A Bayesian approach to the Japanese Black cattle carcass genetic evaluation

A. Arakawa[1#], H. Iwaisaki[1] and K. Anada[2]

[1] Graduate School of Agriculture, Division of Applied Biosciences, Kyoto University, Kyoto 606-8502, Japan
[2] Wagyu Registry Association, Kyoto 604-0845, Japan

Abstract

Volumes of the routine carcass field data used in the official genetic evaluation for carcass traits in Japanese Black cattle are increasing rapidly. The purposes of this paper are to describe a Bayesian approach via Gibbs sampling (GS) to be used in the Japanese Black carcass genetic evaluation, and in particular using a very large carcass field data, to compare breeding values for sires and dams of animals with records estimated using GS and the conventional empirical best linear unbiased prediction (EBLUP) employing the restricted maximum likelihood (REML) estimates of variance components. The total of six carcass traits evaluated were carcass weight, rib eye area, rib thickness, subcutaneous fat thickness, estimated muscle yield percentage and marbling score. As prior distributions for GS, uniform, normal and independently scaled inverted chi-square distributions with degree of belief and scale parameters of –2 and 0 were used for nuisance parameters, breeding values and variance components, respectively. The posterior means were calculated using every 10th sample from a total of 90 000 samples, after 10 000 samples had been discarded. Moment and rank correlations between breeding values estimated by the GS and the REML-EBLUP were all very close to one, and the coefficients and the intercepts for linear regression of the GS on the REML-EBLUP estimates were substantially one and zero, respectively, showing a very good agreement between breeding value estimation by the GS and the REML-EBLUP.

Introduction

In Japan, the genetic evaluation for carcass traits in Japanese Black (Wagyu) cattle, the main native beef breed, using routine carcass field data started in 1996 and has been carried out using an animal model in 42 production areas. The procedure used for predicting breeding values is a two-stage procedure, one that firstly performs the restricted maximum likelihood estimation (REML; Patterson & Thompson, 1971) of variance components using the average information (AI) algorithm (Gilmour *et al.*, 1995) and then implements the empirical best linear unbiased prediction (EBLUP). While the REML procedure has desirable properties, when applied to a large data set, it requires a large memory space. Since the volume of the official data set is rapidly increasing in certain production areas, an alternative approach requiring a less computational burden - in which estimate has desirable properties - needs to be developed.

The Bayesian analysis via Gibbs sampling (GS; Gelfand & Smith, 1990) that has often been used in animal breeding and genetics fields has certain advantages over REML (Van Tassell *et al.*, 1995), especially regarding memory space required for estimating variance components. Arakawa (2009) constructed a GS scheme, like REML, for estimating the relevant variance components and proposed a standard specification of the GS conditions to estimate these variance components using the large data set of carcass traits in Japanese Black cattle. This GS scheme worked well, when compared with the conventional REML-EBLUP procedure.

The purpose of this paper was to describe a Bayesian approach via GS to estimate breeding values for important carcass traits in Japanese Black cattle proposed by Arakawa (2009), and for sires and dams of animals with records in a large data set of Japanese Black carcass field data, to compare breeding values estimated using the GS scheme with those obtained with the conventional REML-EBLUP procedure.

Materials and Methods

Data used were provided by the Wagyu Registry Association, collected from 1987 through to 2007, and consisted of carcass records of 38 1491 Japanese Black cattle being fattened. The data set included

progeny records of 828 sires and 171 945 dams. The number of animals in the pedigree file was 616 743. Carcass traits analysed were carcass weight (CW), rib eye area (REA), rib thickness (RT), subcutaneous fat thickness (SFT), estimated muscle yield percentage (EYP), and marbling score (MS) that were assessed by certified graders of the Japan Meat Grading Association. All the carcass traits except for CW and EYP were measured between the sixth and the seventh rib section. The REA was measured by grid approximation with RT being the distance between the latissimus muscle and pleura membrane measured at the mid-point of the rib ends. The SFT was recorded as the distance between the latissimus muscle and carcass surface measured on the orthogonal line with carcass surface, from the edge of the iliocostalis muscle. The EYP was the estimated ratio of wholesale cuts, from which the surface fat had been trimmed, to CW and calculated using the following equation (JMGA, 1988):

$$EYP (\%) = 69.419 + 0.130 \times REA + 0.667 \times RT - 0.025 \times CW^* - 0.896 \times SFT$$

with CW* being a left-side CW. The MS measured the degree of marbling, especially in the REA and was categorized into 12 grades (zero to three at intervals of 0.33, 4 and 5) according to the Beef Marbling Standards (JMGA, 1988). This unusual classification for MS was used to ensure an approximate normal distribution. A higher MS indicated more intramuscular fat (Oyama *et al.*, 2004).

A single-trait individual animal model was used to describe the records of each trait. The operational linear model included discrete effects of fattening farms, year at slaughter, sex of the animals, animals' breeding values and residuals. Age at slaughter (mo) and the degree of inbreeding of the animal were considered as covariates.

Breeding values for sires and dams of animals with records (estimated using the GS and the conventional REML-EBLUP) were compared by correlation and regression analyses. In a Bayesian setting of the linear model, multivariate normal distributions were used as the priors for fattening farm effects and breeding values of animals. A flat improper prior distribution was used for the nuisance parameters, except for farm effect. For additive genetic and residual variance components, independent scaled inverted chi-square distributions were used, and the values of degree of belief and scale parameters were set at −2 and 0, respectively, since the prior distributions for the variance components were assumed to be non-informative. Samples from marginal distribution for all the effects in the model were generated by a single-site Gibbs sampler (Wang *et al.*, 1994). A total chain length of 100 000 iterations with a burn-in of 10 000 iterations was run, and the samples of the parameters of concern were saved for each of 10 iterations after the burn-in period. The samples then obtained were used to estimate features of the marginal posterior distributions such as the mean and standard deviation. In the REML-EBLUP, an AI-REML procedure in which fattening farm effect, animal's breeding value and residual were regarded as random, was used. Convergence of the REML estimates was reached when changes in the variance ratios between two consecutive rounds were all less than 10 − 8.

Results and Discussion

Means and standard deviations for the six traits were 426.4 ± 55.3 kg for CW, 51.2 ± 7.6 cm^2 for REA, 7.4 ± 0.9 cm for RT, 2.7 ± 0.9 cm for SFT, 73.2 ± 1.4% for EYP and 1.4 ± 0.7 grade point for MS. Heritability estimates and the standard error obtained by the GS was: 0.43 ± 0.01 for CW, 0.52 ± 0.01 for REA, 0.38 ± 0.01 for RT, 0.53 ± 0.01 for SFT, 0.59 ± 0.01 for EYP and 0.63 ± 0.01 for MS, respectively. The estimates were generally high, and the fractions of the variance due to fattening farm to phenotypic variance were in the range of seven for REA and EYP to 21% for CW.

Table 1 sets out moment and rank correlations for breeding values of sires and dams with records estimated using the GS and the conventional REML-EBLUP approaches and the results of linear regression of the GS on the REML-EBLUP estimates for the six carcass traits. Figure 1 illustrates a scatter plot of breeding value estimates of the sires and the dams for CW obtained using both the approaches. The moment and rank correlations were all found to be very close to one. The linear regression coefficients and the intercepts were all substantially one and zero, respectively, showing a very good relationship between breeding values estimated by the GS and the REML-EBLUP. A limited number of papers have been published on the comparison of Bayesian analyses and the mixed model method from the viewpoint of breeding value estimation (e.g., Wang *et al.*, 1994; Schenkel *et al.*, 2002; Ben Gara *et al.*, 2006). So for instance, using approximately 110 thousand records of milk yield in Holstein-Friesian cows, with 3 000

samples used to estimate the marginal distribution of parameters, rankings of bulls and cows according to the BLUP breeding values and the Bayesian estimates were very comparable (Ben Gara *et al.*, 2006).

With the previous numerical and the current results, it appears that the Bayesian and the REML-EBLUP implementations have essentially the same ability to rank animals. At this point, however, one shortcoming of EBLUP is that errors in the estimation of dispersion components are not taken into account when predicting the breeding values (Wang *et al.*, 1994). Contrary to this, Gibbs sampler can take it into account, by considering the marginal posterior distributions of breeding values. The uncertainty about the dispersion components is accounted for in the Bayesian probability intervals of predicted breeding values (Wang *et al.*, 1994).

Thus further detailed investigations on the current results may be necessary for the purpose of sampling variances of the estimated breeding values.

Table 1 Comparison between posterior means obtained with GS and estimates with REML-EBLUP for breeding values of carcass traits for parents of animals with records

Trait[1]	Correlation[2]		Regression[3]	
	Pearson	Spearman	Slope	Intercept
Sires[4]				
CW, kg	0.99998	0.99996	1.00020	0.13180
REA, cm^2	0.99998	0.99995	1.00033	0.01977
RT, cm	0.99998	0.99996	1.00030	0.00232
SFT, cm	0.99997	0.99994	0.99978	0.00287
EYP, %	0.99998	0.99996	1.00023	0.00390
MS, 0 to 5	0.99999	0.99997	1.00043	0.00143
Dams[4]				
CW, kg	0.99994	0.99993	1.00020	0.12250
REA, cm^2	0.99993	0.99992	0.99984	0.02176
RT, cm	0.99994	0.99994	1.00037	0.00212
SFT, cm	0.99991	0.99989	0.99974	0.00268
EYP, %	0.99994	0.99994	0.99989	0.00409
MS, 0 to 5	0.99997	0.99997	1.00027	0.00154

1) CW - carcass weight; REA - rib eye area; RT - rib thickness; SFT - subcutaneous fat thickness; EYP = estimated muscle yield percentage; MS - marbling score.
2) Pearson: Pearson's moment correlation, Spearman: Spearman's rank correlation.
3) Slope and intercept for the linear regression of posterior means with GS on estimates with REML-EBLUP.
4) Sires and dams: sires and dams of animals with record.
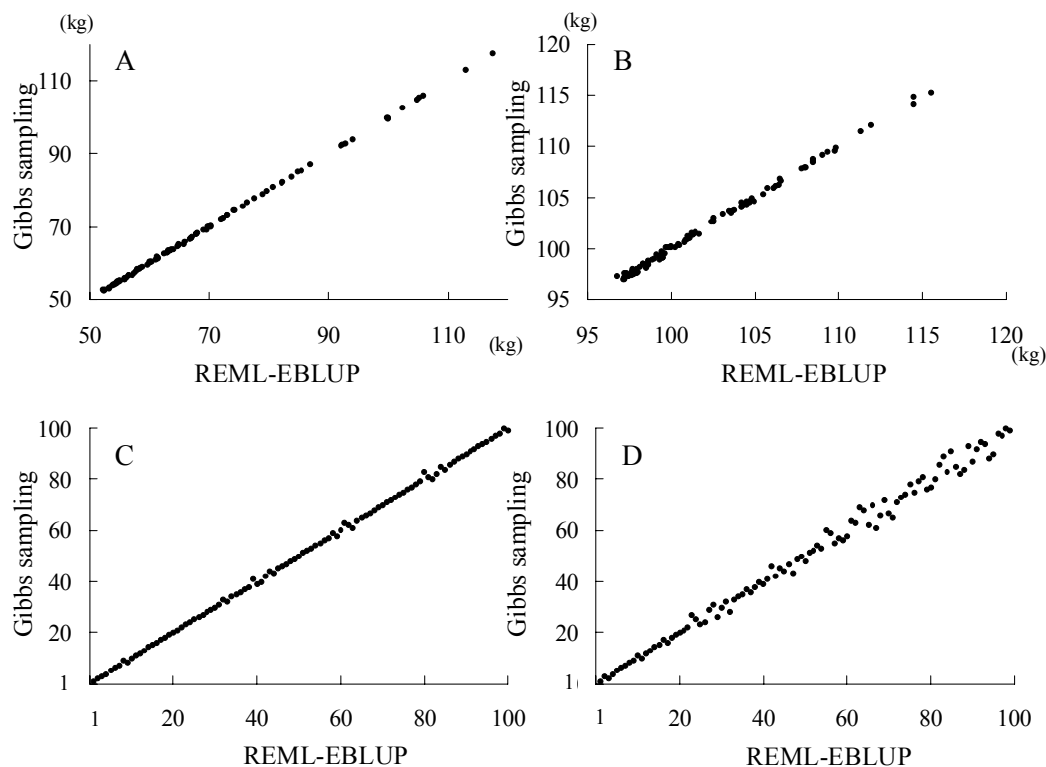
Figure 1 Scatter plots of values of the top 100 sires (A) and 100 dams (B) for breeding values of carcass weight estimated using the GS and the REML-EBLUP, and of the rankings for the sires (C) and the dams (D).

References

Arakawa, A., 2009. A Bayesian approach to genetic evaluation for carcass traits in Japanese Black cattle. Ph.D. thesis, Niigata University, Japan.

Ben Gara, A., Rekik, B. & Bouallègue, M., 2006. Genetic parameters and evaluation of the Tunisian dairy cattle population for milk yield by Bayesian and BLUP analyses. Livest. Sci. 100, 142-149.

Gelfand, A.E. & Smith, A.F.M., 1990. Sampling-based approaches to calculating marginal densities. J. Am. Statist. Assoc. 85, 398-409.

Gilmour, A.R., Thompson, R. & Cullis, B.R., 1995. Average information REML: An efficient algorithm for variance parameter estimation in linear mixed models. Biometrics 51, 1440-1450.

JMGA, 1988. New beef carcass grading standards. Japan Meat Grading Association, Tokyo, Japan.

Oyama, K., Katsuta, T., Anada, K. & Mukai, F., 2004. Genetic parameters for reproductive performance of breeding cows and carcass traits of fattening animals in Japanese Black (Wagyu) cattle. Anim. Sci. 78, 195-201.

Patterson, H.D. & Thompson, R., 1971. Recovery of inter-block information when block sizes are unequal. Biometrika 58, 545-554.

Schenkel, F.S., Schaeffer, L.R. & Boettcher, P.J., 2002. Comparison between estimation of breeding values and fixed effects using Bayesian and empirical BLUP estimation under selection on parents and missing pedigree information. Genet. Sel. Evol. 34, 41-59.

Van Tassell, C.P., Casella, G. & Pollak, E.J., 1995. Effects of selection on estimates of variance components using Gibbs sampling and restricted maximum likelihood. J. Dairy Sci. 78, 678-692.

Wang, C.S., Rutledge, J.J. & Gianola, D., 1994. Bayesian analysis of mixed linear models via Gibbs sampling with an application to litter size in Iberian pigs. Genet. Sel. Evol. 26, 91-115.