Comparison of Different Procedures to Account for Heterogeneity of Variance within Country

Fabiola Canavesi and Martino Cassandro
Research and Development, ANAFI, Via Bergamo 292, 26100 Cremona, Italy
fabiolacanavesi@anafi.it

Introduction

Heterogeneity of variances across herds in dairy cow populations has been found in different countries (Hill et al., 1983; Short et al., 1990; Stanton et al., 1991; Van der Werf et al., 1994; Wiggans and Van Raden, 1991). Bias in selecting bull dams and in evaluation of bulls, if not randomly distributed across herds, may result when homogeneity is assumed in statistical models used for genetic evaluations (Everett et al., 1982; Hill, 1984).

Different methods have been studied and proposed to account for heterogeneity in Animal Model evaluations (Mirande and Van Vleck, 1985; Hill, 1984; Gianola, 1986; Visscher et al., 1991; Garrick and Van Vleck, 1987; Kachman and Everett, 1992; Meuwissen et al., 1996), however, only a few of them have been applied. Current applied adjustments include mainly two approaches: preadjustment of data (Italy, United Kingdom, United States, Australia and Canada) and adjustment in the model (The Netherlands) to account for heterogeneity.

The objective of the present study is to compare different methods of adjusting for heterogeneous variances in the estimation of breeding values from real data.

Material and Methods

Data from the Italian official genetic evaluation of September 1997 were used to compare different methods of adjustment. Three genetic evaluations were computed using different approaches and then compared with results from the official run. There were three methods involved in the comparison: 1) no adjustment (**NOHV**); 2) preadjustment applied in Italy in the official evaluation (**OFF**) and 3) the preadjustment used in Canada (**CAN**).

In Italy and in Canada data are pre-adjusted at the phenotypic level following Hill's approach (1984): in Italy the adjustment is made on a herd-year basis (Bagnato and Jansen, personal communication), while in Canada the adjustment is made on a herd-year-parity basis (Robinson et al., 1994).

A total of three sets of estimated breeding values (EBV) were analysed in order to verify the efficiency of the different adjustment methods. Data consisted of 8,650,000 lactations from 1974 to 1997, from approximately 3,000,000 cows in 26,000 herds.

In order to compare the different methods on EBV, the following three effects were investigated: 1) genetic trend of bulls and cows; 2) deviation of progeny tested bull EBV from their parent average (PA) and the variation of this difference; 3) impact on MACE evaluations of different approaches considering frequencies of top 100 bulls and rank correlation of all bulls and of top 100 bulls. For more information on the last two comparisons, see Cassandro et al., 1997.

Results and Discussion

The trend of EBV for Italian proven bulls and cows population are reported in Figure 1 and 2, respectively. In each situation, the change from the current official method in Italy (**OFF**) shows an increase in EBV trend, both for bulls and cows. Similar results are reported for standard deviation of EBV (Figure 3) showing a higher variance over time when **NOHV** and **CAN** methods are applied to Italian data.

Number of animals by birth year are reported in Table 1. The table demonstrates that the percentage of imported bulls at the beginning of the ›80s is quite high in the Italian population, compared with other countries (around 50% of total bulls are imported). The national breeding scheme began in 1986 with the creation of the Genetic Centre in ANAFI (Cremona), and presently about 300 bulls are tested per year.

The average difference between EBV and Parent Average (PA) in the different scenarios is reported in Table 2. Results are similar for milk, fat and protein yield. Milk yield, which is measured on a larger scale, is used for a clearer comparison. Accurate evaluations will result in small average difference and variance. The average negative difference between EBV and PA is probably due to an overestimation of bull dams which tends to inflate PA. Since bias, due to the failure of accounting for heterogeneity of variance, is usually higher for those dams that perform in just one environment, the efficiency of the adjustment can be measured by considering the reduction in average difference between EBV and PA, and consequently in its variation. A small difference with a reduced variation may be an indication of a small bias. The highest difference is reported for **NOHV** method, whereas the lowest difference seems evident for **OFF** and **CAN** methods. The official method (**OFF**) is slightly better due to the fact that it derived from an analysis of Italian data and therefore was set up specifically for the Italian production system.

Table 4 reports the effect of the different methods to adjust for heterogeneity of variance on the rank of the top 100 bulls from MACE evaluations on the Italian scale for protein yield. Current method used in Italy (**OFF**) uses the same approach as **CAN** and therefore the two methods show similar results. Rank correlation among methods on top 1% bulls (Table 5) confirms these result, showing the lowest value between **NOHV** and **OFF** (.8180).

Conclusions

Different methods, used to adjust for heterogeneous variances across herds, show a clear effect on the genetic trend and on the top 100 bull ranking, resulting from MACE evaluation. This may be due to the effect on the scale of variability of EBV. This also raise the issues of proper definition of factors determining the heterogeneity to define the proper adjustment procedure. In particular, a method studied for a specific population and data structure seems to require a more comprehensive study on the adaptability of the method to different conditions. Further research needs to be conducted on the effects using different methods to account for heterogeneity of herd variances on genetic trend and the impact on international comparisons, as well as the effectiveness of different methodologies to account for heterogeneity on accuracy of EBV.

References

Cassandro, M., Miglior, F., Carnier, P., Bittante, G., Canavesi, F., Santus, E. & Banos, G. 1997. Proceedings of the 1997 Interbull Meeting, Vienna, Austria, August 28-29, *Bulletin No. 16*, 16-20.

Everett, R.W., Keown, J.F. & Taylor, J.F. 1982. The problem of heterogeneous within herd error variances when identifying elite cows. *J. Dairy Sci. 65 (Suppl. 1)*, 100. (Abstr.)

Garrick, D.J. & Van Vleck, L.D. 1987. Aspects of selection for performance in several environments with heterogeneous variance. *J. Anim. Sci. 65*, 409-421.

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

Hill, W.G. 1984. On selection among groups with heterogeneous variance. *Anim. Prod. 36*, 473-477.

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

Kachman, S.D. & Everett, R.W. 1992. A multiplicative mixed model when the variances are heterogeneous. *J. Dairy Sci. 76*, 859-867.

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

Mirande, S.L. & Van Vleck, L.D. 1985. Trends in genetic and phenotypic variances for milk production. *J. Dairy Sci. 68*, 2278-2286.

Robinson, J.A.B., Chesnais, J.P., Abraham, P.A., Osterhout, T.D. & Luu, D.K. 1994. Canadian genetic evaluation for dairy cattle production traits. *Proc. of the 5th WCGALP, Guelph*, Ont., Canada.

Short, T.H., Blake, R.W., Quaas, R.L. & van Vleck, L.D. 1990. Heterogeneous within-herd variance 1. Genetic parameters for first and second lactation milk yield of grade Holstein cows. *J. Dairy Sci. 73*, 3312-3320.

Stanton T. L., Blake, R.W., Quaas, R.L., Van Vleck, L.D. & Carabano, M.J. 1991. Genotype by environment interaction for Holstein in milk yield in Colombia, Mexico and Puerto Rico. *J. Dairy Sci. 74*, 1700-1714.

Van der Werf, J.H.J., Meuwissen, T.H.E. & de Jong, G. 1994. Effects of correction for heterogeneity of variance on bias and accuracy of breeding value estimation in Dutch dairy cattle. *J. Dairy Sci. 77*, 3174.

Visscher, P.M., Thompson, R. & Hill, W.G. 1991. Estimation of genetic and environmental variances for fat yield in individual herds and an investigation into heterogeneity of variance between herds. *Livest. Prod. Sci. 78*, 273-290.

Wiggans, G.R. & Van Raden, P.M. 1991. Method and effect of adjustment for heterogeneous variance. *J. Dairy Sci. 74*, 4350-4357.

Table 1. Italian, foreign bull and cow frequencies by birth year.

Birth year	Proven bulls	Foreign bulls	Cows
1980	41	44	114,948
1981	34	52	123,761
1982	61	43	131,059
1983	74	29	138,929
1984	91	28	152,350
1985	96	35	158,165
1986	212	43	174,104
1987	246	23	183,626
1988	248	26	191,762
1989	291	4	198,509
1990	326	-	204,684
1991	282	-	198,401
1992	208	-	200,380
1993	8	-	199,767
1994	-	-	166,213
1995	-	-	21,733

Table 2. Average difference between EBV and Parent Average (PA) for Italian proven bulls.

Birth year	OFF	NOHV	CAN
1980	127.353	143.147	128.706
1981	-102.926	-120.103	-112.750
1982	-13.098	-23.213	-19.516
1983	74.865	60.885	70.419
1984	-13.357	-29.670	-22.088
1985	-67.938	-88.615	-80.672
1986	-25.708	-50.491	-37.847
1987	-44.411	-84.866	-57.813
1988	-16.760	-53.679	-30.381
1989	26.820	-6.442	17.790
1990	13.571	-14.597	4.324
1991	-42.523	-81.482	-56.108
1992	-71.650	-115.060	-83.727
1993	-75.938	-117.500	-90.438

Table 3. Variance of average difference between EBV and Parent Average (PA) for Italian proven bulls (H1000 kg).

Birth year	OFF	NOHV	CAN
1980	115	146	131
1981	86	110	95
1982	120	152	127
1983	98	126	108
1984	158	208	172
1985	115	157	126
1986	110	153	123
1987	108	150	120
1988	123	172	136
1989	102	148	114
1990	91	134	104
1991	100	148	118
1992	77	116	93
1993	94	133	118

Table 4. Effect of different methods for adjusting heterogeneity of variance on ranking from MACE evaluations on Italian scale for protein yield (top 100 bulls) .

Method	From			
of adjustment	ITA	USA	NLD	DEU
OFF	1	67	21	11
NOHV	2	70	19	9
CAN	1	71	18	10

OFF = official run used in Italy; NOHV = official run without correction for heterogeneity of variance HV); CAN = Canadian method used for correction of HV.

Table 5. Rank correlation among different methods for adjusting heterogeneity of variance on Italian scale. Above diagonal all bulls (No. 28,186), below diagonal top 1% of bulls (No. 300).

	NOHV	OFF	CAN
NOHV	-	0.9978	0.9993
OFF	.8180	-	.9990
CAN	.9609	.9162	-

OFF = official run used in Italy; NOHV = official run without correction for heterogeneity of variance (HV); CAN = Canadian method used for correction of HV.

Figure 1 - Bulls EBVs trend

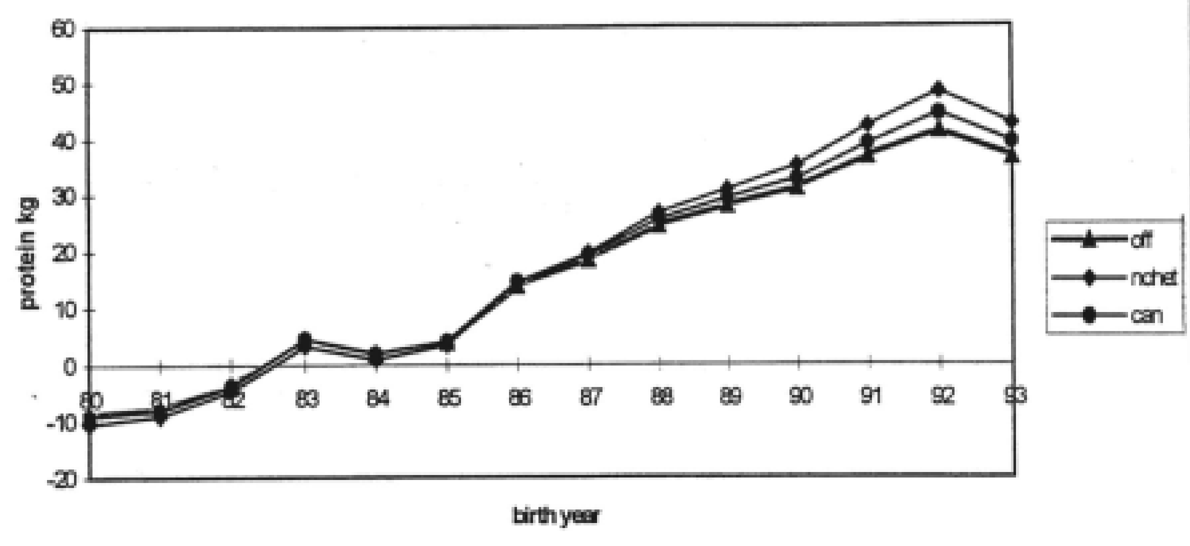

Figure 2 - Cow EBVs trend

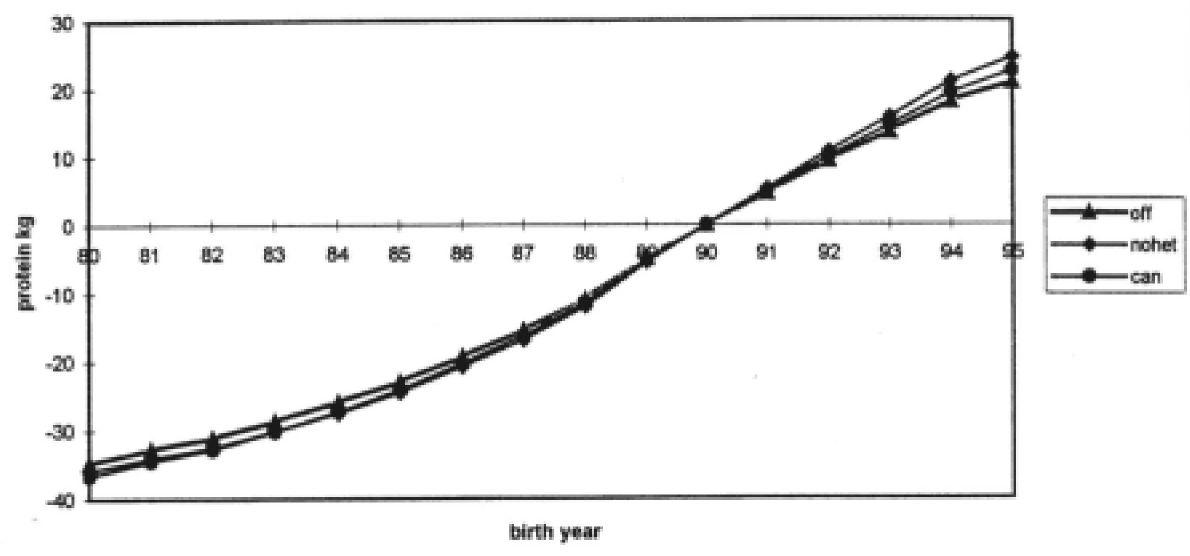

Figure 3 - Italian proven bulls EBVs standard deviations