Comparison of a genetic group and unknown paternity models for growth traits in Nellore cattle

L. Shiotsuki,* F. F. Cardoso,† J. A. II. V. Silva,‡ and L. G. Albuquerque§2,3

*Embrapa Goats and Sheep, Sobral, CE 62010-970, Brazil; †Embrapa Southern Region Animal Husbandry, Bagé, 96401-970, Brazil; ‡Department of Animal Science, São Paulo State University (UNESP), Botucatu, SP 18610-307, Brazil; and §Department of Animal Science, São Paulo State University (UNESP), Jaboticabal, SP 14884-900, Brazil

ABSTRACT: The aim of the present study was to compare a model assuming unknown paternity and a model using genetic grouping to indicate the most adequate statistical procedure for the estimation of breeding values for animals with uncertain paternity. After data consistency, 62,212 Nellore animals, offspring of 581 bulls and 27,743 cows, were used in the analyses. The pedigree file contained 75,088 animals, including 22,810 (30.18%) offspring of multiple sires and 12,876 animals belonging to the base population with unknown parents. Three different approaches were adopted to deal with uncertain paternity of multiple-sire (MS) offspring. In the model of unknown paternity, the MS groups were ignored, and the sires of MS offspring were considered to be unknown and to belong to a single base population. In the genetic group approach, 2 definitions were used. In the first definition (GGa), “phantom parents” for animals with uncertain paternity were attributed, defining the genetic group as the MS group. In the other approach, GGb, phantom parents for animals with uncertain paternity were also attributed; however, MS offspring were clustered in genetic groups according to their year of birth, every 3 yr, on the basis of the average of male generation interval. Univariate analyses were performed under the Bayesian approach via Markov chain Monte Carlo methods. Models were compared by deviance information criteria and the conditional predictive ordinate. According to the choice criteria results, the genetic group model defined by the generation interval of males was more appropriate for predicting the genetic merit of animals with uncertain paternity. Therefore, the use of this model is recommended for the prediction of genetic merit and classification of offspring of multiple sires.

Key words: Bayesian inference, genetic group model, uncertain paternity

INTRODUCTION

Multiple-sire (MS) mating is a system adopted by most beef cattle producers because of its advantages in terms of labor requirements and low cost compared to other systems. Although commonly applied, the system of exposing 1 group of cows to multiple bulls has a great disadvantage in terms of genetic evalua-
fined as the average breeding values of the MS group, might be confounded with other effects of the model such as contemporary group (CG) effects, especially if a CG exclusively consists of all offspring of a certain MS group or if very small groups are formed (Quaas, 1988; Cardoso and Tempelman, 2004).

The assignment of phantom parents to animals with uncertain paternity represents a simple and effective solution to increase the accuracy of genetic evaluations, especially when the possible sires within the MS group are unknown, as required by more complex methods (Henderson, 1988; Cardoso and Tempelman, 2003). Therefore, the aim of the present study was to compare a model assuming unknown paternity and a model using genetic grouping to indicate the most adequate statistical procedure for the estimation of breeding values in animals with uncertain paternity.

**MATERIALS AND METHODS**

Animal Care and Use Committee approval was not obtained for this study because the data were obtained from an existing database.

**Nellore Data**

The data used refer to Nellore animals born between 1984 and 2006 and were obtained from the database of Agropecuária Jacarezinho Ltda. (Valparaíso, São Paulo, Brazil). Bulls and cows were raised on pasture and had ad libitum access to mineral salt. There were 2 breeding seasons. The first, the early season, comprised the period from February to April and includes only heifers aged 16 to 18 mo. The second breeding started in November and lasted approximately 70 d, including cows and 2-yr-old heifers that failed to conceive during the early season. All heifers were exposed, irrespective of weight and body condition. The mating systems used were artificial insemination, controlled breeding, and MS mating (a bull:cow ratio of 1:30). Multiple-sire mating is defined as the simultaneous exposure of a group of females to more than 1 bull during the same breeding season. The births of the calves are concentrated between August and October and between November and January, and the calves stay with their mothers on pasture until 7 mo of age. Heifer pregnancy is evaluated by rectal palpation approximately 60 d after the end of the early breeding season.

**Traits**

The traits studied in this work were yearling weight adjusted to 550 d of age (YW) and postweaning weight gain adjusted to 345 d (PWG), which considered the weight increase in kilograms between weaning weight adjusted to 205 d and YW. The mean and SD of PWG and YW were 98.40 ± 33.33 kg and 274.34 ± 42.02 kg, respectively. At yearling, because of the presence of 2 breeding seasons, the animals ranged in age from 420 to 640 d. Nevertheless, all YW were adjusted to a standard and intermediate age of 550 d, regardless of whether the animals were born in the early or regular season.

**Data Consistency**

During data consistency testing, data from CG containing fewer than 3 animals and those that differed more than were 3.5 SD above or below the mean of the CG for PWG and YW were excluded. In addition, at least the dam had to be known for all animals. Contemporary groups were formed by the combination of sex, farm, year and season of birth, month, and management group at weaning and farm, month, and management group at yearling.

Connectedness between CG was tested on the basis of the total number of genetic ties using the AMC program (Roso and Schenkel, 2006). This program consists of various options to test connectedness, such as type of model (sire and dam, animal, etc.), minimum number of genetic ties, and observations necessary to consider a CG connected. The animal model was used in the present study, and only CG with a minimum number of 10 genetic ties with other groups were accepted.

The description of the data set, pedigree file, and multiple-sire groups are presented in Table 1. After data consistency testing, 62,212 Nellore animals, offspring of 581 bulls and 27,743 cows, were used. The pedigree file contained 75,088 animals, including 22,810 (30.38%) MS offspring and 12,876 animals belonging to the base population with unknown parents.

**Statistical Analysis**

The PWG and YW were modeled as a linear function of fixed effects (effects with a bounded uniform priori distribution) and random effects (effects with a normal multivariate a priori distribution). Random effects included
direct additive genetic effect of the animal, CG effect, maternal genetic effect, and maternal permanent environmental effect. The last 2 effects were only included for YW. For both traits, the effect of age of the animal at yearling (420 to 640 d) and the effect (linear and quadratic) of age of cow (2 to 16 yr) were included as fixed effects.

Three approaches were adopted for the treatment of unknown paternity of multiple-sire offspring: 1) In the model of unknown paternity, the MS groups were ignored, and the sires of MS offspring were considered to be unknown and to belong to a single base population. In the genetic group approach, 2 definitions of genetic groups were used for comparison with the unknown paternity model: 2) In the first definition of genetic group model (GGa), phantom parents were assigned to animals with paternity uncertainty, with each genetic group being formed by the MS group (444 groups). In this case, the genetic group effect is an estimate of the average breeding value of bulls belonging to the MS group. 3) In the second definition (GGb), phantom parents were also assigned to animals with uncertainty of paternity. However, animals resulting from MS mating were grouped according to year of birth in intervals of 3 yr on the basis of the generation interval of MS males since most of these sires were 2-yr-old bulls. In this case, 6 genetic groups were formed.

Although the number of animals within the genetic group was high in the first definition, the latter definition was used to minimize possible problems of confoundedness between genetic groups and CG and the small number of animals within the genetic groups as reported by Quaas (1988).

According to the mixed model equations proposed by Henderson (1984), the unknown paternity model can be described as follows:

\[ y = Xb + Zu + e, \]

where \( y \) is the vector of performance, \( b \) is the vector of fixed effects, \( e \) is the vector of residual effects, \( u \) is used as the vector of the random effects, and \( Z \) is the incidence matrix of these effects, where there were 2 different models, one with just the additive genetic and CG effects for PWG and the other also having maternal genetic and permanent environment effects for long yearling weight (LYW); \( X \) is incidence matrices that relate the observations to fixed effects. It was assumed that

\[
p(y) ~ N(Xb + Zu, I\sigma^2),
\]
\[
p(b) = \text{const},
\]

Here for PWG we have that

\[
u = \begin{bmatrix} a \\ gc \end{bmatrix}, \quad G = \begin{bmatrix} A\sigma^2_a & A\sigma_{am} & 0 & 0 \\ A\sigma_{am} & A\sigma^2_m & 0 & 0 \\ 0 & 0 & \sigma^2_p & 0 \\ 0 & 0 & 0 & \sigma^2_{pm} \end{bmatrix},
\]

where \( a \) is the vector of additive genetic effects, \( gc \) is the vector of CG effects, \( A \) is the average relatedness matrix, \( I \) is the identity matrix, \( \sigma^2_a \) is the residual variance, \( \sigma^2_g \) is the additive genetic variance, and \( \sigma^2_{mg} \) is the CG variance.

For LYW this part of the model is defined as

\[
u = \begin{bmatrix} a \\ m \\ p \\ gc \end{bmatrix}, \quad G = \begin{bmatrix} A\sigma^2_a & A\sigma_{am} & 0 & 0 \\ A\sigma_{am} & A\sigma^2_m & 0 & 0 \\ 0 & 0 & \sigma^2_p & 0 \\ 0 & 0 & 0 & \sigma^2_{pm} \end{bmatrix},
\]

where, in addition to the terms described above, \( m \) is the vector of maternal genetic effects, \( p \) is the vector of maternal permanent environmental effects, \( \sigma^2_m \) is the maternal genetic variance, \( \sigma^2_p \) is the permanent environmental variance, and \( \sigma^2_{pm} \) is the covariance between additive and maternal genetic effects.

The prior distributions assigned to the variance components were

\[
p(\sigma^2_a) \sim IG, \quad \text{for PWG},
\]
\[
p(\sigma^2_p) \sim IG,
\]
\[
p(\sigma^2_m) \sim IG,
\]
\[
p\left(\begin{bmatrix} \sigma^2_a & \sigma_{am} \\ \sigma_{am} & \sigma^2_m \end{bmatrix}\right) \sim IW,
\]
\[
p(\sigma^2_p) \sim \text{for LYW},
\]

for LYW. Here IW is an inverted Wishart density and IG is an inverted \( \gamma \) density.

The equations of the mixed model including genetic groups (Westell et al., 1988) can be described as follows:

\[
y = Xb + Zu + ZgQ_g + e,
\]

where \( g \) is a vector of fixed additive (and maternal for LYW) genetic group effects, \( ZQ \) is an incidence matrix relating records to genetic effects, and \( Q \) is the incidence matrix relating animal genetic effects to group genetics effects. Therefore, using the more general case for LYW as an example, represents the additive \( (a*) \) and maternal \( (m*) \) effects adjusted for the genetic group effects in \( g \). It was assumed that \( p(g) = \text{const.} \)

**Bayesian Inference**

Single-trait Bayesian analyses were performed for all models using Markov chain Monte Carlo (MCMC) algorithms. A total of 400,000 cycles were run, and the first 10,000 cycles were considered as a burn-in. Convergence was monitored using the initial monotone sequence (Geyer, 1992). For each variance component,
the initial monotone sequence was used to calculate the effective sample size, which is an estimate of the number of independent samples with information content equivalent to that contained within the dependent samples (Sorensen et al., 1995). Chain length was determined so that all parameters had at least 100 independent samples.

**Model Choice**

The deviance information criterion (DIC; Spiegelhalter et al., 2002) and a deviance based on conditional predictive ordinates (CPO; Gelfand, 1996) were used to identify the model with the best goodness of fit. The CPO criterion was defined as an MCMC approximation for the conditional predictive ordinate for models \( M \) with parameters \( \theta \) obtained by a harmonic mean of \( R \) MCMC cycles:

\[
p(y_i | y_{-i}, M) \approx \frac{1}{R} \sum_{r=1}^{R} p^{-1}(y_i | \theta^{(r)}, M),
\]

where \( p(y_i | y_{-i}, M) \) is the CPO for observation \( y_i \), which refers to the cross-validation density and which suggests the likelihood values of \( y_i \) when the model is adjusted to all other observations \( y_{-i} \) except for \( y_i \). Finally, a deviance for model \( M \) based on CPO is obtained by

\[
-2 \sum_{i=1}^{n} p(y_i | y_{-i}, M).
\]

The DIC is a measure of overall fit that consists of the posterior mean deviance and a penalization for the complexity of the model. The mean deviance of model \( M \) was estimated by

\[
\text{DEV} = \frac{1}{R} \sum_{r=1}^{R} -2 \log p(y_i | \theta^{(r)}, M)
\]

as proposed by Spiegelhalter et al. (2002). The complexity of the model was evaluated on the basis of the effective number of parameters \( p_\theta = \text{DEV} - \text{DEV}(\bar{\theta}) \), where \( \text{DEV}(\bar{\theta}) = -2 \log p(y | \bar{\theta}, M) \), with \( \bar{\theta} \) corresponding to the posterior mean of \( \theta \); that is, \( p_\theta \) represents the difference between the posterior mean deviance and the deviance based on the posterior mean of the parameters in the model. Finally, the DIC of the model was determined as \( \text{DIC} = \text{DEV} + p_\theta \).

For comparison, lower DIC and deviance based on CPO values indicate a better goodness of fit of the model. A DIC difference higher than 7 indicates an important difference in the goodness of fit between models (Spiegelhalter et al., 2002).

**Genetic Predictions**

Spearman correlations between posterior means of additive genetic effects obtained with the different models for PWG and YW were calculated for the following data sets: 1) data set containing only bulls and bulls ranked as the 5%, 10%, 20%, and 50% best bulls for each trait and 2) data set containing the MS offspring and MS offspring classified as described for bulls in data set 1. In addition, the number of coinciding animals in the data sets containing bulls and animals with uncertain paternity for the described percentages was determined.

For MS offspring, the genetic trends were estimated from the regression of the genetic value for PWG and YW on birth year for the 3 models, obtained by PROC REG in the SAS software (SAS Inst. Inc., Cary, NC). The statistic model \( Y_{ij} = b_0 + b_1 x_i + e \) was considered, where \( Y_{ij} \) is the genetic value of the traits, with \( j \) identifying the animal and \( i \) indicating the year associated with \( j \)th animal, \( b_0 \) is the intercept, \( b_1 \) is the slope coefficient, \( x_i \) is the year of birth, and \( e \) is error.

**Simulation Study**

To verify the accuracy of the model selection criterion, simulated data sets were generated by first deleting records of MS offspring such that all remaining 38,965 records belonged to animals with known sires. Thereafter, uncertain paternity was generated by deleting at random 30% and 50% of paternal pedigree information and assigning these animals with an omitted pedigree to MS groups. For the GGa model, simulated MS offspring were grouped within year of birth according to their omitted sire, generating 5 to 22 MS groups per year, with the number of candidate sires ranging from 2 to 14 bulls within MS group and MS group progeny size ranging from 3 to 315 animals. Otherwise, for the GGb model, MS groups were formed by grouping MS offspring by their year of birth, with 1 group every 3 yr. This procedure was repeated to generate 5 data sets for each percentage.

In these simulated data, there were 1,755 contemporary groups. The file with 30% of paternal pedigree information contained 14,810 animals belonging to base populations, 193 MS groups, 27,275 animals with both parents known, and 11,690 offspring originating from MS mating. The file with 50% of paternal pedigree information contained 14,746 animals belonging to base populations, 258 MS groups, 19,476 animals with both parents known, and 19,489 offspring originating from MS mating. The 3 models described above were applied and compared using DIC and CPO values obtained when the data had complete pedigree information and with 30% or 50% of the paternal pedigree information lost to validate model choices for uncertain paternity data. The DIC and CPO were calculated individually for each generated data set and then were averaged over the 5 replicates to make comparisons.
RESULTS AND DISCUSSION

Model Choice

The model selection criteria used indicated both genetic group modeling strategies used (GGa and GGb) provided a better fit to PWG and YW complete data when compared to the model that did not include genetic groups (Table 2). Most studies comparing models with and without genetic groups reported that inclusion of these effects allows a more accurate prediction of the additive genetic merit of animals (Famula and Van Vleck, 1982; Pieramati and Van Vleck, 1993; Golden et al., 1994). However, Famula and Van Vleck (1982) concluded that the most appropriate definition of a genetic group depends on the selection criterion and that there is no single definition of genetic groups.

For the 2 traits analyzed, lower mean deviations were obtained for the model in which the genetic groups were formed by bulls belonging to the MS group (GGa). However, the DIC takes into account both mean deviations and the complexity of the model (or the effective number of parameters). As a consequence, a higher DIC was obtained for the GGa model because of its complexity (Table 2). Thus, the DIC indicated that the genetic group model defined on the basis of the 3-yr male generation length (GGb) was more appropriate for predicting breeding values for both PWG and YW than the other models studied (Table 2).

The CPO revealed greater differences between models than the DIC but still indicated model GGb to be the most appropriate for the prediction of breeding values of animals with uncertain paternity for the 2 traits. On the basis of the results of the model selection criteria, it is expected that models including appropriate genetic groups permit us to more accurately predict the breeding value of animals when some pedigree information is lost, as reported by Famula and Van Vleck (1982) and Westell et al. (1988). In addition, Pieramati and Van Vleck (1993) concluded that the inclusion of genetic groups in the model is an alternative for predicting breeding values when pedigree information is lost.

Simulation Study

Similar to the complete data, the DIC and CPO values calculated for the simulation study pointed out that using genetic groups, regardless of the grouping strategy (GGa or GGb), provides better data adjustment compared to the unknown paternity model for both levels of uncertain paternity tested, 30% (Table 3) and 50% (Table 4). These results ensure the reliability of the selection criteria used in the present work to demonstrate the utility of genetic grouping for genetic evaluation of MS offspring. Moreover, the employed criteria confirmed the 3-yr male generation length genetic grouping (GGb) as the best option among the 3 models evaluated and the closest to full pedigree for PWG when 30% of the pedigree information was randomly lost (Table 3), which is a similar situation to the 30% of MS offspring observed in the original data set. On the other hand, for YW, which also includes maternal genetic groups, the DIC favored the more complex GGa model, whereas the deviation based on CPO confirmed the results found for PWG in the present work when 30% of the pedigree information was randomly deleted.

Conversely, when 50% of the paternal pedigree was deemed unknown, both criteria concurred to indicated the GGa model based on MS groups as the closest model to a full pedigree analysis for PWG and YW (Table 4), implying that more complex modeling is desirable when a larger proportion of animals have uncertain paternity.
Table 4. The average deviance (DEV), penalty for effective number of parameters (PD), deviance information criterion (DIC), and conditional predictive ordinate (CPO) for postweaning gain (PWG) and long yearling weight (LYW) using full pedigree and a simulation with 50% of offspring of multiple sires averaged over 5 randomly generated replicates

<table>
<thead>
<tr>
<th>Trait</th>
<th>Data and model1</th>
<th>PD</th>
<th>DIC</th>
<th>CPO</th>
</tr>
</thead>
<tbody>
<tr>
<td>PWG</td>
<td>Complete</td>
<td>326,584</td>
<td>7,143</td>
<td>333,727</td>
</tr>
<tr>
<td></td>
<td>Unknown paternity</td>
<td>327,227</td>
<td>6,997</td>
<td>334,224</td>
</tr>
<tr>
<td></td>
<td>GGa</td>
<td>326,543</td>
<td>7,521</td>
<td>334,064</td>
</tr>
<tr>
<td></td>
<td>GGb</td>
<td>326,905</td>
<td>7,226</td>
<td>334,131</td>
</tr>
<tr>
<td>LYW</td>
<td>Complete</td>
<td>333,886</td>
<td>12,019</td>
<td>345,905</td>
</tr>
<tr>
<td></td>
<td>Unknown paternity</td>
<td>331,798</td>
<td>13,741</td>
<td>345,539</td>
</tr>
<tr>
<td></td>
<td>GGa</td>
<td>324,531</td>
<td>17,900</td>
<td>342,431</td>
</tr>
<tr>
<td></td>
<td>GGb</td>
<td>325,387</td>
<td>17,451</td>
<td>342,838</td>
</tr>
</tbody>
</table>

1Complete = all animals have known parents; GGa = the genetic group model formed by the multiple-sire group; GGb = the genetic group model formed with animals grouped according to year of birth (interval of 3 years based on the generation length in males).

In the unknown paternity model, the sire of MS offspring was considered to be unknown and to belong to a single base population. That means that animals belonging to the base population and bulls participating in the MS mating system will have the same average breeding values, a fact that might underestimate the breeding values of these bulls and, consequently, limit their contribution to the next generation since base population animals might belong to different regions and/or generations compared to selected young sires that participate in MS groups. Moreover, combining animals from different origin and generation in the same base population may yield biased estimates genetic variances (Pieramati and Van Vleck, 1993).

**Genetic Parameters**

In the present study, the effective sample size higher than 200 for all variance component parameters indicated that the chain length for the field data analyses was adequate.

The estimates of all parameters and their respective standard deviations for PWG obtained with the different models were closely similar (Table 5). Thus, model choice should not interfere with the estimation of genetic parameters for this trait. The estimates of heritability, which did not include the CG variance to make them comparable to other studies that had treated CG as fixed effects, and additive genetic and residual variance for PWG were similar to those reported in the literature (Eler et al., 1996; Marcondes et al., 1999; Paneto et al., 2002; Cardoso et al., 2004; Malhado et al., 2005; Boligon et al., 2010). Fernandes and Ferreira (2000), comparing statistical models for PWG in Charolais animals, obtained higher heritability estimates (0.36 to 0.37) than those found in the present study. However, the estimates obtained here are higher than that reported by Holanda et al. (2004) for Nellore animals from Pernambuco, Brazil (0.06).

For YW, additive genetic variances estimated with the models that included genetic group were similar, whereas lower estimates were obtained with the unknown paternity model (Table 6). On the other hand, Pieramati and Van Vleck (1993) obtained lower estimates of additive genetic variance with models including genetic group. However, in the present study the unknown paternity model provided the worst fit to the data compared to the models that included genetic group effects, overestimating residual variance (Table 6).

The estimates of covariance and correlation between direct and maternal genetic effects were positive (Table 4). Similar results have been obtained by Meyer (1992) for Angus animals. However, many studies in the literature report high and negative correlations between direct and maternal genetic effects (Gunsky et al., 2001; Ribeiro et al., 2001; Boligon et al., 2008). In the present study, the covariance and correlation estimates between direct and maternal genetic effects differed between models, with higher estimates being obtained with the unknown paternity model. However, the 95% credibility interval of the models overlapped (data not shown), indicating that the choice of the model may not interfere with the estimation of this parameter. The low magnitude of the correlation between direct and maternal genetic effects suggests that this effect is not important from a biological point of view (Eler et al., 1995).

An increase in direct heritability was observed when the model included genetic group. This finding may be attributed in part to the better fit of the model including groups effects (Table 2). The heritabilities of direct effects estimated in the present study (Table 6) are similar to those reported in the literature for Zebu breeds, with estimates ranging from 0.22 to 0.48 (Mercadante

Table 5. Posterior mean and standard deviation for genetic parameters of postweaning gain obtained by unknown paternity and genetic group models

<table>
<thead>
<tr>
<th>Parameter1</th>
<th>Unknown paternity</th>
<th>Model2</th>
<th>GGa</th>
<th>GGb</th>
</tr>
</thead>
<tbody>
<tr>
<td>σ²_a</td>
<td>52.85 ± 2.91</td>
<td>53.75 ± 2.87</td>
<td>54.19 ± 2.95</td>
<td></td>
</tr>
<tr>
<td>σ²_e</td>
<td>252.04 ± 2.70</td>
<td>250.51 ± 2.68</td>
<td>250.88 ± 2.72</td>
<td></td>
</tr>
<tr>
<td>σ²_g</td>
<td>681.91 ± 20.82</td>
<td>674.72 ± 20.76</td>
<td>687.76 ± 21.21</td>
<td></td>
</tr>
<tr>
<td>h²_d</td>
<td>0.17 ± 0.01</td>
<td>0.18 ± 0.01</td>
<td>0.18 ± 0.01</td>
<td></td>
</tr>
</tbody>
</table>

1Parameters are σ²_a = additive genetic variance; σ²_e = residual variance; σ²_g = contemporary group variance; h²_d = direct heritability.
2GGa = the genetic group model formed by the multiple-sire group; GGb = the genetic group model formed with animals grouped according to year of birth (interval of 3 years based on the generation length in males).
Rank Correlations

Spearman correlations between breeding values obtained with the unknown paternity model and the models using genetic grouping (GGa and GGb) for the data set including only sires were high (>0.80) for the 2 traits studied (PWG and YW; Table 7). These results indicate that the use of these models will not cause important changes in sire ranking. This was confirmed by the large number of coinciding bulls (in parentheses in Table 7) observed at all selection intensities (>78%). The breeding values predicted with the unknown paternity, GGa, and GGb models showed that 23 to 24 bulls would be selected in common when 5% of the best bulls are evaluated (the whole data set contained 26 bulls; Table 7).

If only MS offspring were included, the correlations between breeding values for PWG predicted with the unknown paternity and GGb models were all higher than 0.80. The model choice between unknown paternity and GGb did not cause substantial changes in animal ranking. This is confirmed by the large number of MS offspring (in parentheses in Table 8) selected in common by these 2 models. However, the correlations between predicted breeding values considering the other model combinations (unknown paternity × GGa and GGa × GGb) were less than 0.63. These lower correlations may be related to biased estimates of genetic group effects and, consequently, of genetic values of MS offspring in the GGa model, likely due to confounding of genetic groups and CG effects (Quaas, 1988). As result, the use of this model may cause important change in the ranking of offspring of multiple sires.

For YW, the correlations between breeding values predicted by the unknown paternity model and the models using genetic grouping (GGa and GGb) were moderate (<0.70). In addition, for this trait the highest correlations were observed between the models using genetic grouping (Table 8). These low correlations between breeding values suggest that models including genetic group (GGa and GGb) are better for correct ranking of offspring of multiple sires compared to the unknown paternity model when selection is based on YW. It should be pointed out that model GGb best fit the data observed.

Table 6. Posterior mean and SD for genetic parameters of yearling weight obtained by unknown paternity and genetic group models

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Unknown paternity</th>
<th>GGa</th>
<th>GGb</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma_a^2$</td>
<td>111.58 ± 7.82</td>
<td>134.74 ± 8.64</td>
<td>131.25 ± 8.25</td>
</tr>
<tr>
<td>$\sigma_p^2$</td>
<td>305.10 ± 5.07</td>
<td>290.55 ± 5.47</td>
<td>293.19 ± 5.29</td>
</tr>
<tr>
<td>$\sigma_e^2$</td>
<td>10.70 ± 1.42</td>
<td>9.77 ± 1.38</td>
<td>9.67 ± 1.34</td>
</tr>
<tr>
<td>$\sigma_m^2$</td>
<td>8.06 ± 2.68</td>
<td>2.26 ± 3.03</td>
<td>4.02 ± 2.71</td>
</tr>
<tr>
<td>$\sigma_h^2$</td>
<td>20.12 ± 1.42</td>
<td>20.21 ± 1.46</td>
<td>20.09 ± 1.41</td>
</tr>
<tr>
<td>$\sigma_g^2$</td>
<td>455.54 ± 3.26</td>
<td>457.59 ± 3.24</td>
<td>458.09 ± 3.61</td>
</tr>
<tr>
<td>$\sigma_{2g}^2$</td>
<td>1,015 ± 31.48</td>
<td>920.24 ± 25.57</td>
<td>983.49 ± 32.08</td>
</tr>
<tr>
<td>$h_a^2$</td>
<td>0.24 ± 0.016</td>
<td>0.29 ± 0.018</td>
<td>0.29 ± 0.017</td>
</tr>
<tr>
<td>$h_g^2$</td>
<td>0.02 ± 0.002</td>
<td>0.02 ± 0.003</td>
<td>0.02 ± 0.003</td>
</tr>
<tr>
<td>$r_{gdm}$</td>
<td>0.24 ± 0.085</td>
<td>0.07 ± 0.086</td>
<td>0.12 ± 0.079</td>
</tr>
</tbody>
</table>

1Parameters are $\sigma_a^2$ = additive genetic variance; $\sigma_p^2$ = residual variance; $\sigma_m^2$ = maternal variance; $\sigma_e^2$ = covariance between maternal and direct effects; $\sigma_g^2$ = maternal permanent environmental variance; $\sigma_{2g}^2$ = phenotypic variance; $\sigma_{2g}^2$ = contemporary group variance; $h_a^2$ = direct heritability; $h_g^2$ = maternal heritability; $r_{gdm} = $ correlation between direct and maternal genetic effects.

1GGa = the genetic group model formed by the multiple-sire group; GGb = the genetic group model formed with animals grouped according to year of birth (interval of 3 years based on the generation length in males).

Table 7. Number of coincident bulls (in parentheses) and Spearman correlation between posterior means of additive genetic effects predicted by unknown paternity and genetic group models for postweaning gain (PWG) and long yearling weight (LYW)

<table>
<thead>
<tr>
<th>Trait</th>
<th>10% (52)</th>
<th>20% (104)</th>
<th>50% (260)</th>
<th>100% (519)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PWG</td>
<td>Unknown paternity × GGa (49) 0.87 (98)</td>
<td>0.96 (250) 0.98 (519)</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td>Unknown paternity × GGb (50) 0.94 (100) 0.98 (252)</td>
<td>0.99 (519)</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td>GGa × GGb (51) 0.96 (100) 0.99 (254)</td>
<td>0.99 (519)</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>LYW</td>
<td>Unknown paternity × GGa (41) 0.84 (89) 0.82 (244)</td>
<td>0.91 (519)</td>
<td>0.97</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Unknown paternity × GGb (44) 0.84 (94) 0.86 (245)</td>
<td>0.93 (519)</td>
<td>0.98</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GGa × GGb (49) 0.94 (97) 0.94 (254)</td>
<td>0.97 (519)</td>
<td>0.99</td>
<td></td>
</tr>
</tbody>
</table>

1GGa = the genetic group model formed by the multiple-sire group; GGb = the genetic group model formed with animals grouped according to year of birth (interval of 3 years based on the generation length in males).
Probably because of the smaller number of effects included in the model for PWG, the Spearman correlations between breeding values were slightly higher for this trait compared to YW.

**Genetic Trends**

The average breeding values for PWG estimated with the 3 models were close to 0 until 1995 (Fig. 1). After this year, greater trends were observed, especially for the model including genetic groups formed on the basis of male generation length, which provided the highest annual mean genetic trend for PWG (0.22 kg/yr).

For YW, until 1998 the lowest average breeding values were estimated with the models including genetic groups (GGa and GGb). However, a higher genetic trend was obtained with the model including genetic groups formed by the MS groups (1.22 kg/yr). On the other hand, the annual average breeding values obtained with the GGa model for the 2 traits varied widely and did not follow a linear trend. This finding might be due to the contamination of breeding values with environmental effects in this model as a consequence of confounding between MS groups and CG effects. Annual means estimated with model GGb were higher than those estimated with the unknown paternity model, with genetic trends of 1.020 and 1.016 kg/yr, respectively (Fig. 1). By assuming that bulls included in the MS groups belong to the base population, the unknown paternity model underestimates genetic gain because these bulls are selected and it is expected that they present a higher average breeding value than the base population.

Paternity in beef cattle is known because of controlled breeding and artificial insemination. Paternity testing by blood typing and DNA methods is often expensive for producers, especially those working with large herds. Many breeding programs do not include animals with uncertain paternity. However, Sullivan (1995) observed that the exclusion of performance data of animals with uncertain paternity from the analysis can account for a reduction of 24% in the response to selection when comparing a model excluding performance data and models using genetic grouping. The author concluded that the loss of the response to selection depends on the percentage of animals with uncertain paternity. Therefore, the application of models that include genetic groups is an alternative approach to improve the prediction of genetic merit of animals with uncertain paternity.

In conclusion, the genetic group model grouping animals according to male generation length was more adequate for predicting the genetic merit of animals with uncertain paternity. The use of this model is recommended for the prediction of genetic merit and classification of offspring of multiple sires, particularly when more complex modeling is not possible because of the absence of infor-

---

**Table 8.** Number of offspring of multiple sires (MS; in parentheses) and Spearman correlation between posterior means of additive genetic effects predicted by unknown paternity and genetic group models for postweaning gain (PWG) and long yearling weight (LYW)

<table>
<thead>
<tr>
<th>Trait</th>
<th>10% (2,281)</th>
<th>20% (4,562)</th>
<th>50% (11,405)</th>
<th>100% (22,810)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PWG</td>
<td>0.41 × GGa</td>
<td>0.49 × GGa</td>
<td>0.57 × GGa</td>
<td>0.74 GGa</td>
</tr>
<tr>
<td></td>
<td>(1,176)</td>
<td>(2,931)</td>
<td>(9,086)</td>
<td>(22,810)</td>
</tr>
<tr>
<td></td>
<td>0.82 × GGb</td>
<td>0.85 × GGb</td>
<td>0.89 GGb</td>
<td>0.95 GGb</td>
</tr>
<tr>
<td></td>
<td>(1,862)</td>
<td>(3,940)</td>
<td>(10,543)</td>
<td>(22,810)</td>
</tr>
<tr>
<td></td>
<td>GGa × GGb</td>
<td>0.52 GGa × GGb 0.57 GGa × GGb</td>
<td>0.62 GGa × GGb 0.76 GGa × GGb</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1,391)</td>
<td>(3,127)</td>
<td>(9,259)</td>
<td>(22,810)</td>
</tr>
<tr>
<td>LYW</td>
<td>0.29 × GGa</td>
<td>0.35 × GGa</td>
<td>0.44 × GGa</td>
<td>0.63 GGa</td>
</tr>
<tr>
<td></td>
<td>(1,142)</td>
<td>(2,685)</td>
<td>(8,555)</td>
<td>(22,810)</td>
</tr>
<tr>
<td></td>
<td>0.53 × GGb</td>
<td>0.59 × GGb</td>
<td>0.69 GGb</td>
<td>0.85 GGb</td>
</tr>
<tr>
<td></td>
<td>(1,489)</td>
<td>(3,252)</td>
<td>(9,402)</td>
<td>(22,810)</td>
</tr>
<tr>
<td></td>
<td>GGa × GGb</td>
<td>0.70 GGa × GGb 0.71 GGa × GGb</td>
<td>0.73 GGa × GGb 0.80 GGa × GGb</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1,518)</td>
<td>(3,298)</td>
<td>(9,500)</td>
<td>(22,810)</td>
</tr>
</tbody>
</table>

1 GGa = the genetic group model formed by the multiple-sire group; GGb = the genetic group model formed with animals grouped according to year of birth (interval of 3 years based on the generation length in males).
mation on candidate sires (Henderson, 1988; Cardoso and Tempelman, 2003; Shiotsuki et al., 2012).

**LITERATURE CITED**


