Application of Random Regression Models for Herd Life of Holstein Cattle in Japan

O. Sasaki[*], M. Aihara[†], K. Hagiya[‡], A. Nishiura[*], K. Ishii[*] and M. Satoh[*]

Introduction

The proportional hazard (PH) model (Ducrocq and Sölkner (1998)) and multiple-trait (MT) model (Madgwick and Goddard (1989); Veercamp *et al.* (2001)) are widely used for genetic evaluation of longevity. Initially, the PH approach was used only for single-trait sire models. In MT model, censored records are either not included in the analysis or treated as missing data (Jamrozik *et al.* (2008)). Veerkamp *et al.* (1999) showed that a random regression (RR) model for survival traits is relatively robust to the censoring of the data. Furthermore, time-dependent variables can be included in this model. The RR model allows the inclusion of animals that have not yet been culled and can be used for the genetic evaluation of the survival traits of cows (Schaeffer (2004)). The results of a simulation study show that genetic evaluation using RR model is less accurate than that using PH model (Jamrozik *et al.* (2008)). However, the RR model showed a slight superiority over MT and PH models in predicting the proportion of a sire's daughters that survive to the five different endpoints after the first calving. This finding suggests that a RR model can be used for genetic evaluation of the survival traits of dairy cattle.

The objective of this study is to estimate genetic parameters related to herd life by using a RR model. The predictive ability of RR model was assessed by correlating the estimated breeding values (EBVs) of the two data sets containing data obtained at different periods.

Material and methods

Data on the first 10 lactation periods were obtained from the Livestock Improvement Association of Japan. Data extraction was subject to the following conditions: (1) the herds had been enrolled in the herd-performance test from January 1991 to December 2005, and the number of parturitions per year was greater than or equal to 10; (2) the cow was 18–35 months old at first calving; and (3) the calving interval was between 305 and 730 days. The length of the herd life of the cow was considered as the number of days from the day of first calving to the day when the cow was culled. After every month (30 days), the animals were labeled as alive (1) or dead (0). The observation period for each cow began at its first calving and ended 80 months later. The cows underwent monthly observations until they were culled or until the end of the observation period. On the basis of the observation periods, the survival records were divided into two data sets. DATA1 included the survival records for

---

[*] National Institute of Livestock and Grassland Science, Tsukuba, Japan
[†] Livestock Improvement Association of Japan, Inc., Tokyo, Japan
[‡] National Livestock Breeding Center, Nishigo, Japan

the period 1991–2000; this data set contained data on the cows whose first calvings were recorded in the period 1991–1999. DATA2 contained survival records for the observation period 5 years after that of DATA1. Thus, DATA2 contained records of observations and first calving data from the periods 1996–2005 and 1996–2004, respectively. The survival records of 70,172 cows (DATA1) and 78,404 cows (DATA2) from 682 herds and the relationship records for three generations of 121,878 cows (DATA1) and 138,559 cows (DATA2) were included in the two data sets.

The following model was used in the analysis:

$$y_{ijklt} = \sum_{m=0}^{1} \theta_{im} z_{tm} + \sum_{m=0}^{1} \eta_{jm} z_{tm} + \sum_{m=0}^{1} \beta_{km} z_{tm} + \sum_{m=0}^{1} \alpha_{lm} z_{tm} + \sum_{m=0}^{1} \rho_{lm} z_{tm} + e_{ijklt}$$

where, y_{ijklt} is the survival value (0 or 1) for a cow l; this value was obtained t months after the first calving; θ_{im} are the fixed regression coefficients specific to the i^{th} age group at the first calving (≤ 24 months, ≤ 26 months, ≤ 28 months, ≤ 30 months, ≤ 32 months, and > 32 months); η_{jm} are the fixed regression coefficients specific to the j^{th} class of the 305-day milk yield (low, medium, and high) within a herd-year-parity class. There were 3 parity classes (first, second, and third and higher classes); β_{km} are the random herd-first calving year (HFY) coefficients, and these coefficients are specific to HFY class k; α_{lm} are the random additive genetic coefficients specific to the cow l; ρ_{lm} are the permanent random environmental coefficients specific to the cow l; e_{ijklt} is the random residual effect for each observation; and z_{tm} are the covariates. Vectors **b**, **a**, **p**, and **e** correspond to β_{km}, α_{lm}, ρ_{lm}, and e_{ijklt}, respectively. Further,

$$\operatorname{var}\begin{bmatrix}\mathbf{b}\\ \mathbf{a}\\ \mathbf{p}\\ \mathbf{e}\end{bmatrix} = \begin{bmatrix}\mathbf{B}\otimes\mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{G}\otimes\mathbf{A} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{P}\otimes\mathbf{I} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I}\sigma_{\mathbf{e}}^2\end{bmatrix},$$

where **B** is the variance-covariance matrix of HFY random regression coefficients of order m + 1; **G** is the variance-covariance matrix of the additive genetic random regression coefficients of order m + 1; **P** is the variance-covariance matrix of the permanent environmental random regression coefficients of orders m + 1; **A** is the additive genetic relationship matrix for the animals; **I** is an identity matrix; and σ_e^2 is the residual variance.

Variance components and regression coefficients of the RR model were estimated using VCE ver.6.02 (Groeneveld *et al.* 2008).

Results and discussion

For DATA1, in the 80-month period after the first calving, HFY variance, additive genetic variance, and permanent environmental variance increased from 9.4×10^{-5} to 1.8×10^{-3}, 1.7×10^{-4} to 3.7×10^{-3}, and 2.5×10^{-4} to 5.2×10^{-3}, respectively. The variances for DATA2 were similar to those for DATA1 and showed the same pattern. For DATA2, HFY variance, additive genetic variance, and permanent environmental variance increased from 7.9×10^{-5} to 1.8×10^{-3}, 1.5×10^{-4} to 3.7×10^{-3}, and 2.7×10^{-4} to 5.9×10^{-3}, respectively. The residual variance of DATA1 and DATA2 were 0.018 and 0.017, respectively. Heritabilities increased from 0.009 to 0.123 in DATA1 and from 0.009 to 0.132 in DATA2 (Figure 1). The heritability of herd life at an age of 84 months was estimated to be 0.09–0.18 by using linear models (Tsuruta *et al.* 2005; Hagiya *et al.* 2005). Using a MT model, Sewalem *et al.* (2007) showed that the heritability increased from 0.01 to 0.05 at the third calving. The heritability values obtained in this study were within the ranges presented in these reports.

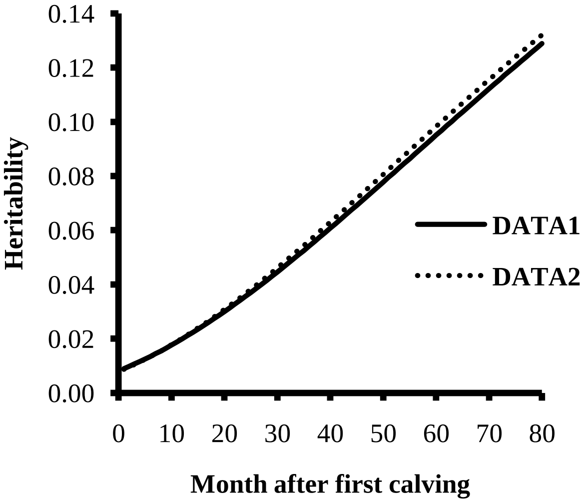

Figure 1: Trend of heritabilities of herd life at different months after the first calving.

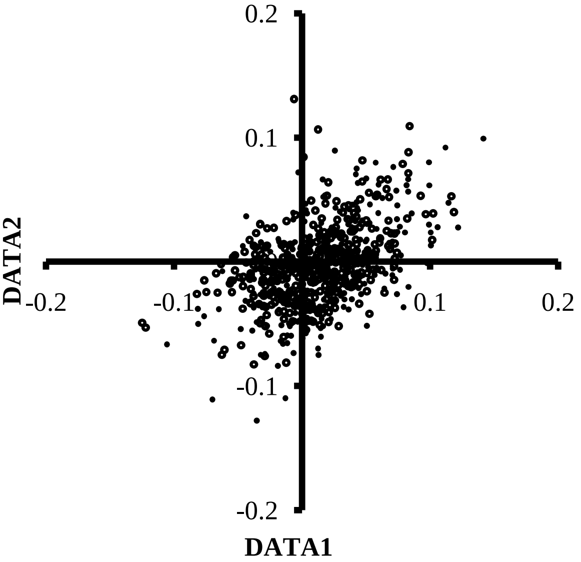

Figure 2: Correlation between the estimated breeding values of sires included in DATA1 and DATA2. The sires had 10 or more daughters with uncensored data. The sires were born between 1982 and 1992 (○) and between 1993 and 1995 (●).

EBVs of the sires at 80 months after the first calving were calculated for both data sets. DATA1 and DATA2 both included 650 sires that were born between 1983 and 1995 and had 10 or more daughters. The correlation coefficient of EBVs of these sires between DATA1 and DATA2 was 0.49, and the regression coefficient of DATA1 to DATA2 was 0.44 (Figure 2). The correlation and regression coefficients of EBVs of the 525 sires born between 1983 and 1992 were 0.47 and 0.38, and those of the 125 sires born between 1993 and 1995 were 0.55 and 0.54, respectively. By using a linear model, Caraviello *et al.* (2004) showed that the correlation coefficient of EBVs of sires for two data sets that were split into 2 samples by the random selection of herds was 0.76. The correlation coefficients obtained in this study were lower than those obtained using the analytical linear model; hence, further study is necessary for the accurate estimation of EBV by using RR model.

Conclusion

A linear model was used for the genetic evaluation of productive life of Holstein cattle in Japan. In a linear model, the analysis of productive life using censored records is difficult (Caraviello *et al.* 2004). By using a RR model, EBV for the herd life of young sires can be estimated as accurately as that of the old sires. This is because the EBV of DATA1 are correlated with those in DATA2, which includes the additional data of DATA1. Therefore, we expect that a RR model can be potentially used to evaluate herd life.

References

Caraviello, D., Weigel, K. A., and Gianola, D. (2004). *J. Dairy Sci.*, 87:3518–3525.

Ducrocq, V., and Sölkner, J. (1998). In *Proc. 6th WCGALP*, volume 27, pages 447–448.

Groeneveld, E., Kovac, M., and Mielenz, N. (2008). *VCE User's Guide and Reference Manual Version 6.0*.

Hagiya, K., Atagi, Y., Shirai, T. *et al.* (2005). *Nihon Chikusan Gakkaiho*, 76:159−165.

Jamrozik, J., Fatehi, J., and Schaeffer, L. R. (2008). *J. Anim. Breed. Genet.*, 125:75–83.

Madgwick, P. A. and Goddard, M. E. (1989). *J. Dairy Sci.*, 72:2624−2632.

Schaeffer, L. R. (2004). *Livest. Prod Sci.*, 86:35−45.

Sewalem, A., Miglior, F., Kistemaker , G. J. *et al.* (2007). *J. Dairy Sci.*, 90:2025−2028.

Tsuruta, S., Misztal, I., and Lawlor, T. J. (2005). *J. Dairy Sci.*, 88:1156−1165.

Veerkamp, R. F., Brotherstone, S., Engel, B. *et al.* (1999). *Interbull Bulletin.*, 21:36−40.

Veerkamp, R. F., Brotherstone, S., Engel, B. *et al.* (2001). *Anim. Sci.*, 72:1–10.