ABSTRACT

The nonlinear effects of inbreeding were studied by comparing linear and curvilinear regression models of phenotypic performances on inbreeding coefficients for production traits (milk, fat, and protein yields) of Holstein cows in their first lactation. Three different regression models (linear, quadratic, and cubic) were introduced separately into a single-trait, single-lactation, random regression test-day model. The significance of the different regression coefficients was studied based on a t-test after estimation of error variances and covariances associated with the different regression coefficients. All of the tested regression coefficients were significantly different from 0. The traditional regression coefficients of milk, fat, and protein yields on inbreeding were, respectively, $-22.10$, $-1.10$, and $-0.72$ kg for Holstein cows in their first lactation. However, the estimates of 305-d production losses for various classes of animals based on inbreeding coefficients showed that the effect of inbreeding was not a linear function of the percentage of inbreeding. The 305-d milk yield loss profiles attributable to inbreeding, obtained by the various regression models, were different. However, for inbreeding coefficients between 0 and 10%, these differences were small.

Key words: inbreeding depression, curvilinear regression, confidence interval, genetic evaluation

INTRODUCTION

Matings between related individuals lead to the reduction of the population mean phenotypic value for quantitative traits such as size, fertility, yield, vigor, and fitness. This phenomenon is better known as inbreeding depression. The loss of the mean value for a trait with the degree of inbreeding depends on the mechanism for the genetic basis of inbreeding depression. Two major hypotheses have been put forward to explain the genetic basis of inbreeding depression. According to the first hypothesis, the partial dominance hypothesis (Davenport, 1908; Bruce, 1910; Keeble and Pellew, 1910; Jones, 1917), inbreeding increases the frequency of homozygotes; thus, inbreeding depression is caused by the expression of deleterious recessive alleles in homozygous individuals. The second hypothesis, the overdominance hypothesis (East, 1908), supposes that heterozygotes are superior to both homozygotes. A reduction in the frequency of heterozygotes attributable to inbreeding will reduce opportunities to express this overdominance (Charlesworth and Charlesworth, 1999). Under these hypotheses, the inbreeding depression for a trait is a linear function of the inbreeding coefficient (Lynch and Walsh, 1998). In addition to these 2 hypotheses, the epistatic interactions between genes can also be considered as a third mechanism to explain the genetic basis of inbreeding depression. If such interactions are involved, the decline of the mean phenotypic value can be considered a nonlinear function of the inbreeding coefficient (Lynch and Walsh, 1998).

Inbreeding depression represents the main economically important consequence of the increase in inbreeding. Indeed, several studies have shown an unfavorable association between performances for production traits and inbreeding.

Because of inbreeding depression, production traits have been shown to decrease significantly, with estimates ranging from approximately 9 to 26 kg of milk, 0.55 to 1.09 kg of fat, and 0.69 to 0.97 kg of protein per lactation and for each 1% increase in inbreeding (Casanova et al., 1992; Miglior et al., 1992, 1995; Wiggans et al., 1995; Smith et al., 1998; Croquet et al., 2006). The effects of inbreeding on mean phenotypic performances can be accounted for in genetic evaluations by regressions on a measure of inbreeding called the inbreeding coefficient, which is estimated a priori from the pedigree (Wiggans et al., 1995). However, doing this will indirectly favor sires with a higher relationship to the dam population. A correction for the contri-
bution of future inbreeding in each mating is then required. Usually, a linear regression of phenotypic performances on inbreeding coefficients is used to estimate inbreeding depression for a trait in genetic evaluation models. However, Thompson et al. (2000) and Mignior et al. (1992) showed the potential nonlinearity of inbreeding depression for production traits when the effect of inbreeding was included in the genetic evaluation model as a classification effect. For Mignior et al. (1992), the intensity of inbreeding depression for production traits of Canadian Jerseys was higher when inbreeding coefficients were greater than 12.5%. Thompson et al. (2000) showed that milk production losses caused by inbreeding for US Holsteins were generally 35 kg per 1% increase in inbreeding but increased to 55 kg per 1% increase in inbreeding from inbreeding coefficients of between 7 and 10%. Despite these results, few studies have tested linear and curvilinear regressions in real-life situations and on large data sets. The principal objective of this paper was to study the potential curvilinearity of inbreeding depression as a function of the inbreeding coefficient. A comparison was made between the different regression models of phenotypic performances on inbreeding coefficients for milk production traits for Holstein cows in their first lactation.

MATERIALS AND METHODS

Pedigree File

The pedigree information used for this study included data from the official Walloon genetic evaluations of February 2004. The pedigree file contained information for 956,516 animals born between 1913 and 2004. The pedigree file was improved by estimating 61,501 unknown birth years and 12,354 unknown breed compositions from known birth years and breed compositions of ancestors or descendants. In total, the pedigree file contained 57.6% Holsteins, 18.1% Dual-Purpose Belgian Blues, and 24.3% other breeds and crossbreeds. More details about this pedigree can be found in Croquet et al. (2006).

Production Data

Only data for Holstein cows in their first lactation were analyzed in this study. A total of 3,491,814 first-lactation test-day records, corresponding to 403,582 Holstein animals, were obtained from the Walloon Breeding Association (AWE), which manages performance recording data in the Walloon Region of Belgium. The period of time covered by the data was 1974 to 2004. These data are part of the production data set used for the official Walloon genetic evaluations of February 2004. Basic statistics are shown in Table 1.

Table 1. Data characteristics for production traits of first-parity Holstein cows included in the Walloon genetic evaluations of February 2004

<table>
<thead>
<tr>
<th>Production trait</th>
<th>Records</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Milk 305-d, kg</td>
<td>3,491,814</td>
<td>5,504.55</td>
<td>1,923.59</td>
</tr>
<tr>
<td>Fat 305-d, kg</td>
<td>3,441,821</td>
<td>224.60</td>
<td>81.46</td>
</tr>
<tr>
<td>Protein 305-d, kg</td>
<td>3,438,833</td>
<td>179.84</td>
<td>60.77</td>
</tr>
</tbody>
</table>

Estimation of Inbreeding Depression

The standard procedure used to estimate inbreeding depression for a trait is regression of individual performances on the individual pedigree inbreeding coefficients. The inbreeding coefficients were computed according to the methodology previously described in Croquet et al. (2006). This method allows an inbreeding coefficient to be attributed to individuals with unknown parents based on the mean inbreeding coefficient of contemporary and genetically similar individuals with known parents. The effects of inbreeding were estimated from a single-lactation, single-trait random regression test-day model. In matrix notation, the general single-trait model was

\[ y = X\beta + Q(W + Za + Zp) + e \]

where \( y \) is a vector of milk, fat, or protein test-day records; \( \beta \) is a vector of fixed effects (herd \times test date, stage of lactation, stage of lactation \times age at calving \times season of calving, gestation stage); \( h \) is a vector of herd \times period of calving random regression coefficients; \( a \) is a vector of additive genetic random regression coefficients; \( p \) is a vector of permanent environmental random regression coefficients; \( e \) is a vector of random residuals; \( X, W, \) and \( Z \) are incidence matrices; and \( Q \) is the covariate matrix for the second-order Legendre polynomials. To adjust for maturity differences within lactation and breed composition, data were preadjusted to account for environmental effects of age nested within lactation \times stage \times breed classes. More details are available in Auvray and Gengler (2002).

The estimates of inbreeding depression were obtained by including the inbreeding coefficients of animals as linear, quadratic, and cubic covariates in the general single-trait model. A total of 9 models (3 regression models \times 3 traits) were tested. The equations were solved by using the iterative method of the preconditioned conjugate gradient (Tsuruta et al., 2001). In addition, the effect of inbreeding was included in the same genetic evaluation model as a classification variable in
order to provide an alternative way to check the linearity.

**Estimation of Standard Errors and Model Validation**

A regression model was validated by conducting a significance test of the regression coefficients and by calculating the confidence interval around the regression curves. The validation step first required determination of the error variances and covariances associated with the different regression coefficients. The method used to approximate these variances and covariances was presented by Gengler and Croquet (2005) and is explained in more detail in Croquet et al. (2006). This method is based on mixed-model conjugate normal equations associated with a preconditioned conjugate gradient solver and allows the variances associated with given effects (in this study, fixed regression coefficients on inbreeding) to be obtained. A t-test was then used to test the significance of this regression coefficient:

\[ t_{obs} = \frac{k' \hat{\beta}}{\sqrt{k'C^{-1}k}} \]

where \( \hat{\beta} \) is the vector containing estimated fixed effects, and \( k \) is a contrast vector replacing the right-hand-side vector. A property of mixed models \( V(\hat{\beta}) = C^{\beta} \) was used, where \( C^{\beta} \) is the block of \( C^{-1} \) corresponding to fixed effects. The number of degrees of freedom of the \( t \) variable can be approximated by N-rank(\( X \)) under the hypothesis that exact residual and genetic variances are known (L. D. Van Vleck, University of Nebraska–Lincoln, personal communication).

**RESULTS AND DISCUSSION**

**Distribution of Animals per Class of Inbreeding Coefficient**

The distribution of Holstein cows in their first lactation per class of inbreeding and basic statistics on inbreeding coefficients for the different classes of animals are presented in Table 2. Classes were chosen to let the average inbreeding of every class, except the first one, reflect levels as 3.125%, 6.25%, 12.5%, and 25%. This table shows that the distribution of animals was not homogeneous by class of inbreeding. Most of the cows (more than 99%) were in the lower classes, with an inbreeding coefficient of <10%. Different hypotheses could explain the small percentage of animals with a higher inbreeding coefficient. The first one is the desire of many breeders to avoid mating close relatives. The second hypothesis is the higher susceptibility of highly inbred animals to early culling. The third hypothesis is related to the incompleteness of the pedigree file. Indeed, Table 2 presents the mean equivalent number of known generations by class of inbreeding. This parameter is the sum of the known proportions of ancestors throughout the generations (8 generations in this study) and constitutes an indicator of the pedigree depth for animals with known parents (Boichard et al., 1996). The analysis in Table 2 shows that the mean equivalent number of known generations is lower for the first 2 classes of inbreeding. Despite our method for estimating inbreeding coefficients, the poor quality of genealogical information could partially explain the results. The maximum inbreeding coefficient of a Holstein cow in her first lactation was 39.18%. This value corresponds to the inbreeding coefficient of an animal that came from a family with atypical mating because her sire was also her maternal grandsire and her maternal great grandsire.

Recognizing the distribution of animals in the different classes of inbreeding is important. Indeed, it will influence the range of inbreeding where the interpretation of the results can be considered relevant. For this study, we considered a range of inbreeding between 0 and 10%.

**Comparison of the 3 Regression Models**

Table 3 shows estimates of 305-d production losses attributable to inbreeding for various classes of Holstein cows in their first lactation, distributed on the basis of inbreeding coefficients. The inbreeding classes in Table 3 correspond exactly to the inbreeding classes presented in Table 2. The results for milk, fat, and protein in Table 3 were obtained when the effect of inbreeding was included in the genetic evaluation model as a classification variable. These results are in accordance with the results obtained by Thompson et al. (2000) and Miglior et al. (1992), which showed the nonlinearity of inbreeding depression for production traits for US Holsteins and Canadian Jerseys, respectively. However, given the small percentage of animals with inbreeding coefficients greater than 10%, as shown in Table 2, one should be cautious not to overinterpret results for inbreeding depression for the last 2 classes of inbreeding. Table 4 presents the different regression coefficients and their relative estimated \( t \)-values for the 3 regression models and for the 3 production traits. The estimated \( t \)-values were used to test the level of significance. One can observe from this table that all of these coefficients were very highly significant except the second-degree regression coefficient of the cubic model describing the impact of inbreeding on fat. The results in Table 4 also show that, even if all the different
regression coefficients were significantly different from 0. t-values for the traditional linear regression coefficients for the 3 production traits were higher than t-values for other regression coefficients. This observation could suggest that the traditional linear model is the best model to estimate the impact of inbreeding on production traits.

Because the decline of the mean phenotypic values in the function of the inbreeding coefficient have the same trend for the 3 production traits, only figures for milk yield are presented in this study. Figures 1, 2, and 3 present the cumulative losses, expressed in 305-d milk yields for the first lactation of Holstein cows, in the function of the inbreeding coefficient. For each of these regression models, a confidence interval, computed using the theoretic t-value corresponding to a 95% significance level, is also shown. This allows the dispersion of the estimated regression curves to be visualized. Figures 1, 2, and 3 also show that the linear and the 2 studied curvilinear regressions of phenotypic performances for milk yield on the inbreeding coefficient were significant, because the confidence intervals never crossed the x-axis between 0 and 25% inbreeding.

The underlying hypothesis in Figure 1 is that the 305-d milk yield losses per 1% increase in inbreeding are the most important for inbreeding coefficients below 5%. The losses in 305-d milk yield per 1% increase in inbreeding diminished when the inbreeding coefficient increased above 5%. For the cubic regression model presented in Figure 3, the 305-d milk yield losses per 1% increase in inbreeding first increased with the inbreeding coefficients and then decreased after 8% of inbreeding. However, Figure 4, which presents inbreeding depressions in milk yield estimated by the 3 regression models on the same graph, shows that between 0 and 10% of inbreeding, the differences between the different regression models are small. One can observe from Figure 4 that below 6% of inbreeding, the inbreeding depression estimated from a quadratic regression seemed to be higher and the inbreeding depression estimated from a cubic regression seemed to be smaller than the inbreeding depression estimated from a traditional linear regression. For inbreeding coefficients ranging between 6 and 10%, the inbreeding depression estimated from a cubic regression seemed to be higher than the inbreeding depression estimated from quadratic and linear regressions.

In this study, the weak differences observed between the 3 regression models of phenotypic performances for inbreeding coefficients between 0 and 10% could be due to the low number of animals with an inbreeding coefficient greater than 10%. Because animals with an inbreeding coefficient higher than 10% have a greater impact on the curvilinear regression, one could suppose that the differences between the 3 regression models would be accentuated if such animals were more numerous.

However, the shape of the loss of the mean value for a trait with the degree of inbreeding is difficult to relate directly to the different mechanisms on the genetic ba-

---

**Table 2.** Number and percentage of Holstein cows in their first lactation with associated descriptive statistics (expressed in %) by class of inbreeding and average equivalent number of known generations by class of inbreeding.

<table>
<thead>
<tr>
<th>Class of inbreeding</th>
<th>Number of animals</th>
<th>Cumulative percentage</th>
<th>Average inbreeding, %</th>
<th>SD, $^1$ %</th>
<th>Average equivalent number of known generations</th>
</tr>
</thead>
<tbody>
<tr>
<td>$F = 0$</td>
<td>104,263</td>
<td>25.8</td>
<td>0.000</td>
<td>0.000</td>
<td>3.85</td>
</tr>
<tr>
<td>$0 &lt; F &lt; 2$</td>
<td>180,060</td>
<td>70.5</td>
<td>0.674</td>
<td>0.575</td>
<td>2.97</td>
</tr>
<tr>
<td>$2 \leq F &lt; 5$</td>
<td>98,634</td>
<td>94.9</td>
<td>3.060</td>
<td>0.787</td>
<td>5.46</td>
</tr>
<tr>
<td>$5 \leq F &lt; 10$</td>
<td>18,841</td>
<td>99.6</td>
<td>6.588</td>
<td>1.220</td>
<td>8.60</td>
</tr>
<tr>
<td>$10 \leq F &lt; 20$</td>
<td>1,550</td>
<td>99.9</td>
<td>12.426</td>
<td>2.061</td>
<td>8.54</td>
</tr>
<tr>
<td>$F \geq 20$</td>
<td>234</td>
<td>100.0</td>
<td>26.032</td>
<td>1.761</td>
<td>6.56</td>
</tr>
</tbody>
</table>

$^1$Standard deviation of the average inbreeding.

---

**Table 3.** Estimate of inbreeding depression for milk, fat, and protein yields of Holstein cows in their first lactation for different classes of inbreeding.

<table>
<thead>
<tr>
<th>Class of inbreeding (%)</th>
<th>Inbreeding depression (305-d), kg</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Milk</td>
</tr>
<tr>
<td>0 &lt; $F &lt; 2$</td>
<td>−28.49</td>
</tr>
<tr>
<td>2 ≤ $F &lt; 5$</td>
<td>−65.91</td>
</tr>
<tr>
<td>5 ≤ $F &lt; 10$</td>
<td>−168.52</td>
</tr>
<tr>
<td>10 ≤ $F &lt; 20$</td>
<td>−285.80</td>
</tr>
<tr>
<td>$F \geq 20$</td>
<td>−314.73</td>
</tr>
</tbody>
</table>
Table 4. Estimates of coefficients for linear, quadratic, and cubic regressions describing the impact of inbreeding on milk, fat, and protein and associated t-values

<table>
<thead>
<tr>
<th>Tested regression coefficient</th>
<th>Milk Regression coefficient, $\times 10^{-2}$</th>
<th>Estimated t-value</th>
<th>Fat Regression coefficient, $\times 10^{-2}$</th>
<th>Estimated t-value</th>
<th>Protein Regression coefficient, $\times 10^{-2}$</th>
<th>Estimated t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cubic</td>
<td>6.966 $\pm$ 1.278</td>
<td>5.449</td>
<td>0.232 $\pm$ 0.0537</td>
<td>4.315</td>
<td>0.206 $\pm$ 0.039</td>
<td>5.243</td>
</tr>
<tr>
<td>Quadratic</td>
<td>$-166.631 \pm 39.934$</td>
<td>4.173</td>
<td>$-5.020 \pm 1.676$</td>
<td>2.996</td>
<td>$-4.682 \pm 1.226$</td>
<td>3.820</td>
</tr>
<tr>
<td>Linear</td>
<td>$-1,634.687 \pm 258.041$</td>
<td>6.335</td>
<td>$-89.973 \pm 10.814$</td>
<td>8.320</td>
<td>$-58.148 \pm 7.919$</td>
<td>7.342</td>
</tr>
<tr>
<td>(Model with cubic regression)</td>
<td>(Model with quadratic regression)</td>
<td></td>
<td>(Model with linear regression)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quadratic</td>
<td>42.297 $\pm$ 10.699</td>
<td>3.953</td>
<td>1.931 $\pm$ 0.448</td>
<td>4.309</td>
<td>1.503 $\pm$ 0.328</td>
<td>4.582</td>
</tr>
<tr>
<td>Linear</td>
<td>$-2,727.475 \pm 159.078$</td>
<td>17.146</td>
<td>$-126.390 \pm 6.664$</td>
<td>18.966</td>
<td>$-90.692 \pm 4.880$</td>
<td>18.585</td>
</tr>
<tr>
<td>(Model with quadratic regression)</td>
<td>(Model with linear regression)</td>
<td></td>
<td>(Model with linear regression)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linear</td>
<td>$-2,212.680 \pm 93.970$</td>
<td>23.547</td>
<td>$-102.914 \pm 3.937$</td>
<td>26.143</td>
<td>$-72.394 \pm 2.882$</td>
<td>25.123</td>
</tr>
</tbody>
</table>

sis of inbreeding depression. Indeed, as signaled by Kristensen and Sorensen (2005), the nonlinearity of inbreeding depression is difficult to interpret because different elements exist that are not directly related to inbreeding. Thus, according to Kristensen and Sorensen (2005), the first hypothesis that may explain the deviation from the linear regression may be a shift in the environment, especially if the distribution of animals among the different inbreeding classes is influenced by time. The use of a mixed model and the best linear unbiased estimator property of our inbreeding depression estimates should allow us to avoid confusion between inbreeding and time trends. However, as we regress on estimated inbreeding, the conditions for best linear unbiased estimator are not totally fulfilled and we can therefore not totally exclude this hypothesis. The second element that may disrupt the interpretation of the curvilinearity of inbreeding depression is the assumption of independence and homoscedasticity across inbreeding coefficients (Lynch, 1988). A third element, pointed out by Weir et al. (1980), is the deviation of the expected proportion of loci identical by descent from this expectation. According to these authors, if inbreeding depression affects survival in the population, the inbreeding coefficient calculated from the pedigree can be expected to be larger than the actual proportion of loci identical by descent. If true, this would then lead to the nonlinearity of inbreeding depression. Finally, the nonlinearity of 305-d milk yield losses per 1% increase in inbreeding could also be explained by the survival of superior animals with high inbreeding coefficients.

CONCLUSIONS

This paper studied the nonlinearity of inbreeding effects by comparing linear, quadratic, and cubic regres-

Figure 1. Losses in milk yield attributable to inbreeding depression calculated by a linear regression of phenotypic performances on the inbreeding coefficient and confidence interval.

Figure 2. Losses in milk yield attributable to inbreeding depression calculated by a quadratic regression of phenotypic performances on the inbreeding coefficient and confidence interval.
Figure 3. Losses in milk yield attributable to inbreeding depression calculated by a cubic regression of phenotypic performances on the inbreeding coefficient and confidence interval.

sion models of phenotypic performances on inbreeding coefficients for production traits (milk, fat, and protein yields) of Walloon Holstein cows in their first lactation. All of the regression coefficients appearing in the studied regression models are at least highly significant; thus, each of these regression models can be used to explain the inbreeding depression for production traits. The 305-d milk yield losses attributable to inbreeding estimated by the 3 regression models are different, especially for coefficients of inbreeding higher than 10%. However, to interpret the results for inbreeding coefficients higher than 10% correctly, it would be necessary to have a much larger number of animals with inbreeding coefficients higher than 10%. The differences among the 305-d milk production profiles of inbreeding depression may be directly related to the way in which the loss of dominance for the loci involved in the milk production varies when inbreeding increases. However, it is necessary to be very cautious in this interpretation because other very different reasons could also explain deviation from linearity, such as artifacts from unknown pedigrees or simplifications of computations. Between 0 and 10% of inbreeding, where the majority of data were situated, differences existed between the 3 regression models in the profiles of 305-d milk yield losses attributable to inbreeding depression, but they were small. These small differences may be related to the low number of animals with an inbreeding coefficient higher than 10%. A greater number of highly inbred animals may accentuate the differences between the 3 regression models. However, in practice, to facilitate the interpretation and calculation of inbreeding depression, a linear regression model is often used, and this study showed that its use is reasonable for Walloon breeders in a normal range of inbreeding (e.g., between 0 and 10%).

ACKNOWLEDGMENTS

Coraline Croquet, research fellow, and Nicolas Gengler, research associate, with the National Fund for Scientific Research (Brussels, Belgium), acknowledge their support. Additional support was provided through Grant 2.4507.02F (2) of the National Fund for Scientific Research. The authors gratefully acknowledge the support of the Walloon Breeding Association (AWE) and the Walloon Regional Ministry of Agriculture (Projects RW1009 and D31-1039).

REFERENCES


Gengler, N., and C. Croquet. 2005. Standard errors of solutions in large scale mixed models, application to linear and curvilinear...
effects of inbreeding on production traits. J. Dairy Sci. 88(Suppl. 1):74. (Abstr.)