J. Dairy Sci. 92:4063–4071
doi:10.3168/jds.2008-1918

Comparison between a sire model and an animal model for genetic evaluation of fertility traits in Danish Holstein population

C. Sun,*† P. Madsen,* U. S. Nielsen,‡ Y. Zhang,† M. S. Lund,* and G. Su*[1]
*Department of Genetics and Biotechnology, Faculty of Agricultural Sciences, Aarhus University, DK-8830, Tjele, Denmark
†College of Animal Science and Technology, China Agricultural University, Beijing, China
‡Danish Agricultural Advisory Service, DK-8200 Aarhus N, Denmark

ABSTRACT

Comparisons between a sire model, a sire-dam model, and an animal model were carried out to evaluate the ability of the models to predict breeding values of fertility traits, based on data including 471,742 records from the first lactation of Danish Holstein cows, covering insemination years from 1995 to 2004. The traits in the analysis were days from calving to first insemination, calving interval, days open, days from first to last insemination, number of inseminations per conception, and nonreturn rate within 56 d after first service. The correlations between sire estimated breeding value (EBV) from the animal model and the sire-dam model were close to 1 for all the traits, and those between the animal model and the sire model ranged from 0.95 to 0.97. Model ability to predict sire breeding value was assessed using 4 criteria: 1) the correlation between sire EBV from 2 data subsets ($\mathrm{DATA_A}$ and $\mathrm{DATA_B}$); 2) the correlation between sire EBV from training data ($\mathrm{DATA_A}$ or $\mathrm{DATA_B}$) and yield deviation from test data ($\mathrm{DATA_B}$ or $\mathrm{DATA_A}$) in a cross-validation procedure; 3) the correlation between the EBV of proven bulls, obtained from the whole data set ($\mathrm{DATA_T}$) and from a reduced set of data ($\mathrm{DATA_{C1}}$) that contained only the first-crop daughters of sires; and 4) the reliability of sire EBV, calculated from the prediction error variance of EBV. All criteria used showed that the animal model was superior to the sire model for all the traits. The sire-dam model performed as well as the animal model and had a slightly smaller computational demand. Averaged over the 6 traits, the correlations between sire EBV from $\mathrm{DATA_A}$ and $\mathrm{DATA_B}$ were 0.61 (sire model) versus 0.64 (animal model), the correlations between EBV from $\mathrm{DATA_T}$ and $\mathrm{DATA_{C1}}$ for proven bulls were 0.59 versus 0.67, the correlations between EBV and yield deviation in the cross-validation were 0.21 versus 0.24, and the reliabilities of sire EBV were 0.42 versus 0.46. Model ability to predict cow breeding value was measured by the reliability of cow EBV, which increased from 0.21 using the sire model to 0.27 using the animal model. All the results suggest that the animal model, rather than the sire model, should be used for genetic evaluation of fertility traits.

Key words: animal model, female fertility, genetic evaluation, model validation

INTRODUCTION

Intensive selection for yield and unfavorable genetic correlation between yield and fertility traits has resulted in a downward genetic trend in the fertility of dairy cows (Weller and Ezra, 1997; Wall et al., 2003; Pryce et al., 2004). Poor fertility decreases economic efficiency in dairy cattle production because reproduction problems are often followed by extra inseminations and veterinary treatment costs, prolonged calving intervals, and greater rates of involuntary culling. Many studies have reported that poor fertility could be a major reason for involuntary culling of dairy cows (Westell et al., 1992; Esslemont, 1993; Olori et al., 2002). Therefore, fertility is considered one of the most important traits in the breeding goal after milk production and mastitis.

Female fertility is a combination of many factors. The traits often used in genetic evaluation of female fertility are those that reflect the ability to return to cycling after calving [days from calving to first insemination (**ICF**), the ability to conceive following insemination (days from the first to last insemination; **IFL**), number of inseminations per conception (**AIS**), and nonreturn rate within 56 d after first service (**NRR56**)] and the combination of these abilities [calving interval (**CI**) and days open (**DO**)] (Jorjani, 2006, 2007). These traits have a low heritability (Dematawewa and Berger, 1998; Pryce et al., 2004), and the data are not normally distributed and often include censored records (e.g., the records for cows that have not cycled or conceived within the inspecting period).

More sophisticated models (e.g., proportional hazards model, censored Gaussian model, censored threshold

---

Received November 21, 2008.
Accepted April 8, 2009.
[1] Corresponding author: guosheng.su@agrsci.dk

model, and threshold-linear model) have been proposed for genetic evaluation of fertility traits, with the concern on censoring and data distribution (Schneider et al., 2005; González-Recio et al., 2006; Urioste et al., 2007; Hou et al., 2009). However, the more sophisticated models require more computational resources and are more complicated to implement. Currently, a linear sire model (**SM**) is used for genetic evaluation of fertility traits in most countries, including Denmark (Interbull, 2009). The SM is theoretically inferior to the animal model (**AM**) in the estimation of variances and other genetic parameters (Everett et al., 1979; Schaeffer, 1983; Hudson and Schaeffer, 1984), but the superiority of the linear AM over the linear SM in the ability to predict breeding values has received less attention (Ramirez-Valverde et al., 2001).

Therefore, the aim of this study was to test the superiority of the AM, sire-dam model (**SDM**), and SM for genetic evaluation of fertility traits, based on the data from first lactation in the Danish Holstein population, with regard to the ability to predict sire breeding value (i.e., the future fertility performance of the daughters of sires). In addition, effects of the models on the reliability of cow EBV were investigated.

MATERIALS AND METHODS

Data

Female fertility data on the Danish Holstein population were obtained from the Danish Cattle Federation (Aarhus, Denmark). Detailed information on the breeding scheme for this population can be found at online (http://www.nordicebv.info/BreedingWork/Breedingwork.htm). The raw data included records from heifers and from the first 3 lactations of cows, covering insemination years from 1992 to 2006. The data from only the first lactation during insemination years 1995 to 2004 were used in the present study. The restriction to the period from 1995 to 2004 was imposed to exclude left-censored cows and cows with fertility events in progress (still undergoing inspection of fertility events). The traits in the analysis were ICF, CI, DO, IFL, AIS, and NRR56. The raw data were edited using the following 3 steps.

- ***Step 1***. Editing criteria in this step were as follows: 1) Age at first insemination as a heifer should be between 270 and 900 d, and age at first calving should be between 550 and 1100 d. 2) Herds should have records in each year from 1995 to 2004, and on average should have at least 50 records (sum of the number of records across heifers and the first 3 parities) per year. 3) Sires of cows should be known.
- ***Step 2***. Data were further edited for each particular trait in the first lactation. First, data from cows that were not inseminated and cows with ICF <20 d were deleted. Second, herd-year subclasses were required to have a minimum of 5 records, and sires should have at least 5 daughters with records.
- ***Step 3***. The aim of this step was to handle censored and extreme records. Approximately 15% of the cows had no known date of confirmed successful insemination. For these cows, the last insemination was taken as an unsuccessful insemination, and the corresponding records were taken as censored records. Many strategies have been used to handle censored records of fertility traits (Donoghue et al., 2004; González-Recio et al., 2006; Urioste et al., 2007; Hou et al., 2009). In this study, a penalty of 21 d was added to censored IFL, DO, and CI, and a penalty of 1 was added to censored AIS. This was a simple approach, although not satisfactory, to deal with censoring in a linear model setting. For cows without a subsequent calving, CI was calculated as DO plus 280 d. The following upper limits were imposed: 200 d for ICF, 600 d for CI, 320 d for DO, 230 d for IFL, and 8 for AIS. Records with values larger than the upper limit were replaced with the upper limit.

After editing, the whole data set (**DATA$_\mathbf{T}$**) contained 471,742 first-lactation records from 6,887 sires and 1,899 herds. Pedigrees for the 3 models were built by tracing the ancestors back as far as possible by using the sire-dam structure. Consequently, the pedigrees included 928,665; 645,444; and 23,744 individuals for the AM, SDM, and SM, respectively.

Another data set was created from the raw data by the above editing procedure, with the exception that herds were not required to have records in each year or were not required to have at least 50 per year. This approach resulted in a large data set with 1,050,494 records (**DATA$_\mathbf{DH}$**), which was used to investigate the reliabilities of EBV of fertility traits for first parity in the Danish Holstein population. In addition, a reduced data set (**DATA$_\mathbf{DHR}$**) was created from DATA$_\mathrm{DH}$ by leaving out the records of the last 2 insemination years, to estimate the reliabilities of EBV of the cows without their own records.

For the purpose of model validation, 3 data subsets were created from DATA$_\mathrm{T}$. Subsets **DATA$_\mathbf{A}$** and **DATA$_\mathbf{B}$** were created by a division of the whole data set (DATA$_\mathrm{T}$) randomly by herds. Subset **DATA$_\mathbf{C1}$** consisted of the records from first-crop daughters of