Effectiveness of adjusting for heterogeneity of variance in genetic evaluation of Japanese Black ...
Nakaoka, H;Narita, A;Ibi, T;Sasae, Y;et al
Journal of Animal Science; Oct 2007; 85, 10; ProQuest
pg. 2429

Effectiveness of adjusting for heterogeneity of variance in genetic evaluation of Japanese Black cattle[1]

H. Nakaoka,[*2] A. Narita,[*3] T. Ibi,[†] Y. Sasae,[†] T. Miyake,[*] T. Yamada,[*] and Y. Sasaki[*]

*Laboratory of Animal Breeding and Genetics, Graduate School of Agriculture, Kyoto University, Kyoto, 606-8502, Japan; and †Department of Animal Research, Agura Farm, Nasushiobara, 325-0033, Japan

ABSTRACT: Heterogeneity of variance among subclasses of an effect is a potential source of bias in genetic evaluation. The objectives of this study were to quantify the heterogeneity of variance in carcass weight in Japanese Black cattle, to develop an adjustment method to account for the heterogeneity, and to evaluate the effectiveness of the method. A total of 96,950 records were collected from steers and heifers slaughtered from 1997 to 2005. These records were grouped into 2,767 farm-market-year-sex subclasses. Fourteen log-linear models for the variances were set up to estimate the heterogeneous phenotypic variances within subclasses. Schwarz's Bayesian information criterion was used for model selection. The preadjustment of records to a baseline variance was based on maximum likelihood estimates obtained from the selected model. As a result of adjustment, the SD, the CV, and the Gini coefficient for the phenotypic variance decreased by 68.6, 69.8, and 70.1%, respectively. When the top 5% of sires and top 1% of dams were selected, Spearman's rank correlation coefficients between the adjusted and unadjusted data were 0.95 for the selected sires and 0.78 for the selected dams. The effectiveness of the adjustment was evaluated in terms of the ability to predict breeding values, using the results of the successive genetic evaluations. Mean squared error between the parent averages and actual predicted values of the genetic merit for the sires whose progeny had a carcass record only from 2003 to 2005 was significantly reduced by the adjustment ($P < 0.05$). The results suggest that the genetic evaluation becomes more accurate by adjusting the data using the procedure developed in this study.

Key words: beef cattle, carcass weight, genetic evaluation, heterogeneity of variance, model selection, statistical validation

INTRODUCTION

Many applications of the BLUP procedure (Henderson, 1973) assume that variance components are constant across subclasses of effects considered. However, genetic, environmental, and phenotypic variances seem to be heterogeneous in practice (Brotherstone and Hill, 1986). The bias caused by ignoring the heterogeneity of variance is severe, especially in the selection of dams of future sons, because progeny of a dam tend to make records in the same herd (Vinson, 1987).

Japanese Black cattle are a main beef breed in Japan. Genetic evaluation of sires and dams has been carried out, and dramatic genetic gains for carcass traits have been achieved within prefectures (Sasaki et al., 2006). Recently, a national genetic evaluation has provided the possibility of accelerating the genetic improvement. The management and environmental conditions of the fattening farms vary greatly from the north to the south of Japan; these differences may cause heterogeneity of variance and bias in the national cattle evaluation.

Several procedures for accounting for heterogeneous variances have been presented. Among them, the procedure using standardization of records before solving the mixed model equations (i.e., 2-step procedure) may work well when the size of the subclasses is relatively small and have applicability to BLUP with the animal model (Hill, 1984). Modeling of the heterogeneity of variance seems to be important in determining the effectiveness of the 2-step procedure; therefore, a 2-step procedure that allows setting and comparing various methods of modeling the heterogeneity of variance is needed.

The objectives of this study were 1) to quantify the heterogeneity of variance in Japanese Black cattle, 2) to develop a 2-step procedure that allows setting and comparing various methods of modeling the heterogeneity of variance, and 3) to evaluate the effectiveness of the method developed using field data from Japanese Black cattle.

MATERIALS AND METHODS

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

Data

Records of carcass weights of Japanese Black cattle were collected from 26 carcass markets from 1997 to 2005. All animals used for this study were fattened at 130 farms of Agura Kyosai Farm Ltd. across Japan. Odani et al. (2004) demonstrated that the fattening farms of Agura Kyosai Farm Ltd. across Japan were genetically connected, and a mathematical model containing the combination of fattening farm, market, year, and sex provided the best fit. Accordingly, farm-market-year-sex (**FMYS**) subclasses were created. The FMYS subclasses with less than 5 observations were excluded from the analysis. After data editing, a total of 96,950 carcass weight records with 2,767 FMYS subclasses were available for this study. The sample size of the FMYS subclasses ranged from 5 to 553. The average number of observations per FMYS subclass was 35.4. The number of FMYS subclasses with less than or equal to 10 observations was 425. Of the 96,950 total records, 57,461 records collected from 1997 to 2002 with 1,591 FMYS subclasses were extracted and denoted as data set I, which was used to quantify the heterogeneity of variance and to evaluate the method developed for its adjustment. The whole data set was denoted as data set II and was used to evaluate the effectiveness of the method developed in terms of the ability to predict breeding values.

Degree of Heterogeneity of Phenotypic Variance in Japanese Black Cattle

Within-FMYS phenotypic variance was estimated by the following expression:

$$\hat{s}_i^2 = (\mathbf{y}_i'\mathbf{y}_i - n_i\bar{y}_i^2)/(n_i - 1), \qquad [1]$$

where $\mathbf{y}_i$ = the $n_i \times 1$ vector of carcass weights; $\bar{y}_i$ = the mean of carcass weight; and n_i = the number of animals in the ith FMYS subclass.

Heterogeneity of the within-FMYS phenotypic variances was measured by Bartlett's test (Kendall and Stuart, 1979).

BLUP Evaluation

Breeding values were predicted using the following animal model:

$$y_{ij} = FMYS_i + b_1(t_{ij} - \bar{t}) + b_2(t_{ij} - \bar{t})^2 + u_j + e_{ij}, \quad [2]$$

where y_{ij} = the carcass weight of animal j in the ith FMYS subclass; $FMYS_i$ = the fixed effect of the ith FMYS subclass; b_1 and b_2 = linear and quadratic regression coefficients, respectively, on fattening period (i.e., period from start of fattening to shipping to market in number of days); t_{ij} and $\bar{t}$ = the fattening period of a particular animal and the arithmetic mean of the fattening period, respectively; u_j = the breeding value of animal j; and e_{ij} = the random residual. Ancestors were traced back 2 generations from the fattened steers and heifers and added as pedigree animals. Consequently, the total number of animals included in the analysis was 128,505 and 197,883 for data set I and II, respectively. Number of sires of steers and heifers included in the pedigree file was 739 and 911 for data set I and II, respectively. Variance components and breeding values were estimated using the set of multiple-trait, derivative-free, REML (MTDFREML) programs (Boldman et al., 1993).

Adjustment for Heterogeneity of Phenotypic Variance

Maximum Likelihood Estimation. The estimates of variance obtained from a maximum likelihood method with a log-linear model were used for the adjustment. The maximum likelihood estimation discussed by Harvey (1976) and Aitkin (1987) was applied. The general principle is outlined below:

$$y_i = \mathbf{x}_i'\beta + e_i, \ (i = 1, \ldots, n), \qquad [3]$$

where y_i = the carcass weight of animal i; β = the $l \times 1$ vector of parameters of fixed effects; and $\mathbf{x}_i$ = the $l \times 1$ vector of explanatory variables linking y_i and β. The error terms (e_i) are independently distributed as $N(0, \sigma_i^2)$. The variance model is the log-linear form:

$$\log \sigma_i^2 = \mathbf{z}_i'\lambda \ (i = 1, \ldots, n), \qquad [4]$$

where λ = the $m \times 1$ vector of parameters of fixed effects and $\mathbf{z}_i$ = the $m \times 1$ vector of explanatory variables linking $\log \sigma_i^2$ and λ. In addition, $\mathbf{z}_i$ may contain some or all of the explanatory variables included in $\mathbf{x}_i$ and other explanatory variables not included in $\mathbf{x}_i$.

Let $\mathbf{y}$ be the $n \times 1$ vector of observations. The log-likelihood under models [3] and [4] is:

$$\log L(\beta, \lambda; \mathbf{y}) = -\frac{1}{2}\left\{\sum_{i=1}^{n} \log \sigma_i^2 + \sum_{i=1}^{n} \frac{(y_i - \mathbf{x}_i'\beta)^2}{\sigma_i^2}\right\} =$$

$$-\frac{1}{2}\left(\sum_{i=1}^{n}\mathbf{z}_i'\lambda + \sum_{i=1}^{n} \frac{d_i}{\sigma_i^2}\right), \qquad [5]$$

where $d_i = (y_i - \mathbf{x}_i'\beta)^2$. The iterative equations for estimating β and λ are written separately as:

$$\hat{\beta}_{(t+1)} = (\mathbf{X}'\Sigma_t^{-1}\mathbf{X})^{-1}\mathbf{X}'\Sigma_t^{-1}\mathbf{y}, \qquad [6]$$

$$\hat{\lambda}_{(t+1)} = \lambda_t + (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'(\Sigma_t^{-1}\mathbf{d} - \mathbf{1}_n) = \qquad [7]$$

$$(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'(\Sigma_t^{-1}\mathbf{d} - \mathbf{1}_n + \mathbf{Z}\lambda_t),$$

where $\mathbf{d}$ = the $n \times 1$ vector whose ith element is d_i; $\mathbf{X}$ and $\mathbf{Z} = n \times l$ and $n \times m$ matrices of the explanatory variables, respectively; and Σ = the diagonal variance matrix whose ith diagonal element is σ_i^2. $\mathbf{1}_n$ = the $n \times 1$ vector of unit elements; and t = the tth iteration.

The iteration begins by taking $\sigma_i^2 \equiv \sigma^2$. The deviance ($-2\log L$) is calculated for each iteration. Equations [6] and [7] are alternated until the deviance converges. Convergence was considered to have been reached when the absolute difference between the deviances of 2 consecutive iterations was less than 10^{-8}.

Definition of the Mathematical Model in Maximum Likelihood Estimation. Only 1 mathematical model was formed as a mean model, Eq. [3], where β contains fixed FMYS subclass effects and fixed linear and quadratic regression coefficients on fattening period of the fattened animal.

On the other hand, various mathematical models were considered as a variance model, Eq. [4]. Model 1 was for the homogeneous modeling. For the heterogeneous modeling, 2 alternative methods of modeling were considered.

First, management-group modeling was considered. According to previous studies on dairy populations (Hill et al., 1983; Weigel et al., 1993; Ibáñez et al., 1996; Urioste et al., 2001), production levels of herds largely affect the heterogeneity of variance due to differences in management and due to scale effects, and factors such as period, herd size, and management practices were also reported as sources of heterogeneity of variance. In this study, the management factors (farm, market, year, and sex), the size of FMYS subclass (classified into 4 levels: 5 to 12, 13 to 23, 24 to 46, and >46 animals), and the production level of FMYS subclass (classified into 4 levels: <378 kg, 378 to 405 kg, 405 to 432 kg, and >432 kg) seemed to cause the heterogeneity of variance. Screening of these 6 factors as possible causes of heterogeneous phenotypic variances was performed using Levene's test (Levene, 1960). As a result, significant heterogeneity was detected across levels of all factors ($P < 0.001$), except for the size of FMYS subclass. These 5 factors were used to form the fixed effect part in Eq. [4]. Combining some of the 5 factors might better describe the heterogeneity of variance. Therefore, 6 different combinations of the 5 factors were set up as follows:

Model 2: F + M + Y + S + P;

Model 3: FY + M + S + P;

Model 4: FM + Y + S + P;

Model 5: FMS + Y + P;

Model 6: FMY + S + P; and

Model 7: FMYS + P,

where F = farm; M = market; Y = year; S = sex; and P = production level of FMYS subclass.

Alternatively, cluster modeling was considered to reduce the number of parameters fitted. Cluster modeling assumes that the variance is homogeneous among the subclasses of the same cluster and is heterogenous between subclasses of different clusters. The FMYS subclasses were grouped into clusters based on the simultaneous similarity of the phenotypic mean and variance. For the determination of the clustering, cluster analysis was performed using the SAS FASTCLUS procedure (SAS Inst. Inc., Cary, NC). Phenotypic means and variances were standardized using the STANDARD procedure of SAS when the distances between subclasses were calculated. Each cluster had at least 2 subclasses. The 7 models (models 8, 9, 10, 11, 12, 13, and 14) differed in number of clusters. The maximum numbers of clusters allowed in the clustering step for these 7 models were 10, 30, 50, 100, 200, 300, and 500, respectively. The factor of the production level of FMYS subclass was included as a fixed effect. The fixed effect part in Eq. [4] was:

$$CLFMYS + P,$$

where CLFMYS = cluster consisting of various numbers of FMYS subclasses and P = production level of the FMYS subclass (4 levels).

Model Comparison. Generally, as the number of parameters in the model increases, or the model becomes more complex, the fit is improved (the deviance decreases). A balance between fit and complexity is needed. Therefore, 14 models were compared, taking into account the balance. Schwarz's Bayesian information criterion (**BIC**; Schwarz, 1998) is a criterion based on parsimony and imposes a penalty on more complicated models. A smaller value of BIC indicates a better model. The choice of an optimal model was based on BIC. The BIC was calculated as:

$$BIC_K = -2\log L_K + p_K\log(n),$$

where L_K = the maximum of the likelihood within model K; p_K = the number of parameters in model K; and n = the number of observations.

Adjustment of Observations to Constant Phenotypic Variance. To account for the heterogeneity of the within-FMYS phenotypic variances, a 2-step procedure was adopted. Carcass weights were standardized to a baseline SD as follows,

$$y_{ij}^* = FMYS_i^{ML} + \frac{y_{ij} - FMYS_i^{ML}}{\sqrt{\exp(\mathbf{z}_{ij}'\hat{\lambda}^{ML})}} \sigma_{BASE}, \qquad [8]$$

where y_{ij}^* = the adjusted observation of the fattened animal j in the ith FMYS subclass; $FMYS_i^{ML}$ and $\hat{\lambda}^{ML}$ = maximum likelihood estimates of the fixed effect of the ith FMYS subclass and λ, respectively; and σ_{BASE} = the baseline SD. In this study, σ_{BASE}^2 was defined as 1,788.4 kg^2. After adjustment, within-FMYS phenotypic variances (s_i^{*2}) were estimated using Eq. [1], with $\mathbf{y}_i^*$ and $\bar{y}_i^*$, which are the adjusted observation and its mean, respectively.

Assessment of the Adjustment Based on Reduction in Heterogeneity of Phenotypic Variance

The effect of adjusting the observation for heterogeneity of within-FMYS phenotypic variances was visualized by Lorenz curves (Marshall and Olkin, 1979) and measured by 3 indices: SD of variances, CV of variances, and Gini coefficients (Urioste et al., 2001). In Lorenz curves, the cumulative proportion of the ordered subclass (x-axis) was plotted against the cumulative proportion of their variances (y-axis). If a Lorenz curve downwardly separates from the line of perfect equality (y = x), then a large degree of heterogeneity exists.

Validation of the Adjustment Based on Predicted Breeding Value

The effect of the adjustment on BLUP evaluation was investigated by comparing the predicted breeding values obtained using the unadjusted and adjusted data from data set I. The breeding values with the adjustment for heterogeneity of variance were obtained, replacing y_{ij} by y_{ij}^* in Eq. [2]. Spearman's rank correlations between predicted breeding values of the adjusted and unadjusted data were calculated to assess the reranking of animals.

Furthermore, the effect of the adjustment on the ability to predict breeding values was investigated by comparing the results of the successive genetic evaluations with data set I and data set II. The parent average (**PA**) for each of the sires whose progeny had a carcass record only in data set II was calculated from the predicted breeding values of their parents in data set I. The number of sires with more than 10 progeny that had a carcass record only in data set II was 47. Average number of progeny of these sires was 87.4. The PA values for these sires were compared with their predicted breeding values in data set II ($\hat{u}$). Parent average and $\hat{u}$ were considered as the expected values and actual predicted values of the genetic merit for the sires, respectively. Hence, the ability to predict breeding values was assessed by 3 statistics: 1) mean squared error (**MSE**) between PA and $\hat{u}$, 2) correlation coefficient be-

tween PA and $\hat{u}$ ($r_{\hat{u},PA}$), and 3) regression coefficient of $\hat{u}$ on PA ($b_{\hat{u}|PA}$). A small value of MSE indicates consistency between the expected values as parental average and the actual predicted values with performance progeny records and also indicates the ability to predict breeding values of future animals. The correlation $r_{\hat{u},PA}$ depends on both the precision and bias of the evaluations, and a value of unity is preferable. The deviation of $b_{\hat{u}|PA}$ from 1 indicates that bias exists (Reverter et al., 1994).

Least-squares ANOVA was performed using the GLM procedure of SAS to compare the unadjusted procedure and the heterogeneity-adjusted procedure based on MSE. The squared errors between PA and $\hat{u}$ for each of the sires derived from the 2 procedures were treated as dependent variables (94 observations), and the procedure (2 levels) and the individual sire (47 levels) were treated as fixed and random effects, respectively. The correlation coefficients and the regression coefficients were calculated using the CORR and REG procedures of SAS, respectively. The test of significance of the difference between 2 correlation coefficients and the t-test of 2 regression coefficients were performed (Snedecor and Cochran, 1980).

RESULTS AND DISCUSSION

The within-FMYS phenotypic variances of the unadjusted carcass weights ($\hat{s}^2$) were estimated, and 6 FMYS subclasses, for which $\hat{s}^2$ or sample size were extreme, were extracted and listed in Table 1. The largest $\hat{s}^2$ was about 85 times as large as the smallest one. Even when subclasses with more than 100 observations were considered, the $\hat{s}^2$ ranged from 786.8 to 3,496.6 kg^2. The Bartlett's test rejected homogeneity of the within-FMYS phenotypic variances ($P < 0.001$), indicating the necessity of accounting for heterogeneity of variance in genetic evaluation of Japanese Black cattle.

Several procedures of accounting for heterogeneous variances have been presented, using a 2-step procedure (Hill, 1984; Wiggans and VanRaden, 1991; Weigel and Gianola, 1993; Weigel and Lawlor, 1994; Dodenhoff and Swalve, 1998; Urioste et al., 2001), a multiplicative mixed model (Meuwissen et al., 1996; Robert-Granié et al., 1999), an empirical Bayesian method (Gianola et al., 1992), a structural model approach (Foulley et al., 1992; San Cristobal et al., 1993), and a multiple-trait approach (Henderson, 1984; Gianola, 1986; See, 1998).

The sizes of fattening farms and carcass markets of Japanese Black cattle are generally small (Sasaki, 1992, 2001); as a result, the sizes of the FMYS subclasses became relatively small. Therefore, the 2-step procedure seems to meet the conditions of this study. Modeling of the heterogeneity of variance was of interest in some of the 2-step procedures (Weigel and Lawlor, 1994; Urioste et al., 2001). A model containing management factors was fitted to the phenotypic variances within subclass, and the resulting solutions were used to estimate the prior variance for each subclass. The

Table 1. Comparison of the estimates of within-farm-market-year-sex (FMYS) phenotypic variances (kg^2) for carcass weight for extreme examples of FMYS subclass with respect to the phenotypic variance with the unadjusted data ($\hat{s}^2$) or sample size

Example FMYS subclasses	Observations per subclass	$\hat{s}^2$	Management-group modeling			Cluster modeling		$\hat{s}^{*2}$ [1]
			Model 2	Model 3	Model 7	Model 10	Model 14	
A	8	7,087.5	1,483.7	1,877.6	5,933.5	5,387.1	4,724.1	2,352.9
B	6	83.5	1,926.2	1,984.6	67.0	681.0	263.8	219.2
C	101	3,496.6	2,348.7	3,136.7	3,406.5	3,192.0	3,472.3	1,959.0
D	102	786.8	1,092.2	967.4	741.6	837.1	770.7	1,680.8
E	553	2,143.4	1,957.6	2,020.0	2,087.0	2,237.8	2,108.9	1,967.3
F	5	3,114.0	1,382.4	1,208.7	2,498.4	2,830.8	3,046.5	1,713.0

[1]The observations were adjusted based on the estimates with the best log-linear model (model 10). After the adjustment, $\hat{s}^{*2}$ were estimated.

prior variances and phenotypic variances within subclass were combined using Bayesian methods to obtain the posterior variances, and the posterior variances, in turn, were used to standardize the observations. In the series of steps, the method of modeling the heterogeneity to obtain the prior variance has not been fully discussed by previous authors. Therefore, a 2-step procedure that allows setting and comparing various methods of modeling the heterogeneity of variance was developed in this study.

Fourteen log-linear models were set and compared to obtain maximum likelihood estimates of the heterogeneous within-FMYS variances. The model selection criteria of the number of parameters, the deviance, and the BIC for each model are shown in Table 2. With respect to management-group modeling, the deviance decreased (the fit was improved), as more factors were combined in the model. In addition, estimates of the within-FMYS phenotypic variances obtained with models 2 and 3 were largely different from the $\hat{s}^2$ for small subclasses (example A, B, and F in Table 1), whereas those obtained with models 4 to 7, where a larger number of factors were combined, were comparatively close to the $\hat{s}^2$. Models 2 and 3 seem to be unsuitable to estimate the variances, especially for small subclasses. Models 4 to 7 fitted well, but a drawback of these models was that the extremely large number of parameters (overparameterization) might reduce the accuracy of the estimates. As it turned out, all of the models using management-group modeling were worse than homogeneous modeling in terms of BIC.

As a more simplified model, cluster modeling was considered. Even if a very small number of parameters were considered, the deviances for clustering modeling were relatively good. The deviances for models 9 to 14 were similar to each other. All of the models using cluster modeling were better than homogeneous modeling in terms of BIC. Consequently, model 10 seems to be best considering both fit and complexity. Then, the observations were adjusted with Eq. [8], using the maximum likelihood estimates with model 10. After adjustment, the within-FMYS phenotypic variances ($\hat{s}^{*2}$) were estimated with the adjusted data and included in Table

1. The $\hat{s}^{*2}$ were more homogeneous than the $\hat{s}^2$, at least in the case of the 6 subclasses shown in Table 1.

The distributions of the estimates of the within-FMYS phenotypic variances for carcass weight obtained from the unadjusted ($\hat{s}^2$) and adjusted data ($\hat{s}^{*2}$) are displayed as box-and-whisker plots (Figure 1). Box-and-whisker plots show the median, upper and lower quartiles, and minimum and maximum values of the estimates of the within-FMYS phenotypic variances. Although the distribution of the $\hat{s}^2$ had a long upper tail, that of the $\hat{s}^{*2}$ was more symmetric. The reduction in range and difference between the upper and lower quartiles were 55.0 and 69.0%, respectively, indicating that most of the subclasses moved closer to the median after the adjustment.

Table 2. Model selection criteria for the 14 log-linear variance models[1]

Variance model[2]	Parameters[3]	Deviance[4]	BIC
Homogeneous modeling			
Model 1	1	0	0
Heterogeneous modeling			
Management-group modeling			
Model 2	168	−1,201	630
Model 3	445	−1,883	2,983
Model 4	772	−2,370	6,080
Model 5	998	−2,713	8,213
Model 6	1,295	−3,354	10,827
Model 7	1,595	−3,788	13,681
Cluster modeling			
Model 8	11	−2,367	−2,257
Model 9	27	−3,145	−2,860
Model 10	39	−3,339	−2,922
Model 11	70	−3,545	−2,788
Model 12	129	−3,660	−2,257
Model 13	184	−3,682	−1,676
Model 14	254	−3,694	−921

[1]Deviance = $-2\log L$ and BIC = Schwarz's Bayesian information criterion.
[2]Formulas for each model are described in the text.
[3]The number of parameters of the corresponding log-linear variance model are listed; an additional 1,593 parameters of the mean model are actually fitted in the maximum likelihood estimation.
[4]Deviance and BIC are relative to the values of deviance and BIC for homogeneous modeling, which were 487,787 and 505,255, respectively.
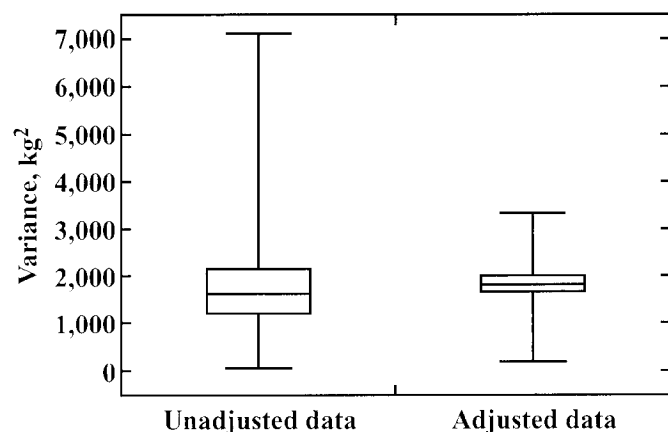
Figure 1. Box-and-whisker plots of within-farm-market-year-sex phenotypic variances for carcass weight with the unadjusted and adjusted data.

Table 3. Change in degree of heterogeneity of within-farm-market-year-sex (FMYS) phenotypic variances for carcass weights between unadjusted data and adjusted data

Parameter	Unadjusted data (A)	Adjusted data (B)	Reduction,[1] %
SD	839.1	263.1	68.6
CV	47.7	14.4	69.8
Gini coefficient	25.3	7.6	70.1

[1]Reduction of the heterogeneity of within-FMYS phenotypic variances was calculated as: $100 \times (A - B)/A$.

Lorenz curves of the within-FMYS phenotypic variances for carcass weight depict the existence of heterogeneity in the variances (Figure 2). The Lorenz curve moved closer to the line of perfect homogeneity after adjustment. Table 3 shows the change in heterogeneity due to the adjustment in terms of SD, CV, and Gini coefficients for the phenotypic variances. After the adjustment, all 3 measures of heterogeneity decreased by 68.6 to 70.1%. The adjustment procedure significantly reduced the heterogeneity of within-FMYS phenotypic variances.

As a reference, if the observations had been adjusted using the maximum likelihood estimates obtained with model 2, the 3 measures of heterogeneity would have been reduced by only 7.8 to 12.1%. Judging from the results of this study and Urioste et al. (2001), the procedure of modeling heterogeneity is especially important in determining the efficiency of the adjustment.
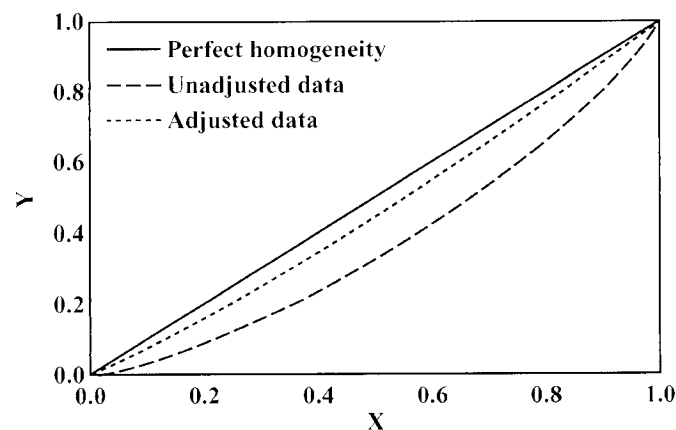
Spearman's rank correlation coefficients based on predicted breeding values obtained with the unadjusted and adjusted data were calculated within sex. The overall correlation was high: 0.997 for sires and 0.994 for dams. When only the top 1% of dams was used, however, the correlation dropped to 0.780, indicating reranking of the elite dams due to the adjustment. On the other hand, the correlation was still high (0.947) when only the top 5% of sires was used. Therefore, the reranking was greater for dams than for sires.

Sire ranking may also be affected when progeny are nonrandomly distributed between high and low variance environments (Vinson, 1987). However, the sire evaluation seems to be robust against heterogeneity of variance, because progeny of a sire are distributed across many subclasses (an average of 28.5 subclasses in this study). On the other hand, progeny of dams existed in only 1.6 subclasses on average; therefore, the predictions of dams were more likely to be biased by heterogeneity of variance. It is suggested that the effect of the adjustment on bull-dam selection is large. This finding is in agreement with previously reported results (Vinson, 1987; Meuwissen and Van der Werf, 1993; Urioste et al., 2003).

The ability to predict genetic values estimated by BLUP was compared between the unadjusted procedure and the heterogeneity-adjusted procedure based on the MSE between PA with data set I and their predicted breeding values obtained with data set II, the correlation coefficient ($r_{\hat{u},PA}$) between PA and $\hat{u}$, and the regression coefficient ($b_{\hat{u}|PA}$) of $\hat{u}$ on PA. The resulting statistics of the 3 criteria are shown in Table 4. The MSE obtained with the heterogeneity-adjusted procedure was smaller than that obtained with the unadjusted procedure, and the effect of the procedure on the MSE between PA and $\hat{u}$ was significant ($P = 0.026$). The values of $r_{\hat{u},PA}$ and $b_{\hat{u}|PA}$ obtained with the heterogeneity-adjusted procedure were slightly closer to 1.0 than those obtained with the unadjusted procedure, although differences between the procedures were not statistically significant. The significant reduction in MSE suggests that the genetic evaluation becomes more accurate using the adjustment for heterogeneity of variance. The consistency between the expected values as PA and the actual predicted values obtained with progeny perfor-

Figure 2. Lorenz curves of within-farm-market-year-sex phenotypic variances for carcass weight with the unadjusted and adjusted data and the line of perfect homogeneity (where y = x).

Table 4. Comparison between the unadjusted procedure and the heterogeneity-adjusted procedure regarding the ability to predict breeding values based on MSE, $r_{\hat{u},PA}$, and $b_{\hat{u}|PA}$[1]

Procedure	MSE	$r_{\hat{u},PA}$	$b_{\hat{u}|PA}$
Unadjusted	500.9	0.657	0.855
Adjusted	452.9	0.664	0.858
P[2]	*	NS	NS

[1]MSE = mean squared error between parent average (PA) with data set I and their predicted breeding values obtained with data set II ($\hat{u}$) for sires whose progeny had a carcass record only in data set II; $r_{\hat{u},PA}$ = the correlation coefficient between PA and $\hat{u}$; and $b_{\hat{u}|PA}$ = the regression coefficient of $\hat{u}$ on PA.

[2]Results of significance testing for the statistics for ability to predict breeding values between the unadjusted procedure and the heterogeneity-adjusted procedure.

*$P < 0.05$.

mance data is an important property of genetic evaluation, because the genetic performance of future animals is of interest to breeders and farmers. Consequently, the heterogeneity-adjusted procedure developed in this study proved to be effective in accounting for heterogeneity of variance.

In conclusion, an adjustment procedure based on a maximum likelihood method with a log-linear model was developed to account for heterogeneity of within-FMYS phenotypic variances in genetic evaluation. Application of the developed procedure to the analysis of carcass weights of Japanese Black cattle significantly reduced the heterogeneity of variance. The effectiveness of the procedure was evaluated in terms of the MSE between the expected values and actual predicted values of the genetic merit for the sires using the results of successive genetic evaluations. The procedure is simple to use and applicable to other quantitative traits or other breeds. This study was based on data collected from various farms owned by a commercial company using a unified recording system. If the data from different recording systems are pooled for national genetic evaluation of Japanese Black cattle, the benefits of accounting for heterogeneous variances could become more substantial.

LITERATURE CITED

Aitkin, M. 1987. Modelling variance heterogeneity in normal regression using GLIM. Appl. Stat. 36:332–339.

Boldman, K. G., L. A. Kriese, L. D. Van Vleck, and S. D. Kachman. 1993. A manual for use of MTDFREML. A set of programs to obtain estimates of variances and covariances. [Draft]. USDA-ARS, Clay Center, NE.

Brotherstone, S., and W. G. Hill. 1986. Heterogeneity of variance among herds for milk production. Anim. Prod. 42:297–303.

Dodenhoff, J., and H. H. Swalve. 1998. Heterogeneity of variances across regions of northern Germany and adjustment in genetic evaluation. Livest. Prod. Sci. 53:225–236.

Foulley, J. L., M. San Cristobal, D. Gianola, and S. Im. 1992. Marginal likelihood and Bayesian approaches to the analysis of heterogeneous residual variances in mixed linear Gaussian models. Comput. Stat. Data Anal. 13:291–305.

Gianola, D. 1986. On selection criteria and estimation of parameters when the variance is heterogeneous. Theor. Appl. Genet. 72:671–677.

Gianola, D., J. L. Foulley, R. L. Fernando, C. R. Henderson, and K. A. Weigel. 1992. Estimation of heterogeneous variances using empirical Bayes method: Theoretical considerations. J. Dairy Sci. 75:2805–2823.

Harvey, A. C. 1976. Estimating regression models with multiplicative heteroscedasticity. Econometrica 41:461–465.

Henderson, C. R. 1973. Sire evaluation and genetic trends. Pages 10–41 in Proc. Anim. Breed. Genet. Symp. Honor Dr. J. L. Lush. Am. Soc. Anim. Sci and Am. Dairy Sci. Assoc., Champaign, IL.

Henderson, C. R. 1984. Application of Linear Models in Animal Breeding. Univ. Guelph, Guelph, Ontario, Canada.

Hill, W. G. 1984. On selection among groups with heterogeneous variance. Anim. Prod. 39:473–477.

Hill, W. G., M. R. Edwards, M.-K. A. Ahmed, and R. Thompson. 1983. Heritability of milk yield and composition at different levels and variability of production. Anim. Prod. 36:59–68.

Ibáñez, M. A., M. J. Carabaño, J. L. Foulley, and R. Alenda. 1996. Heterogeneity of herd-period phenotypic variances in the Spanish Holstein-Friesian cattle: Sources of heterogeneity and genetic evaluation. Livest. Prod. Sci. 45:137–147.

Kendall, M. G., and A. Stuart. 1979. The Advanced Theory of Statistics. Volume 2. Inference and Relationship. 4th ed. Charles Griffin, London, UK.

Levene, H. 1960. Robust test for equality of variances. Pages 278–292 in Contributions to Probability and Statistics. I. Olkin, ed. Stanford Univ. Press, Palo Alto, CA.

Marshall, A. W., and T. Olkin. 1979. Inequalities: Theory of Majorization and its Applications. Acad. Press, New York, NY.

Meuwissen, T. H. E., G. De Jong, and B. Engel. 1996. Joint estimation of breeding values and heterogeneous variances of large data files. J. Dairy Sci. 79:310–316.

Meuwissen, T. H. E., and J. H. J. Van der Werf. 1993. Impact of heterogeneous within herd variances on dairy cattle breeding schemes: A simulation study. Livest. Prod. Sci. 33:31–41.

Odani, M., H. Nakaoka, A. Narita, T. Ibi, Y. Sasae, and Y. Sasaki. 2004. A national genetic evaluation of carcass traits for the Japanese Black cattle is possible. Nihon Chikusan Gakkaiho 75:353–361.

Reverter, A., B. L. Golden, R. M. Bourdon, and J. S. Brinks. 1994. Detection of bias in genetic predictions. J. Anim. Sci. 72:34–37.

Robert-Granié, C., B. Bonaïti, D. Boichard, and A. Barbat. 1999. Accounting for variance heterogeneity in French dairy cattle genetic evaluation. Livest. Prod. Sci. 60:343–357.

San Cristobal, M., J. L. Foulley, and E. Manfredi. 1993. Inference about multiplicative heteroskedastic components of variance in a mixed linear Gaussian model with an application to beef cattle breeding. Genet. Sel. Evol. 25:3–30.

Sasaki, Y. 1992. The effectiveness of the best linear unbiased prediction of beef sires using field data collected from small farms. J. Anim. Sci. 70:3317–3321.

Sasaki, Y. 2001. Beef cattle breeding and its trend in Japan. Asian-australas. J. Anim. Sci. 14:111–122.

Sasaki, Y., T. Miyake, C. Gaillard, T. Oguni, M. Matsumoto, M. Ito, T. Kurahara, Y. Sasae, K. Fujinaka, S. Ohtagaki, and T. Dougo. 2006. Comparison of genetic gains per year for carcass traits among breeding programs in the Japanese Brown and the Japanese Black cattle. J. Anim. Sci. 84:317–323.

Schwarz, G. 1998. Estimating the dimension of the model. Ann. Stat. 6:127–132.

See, M. T. 1998. Heterogeneity of (co)variance among herds for backfat measures of swine. J. Anim. Sci. 76:2568–2574.

Snedecor, G. W., and W. G. Cochran. 1980. Statistical Methods. 7th ed. Iowa State Univ. Press, Ames.

Urioste, J. I., D. Gianola, R. Rekaya, W. F. Fikse, and K. A. Weigel. 2001. Evaluation of extent and amount of heterogeneous variance for milk yield in Uruguayan Holsteins. Anim. Sci. 72:259–268.

Urioste, J. I., R. Rekaya, D. Gianola, F. Fikse, and K. A. Weigel. 2003. Model comparison for genetic evaluation of milk yield in Uruguayan Holsteins. Livest. Prod. Sci. 84:63–73.

Verbyla, A. P. 1993. Modelling variance heterogeneity: Residual maximum likelihood and diagnostics. J. R. Stat. Soc. B 55:493–508.

Vinson, W. E. 1987. Potential bias in genetic evaluation from differences in variation within herd. J. Dairy Sci. 70:2450–2455.

Weigel, K. A., and D. Gianola. 1993. A computationally simple Bayesian method for estimation of heterogeneous within-herd phenotypic variances. J. Dairy Sci. 76:1455–1465.

Weigel, K. A., D. Gianola, B. S. Yandell, and J. F. Keown. 1993. Identification of factors causing heterogeneous within-herd variance components using a structural model for variances. J. Dairy Sci. 76:1466–1478.

Weigel, K. A., and T. J. Lawlor. 1994. Adjustment for heterogeneous variance in genetic evaluations for conformation of United States Holsteins. J. Dairy Sci. 77:1691–1701.

Wiggans, G. R., and P. M. VanRaden. 1991. Method and effect of adjustment for heterogeneous variance. J. Dairy Sci. 74:4350–4357.