

## Evaluation of an average numerator relationship matrix model and a Bayesian hierarchical model for growth traits in Nellore cattle with uncertain paternity

L. Shiotsuki <sup>a</sup>, F.F. Cardoso <sup>b,e</sup>, J.A.II.V. Silva <sup>c</sup>, G.J.M. Rosa <sup>d</sup>, L.G. Albuquerque <sup>e,\*</sup>

<sup>a</sup> Brazilian Agricultural Research Corporation- Goat and Sheep, Sobral, CE, 62.010-970, Brazil

<sup>b</sup> Embrapa Southern Region Animal Husbandry, Bagé, RS, 96401-970, Brazil

<sup>c</sup> Department of Animal Science, São Paulo State University (UNESP), Botucatu, SP, 18610-307, Brazil

<sup>d</sup> Department of Dairy Science, University of Wisconsin, Madison, WI, 53706, United States

<sup>e</sup> National Council for Scientific and Technological Development (CNPq), Brazil

### ARTICLE INFO

#### Article history:

Received 12 January 2011

Received in revised form 26 August 2011

Accepted 1 November 2011

#### Keywords:

Bayesian inference  
Beef cattle  
Multiple-sires  
Uncertain paternity

### ABSTRACT

The objective of this work was to compare a model based on the use of an average numerator relationship matrix (ANRM) and a hierarchical animal model (HIER) to indicate the most appropriate statistical procedure to better estimate the genetic value of Nellore animals that have unknown paternity. The data set contained records of 62,212 Nellore animals. The pedigree file contained a total of 75,088 animals. Two approaches were adopted for the treatment of uncertain paternity. In the model based on the use of the ANRM probabilities were attributed to each of the possible parents of the animals with uncertain paternity. The other method adopted in the present study, i.e., the HIER, considers uncertainty in the assignment of paternity of animals participating in the multiple-sire (MS) system. Within this context, a priori probabilities are assigned to each possible sire of animals with uncertain paternity, which are altered according to information present in the data for the generation of posterior probabilities. Univariate analyses were carried out under Bayesian approach via Markov Chain Monte Carlo (MCMC) methods, implementing a chain of 400,000 rounds where the first 10,000 rounds were discarded (burn-in period). Models were compared by deviance information criteria (DIC) and pseudo Bayes factors (PBF). The model that best fits the data for estimating genetic parameter of animals with uncertain paternity is the Bayesian hierarchical model. Nevertheless, for genetic evaluation, the choice between these models would have no impact on genetic value classification of animals for selection.

© 2011 Elsevier B.V. All rights reserved.

### 1. Introduction

Different statistical models have been proposed for the prediction of genetic merit of animals with uncertain paternity. The average numerator relationship matrix (ANRM) model proposed by [Henderson \(1988\)](#) is a model used for the genetic

evaluation of animals with uncertain paternity. This procedure is based on knowledge of the probability of each sire being the true sire of a certain animal. The ANRM specify the correct genetic (co)variance matrix when these probabilities are presumed to be known, thus facilitating best linear unbiased prediction of genetic merit.

A hierarchical animal model (HIER), proposed by [Cardoso and Tempelman \(2003\)](#), is another procedure used for the genetic evaluation of animals with uncertain paternity. The HIER combines data and prior information to determine posterior probabilities of sire assignments to infer genetic merit of individuals with uncertain paternity and their sire assignments.

\* Corresponding author at: Department of Animal Science, São Paulo State University (UNESP), Jaboticabal, SP, 14.884-900, Brazil. Tel.: +55 16 3209 2678; fax: +55 16 3209 2685.

E-mail address: [lgalb@fcav.unesp.br](mailto:lgalb@fcav.unesp.br) (L.G. Albuquerque).

The main advantage of HIER model is that it can properly account for smaller precision on genetic merit inference due to uncertainty on sire assignments. Cardoso and Tempelman (2004) presented that the HIER model was consistently favored based on Bayesian model choice criteria over the ANRM. Moreover, there is literature that suggests that inclusion of phenotypic information to ascertain paternity can enhance the accuracy of breeding value predictions (Cardoso and Tempelman, 2003, 2004; Sapp et al., 2007).

The limitation with using phenotypic data for ascertaining paternity is that the relationship matrix has to be reconstructed for every possible combination of offspring and sire which may not be computationally feasible for large data sets. Thus, Sapp (2005) presented a method for predicting breeding values that does not require construction of the inverse matrix. This method can have smaller computation time and the estimates tend to be lower than reported by Cardoso and Tempelman (2003), although the differences are not significant (Sapp et al., 2007). Therefore, the objectives of the present study were to compare the average numerator relationship matrix (ANRM) and hierarchical animal models (HIER) to indicate the most appropriate statistical procedure to better estimate the genetic value of Nelore who have uncertain paternity.

## 2. Material and methods

### 2.1. General

Records of 62,212 Nelore animals born between 1984 and 2006 were used. They belong to Agropecuária Jacarezinho Ltda., municipality of Valparaíso, São Paulo, Brazil.

The artificial insemination, controlled mating and multiple-sire (MS) mating (with a bull:cow ratio of 1:30) were used as mating system. There were two breeding seasons. In the first, called anticipated breeding season, all heifers are mated between February and April, irrespective of weight and body condition. This first breeding season lasts approximately 60 days. The second, called normal breeding season, starts in the second half of November and lasts approximately 70 days. In this season, heifers that did not conceive during the anticipated breeding season and all cows of the herd are mated. The birth periods of the calves are concentrated between August and October and between November and January. The calves are kept with their mothers on pasture until 7 mo of age.

### 2.2. Traits

The traits considered in this work were postweaning gain adjusted to 345 days (PWG), which considered the gain from weaning (205 days) to long-yearling (550 days) and yearling weight (YW) at 550 days, obtained at a minimum and maximum age of 420 and 640 days, respectively. The mean  $\pm$  standard deviations were  $98.40 \pm 33.33$  kg and  $274.34 \pm 42.02$  kg for PWG and YW, respectively.

### 2.3. Data consistency

Data from contemporary groups (CGs) with less than three animals and those which differed more than 3.5SD from the mean, respectively, for each trait were discarded. Contemporary groups were formed by combining the information relative

to sex, farm, year and season of birth, month and CG at weaning and farm, month and management group at yearling. At least the dam was known for all animals.

Connectedness between CGs was tested based on the total number of genetic ties (minimum of 10) in the animal model using the AMC program (Roso et al., 2006). The program offers various options to test CG connectedness, with the possibility of choosing the model (sire and cow, animal, etc.) and the minimum number of direct genetic links and observations to define a connected CG. The program identifies the CG groups with the largest number of genetic connections and subsequently all other groups connected to it (archipelago 1). Next, the program identifies the group with the largest number of connections among those not connected to archipelago 1 and the remaining groups connected to it to form archipelago 2, and so forth until only unconnected groups remain.

A summary of data structure is presented in Table 1. After data edition, 62,212 animals were maintained in the file. The pedigree file contained a total of 75,088 animals. The multiple-sire mating generally is cows exposed to more than one male within the same breeding season. The sires should be identified in the multiple-sire groups. In the present work, there were 1024 MS groups, however, only 177 groups possessed an identification of the participating sires, i.e., the 847 MS groups have no identification about participating sires. Among the 177 MS groups, 153 were connected and possessed performance or progeny information and no progeny or performance information was available for 24. A genetic group model was adopted for MS groups without sire identification, in which each MS group was identified as a genetic group. The genetic groups were formatted by assigning “phantom parents” for animals with uncertain paternity, however, MS offspring were clustered in genetic groups according to their year of birth, every three years (based on generation interval for males).

In the file with identified sires and their respective groups of sires (177 groups), 559 sires were identified in groups of specific sires. For 40 sires, no performance data or known progeny were available. Only six sires of the MS groups possessed no performance data but had known progeny, 351 possessed performance data but had no known progeny, and 162 had performance data and known progeny. The average size of the MS groups was 4.8 sires (ranging from 2 to 16).

### 2.4. Statistics

PWG and YW were modeled as a linear function of fixed effects (effects with a bounded uniform priori distribution)

**Table 1**

Data structure of datafile, pedigree file and multiple sire (MS) file after data edition.

Number of animals in datafile	62,212
Number of sires	581
Number of dams	27,743
Number of animals in a pedigree file	75,088
Number of animals belonging base population	12,876
Number of animals with parents known	39,402
Number of offspring originated MS mating	22,810
Number of MS group	1024
Groups possessed an identification of the participating sires	177
Groups have no identification about participating sires	847

and random effects (effects with a normal multivariate a priori distribution). Direct genetic effects of the animals, CG effects, maternal genetic effects and permanent environmental effects were included as random effects. The last two effects were only considered for the analysis of YW. Effects of animal age at yearling and dam age (2–16 yr) (linear and quadratic effects) were included as fixed effects for both traits.

Two approaches were adopted for the treatment of uncertain paternity. In the model based on the use of the ANRM proposed by Henderson (1988) equal probabilities were attributed to each of the possible parents of the animals with uncertain paternity. These probabilities are used to obtain the additive relationship coefficient of the animal with uncertain paternity, which is calculated as follows:

$$a_{ij} = \frac{1}{2} \left[ a_{is}^{(j)} p_i^{(j)} + \dots + a_{is_g}^{(j)} p_g^{(j)} + a_{id}^{(j)} \right] = \frac{1}{2} \left[ \sum_{k=1}^g a_{is_k}^{(j)} p_k^{(j)} + a_{id}^{(j)} \right]$$

where  $a_{ij}$  is the additive relationship between animal  $i$  and  $j$ ;  $a_{is_k}^{(j)}$  is the additive relationship between animal  $i$  and each of the possible parents ( $k = 1, 2, \dots, g$ ) of animal  $j$ ;  $p_k^{(j)}$  is the probability of sire  $k$  to be father of animal  $j$ ; and  $a_{id}^{(j)}$  is the additive relationship between animal  $i$  and the mother of animal  $j$ .

The other method adopted in the present study, i.e., the HIER proposed by Cardoso and Tempelman (2003), considers uncertainty in the assignment of paternity of animals participating in the MS system. Within this context, equal priori probabilities are assigned to each possible sire of animals with uncertain paternity, which are altered according to information present in the data for the generation of posterior probabilities.

This model can be divided into four stages:

**1st Stage:** The performance of animals is described in this stage as a linear function of genetic and non-genetic effects, resulting in the following conditional distribution of all parameters:

$$y | a, b, m, gc, pe, \sigma_e^2 \sim N(Xb + Z_1 a + Z_2 m + Z_3 gc + Z_4 pe, I \sigma_e^2);$$

where  $y$  is the vector of data,  $b$  is the vector of non-genetic effects,  $a$  is the vector of additive genetic effects;  $m$  is the vector of maternal genetic effects and  $pe$  is the vector of maternal permanent environmental effect, both were considered only for the YW;  $gc$  is the vector of contemporary group effects;  $e$  is the vector of residual, with  $e | \sigma_e^2 \sim N(0, I \sigma_e^2)$ ;  $X, Z_1, Z_2, Z_3$  and  $Z_4$  are incidence matrices relating observations to non-genetic, additive genetic effects, maternal genetic effects, contemporary group and maternal permanent environmental, respectively;  $\sigma_e^2$  is the residual variances; and  $I$  is identity matrix.

**2nd Stage:** The prior distribution for non-genetic and additive genetic effects is described in this stage:

Non-genetic effects ( $b$ ):  $b \sim N(b_0, V_b)$ .  
Breeding values ( $a$ ):  $a | s, \sigma_a^2 \sim N(0, A_s \sigma_a^2)$ .

$$\begin{bmatrix} a \\ m \end{bmatrix} | S, G_0 \sim N(0, G)$$

$$G = A \otimes G_0$$

$$G_0 = \begin{bmatrix} \sigma_a^2 & \sigma_{am}^2 \\ \sigma_{am}^2 & \sigma_m^2 \end{bmatrix}.$$

Contemporary group effects:  $gc \sim N(0, I \sigma_{gc}^2)$ .

Maternal permanent environmental effects:  $pe \sim N(0, I \sigma_{pe}^2)$ .

To calculate the animal genetic values, additive genetic (co)variances are based on the genetic relationship matrix ( $A$ ), designated sires ( $s$ ) and genetic variance ( $\sigma_a^2$ ).  $\sigma_{gc}^2$  is the contemporary group variance;  $\sigma_{pe}^2$  is the maternal permanent environmental variance;  $V_b$  is a diagonal matrix of the a priori of  $b$ , assuming  $V_b \propto \infty$ .

A scale inverse chi-squared a priori distribution is assumed for residual variances ( $\sigma_e^2$ ), which is given by  $\sigma_e^2 \sim S_e^2 \chi_v^{-2}$ , where  $S_e^2$  is the a priori value of residual variance based on literature data and  $v$  is the degree of confidence in this value (Sorensen and Gianola, 2002).

**3rd Stage:** Genetic variances assume the following values:

$$\sigma_a^2 \sim S_a^2 \chi_{v(a)}^{-2},$$

where  $S_a^2$  is the a priori value of genetic variance based on literature data and  $v(a)$  is the degree of confidence in this value.

Moreover, in this stage is a specified probability that  $s_{(j)}$  is being the true sire of animal  $j$ :

$$\pi_j^{(k)} = \text{Prob}(s_j = s_j^{(k)}) \text{ for all candidate sires } k = 1, 2, \dots, v_j;$$

thus, the set of priori probabilities of each sire  $v_j$  be the true sire of animal  $j$ :

$$\pi_j = \left\{ \pi_j^{(1)}, \pi_j^{(2)}, \dots, \pi_j^{(v_j)} \right\} \text{ for } v_j \text{ candidate sires of animal } j.$$

**4th Stage:** The prior distribution for probabilities of paternity is described in this stage. The entire set of probabilities for all non-base animals is rarely known with absolute certainty, and so in this model they are regarded as random quantities from a Dirichlet distribution:

$$p(\pi | \alpha) \propto \prod_{j=q_b+1}^q \prod_{k=1}^{v_j} \left( \pi_j^{(k)} \right)^{\alpha_j^{(k)}};$$

where  $\alpha_j = \{ \alpha_j^{(k)} \}_{k=1}^{v_j}$ ,  $\alpha_j^{(k)} > 0$  to  $k = 1, 2, \dots, v_j$  and  $\pi_j^{(v_j)} = 1 - \sum_{k=1}^{v_j-1} \pi_j^{(k)}$ . Specifications set of hyper-

parameters  $\alpha = \{ \alpha_j \}_{j=q_b+1}^q$  can be based on the reliability of the evaluation of an external source of information about the a priori probabilities assigned to each sire.

A Gibbs sampler in which each chain consisted of 400,000 cycles was used. First 10,000 cycles were

removed before estimating marginal posterior distributions of unknowns, directly from the samples. Convergence was monitored using initial monotone sequence (Geyer, 1992).

The deviance information criterion (DIC) (Spiegelhalter et al., 2002) and conditional predictive ordinate (CPO) (Gelfand, 1996) were used as model choice criteria to determine the best fitting model. CPO was defined as an MCMC approximation for the CPO of model  $M_r$  with parameters  $\theta$  is obtained by a harmonic mean of  $G$  MCMC cycles:

$$p(y_{ij}|y_{(-ij)}, M_r) \approx \frac{1}{\frac{1}{G} \sum_{l=1}^G p^{-1}(y_{ij}|\theta^{(l)}, M_r)}$$

where  $p(y_{ij}|y_{(-ij)}, M_r)$  is the conditional predictive ordinate (CPO) for observation  $y_{ij}$ , which refers to the cross-validation density and describes what values of  $y_{ij}$  are likely when the model is fit to all other observations  $y_{(-ij)}$  except for  $y_{ij}$ .

The DIC consists of a measure of global fit, posterior deviance mean, and a penalization for complexity of the model. According to Spiegelhalter et al. (2002), the model deviance is estimated by  $DEV = \frac{1}{G} \sum_{i=1}^G -\log p(y|\theta^{(i)}, M_r)$ . The complexity of the model is determined by the effective number of parameters  $p_{D(r)} = \overline{DEV} - DEV(\bar{\theta})$ , where  $DEV(\bar{\theta}) = -2 \log p(y|\bar{\theta}, M_r)$ , with  $\bar{\theta}$  being the posterior mean of  $\theta$ , i.e.,  $p_{D(r)}$  corresponds to the difference between the posterior deviance mean and the deviance based on the posterior mean of the parameters of the model. The DIC for the model is determined as:

$$DIC = \overline{DEV} + p_{D(r)}$$

Lower values of DIC are indicative of a better fitting model. Spiegelhalter et al. (2002) have suggested a DIC difference of 7 to be an important difference in the model fit.

Spearman correlation between posterior means of additive genetic effects obtained by the two different models for PWG and YW was calculated for different sets of animals: the first dataset included all animals of the pedigree. The second set included only offspring from MS mating, so animal with uncertain paternity. The third dataset included only sires. Within each dataset, animals were classified by their genetic value, the best animals (TOP) 10, 30 and 50% were used to calculate the correlation between the mean genetic values obtained by HIER and ANRM models to different selection pressures.

The percentage of squared bias (PSB) investigates the overall goodness of fit of all models, as

introduced by Ali and Schaeffer (1987). The PSB for the  $j$ th trait was defined as:

$$PSB_j = \frac{\sum_{i=1}^o (y_i - \hat{y}_i)^2}{\sum_{i=1}^o (y_i)^2} * 100;$$

where  $y_i$  is the observed record,  $\hat{y}_i$  is the predicted record and  $o$  is the number of observations.

### 3. Results and discussion

The DIC consists of the average deviance (DEV) and penalty for effective number of parameters ( $p_d$ ). In this study, the difference between ANRM and HIER was 44 for PWG and 255 for YW (Table 2). Thus, the biggest difference between ANRM and HIER models was observed for YW, which is more heritable than PWG (see Tables 3 and 4). Since smaller values of DIC indicate a more appropriate model for estimating genetic parameters, it was observed that the values of DIC and DEV smallest for the HIER model compared to ANRM (Table 2). Thus, the HIER model was more appropriate to calculate the breeding value for PWG and YW. Although the model HIER has been described as the most appropriate model, the HIER was also the more complex model, indicated by higher effective number of parameters. However, this penalty model was not high enough to change the criteria of choice and DIC.

For the CPO, the difference between the deviances was 224 for PWG and 193 for the YW. Thus, both criteria, DIC and CPO, consistently indicated the HIER model as the most appropriate to predict the genetic value of animals with uncertain paternity for all traits. These results agree with those reported in the literature by Cardoso and Tempelman (2003). These authors also concluded that the HIER model was more appropriate than ANRM model to infer the genetic merit of animals with uncertain paternity, when the variance components were considered known and unknown. Moreover, Cardoso and Tempelman (2004) observed that the HIER model produced estimates of breeding values for animals with uncertain paternity with less bias.

The percentages of squared bias from both models, ANRM and HIER, were low and similar, being 1.996 and 1.986 for PWG and 0.393 and 0.399 for YW, respectively. Nevertheless,

**Table 2**

The average deviance (DEV), penalty for effective number of parameters ( $p_d$ ), deviance information criterion (DIC) and conditional predictive ordinate (CPO) for postweaning gain (PWG) and yearling weight (YW).

Model	DEV	$p_d$	DIC	CPO
PWG				
ANRM	520,267	10,862	531,129	522,486
HIER	520,110	10,975	531,085	522,262
YW				
ANRM	530,093	19,400	549,493	533,822
HIER	529,371	19,867	549,238	533,629

ANRM = model based on the average numerator relationship matrix; HIER = hierarchical animal model.

**Table 3**

Posterior mean (PM), high posterior density at a 95% of probability (HPD 95%), and effective sample size (ESS) for genetic parameters of post weaning gain obtained by a model based on the average numerator relationship matrix (ANRM) and hierarchical animal model (HIER).

Model	Parameter	PM	HPD 95%	ESS
ANRM	$\sigma_a^2$	60.62	[59.19 to 67.29]	1396
	$\sigma_e^2$	250.88	[245.52 to 256.25]	2133
	$h_d^2$	0.19	[0.18 to 0.21]	1406
HIER	$\sigma_a^2$	61.80	[55.24 to 68.54]	1275
	$\sigma_e^2$	250.24	[244.87 to 255.59]	1968
	$h_d^2$	0.20	[0.18 to 0.22]	1284

$\sigma_a^2$  = additive genetic variance;  $\sigma_e^2$  = residual variance;  $h_d^2$  = direct heritability.

this criterion does not take into account the complexity of the model.

The effective sample size (ESS) for each variance component was calculated using the initial sequence of monotonic proposed by Geyer (1992) and was presented in Tables 3 and 4. The ESS estimates the number of independent samples with information equivalent to those contained within the sample dependent (Sorensen et al., 1995). The minimum expected for all components was 100 samples. It was found that in the present study, the number of cycles (400,000) used was adequate for convergence of the chain. For PWG, the ESS was up to 1200 samples for all components of variance, in addition, the HIER model showed a larger number of samples. For YW, all components of variance obtained a much smaller number of samples compared to samples of components of variance of the PWG. And all samples reached the minimum (over 100 samples).

The HIER model considers performance data and a prior probability of candidate sires to infer the posterior probabilities of a particular sire to be the true sire of an animal with uncertain paternity. Thus, this model should infer the genetic merit of animals with greater accuracy compared to the

**Table 4**

Posterior mean (PM), high posterior density at a 95% of probability (HPD 95%), and effective sample size (ESS) for genetic parameters for yearling weight obtained by a model based on the average numerator relationship matrix (ANRM) and a hierarchical animal model (HIER).

Model	Parameter	PM	HPD 95%	ESS	
ANRM	$\sigma_a^2$	147.34	[127.70 to 168.61]	277	
	$\sigma_e^2$	293.84	[282.40 to 305.04]	365	
	$\sigma_m^2$	10.33	[7.97 to 13.44]	210	
	$\sigma_{am}$	3.08	[-4.02 to 9.49]	211	
	$\sigma_{gc}^2$	1013.88	[956.23 to 1075.97]	35,642	
	$h_d^2$	0.31	[0.27 to 0.35]	268	
	$h_m^2$	0.02	[0.01 to 0.03]	215	
	$rg_{dm}$	0.08	[-0.09 to 0.26]	202	
	HIER	$\sigma_a^2$	155.64	[134.65 to 177.61]	231
		$\sigma_e^2$	290.46	[278.50 to 301.93]	301
$\sigma_m^2$		10.35	[8.04 to 13.70]	178	
$\sigma_{am}$		1.27	[-6.26 to 8.21]	138	
$\sigma_{gc}^2$		1013.49	[954.20 to 1076.48]	36,928	
$h_d^2$		0.32	[0.28 to 0.37]	223	
$h_m^2$		0.02	[0.01 to 0.03]	184	
$rg_{dm}$		0.04	[-0.14 to 0.22]	138	

$\sigma_a^2$  = additive genetic variance;  $\sigma_e^2$  = residual variance;  $\sigma_m^2$  = maternal variance;  $\sigma_{am}$  = covariance between maternal and direct effects;  $\sigma_{gc}^2$  = contemporary group variance;  $h_d^2$  = direct heritability;  $h_m^2$  = maternal heritability;  $rg_{dm}$  = correlation between direct and maternal genetic effects.

ANRM model. However, it was found in this study that the range of highest posterior density interval was higher for the HIER model for both traits, PWG and YW (Tables 3 and 4). Moreover, it was found that the HPD 95% of each variance component overlaps in all cases, indicating that there is practically no difference between the models to estimate this parameter.

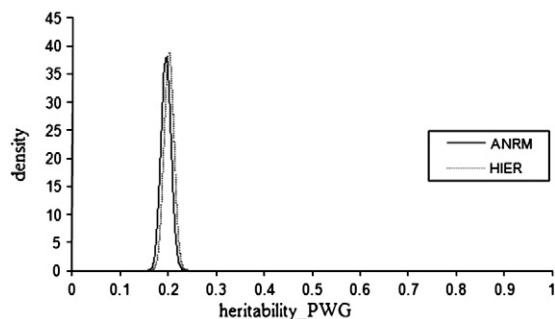
In general, posterior means of variance components estimated by ANRM and HIER models were similar for both traits (Tables 3 and 4). This work is in agreement with the results of Cardoso and Tempelman (2004), who concluded that there were no significant differences for the a posteriori average of variance components estimated by the ANRM and HIER models.

The slightly higher genetic variance posterior means were observed when using the HIER model for PWG and YW. The residual variance estimates were also lower with this model. These results are in agreement with Cardoso and Tempelman (2004), who found that the estimates of additive genetic variance calculated by ANRM and HIER models were close. However, these authors found that the HIER model had a tendency of underestimate the residual variance (Cardoso and Tempelman, 2003).

The average standard deviations (SD) of additive genetic variances were higher when estimated by HIER model (data not shown). These results were similar to those described by Cardoso and Tempelman (2004), who described that the SD of the additive genetic variances were similar between ANRM and HIER models, both considering only the base population animals and animals with certain paternity. However, for individuals with uncertain paternity, the SD of the additive genetic variance were higher when estimated by HIER model.

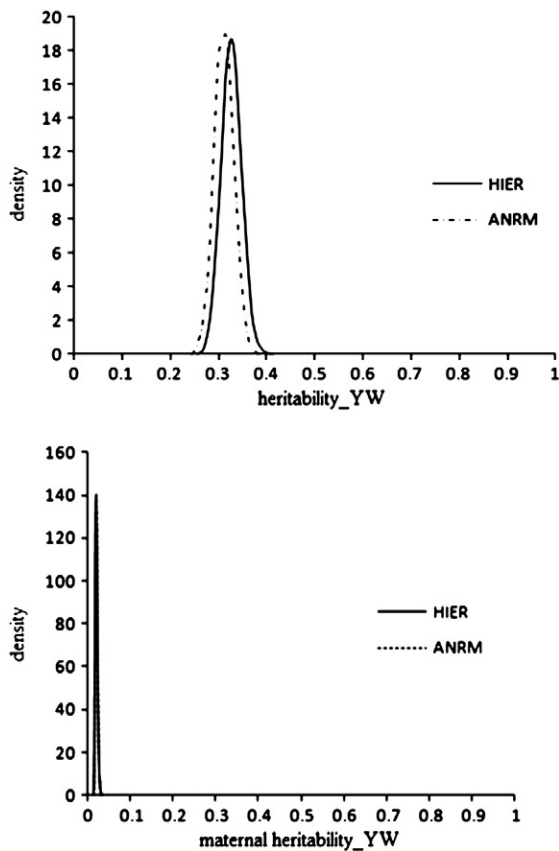
The posterior means the heritability ( $h_d^2$ ) obtained with the ANRM and HIER models overlapped almost completely, indicating that there were no practical differences between the models for this parameter (Fig. 1). These values are close to those reported in the literature for Zebu animals (Eler et al., 1996; Malhado et al., 2005; Marcondes, 1999; Paneto et al., 2002). According to these results, selection for PWG will probably result in long-term genetic gains.

Fig. 2 shows the posterior heritability means for direct genetic and maternal effects for YW. Posterior mean of heritability obtained with the ANRM model was closed when compared to the HIER (Table 4). Moreover, the high posterior



**Fig. 1.** Marginal posterior distribution of heritability for postweaning weight gain obtained with a model based on the average numerator relationship matrix (ANRM) and a hierarchical model (HIER).





**Fig. 2.** Posterior density of direct heritability (above) and maternal heritability (below) for yearling weight obtained with a model based on the average numerator relationship matrix (ANRM) and hierarchical animal model (HIER).

density intervals at 95% overlapped in an important part, indicating no differences between both estimates of this parameter. The posterior means obtained here are within the range of values reported by several authors for zebu animals (Bertazzo et al., 2004; Bittencourt et al., 2002; Boligon et al., 2008), thus confirming that YW will respond to individual selection.

The posterior heritability means for maternal genetic effects obtained with the ANRM and HIER models are close to those described in the literature for Zebu animals (Gunsky et al., 2001; Ribeiro et al., 2001). These results indicate little change of gain by selecting for this purpose in Nellore herd.

Covariances and correlations between genetic and maternal effects estimated with the two models were positive, as those described by Meyer (1992) in Angus. Differently, Boligon et al. (2008), Gunsky et al. (2001) and Ribeiro et al. (2001) also working with Nellore cattle, reported negative estimates for this correlation. The genetic direct and maternal covariance and correlation estimates overlapped in the 95% highest posterior density interval (HPD 95%), confirming that there are, virtually, no difference between the estimates obtained with the two models. In general, the choice between ANRM and HIER models does not interfere in the estimation of variance components for PWG and YW.

**Table 5**

Spearman correlation between means of additive genetic effects obtained in different sets of animals with a model based on the average numerator relationship matrix and a hierarchical animal model for post-weaning weight gain (PWG) and yearling weight (YW).

Animal set		Number of animals	Traits	
			PWG	YW
Animals <sup>a</sup>	Complete	75,088	0.999	0.999
	50%	37,544	0.999	0.998
	30%	22,526	0.998	0.998
	10%	7509	0.997	0.996
Offspring of MS <sup>b</sup>	Complete	22,810	0.999	0.999
	50%	11,169	0.995	0.998
	30%	6012	0.993	0.997
	10%	2281	1.000	0.990
Sires <sup>c</sup>	Complete	519	0.999	0.999
	50%	259	0.994	0.999
	30%	155	0.990	0.997
	10%	52	0.972	0.992

<sup>a</sup> Dataset was composed of all animals.

<sup>b</sup> Animals with uncertain paternity were included.

<sup>c</sup> All the sires were included.

Table 5 shows the rank correlations between breeding values predicted with the ANRM and HIER models for the three sets of animals. The rank correlations between breeding values predicted with the ANRM and HIER models for the two traits were high for all data sets (Table 5), ranging from 0.972 to 1.000. In general, the correlations between models suggest that the choice of the model does not interfere with the classification of animals when these data sets are evaluated. These results agree with those reported by Cardoso and Tempelman (2003), which also founded high correlation between genetic parameters estimates by ANRM and HIER models for traits with medium and high heritability, these authors reported the highest correlations when trait with medium heritability was evaluated.

#### 4. Conclusion

The model that best fits the data for estimating genetic parameter of animals with uncertain paternity is the Bayesian hierarchical model. However, for genetic evaluation, the choice between these models has no difference in classification of genetic value of animals.

#### Acknowledgments

The authors acknowledge the Foundation for Research Support São Paulo (FAPESP, 2006/ 58896-6) and Coordination of Improvement of Higher Education Personnel (CAPES, 4057/08-2) for financial support and Agriculture Jacarezinho for providing the data.

#### References

- Ali, T.E., Schaeffer, L.R., 1987. Accounting for covariances among test day milk yields in dairy cows. *Can. J. Anim. Sci.* 67, 637–644.
- Bertazzo, R.P., Freitas, R.T.F., Gonçalves, T.M., Pereira, I.G., Eler, J.P., Ferraz, J.B.S., Oliveira, A.I.G., Andrade, I.F., 2004. Parâmetros Genéticos de Longevidade e Produtividade de Fêmeas da Raça Nelore. *Rev. Bras. Zootec.* 33, 1118–1127.

- Bittencourt, T.C.C., Rocha, J.C.M.C., Lobo, R.B., Bezerra, L.F., 2002. Estimação de componentes de (co)variâncias e predição de DEP's para características de crescimento pós-desmama de bovinos da raça Nelore, usando diferentes modelos estatísticos. *Arq. Bras. Med. Vet. Zootec.* 54 (3), 303–308.
- Boligon, A.A., Albuquerque, L.G., Rorato, P.R.N., 2008. Associações genéticas entre pesos e características reprodutivas em rebanhos da raça Nelore. *Rev. Bras. Zootec.* 37 (4), 596–601.
- Cardoso, F.F., Tempelman, R.J., 2003. Bayesian inference on genetic merit under uncertain paternity. *Gen. Sel. Evol.* 35, 469–487.
- Cardoso, F.F., Tempelman, R.J., 2004. Genetic evaluation of beef cattle accounting for uncertain paternity. *Liv. Prod. Sci.* 89, 109–119.
- Eler, J.P., Ferraz, J.B.S., Silva, P.R., 1996. Parâmetros genéticos para peso, avaliação e circunferência escrotal na raça Nelore, estimados por modelo animal. *Arq. Bras. Med. Vet. Zootec.* 48 (2), 203–213.
- Gelfand, A.E., 1996. Model determination using sampling-based methods. In: Gilks, W.R., Richardson, S., Spiegelhalter, D.J. (Eds.), *Markov Chain Monte Carlo in Practice*. Chapman & Hall, London, p. 145.
- Geyer, C.J., 1992. Practical Markov chain Monte Carlo. *Stat. Sci.* 7, 473–511.
- Gunsky, R.J., Garneiro, A.D.V., Borjas, A.R.L., Bezerra, L.A.F., Lôbo, R.B., 2001. Estimativas de parâmetros genéticos para características incluídas em critérios de seleção em gado Nelore. *Ciência Rural, Santa Maria* 31 (4), 603–607.
- Henderson, C.R., 1988. Use of an average numerator relationship matrix for multiple-sire joining. *J. Anim. Sci.* 66, 1614–1621.
- Malhado, C.H.M., Martins Filho, R., Lobo, R.N.B., Facó, O., Azevedo, D.M.M.R., Souza, J.C., Oliveira, S.M.P., 2005. Tendências genéticas para características relacionadas à velocidade de crescimento em bovinos Nelore na região Nordeste do Brasil. *Rev. Bras. Zootec.* 34 (1), 60–65.
- Marcondes, C.R., 1999. Análise de alguns critérios de seleção para características de crescimento na raça Nelore. Belo Horizonte: Universidade Federal de Minas Gerais. Master Diss. Universidade Federal de Minas Gerais, Brazil.
- Meyer, K., 1992. Variance components due to direct and maternal effects for growth traits of Australian beef cattle. *Liv. Prod. Sci.* 31, 179–204.
- Paneto, J.C.C., Lemos, D.C., Bezerra, L.A.F., Martins Filho, R., Lôbo, R.B., 2002. Estudo de características quantitativas de crescimento dos 120 aos 550 dias de idade em gado Nelore. *Rev. Bras. Zootec.* 31 (2), 668–674.
- Ribeiro, M.N., Pimenta Filho, E.C., Martins, G.A., 2001. Herdabilidade para efeitos direto e materno de características de crescimento de bovinos Nelore no estado da Paraíba. *Rev. Bras. Zootec.* 30 (4), 1224–1227.
- Roso, V.M., Schenkel, F.S., 2006. AMC – a computer program to assess the degree of connectedness among contemporary groups. *Proc. 8th World Congress on Genetics Applied to Livestock Production*. Belo Horizonte, MG, Brazil, p. 27.
- Sapp, R.L., 2005. Statistical approach for dealing with uncertain paternity, PhD Diss. Univ. Georgia, Athens.
- Sapp, R.L., Zhang, W., Bertrand, J.K., Rekaya, R., 2007. Genetic evaluation in the presence of uncertain additive relationships. I. Use of phenotypic information to ascertain paternity. *J. Anim. Sci.* 85, 2391–2400.
- Sorensen, D., Gianola, D., 2002. Likelihood, Bayesian, and MCMC methods in quantitative genetics. ed. Springer-Verlag, New York.
- Sorensen, D.A., Andersen, S., Gianola, D., et al., 1995. Bayesian-inference in threshold models using Gibbs sampling. *Gen. Select. Evol.* 27 (3), 229–249.
- Spiegelhalter, D.J., Best, N.G., Carlin, B.P., Linde, A., 2002. Bayesian measures of model complexity and fit. *J. R. Stat. Soc. Ser. B Stat. Method.* 64, 583–616.