IMPLEMENTATION OF CORRECTION FOR HETEROGENEOUS WITHIN HERD VARIANCES

G. de Jong, NRS, Arnhem, The Netherlands
T.H.E. Meuwissen, ID-DLO, Zeist, The Netherlands

1. Introduction

In April 1995 a correction for heterogeneous within herd variances was introduced in the genetic evaluation for milk production traits in The Netherlands. The applied method is described in Meuwissen et al (1994 and 1995). Results from research on this method are presented in this paper.
The goal of the research was estimation of two parameters needed for the correction for heterogeneous variances. The two parameters are the correlation between the herd*year heterogenity correction factors (rho) of adjacent years within a herd and the variance of these correction factors (vhy). Further effects on breeding values are also investigated.

2. Data

Milk production data from the national Dutch genetic evaluation run of April 1995 were used for this research. This data set consisted of lactation records with calving dates from July 1978 to December 1994. In total 12,629,403 lactation records of 5,819,607 cows in 42,480 herds were analysed, where 499,608 herd-year-seasons were defined.

The same animal model as used for the national genetic evaluation is used in the analysis, with the same precorrections:

$$Y_{ijklm} = HYS_i + KM_j + PERM_{kl} + A_k + error$$

where:

Y_{ijklm} = 305-days lactation for milk fat and protein yield, with precorrection for:
- age at calving
- days open during the lactation
- heterosis and recombination;

HYS_i = herd-year-season-parity class i, in which the lactation is compared with lactations of herdmates;
KM_j = calving month*year j, when the lactation was started;
$PERM_{kl}$ = permanent environment effect kl of cow k in herd l (repeatability=.55);
A_k = breeding value of cow k (heritability=.35);
error = error term of observation Y_{ijklm}.

For A_k the relation matrix is added, following the procedures of an animal model.

3. Results

3.1 Parameters

Correlations between herd*year heterogenity correction factors (rho) for milk yield vary between 0.967 and 0.998. The highest values are found between the years at the beginning and end of the data set (table 1). The estimates for rho approximately follow a hyperbolic function. The same results are found for fat and protein yield with the correlation for these traits also high. Fat yield varies between 0.959 and 1.00 and protein yield between 0.963 and 0.992. The average correlation between the different years is 0.984 for milk, 0.983 for fat yield and 0.980 for protein yield.

The reason for the higher rho in the middle of the time span of the data is the influence adjacent years have on both sides of the year concerned, while the years at the side of the time span are affected mainly from one side (later or previous years).
Considering the estimated rho, the variance lactation records of different years does not vary much from year to year.

Table 1. Correlation between the within herd herd*year heterogenity correction factors (rho) and the variance of these factors (vhy) per year for the traits milk, fat and protein yield.

Year	rho				vhy			
	n	trait			n	trait		
		kg milk	kg fat	kg prot		kg milk	kg fat	kg prot
oct78-sep79	25126	0.992	0.994	0.990	25126	0.107	0.127	0.147
oct79-sep80	26253	0.991	0.994	0.989	26253	0.108	0.127	0.147
oct80-sep81	26804	0.990	0.993	0.985	27228	0.108	0.123	0.144
oct81-sep82	27404	0.988	0.991	0.982	27734	0.105	0.117	0.139
oct82-sep83	28040	0.984	0.985	0.978	28409	0.102	0.110	0.134
oct83-sep84	28348	0.977	0.973	0.972	28998	0.100	0.103	0.129
oct84-sep85	28564	0,974	0.967	0.969	29148	0.095	0.094	0.123
oct85-sep86	28373	0.967	0.959	0.963	29268	0.092	0.088	0.120
oct86-sep87	27599	0.970	0.963	0.966	28806	0.091	0.085	0.118
oct87-sep88	26913	0.976	0.973	0.972	27956	0.092	0.086	0.120
oct88-sep89	26401	0.982	0.980	0.977	27211	0.095	0.089	0.123
oct89-sep90	25851	0.985	0.984	0.981	26642	0.099	0.094	0.129
oct90-sep91	25017	0.991	0.992	0.986	26025	0.104	0.099	0.136
oct91-sep92	24273	0.992	0.992	0.987	25235	0.110	0.105	0.143
oct92-sep93	23529	0.994	0.995	0.992	24406	0.114	0.109	0.149
oct93-sep94	20179	0.998	1.003	0.989	23529	0.114	0.108	0.149
oct94-sep95		-	-	-	20179	0.112	0.104	0.141
average		0.984	0.983	0.980		0.103	0.104	0.135

The variance of herd*year heterogenity correction factors (vhy) varies between 0.091 and 0.114 for milk yield, with an average of 0.103 (table 1). The lowest estimates are found for the years in the middle of the time span of the data set. For fat yield the vhy varies between 0.085 and 0.127, with an average of 0.104. The estimates for protein yield are higher: varying between 0.118 and 0.149, with an average of 0.135.

3.2 Breeding values

A test to check how the correction method works is to compare parent average (PA) of progeny tested bulls with the realized breeding value (RBV) of these bulls. When comparing the results of no correction for heterogenity and applying the correction method (table 2), it is obvious that the average PA is higher than the average realized breeding value. The difference between average PA and average realized breeding value is smaller (25%) with correction than without correction. The genetic trend in breeding values for tested bulls is similar for both methods, 1049 kg milk without correction and 1086 kg milk with correction, in 10 years (table 3). The same is found for the trend in PA with 1119 kg milk without correction and 1111 kg milk with correction, in 10 years. Taking a closer look where differences between PA and RBV start for both methods (table 2), it can be seen that dams of bulls especially cause differences between PA and RBV (table 3). The dams' breeding values appear to be overestimated values while sires

Table 2. Realized breeding values minus parent average for progeny tested bulls per year of birth.

Year of birth	Number	Without correction	With correction
81	87	-97	-89 (8%)
82	139	-158	-143 (8%)
83	153	-106	-82 (23%)
84	226	-152	-123 (19%)
85	206	-104	-80 (23%)
86	200	-190	-159 (16%)
87	167	-150	-94 (37%)
88	158	-215	-160 (26%)
89	105	-181	-130 (28%)
90	78	-167	-114 (32%)

*(Bulls have at least 75% HF genes and their dams realized a lactation in The Netherlands. The base for the correction for heterogeneous variances is the average correction factor for cows born in 1990.)

of bulls do not show differences for both methods. On average dams are estimated 60 to 80 kg milk higher without correction for heterogenity than with this correction.
The average breeding value of The Netherlands progeny tested bulls (having a first crop of daughters) are 20 kg milk higher with correction than without correction (table 4). Imported bulls (with second crop daughters) received, on average, a 20 kg lower breeding value for milk yield with correction compared to without correction. The absolute differences (about 55 kg milk) and standard deviation of differences (about 70 kg milk) are higher for imported bulls than for progeny tested bulls. Progeny tested bulls showed an absolute difference of breeding values of 30 kg milk and a standard deviation of 30 kg milk for both methods. The changes for imported bulls were twice as large than for progeny tested bulls, although differences were relatively small.

Table 3. Realized breeding values(RBV), parent average(PA) and breeding values for sires and dams for progeny tested bulls per year of birth.

Year of birth	Number	Without correction				With correction			
		RBV	PA	dam	sire	BV	PA	dam	sire
81	87	-596	-499	-734	-263	-618	-529	-813	-246
82	139	-565	-408	-667	-148	-583	-440	-749	-131
83	153	-371	-265	-593	64	-372	-290	-656	76
84	226	-329	-177	-513	160	-323	-199	-572	174
85	206	-278	-174	-482	134	-271	-191	-535	154
86	200	-183	8	-249	264	-168	-9	-302	285
87	167	100	250	23	478	123	217	-39	474
88	158	145	359	194	5225	167	327	131	522
89	105	251	431	313	550	270	400	240	559
90	78	453	620	548	693	468	582	477	686
81-90		1049	1119			1086	1111		

*(Bulls have at least 75% HF genes and their dams realized a lactation in The Netherlands. The base for the correction for heterogeneous variances is the average correction factor for cows born in 1990.)

More changes in breeding values are also found with imported bulls than with progeny tested bulls for fat yield and protein yield (for protein see table 5). The standard deviation of differences for protein yield was 2.2 kg, twice the standard deviation found for progeny tested bulls (1.1 kg protein).
The larger differences of breeding values for imported bulls may be caused by the kind of herd their daughters are producing in. The average production level of herds using imported semen is higher than the average herd production level. As it is known that with the increase of production level, the variation of lactations also increases, it is logical that imported bulls are affected more by the correction method than progeny tested bulls.

Table 4. Comparison of breeding values for milk yield for progeny tested bulls and imported bulls.

Progeny tested bulls						Imported bulls					
year of birth	avg. diff	avg abs diff	stand dev.	max diff	n	year of birth	avg. diff	avg abs diff	stand dev.	max diff	n
85	19.3	29	29.9	179	371	81	-12.8	54.6	74.5	202	73
86	22.1	30	30.2	153	333	82	-12.0	48.9	70.0	249	82
87	24.6	33	34.1	183	323	83	-20.9	55.7	68.0	248	61
88	23.1	32	31.0	114	384	84	-16.4	58.2	76.6	244	45
89	19.8	31	34.4	157	412	85	-18.8	56.0	72.1	221	44
90	18.2	27	27.9	88	290	86	-24.3	55.6	78.4	283	53

*(The difference (diff) is calculated as breeding value with heterogeneous variance correction minus without this correction).

Table 5. Comparison of breeding values for protein yield for progeny tested bulls and imported bulls.

Progeny tested bulls						Imported bulls					
year of birth	avg. diff	avg abs diff	stand dev.	max diff	n	year of birth	avg. diff	avg abs diff	stand dev.	max diff	n
85	0.6	1.0	1.0	3	371	81	0.2	1.7	2.0	5	73
86	0.8	1.0	1.1	5	333	82	0.0	1.4	2.0	6	82
87	0.7	1.0	1.1	5	323	83	0.0	1.6	2.2	6	61
88	0.7	1.0	1.1	4	384	84	0.0	1.4	1.8	4	45
89	0.6	1.0	1.2	5	412	85	0.1	1.6	2.1	5	44
90	0.6	0.9	1.1	4	290	86	-0.1	1.6	2.3	9	53

*(The difference (diff) is calculated as breeding value with heterogeneous variance correction minus without this correction)

Literature

Meuwissen, T.H.E., G. de Jong en B. Engel, 1994. Joint estimation of breeding values and heterogeneous variances by mixed model equations in large national data sets. Proceedings of the open session of Interbull meeting Ottawa Canada, 1994, Interbull Bulletin no 10.

Meuwissen, T.H.E., G. de Jong en B. Engel, 1995. Joint estimation of breeding values and heterogeneous variances in large scale dairy data. Submitted.

95-1286/GdJ/CS