</td>
<td>.37</td>
<td>.27</td>
<td>.46</td>
<td>.35</td>
<td>.0083</td>
</tr>
</tbody>
</table>
nonsignificant, although the yearly increase for the estimate from half-sib correlation for mature equivalent records was about 1.3%, whereas for deviations the increase as shown by the regression coefficient was about 0.2%. The regressions for daughter-dam estimates on time were also positive, although not consistent from mature equivalent to deviation records or from the full model to the reduced model.

The unweighted averages of the heritability estimates show the pattern found by Bradford and Van Vleck (1) and Van Vleck and Bradford (5, 6). For deviations, the daughter-dam estimates are substantially higher than the paternal half-sib estimates. The two methods of estimation from the mature equivalent records provide almost identical estimates only slightly lower than the daughter-dam estimates for deviations. Whether there is enough confounding of sires with season effects to increase the half-sib estimates from about .25 to .36 is unknown. If the sire component is inflated with less than 3% of the total variance, this could easily account for the increase from .25 to .36. Van Vleck and Bradford (6) reported estimates considerably higher than these from half-sib analyses of mature equivalent records. Their analyses did not consider season effects and their data extended over the entire range of that used here.

If the true paternal half-sib estimate is .25, why is the daughter-dam estimate so much higher? A maternal effect is a possible cause and was postulated by Van Vleck and Bradford (6). The analyses of Van Vleck and Hart (7), however, tend to dispute this conclusion, although they presented evidence that environmental correlation probably accounted for no more than .02 of the difference between the paternal half-sib and daughter-dam estimates.

Change in variance may account for part of the difference through the daughter-dam covariance, although change in variance may in some way also affect the estimation of the sire component. The latter is unlikely in the present study because of estimating components of variance for each year of the data. For example, the average yearly change in within herd and sire variance of deviations was 561 for the reduced model. The variance of daughters for 1954 was 7,459. Van Vleck and Hart (7) reported an average time difference between daughter and dam records of 49 months. An estimate of the variance of daughters 4 yr later would be 7,459 + 4(561) = 9,703 (not quite what the 1958 value really was) or 125% of the variance of the dam’s generation. The covariance between the 1958 daughters and their 1954 dams would be accordingly increased by (1.25)^{1/2} = 1.12 over what it would have been if the variance had not changed. A corrected estimate of heritability could be obtained by dividing the biased estimate by 1.12. This could account for about .04 of the difference in heritability estimates. If environmental correlation also accounts for .02-.04 of the difference, the daughter-dam heritability estimate can be reduced to .30-.32, not much different from the half-sib estimate for deviations.

There was evidence for selection of dams based on the higher first-lactation average deviations of the dams over their daughters. This does not agree with the recent summary and results of Miller and Corley (4). The daughter-dam regression method of estimating heritability takes this selection into account. The apparent reduction in variance of dam records contrasted to daughter variance in the reports of Bradford and Van Vleck (1) and Van Vleck and Bradford (5, 6) was probably not entirely due to selection but partially due to the increase in variance with time.

Conclusions
There has been a steady and substantial increase with time in the within herd-sire and total variances for both mature equivalent records and deviations from herd-mate averages. Much of this increase can be accounted for by the increase in mean production. The sire (genetic) component of variance has also increased steadily and relatively, although only slightly, faster than the increase in residual variance. Thus, apparent heritability from paternal half-sib correlation has tended to increase slightly. Variance components due to herd and sire-by-herd interaction effects have changed relatively little. Estimates of heritability from daughter-dam regression have also tended to increase slightly. The daughter-dam regression can be shown to be highly susceptible to bias, due to the change in variance from the time of the dams’ generation to the daughters’ generation. The apparent upward bias is about 10%. Even this bias and a small bias due to environmental correlation between daughter and dam records do not account for all the difference in heritability estimates between daughter-dam regression and paternal half-sib correlation. A general statistical rule is that if the mean is directly related to the variance an appropriate transformation of the data should be made. The original rule applied generally to variables following a Poisson distribution. Whether such a transformation is appropriate or desirable for milk records is probably debatable. Dickerson
(2) has shown, however, that unequal variances in different environments can give misleading estimates of components of variance, due to genetic-by-environmental interaction effects. A suitable transformation may, therefore, reduce the sire-by-herd components found in this study (considered unimportant). The over-all conclusion is that transformations should be tried, to see if scale factors are causing the difference in estimates of heritability from daughter-dam regression and paternal half-sib correlation.

References