Genotype × Environment Interaction for Protein Yield in Dutch Dairy Cattle as Quantified by Different Models

M. P. L. Calus,* A. F. Groen,* and G. de Jong†
*Animal Breeding and Genetics Group, Wageningen Institute of Animal Sciences, Wageningen University, PO Box 338, 6700 AH Wageningen, The Netherlands
†CR-Delta, PO Box 454, 6800 AL Arnhem, The Netherlands

ABSTRACT

Variance components and breeding values for protein yield were estimated with REML without and with correction for heterogeneity of variances. Three different sire models were applied, which all accounted for genotype × environment (G × E) interaction. The first model included a sire × herd-year-season subclass (HYS) interaction. The second model divided all records in four different types of management groups, based on estimated HYS subclass effect. The third model, the reaction norm model, performed a random linear regression on the estimated HYS effect. For comparison, a standard model that did not take G × E interaction into account was also applied. Data consisted of 102,899 305-d first-lactation protein records of Holstein Friesians of 1000 of the largest Dutch dairy herds. All animals calved in 1997, 1998, or 1999. Estimated breeding values (EBV) for 2150 bulls with at least five daughters were calculated. The interaction model detected an interaction variance of 2.5% of the phenotypic variance. The EBV showed a correlation of 1.00 with those of the standard model without interaction. The model with the division in groups showed correlations between groups ranging from 0.73 to 0.86. The EBV showed correlations from 0.84 to 0.91 with the EBV of the standard model. The reaction norm model calculated EBV that had a correlation of 1.00 with the EBV of the standard model. The reaction norm model was not able to detect significant variance of the slope for the protein data corrected for heterogeneity of variances. (Key words: genotype × environment interaction, protein yield, reaction norm model, dairy cattle)

Abbreviation key: G × E = genotype by environment, HYS = herd-year-season, MSE = mean squares of errors.

INTRODUCTION

In the current Dutch breeding value estimation, data are corrected for heterogeneity of variances (Meuwissen et al., 1996; Nederlands Rundvee Syndicaat, 1999). Heterogeneity of variances has two possible origins: scaling of both genetic and residual variances and genotype × environment (G × E) interaction. Scaling means that, for example, herds with higher average production or higher average breeding values also show higher standard deviations. G × E interaction is the phenomenon that different environments do not equally affect performances of different genotypes (Falconer, 1989). In breeding value estimation, G × E interaction is usually not taken into account. If not accounted for, the interaction might bias EBV and reduce effectiveness of selection, because sires might rank differently across different environments.

In this study, three different sire models were applied to quantify G × E interaction. Given phenotype P, genotype G, environment E, and genotype × environment interaction G*E, the three models can be represented as:

1. \[ P = G + E + G*E \]
2. \[ P = (G + G*E) + E \]
3. \[ P = G(E) + E \]

where the underlined terms represent the sires breeding values. These are, respectively, a model with a sire × herd-year-season (HYS) interaction, a model with HYS subclasses divided in different groups based on an environmental parameter and the reaction norm model. In this study, model 1, the interaction model, included the sire by HYS interaction as a variance component and so calculated EBV excluding G × E interaction. Model 2 calculated EBV per group of environments; the G × E interaction becomes part of the breeding value per group. Model 3 calculated EBV as a linear function, the reaction norm, of an environmental parameter. This model is only recently introduced for use in dairy cattle breeding value estimation (Strandberg et al., 2000). A reaction norm is the environmental sensitivity of a ge-
notype in different environments (Falconer, 1990). The EBV of model 3 are built up of two parts: an environment-independent part, the ‘level’, and an environment-dependent part, the ‘slope.’ In fact, the level is the random intercept of the reaction norm, and the slope is the random linear coefficient of the random regression on the environmental parameter. This model was applied by using a random regression model (Strandberg et al., 2000) rather than a repeatability model (Ravagnolo and Misztal, 2000).

The results of all models were compared to results of a standard model that did not account for \( G \times E \) interaction.

The objective of this study was to compare the different models for their ability to account for \( G \times E \) interaction.

**MATERIALS AND METHODS**

**Data**

Data were obtained from the Dutch Cattle Syndicate (NRS). The data included all 305-d heifer protein lactation records of herdbook Holstein Friesians of the 100 largest Dutch dairy herds. All animals had first calving in 1997, 1998, or 1999. Each year was divided in three HYS subclasses: January until May, June until September, and October until December. The minimum number of animals in a HYS subclass was three. Subclasses with lower numbers were discarded. Based on their age at calving, all animals were divided in age classes with length of 1 mo. Age at calving ranged from 16 to 39 mo.

The pedigree information, from 1980 onwards, contained 11,656 sires in total. Sires (\( n = 2150 \)) with at least five daughters with a lactation record in the data were identified. EBV were calculated for all those sires, based on 102,899 records for protein yield. For the sire \( \times \) HYS interaction model, 65,711 sire \( \times \) HYS subclasses were formed.

**Models**

To show the influence of the heterogeneity of variances, some of the models were also applied to the data after correction for heterogeneity of variances. From here, uncorrected and corrected will refer, respectively, to the situations without and with correction on the data for heterogeneity of variances. For the correction, a simple method was used; both within HYS subclass variances and HYS subclass means were standardized for all animals in the data. The within variances for all HYS subclasses were calculated as:

\[
\hat{\sigma}_i^2 = \frac{\sum_{j=1}^{n_i} (y_{ij} - y_{i.})^2}{(n_i - 1)}
\]

where \( y_{ij} \) is the phenotypic value of animal \( j \) in HYS subclass \( i \), \( y_{i.} \) is the mean of HYS subclass \( i \), and \( n_i \) is the number of phenotypic records in HYS subclass \( i \).

The weighted average of all within-HYS standard deviations was arbitrarily used as ‘base’ standard deviation \( \sigma_b \). Also, the weighted average of all HYS averages was arbitrarily used as ‘base’ average \( y_b \). The preadjusted record was (Lofgren et al., 1985; Ibáñez et al., 1996):

\[
y^*_ij = y_{ij} - y_{i.} \hat{\sigma}_i \sigma_b + y_b
\]

This adjustment does not account for fixed effects. Because only first-lactation animals are involved, the within-herd variances were comparable to each other. The variance components of all models were estimated by REML (Gilmour et al., 1999).

The standard model was:

\[
y_{ijkl} = \text{hys}_i\phi_i + \text{age}_j + G_k + e_{ijkl} \quad \text{[Model 0]}
\]

where \( y_{ijkl} \) is the record of daughter \( l \) of sire \( k \), \( \text{hys}_i\phi_i \) is a fixed effect of herd-year-season subclass \( i \), \( \text{age}_j \) is a fixed effect of age class \( j \), \( G_k \) is a random additive genetic effect of sire \( k \), and \( e_{ijkl} \) is a residual effect. This model was applied for reasons of comparison and to estimate environmental effects, corrected for genetic influences, for all HYS subclasses. These \( \text{hys}_i\phi_i \) effects were used as environmental parameters in the reaction norm model.

The interaction model was:

\[
y_{ijkl} = \text{HYS}_i + \text{age}_j + G_k + I_{ik} + e_{ijkl} \quad \text{[Model 1]}
\]

where \( I_{ik} \) is a random effect of \( G \times E \) interaction of sire \( k \) and HYS subclass \( i \).

In matrix algebra models 0 and 1 are:

\[
y = Xb + Ug + Zs + e = V\theta + e \quad \text{[0]}
\]

\[
y = Xb + Ug + Zs + Qt + e = V\theta + e \quad \text{[1]}
\]

where

- \( b \) is a vector with fixed HYS subclass effects,
- \( g \) is a vector of fixed age class effects,
- \( s \) is the vector of random additive sire genetic effects,
Also, for model 1:

\[ E(y) = Xb + Ug, E(s) = 0, E(t) = 0, E(e) = 0 \]

And, for model 0:

\[ E(y) = Xb + Ug, E(s) = 0, E(t) = 0, E(e) = 0 \]

\[
\text{Var} \begin{bmatrix} s \\ t \\ e \end{bmatrix} = \begin{bmatrix} A\sigma_s^2 & 0 & 0 \\ 0 & I_n\sigma_t^2 & 0 \\ 0 & 0 & I_n\sigma_e^2 \end{bmatrix}
\]

Also, for model 1:

\[ E(y) = Xb + Ug, E(s) = 0, E(t) = 0, E(e) = 0 \]

\[
\text{Var} \begin{bmatrix} s \\ t \\ e \end{bmatrix} = \begin{bmatrix} A\sigma_s^2 & 0 & 0 \\ 0 & I_m\sigma_t^2 & 0 \\ 0 & 0 & I_n\sigma_e^2 \end{bmatrix}
\]

where

\[ A \] is the relationship matrix of the sires,
\[ \sigma_s^2 \] is the variance of additive sire effects,
\[ I_m \] is an identity matrix of \( m \times m \) with \( m \) is the number of interaction effects,
\[ \sigma_t^2 \] is the variance of interaction effects, and
\[ \sigma_e^2 \] is the variance of residual effects.

The heritabilities were calculated as \( 4\hat{\sigma}_p^2/(\hat{\sigma}_g^2 + \hat{\sigma}_e^2) \) [0] and \( 4\hat{\sigma}_p^2/(\hat{\sigma}_s^2 + \hat{\sigma}_g^2 + \hat{\sigma}_e^2) \) [1].

In this model, the interaction variance as part of the total phenotypic variance explains the amount of \( G \times E \) interaction.

Model 2 is model 0 applied independently to four different subgroups of the data. The division of the data is based on the estimated HYS effect in model 0 without correction for heterogeneity of variances. Covariances between effects in different groups were not taken into account.

In model 2, calculating the correlation between the sets of EBV shows the \( G \times E \) interaction. The correlation depends on the reliability of the EBV. Reliability is defined as:

\[ r^2 = 1 - \frac{P}{\sigma_p^2} \]

where \( P \) (prediction error variance) is equal to the squared standard error of the EBV and \( \sigma_p^2 = 4\hat{\sigma}_y^2 \). And approximately, \( r^2 = n_e/(n_e + k) \) Wilmink et al. (1983), where \( n_e \) is the effective number of daughters, and

\[ k = \frac{\sigma_e^2}{\sigma_g^2}. \]

The effective number of daughters can be calculated as \( n_e = n^m(n + m) \), where \( n \) is the number of daughters of a sire and \( m \) is the number of daughters of other sires in the same HYS subclasses (Van Vleck, 1987).

In this study, sires with a minimum number of 75 daughters in the protein data were included in the calculations of the correlations between different sets of EBV. This made a total of 107 sires with 75,683 daughters with records.

Correlations between the EBV were transformed to genetic correlations (Blanchard et al., 1982),

\[ \hat{r}_{EBV} = \frac{\sqrt{(\Sigma b_i)\Sigma b'_i}}{\Sigma b_i b'_i} \]

where,

\[ \hat{r}_{EBV} \] is the estimator of the genetic correlation,
\( EBV \) is the EBV of sire \( i \) calculated in one group,
\( EBV' \) is the EBV of sire \( i \) calculated in another group,
\( r_{EBV} \) is the ‘simple’ correlation between the set of EBV and the set of EBV’s,
\( b_i \) is the reliability of the EBV of sire \( i \), and
\( b'_i \) is the reliability of the EBV’ of sire \( i \).

The third model is

\[ y_{ijkl} = HYS_i + age_j + L_k + (hys\phi_i)S_h + e_{ijkl} \quad \text{[Model 3a]} \]

where \( y_{ijkl} \) is the record of daughter \( l \) of sire \( j \), \( HYS_i \) reflects the HYS subclass \( i \), \( age_j \) is a fixed effect of age class \( j \), \( L_k \) is the random level of the additive genetic effect of sire \( k \), \( hys\phi_i \) is the estimated effect of HYS subclass \( i \) as estimated by model 0, \( S_h \) is the value of the random slope of the additive genetic effect of sire \( k \) and \( e_{ijkl} \) is a residual effect.

The same model is applied in a situation in which, instead of taking HYS subclasses into account as a fixed effect, a fixed regression is performed on estimated HYS effects \( hys\phi_i \). So, \( hys\phi_i \) is a fixed covariate of HYS subgroup \( i \) and \( b \) is the coefficient of a fixed regression of \( y \) on \( hys\phi_i \).
where $U$, $X$, $W$ and $Z$ scaling was done with the following formula: 

$$y_{ijkl} = b(h_{\text{sys}}) + age_j + L_k + (h_{\text{sys}})S_k + e_{ijkl}$$  

[Model 3b]

For both models, the estimated HYS effects were scaled between $-1.0$ and $+1.0$. This scaling makes internal calculations in the used ASREML-package easier (Gilmour et al., 1999; M. H. Pool, pers. comm.). This scaling was done with the following formula:

$$\hat{h}_i = (\hat{h}_i - \hat{h}_{\text{min}}) / 0.5 \delta = (\hat{h}_i - \hat{h}_{\text{min}}) / 0.5 \delta - 1$$

where

$$\hat{h}_i = \hat{h}_{\text{max}} - \hat{h}_{\text{min}}$$

$\hat{h}_i$ is the estimated HYS-effect of HYS subclass $i$ as estimated by model 0, $\hat{h}_{\text{max}}$ is the maximum estimated HYS-effect in the data, and $\hat{h}_{\text{min}}$ is the minimum estimated HYS-effect in the data.

For both models:

$$G_k = L_k + (h_{\text{sys}})S_k$$

with $G_k$ is the additive genetic effect of sire $k$.

In matrix algebra model 3a and 3b are:

$$y = Xb + Ug + Ws_L + Zs_S + e = V\theta + e$$  

[3a]

$$y = S\beta + Ug + Ws_L + Zs_S + e = V\theta + e$$  

[3b]

where

$\beta$ is a coefficient of a fixed regression of $y$ on $S$, $s_L$ is a vector of random levels of additive genetic sire effects, $s_S$ is a vector of random slopes of additive genetic sire effects, $U$, $X$, $W$ and $Z$ are incidence matrices, $S$ is a matrix which elements represent estimated HYS-effects, $V$ is $[X, U, W, Z]$ for model 3a, and $[S, U, W, Z]$ for model 3b, $\theta$ is $[b', g', s_L', s_S']$ for model 3a, and $[\beta', g', s_L', s_S']$ for model 3b.

As far as the variances are concerned, models 3a and 3b both have two variants. The first variant ignores the covariance between the variance of the level and the slope. These are referred to as respectively model 3a1 and 3b1. The second variant includes the covariance between level and slope. Those are referred to, respectively, as model 3a2 and 3b2. So,

$$E(y) = Xb + Ug$$

for model 3a1 and 3a2, and

$$E(y) = S\beta + Ug$$

for model 3b1 and 3b2, $E(s_L) = 0$, $E(s_S) = 0$, $E(s_{L,S}) = 0$ (for models 3a2 and 3b2), $E(e) = 0$,

$$\begin{bmatrix} s_L \\ s_S \\ e \end{bmatrix}$$

$$\begin{bmatrix} A\sigma^2_{s_L} & 0 & 0 \\ 0 & A\sigma^2_{s_S} & 0 \\ 0 & 0 & I_n\sigma^2_e \end{bmatrix}$$

for models 3a1 and 3b1, and

$$\begin{bmatrix} A\sigma^2_{s_L} & A\sigma_{s_{L,S}} & 0 \\ A\sigma_{s_{L,S}} & A\sigma^2_{s_S} & 0 \\ 0 & 0 & I_n\sigma^2_e \end{bmatrix}$$

for models 3a2 and 3b2,

where

$\sigma^2_{s_L}$ is the variance of the genetic level of sires, $\sigma^2_{s_S}$ is the variance of the slope of the genetic effect of sires, and $\sigma_{s_{L,S}}$ is the covariance of the level and the slope of the genetic effect of sires.

The obtained variance components of models 3a1 and 3b1 were used as prior values for respectively models 3a2 and 3b2.

Sire variances were calculated as $\hat{\sigma}^2_s = \hat{\sigma}^2_{s_L} + (h_{\text{sys}})^2\hat{\sigma}^2_{s_S}$ in models 3a1 and 3b1, and $\hat{\sigma}^2_s = \hat{\sigma}^2_{s_L} + (h_{\text{sys}})^2\hat{\sigma}^2_{s_S} + 2(h_{\text{sys}})\sigma_{s_{L,S}}$ in models 3a2 and 3b2. Heritabilities were calculated as $4\hat{\sigma}^2_s / (\hat{\sigma}^2_s + \hat{\sigma}^2_e)$.

$(h_{\text{sys}})^2$ is the weighted average of the squared estimated HYS effects and $h_{\text{sys}}$ is the weighted average of the estimated HYS effects.

In this model, the G × E interaction is explained by the variance of the slope and the covariance between the level and the slope. All models are summarized in Table 1.

Spearman's rank correlation coefficients (SAS, 1985) were calculated for all combinations of sets of EBV of the different models. Also mean squares of errors (MSE) were compared,
Table 1. Summary of models by description of the model and the way the genotype by environment (G × E) interaction is estimated in the model.

<table>
<thead>
<tr>
<th>Model</th>
<th>Description</th>
<th>G × E interaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Standard model</td>
<td>Not taken into account</td>
</tr>
<tr>
<td>1</td>
<td>Model with a sire by herd-year-season subclass (HYS) interaction</td>
<td>The interaction variance as part of the total phenotypic variance</td>
</tr>
<tr>
<td>2</td>
<td>Model with one trait treated as different single traits in four different groups of environments</td>
<td>Correlations between estimated breeding values (EBVs) in the different groups</td>
</tr>
<tr>
<td>3a</td>
<td>Reaction norm model with a random linear regression on the estimated HYS effect</td>
<td>Variance of the slope (and the covariance between the level and the slope)²²</td>
</tr>
</tbody>
</table>

²²This refers to models 3a1, 3b1, 3a2, and 3b2.
²²²This refers only to models 3a2 and 3b2.

\[ \text{MSE} = \frac{\sum_{i=1}^{n} e_i^2}{n} = \frac{\sum_{i=1}^{n} (\hat{y}_i - y_i)^2}{n} \]

with

- \( e_i \) is the residual of record \( i \),
- \( \hat{y}_i = v_i^\prime \hat{\theta}_i \) is the sum of the estimates of the relevant fixed and random effects according to equations 0, 1, 2, 3a and 3b,
- \( y_i \) is the observed value, and
- \( n \) is the total number of observations.

RESULTS

Variance Components and Heritabilities

In Table 2, the variance components and heritabilities are given for model 0 and 1. Model 0 shows a slightly higher sire variance than model 1, but the ratio between residual variance and sire variance is almost equal for both models. The interaction variance is 2.5% of the total phenotypic variance.

The heritabilities of the corrected data are comparable to those of the uncorrected data. Due to the correction of the data, both sire and residual variance reduced substantially.

In Table 3, the variance components and heritabilities of model 2 for the uncorrected and the corrected data are shown. The minimum and average estimated HYS effects are given. The heterogeneity of sire variance is clearly visible. The residual variance of the corrected data does not show heterogeneity, due to the correction. The sire variance in general increased more than the residual variance with increasing level of estimated HYS effect. This resulted in higher heritabilities in groups with higher estimated HYS effect.

In Table 4, the estimated genetic correlations between protein production in the different groups for model 2 are shown. The correlations for the corrected data are slightly lower, ranging from 0.73 to 0.80, versus 0.74 to 0.86 for the uncorrected data. For uncorrected protein yield, the corresponding Spearman’s rank correlations are also given. Those are in all cases slightly lower than the genetic correlations.

In Table 5, the variance components and heritabilities of the reaction norm models for protein yield are given. As model 3a2 did not converge, no results are included for this model.

For 3a1 and 3b1 \( \text{var}(s) \) was calculated as: \( \text{var}(s) = \text{var}(s_L) + (0.0517)^2 \text{var}(s_S) \).

For 3b2, the covariance component was added to this: \( \text{var}(s) = \text{var}(s_L) + (2^* - 0.0427)^2 \text{cov}(s_L, s_S) + (0.0517)^2 \text{var}(s_S) \).
Table 3. Estimated sire (\(\sigma_s^2\)) and residual variances (\(\sigma_e^2\)) and heritabilities (\(h^2\), with standard errors as subscript) for both uncorrected (unc) and corrected (corr) protein yield of the model that divides herd-year-season (hys) subclasses in four groups (2).

<table>
<thead>
<tr>
<th>Group</th>
<th>(\sigma_s^2)</th>
<th>(\sigma_e^2)</th>
<th>(\sigma_e^2/h_s^2)</th>
<th>(h^2)</th>
<th>min(hys)</th>
<th>av(hys)</th>
<th># cows</th>
<th># sires</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-unc</td>
<td>40.3</td>
<td>803.0</td>
<td>19.9</td>
<td>0.19</td>
<td>0.03</td>
<td>135</td>
<td>209</td>
<td>25,724</td>
</tr>
<tr>
<td>2-unc</td>
<td>69.8</td>
<td>865.5</td>
<td>12.4</td>
<td>0.30</td>
<td>0.04</td>
<td>225</td>
<td>233</td>
<td>25,728</td>
</tr>
<tr>
<td>3-unc</td>
<td>81.3</td>
<td>913.1</td>
<td>11.2</td>
<td>0.33</td>
<td>0.04</td>
<td>242</td>
<td>250</td>
<td>25,721</td>
</tr>
<tr>
<td>4-unc</td>
<td>98.4</td>
<td>1036.2</td>
<td>10.5</td>
<td>0.35</td>
<td>0.04</td>
<td>260</td>
<td>278</td>
<td>25,726</td>
</tr>
<tr>
<td>1-corr</td>
<td>1.3</td>
<td>28.2</td>
<td>21.2</td>
<td>0.18</td>
<td>0.03</td>
<td>135</td>
<td>209</td>
<td>25,724</td>
</tr>
<tr>
<td>2-corr</td>
<td>2.10</td>
<td>27.9</td>
<td>13.3</td>
<td>0.28</td>
<td>0.04</td>
<td>225</td>
<td>233</td>
<td>25,728</td>
</tr>
<tr>
<td>3-corr</td>
<td>2.38</td>
<td>27.8</td>
<td>11.7</td>
<td>0.32</td>
<td>0.04</td>
<td>245</td>
<td>250</td>
<td>25,721</td>
</tr>
<tr>
<td>4-corr</td>
<td>2.46</td>
<td>27.9</td>
<td>11.4</td>
<td>0.32</td>
<td>0.04</td>
<td>260</td>
<td>278</td>
<td>25,726</td>
</tr>
</tbody>
</table>

1The minimum estimated herd-year-season effect is at the same time the maximum estimated herd-year-season effect of the previous group. The maximum estimated herd-year-season effect of the last group is 382.

2The average estimated herd-year-season effect of all animals in the data.

3Total number of cows in the data.

4Total number of sires of cows in the data.

A comparison between models 3a1 and 3b1 shows the influence of either including HYS as subclasses or as estimated HYS effect is needed. A decrease of both sire and residual variance is observed if HYS is an estimated HYS effect. However, the ratios between sire and residual variance are almost equal in these situations.

To show the distinctiveness of the reaction norm model, EBV of model 3b2 were calculated for all four groups from model 2. The Spearman’s rank correlation coefficients between the groups are shown in Table 6. The correlations are very high between the different groups. The average estimated HYS effect of groups 1 and 4 are \(-0.29\) and \(0.28\), respectively, in the range of \(-1\) to \(+1\). This means that very few HYS subclasses are located in the extreme ends. Spearman’s rank correlation between EBV in the two extreme ends, e.g., for scaled estimated HYS effects of, respectively, \(-1\) and \(+1\), was 0.42.

Correlations

Table 7 shows rank correlations between models 0, 1, 2, and 3 for uncorrected protein yield. For the calculation of the correlation coefficients between models 2 and 3, EBV were calculated for the reaction norm models per group of records. Therefore, the average estimated HYS effects from the groups of model 2 were scaled. The slope was multiplied with those values before adding to the level.

Correlation coefficients between models 0, 1, and 3 were all close to 1.00. Model 2 shows correlations between 0.79 and 0.91 with models 0, 1, and 3. Those correlation coefficients show that EBV estimated in groups with highest estimated HYS effect are in general closer to those of models 0, 1, and 3.

Table 4. Estimated genetic correlations between the estimated breeding values of sires of the model which divides herd-year-season subclasses in four different groups (2) for uncorrected (Unc) and corrected (Corr) protein yield and Spearman’s rank correlation coefficients for uncorrected protein (Unc – Sp).

<table>
<thead>
<tr>
<th>Data</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unc*</td>
<td>1.00</td>
<td>0.86</td>
<td>0.77</td>
<td>0.74</td>
</tr>
<tr>
<td>1</td>
<td>1.00</td>
<td>0.82</td>
<td>0.79</td>
<td>0.79</td>
</tr>
<tr>
<td>2</td>
<td>1.00</td>
<td>0.81</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1.00</td>
<td>0.80</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corr*</td>
<td>1.00</td>
<td>0.77</td>
<td>0.73</td>
<td>0.74</td>
</tr>
<tr>
<td>1</td>
<td>1.00</td>
<td>0.77</td>
<td>0.73</td>
<td>0.73</td>
</tr>
<tr>
<td>2</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.80</td>
</tr>
<tr>
<td>3</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unc – Sp</td>
<td>1.00</td>
<td>0.78</td>
<td>0.70</td>
<td>0.68</td>
</tr>
<tr>
<td>1</td>
<td>1.00</td>
<td>0.72</td>
<td>0.75</td>
<td>0.78</td>
</tr>
<tr>
<td>2</td>
<td>1.00</td>
<td>0.78</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

*Standard errors ranging from 0.03 to 0.05.
The correlation coefficients between the EBV of corrected protein yield were almost equal to those for uncorrected protein yield.

Mean Squares of Errors

Table 8 shows the MSE of the different models for both uncorrected and corrected protein yield. For uncorrected protein yield, model 1 clearly has the lowest MSE. After that, model 2 is second best. MSE of model 2 were calculated as an average of the groups. For uncorrected protein the MSE increases if the estimated herd-year-season effect. The results of model 2 clearly show the heterogeneity in genetic effects between the different environments. The reaction norm model and the reaction norm model hardly seemed to quantify interaction (or slope) variance.

The interaction variance in model 1 was 2.5% of the phenotypic variance. The EBV were, however, close to those of the standard model at the cost of a higher calculation time. Ten Napel and Van der Werf (1992) found comparable results with an interaction variance of 3% of the total phenotypic variance for protein yield.

Judging from the results of model 2, it seems that there is more G × E interaction than model 1 detects. Probably, the sire × HYS interaction accounts only partly for G × E interaction. Therefore, it might be worthwhile to consider other possible definitions for G × E interaction then only sire by HYS. A possibility is dividing herds in different groups of production systems (e.g., in levels of intensity of production) and defining a sire by production system interaction.

The results of model 2 clearly show the heterogeneity of variances in the data. The model with the division of the data in groups showed correlations between 0.74 and 0.86 for EBV of different groups for uncorrected protein. Based on literature, correlations between 0.85 and unity were expected.

The records used in model 2 are divided over four equally sized groups and, therefore, the data are different from the other models. The fact that the correlations of EBV between groups are smaller than correlations with models 0, 1, and 3 might have originated because this model only uses a subset of the data that the other models use. The correlations might also be influenced by the method used to calculate them. Application of a multiple-trait model that uses all data together was tried but was not feasible. The results of such a model might be more comparable to those of the other models.

The reaction norm model has one major difference with model 0: the slope. The reaction norm model divides the standard breeding value in a level and a slope. It was expected that the slope would cause reranking of sires over environments. However, the level had such a great impact, that the slope had very little influence.

### Table 5

<table>
<thead>
<tr>
<th>Model</th>
<th>Data</th>
<th>$\sigma^2_{level}$</th>
<th>$\sigma^2_{level,slope}$</th>
<th>$\sigma^2_{slope}$</th>
<th>$\sigma^2_{level,slope}$</th>
<th>$\sigma^2_e$</th>
<th>$\sigma^2_s$</th>
<th>$\sigma^2_s/\sigma^2_e$</th>
<th>$h^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>3a1</td>
<td>Unc</td>
<td>74.2</td>
<td>. . .</td>
<td>43.6</td>
<td>. . .</td>
<td>76.5</td>
<td>904.3</td>
<td>11.8</td>
<td>0.31</td>
</tr>
<tr>
<td>3b1*</td>
<td>Unc</td>
<td>70.5</td>
<td>. . .</td>
<td>24.5</td>
<td>. . .</td>
<td>71.8</td>
<td>828.4</td>
<td>11.6</td>
<td>0.32</td>
</tr>
<tr>
<td>3b2*</td>
<td>Unc</td>
<td>77.9</td>
<td>49.1</td>
<td>45.5</td>
<td>0.82</td>
<td>76.0</td>
<td>828.0</td>
<td>10.9</td>
<td>0.34</td>
</tr>
<tr>
<td>3a1</td>
<td>Corr</td>
<td>2.21</td>
<td>. . .</td>
<td>0.034</td>
<td>. . .</td>
<td>2.21</td>
<td>27.96</td>
<td>12.6</td>
<td>0.29</td>
</tr>
</tbody>
</table>

*The regression coefficient of the fixed regression on the estimated herd-year-season effects was 123 for model 3b1 and 120 for 3b2.

### DISCUSSION

The aim of this study was to compare the reaction norm model with the interaction model and the model with the division in groups for their ability to correct for G × E interaction. Although the model with the division in groups seemed to be able to make a distinction in genetic effects between the different environments, the interaction model and the reaction norm model hardly seemed to quantify interaction (or slope) variance.

The interaction variance in model 1 was 2.5% of the phenotypic variance. The EBV were, however, close to
Table 7. Spearman’s rank correlations between estimated breeding values of the models 0, 1, 2, 3a1, 3b1, and 3b2 for uncorrected protein yield.

<table>
<thead>
<tr>
<th>Model</th>
<th>0</th>
<th>1</th>
<th>3a1</th>
<th>3b1</th>
<th>3b2</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.00</td>
<td>0.999</td>
<td>0.999</td>
<td>0.999</td>
<td>0.999</td>
</tr>
<tr>
<td>1</td>
<td>1.00</td>
<td>0.999</td>
<td>0.999</td>
<td>0.999</td>
<td>0.998</td>
</tr>
<tr>
<td>3a1</td>
<td>1.00</td>
<td>0.999</td>
<td>0.999</td>
<td>0.998</td>
<td>0.998</td>
</tr>
<tr>
<td>3b1</td>
<td>1.00</td>
<td>0.999</td>
<td>0.993</td>
<td>0.998</td>
<td>0.998</td>
</tr>
<tr>
<td>3b2</td>
<td>1.00</td>
<td>0.999</td>
<td>0.998</td>
<td>0.998</td>
<td>0.998</td>
</tr>
</tbody>
</table>

Table 8. Mean squares of errors (MSE) for the models 0, 1, 2, 3a1, 3b1, and 3b2 for uncorrected and corrected protein.

<table>
<thead>
<tr>
<th>Model</th>
<th>Uncorrected</th>
<th>Corrected</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>822.2</td>
<td>25.3</td>
</tr>
<tr>
<td>1</td>
<td>783.8</td>
<td>24.2</td>
</tr>
<tr>
<td>2</td>
<td>723.4</td>
<td>25.5</td>
</tr>
<tr>
<td>2</td>
<td>779.0</td>
<td>25.2</td>
</tr>
<tr>
<td>2</td>
<td>821.3</td>
<td>25.0</td>
</tr>
<tr>
<td>2</td>
<td>939.7</td>
<td>25.3</td>
</tr>
<tr>
<td>2</td>
<td>816.3</td>
<td>25.2</td>
</tr>
<tr>
<td>3a1</td>
<td>820.1</td>
<td>25.4</td>
</tr>
<tr>
<td>3b1</td>
<td>820.5</td>
<td>25.4</td>
</tr>
<tr>
<td>3b2</td>
<td>820.1</td>
<td>25.4</td>
</tr>
</tbody>
</table>

*Weighted averages of the four groups.
is calculated. Model 2 calculates a specific residual for each group of environments. Model 3 calculates only one residual for all environments. This might explain the fact that the slope is not specific enough. One way to solve this problem is to include a permanent environmental effect as a function of the environmental effect as done by Ravagnolo and Misztal (2000).

CONCLUSIONS

The model with the division in groups seemed able to distinguish between different environments. The correlations of this model indicated the presence of $G \times E$ interaction. However, the reaction norm model theoretically seems to be ideal. It uses all data together and should therefore be able to calculate more reliable EBV that are based on direct responses. EBV are depending on environment and this makes it possible to calculate specific EBV for each environment. This was clearly shown in Figure 1. However, in this research, the reaction norm model was not capable of calculating EBV that diverged from those of the standard model. This originated in the fact that the influence of the slope of the reaction norm on the EBV was limited. Therefore it could be conducted that estimated HYS effect did not cause $G \times E$ interaction. Other environmental parameters and inclusion of differing residual variances across environments should be considered. Also, nonlinear reaction norms might explain the sire variance better.

Further research is needed to explore the characteristics and abilities of the reaction norm model further.

REFERENCES


Falconer, D. S. 1990. Selection in different environments: effects on environmental sensitivity (reaction norm) and on mean performance. Genet. Res. 56:57–70.


