Stochastic simulation of the influence of insemination on the estimation of breeding value and its reliability

Z. Veselá[1], J. Přibyl[1], L. Vostrý[1], L. Štolc[2]

[1]Institute of Animal Science, Praha-Uhříněves, Czech Republic
[2]Czech University of Life Sciences Prague, Prague-Suchdol, Czech Republic

ABSTRACT: Insemination is used less commonly in the population of beef cattle than in dairy cattle. As a consequence, herds are genetically isolated and their comparison is difficult. The objective of the paper was to evaluate the impact of the use of reference sires in insemination on the estimation of breeding value and its reliability by means of stochastic simulation based on a single addition of calves to the basic set. Simulation was divided into three variants according to the connectedness calves through sires (variant 1 – unrelated calves, variant 2 calves related between HYS but unrelated within HYS, variant 3 – calves related within and between HYS). The degree of connectedness was determined by computation of genetic drift variance (GDV). The average value of GDV in the basic set was 0.71. In our study the highest values of GDV were reached in simulated variant 3, in which the relationship increased not only between HYS but also within HYS. The lowest values of GDV were in variant 2, when the relationship increased only between HYS (simulated calves after reference sires did not have any half-sibs in HYS). The single use of reference sires in one generation to make herds connected slightly increased the average coefficient of reliability from 0.121 to 0.145 (by 2.4%) in the whole set. But this increase from 0.24 to 0.33 (by 8.9%) was observed only in animals that were directly connected through reference sires. The increase in reliability was mainly due to an increase in the number of contemporaries. A decrease in the effective number of sires' offspring was an indirect consequence of the increase in the connectedness of the set and higher use of reference sires in insemination. This is the reason why the increase in the connectedness of the set influenced the reliability of breeding value of the whole set less than we would expect. The use of reference sires influenced the fluctuation of breeding values. The use of a higher number of reference sires in insemination with only one offspring in each of the connected herds appeared as the most suitable approach from the aspect of the connectedness of the set and reliability of breeding value estimation. A marked influence of herd connectedness on the estimation of breeding value and its reliability, and on the result of breeding, is to be expected after reference sires have been used in several successive generations.

Keywords: beef cattle; breeding value; connectedness; reference sires; reliability; stochastic simulation

Great progress has been made in the breeding of beef cattle in the Czech Republic in the last years. Since 2000 breeding value in beef cattle in the Czech Republic has been estimated for the results of a field test (calving ease, birth weight and weight at 120, 210 and 365 days of age) by a multi-trait animal model (Přibyl et al., 2003). In 2004 the estimation of breeding value for the own growth of bulls at performance-testing stations by a single-trait animal model was introduced (Přibylová et al., 2004), and in 2005 for a description of the type traits of young animals of beef cattle by a multi-trait animal model (Veselá et al., 2005). In 2006 selection indexes for beef cattle in the Czech Republic were constructed on the basis of these evaluated traits (Šafus et al., 2006).

The estimation of breeding value for the growth of beef cattle is frequently done by a multi-trait

animal model with maternal effect. Fixed effects included in the model are herd-year-season, sex, age of dam and regressions on calf and maternal heterosis. Random effects are direct and maternal genetic and maternal environmental effects (Přibyl et al., 2003). The importance of maternal effects in beef cattle has been widely reported (Robison, 1981; Meyer, 1992; Meyer, 2001). Multi-trait models are applied, increasing the reliability of estimation of breeding values for the particular traits and coping with problems due to missing data (Přibyl et al., 2003).

In the population of beef cattle artificial insemination is used less commonly than in dairy cattle. Therefore the particular herds are genetically isolated and their mutual comparison is difficult (Kennedy and Trus, 1993). A part of the optimum organisation of breeding is such a structure of the mating of sires that creates conditions for reliable estimation of breeding value and subsequent high genetic gain. One of the possibilities is the insemination of reference sires making up a genetic link among the herds. In this context it is possible to study changes in the estimation of breeding value and coefficient of reliability of breeding value estimation and the degree of connectedness of the particular herds.

Minimisation of prediction error variance (PEV) and maximisation of the coefficient of reliability (r^2) of breeding values is often thought to be the goal of genetic evaluation. But deviations of estimated breeding values also participate in the total mean square error (MSE): MSE = (Bias)^2 + PEV. For genetic evaluation of animal i:

$$\text{MSE}_i = (\text{E}(\hat{u}_i - u_i))^2 + \text{V}(\hat{u}_i - u_i)$$

where:
$(\text{E}(\hat{u}_i - u_i)$ = bias of estimated and true breeding value of animal i;
$\text{V}(\hat{u}_i - u_i)$ = prediction error variance of animal i.

In practice, true parameter values for the true model are required to evaluate bias. Limits for bias, however, might be approximated for some models for a comparison with PEV (Van Vleck, 1987).

Van Vleck (1987) reported the influence of the effective number of daughters on the size of PEV. The effective number of daughters is often a term applied to the diagonal elements of the least-squares matrix corresponding for example to a sire after the absorption of HYS equations. In fact, the inverse of the diagonal elements is often used to approximate the prediction error variance. If the coeffi-

cient matrix after absorption is inverted, the diagonal term corresponds to PEV. Because off-diagonal elements also affect the inverse, and are also generated by absorption, the inverse of the effective number of daughters can be considered only as an approximation of PEV. As an example of the effect of the number of contemporaries on the effective number of daughters, assume HYS are treated as fixed effect. Assume a bull has n daughters in HYS with m daughters of other bulls. The daughters of the bull in HYS contribute n^* to the diagonal of the bull equation after the absorption of HYS equation with:

$$n^* = n - \frac{n^2}{n + m} = \frac{m}{n + m}$$

Because PEV is related to the accuracy of evaluation and therefore to the genetic gain from selection, i.e. $PEV = (1 - r^2)\, \sigma_G^2$ (σ_G – standard deviation for genetic effect), the importance of the effective number of daughters by each HYS comparison depends on whether

$$\sum_i n_i^*$$

is large or small (Van Vleck, 1987).

To evaluate the stability of breeding values at repeated evaluations in time it is possible to use relations between the breeding values of related individuals and Mendelian sampling. The evaluated databases usually comprise several generations of individuals, therefore the breeding values of parents and offspring are known. The breeding value of an offspring should approximate the mid-parent breeding value as much as possible. Differences in the expected and estimated breeding value are caused by Mendelian sampling with variance ($0.5\,\sigma_G^2$) and error of breeding value evaluation. Mendelian sampling (MS) may be determined for each individual on the basis of direct parents:

$$MS_P = BV_P - \left(\frac{BV_S - BV_D}{2} \right)$$

where:
MS_P = Mendelian sampling of progeny
BV_P = breeding value of progeny
BV_S = breeding value of sire
BV_D = breeding value of dam

Mendelian sampling may be used to evaluate the suitability of a model. It is also applicable to evaluation by means of a correlation between the breeding value of an individual and the mid-parent breeding value (Miglior and Van Doormaal, 2000).

Original Paper

Czech J. Anim. Sci., 52, 2007 (8): 236–248

In connection with the use of reference sires some papers dealt with the connectedness of a set. Searle (1987) presented a simple geometric method to determine whether data were connected or not. This method is based on the multiplication of matrices *X'Z*, where *X* and *Z* are incidence matrices for management units (herd) and sire. The results of this method document only whether the set is connected but they do not provide any information on the degree of connectedness.

The degree of connectedness is determined by means of genetic drift variance (Kennedy and Trus, 1993). The method is based on the multiplication of matrices *X'ZAZ'X*, where *X* is the incidence matrix of management units (herd-year-season), *Z* is the incidence matrix for animals and *A* is the relationship matrix. The matrix *X'ZAZ'X* measures the sum of genetic relationships within and between management units. Average relationships between and within management units (including the relationship of animal to itself) Ā are obtained by dividing diagonal elements of *X'ZAZ'X* by the square of the number of records in the unit and off-diagonal elements by the product of the number of records in each of the units. The elements of Ā can be interpreted as the genetic components of drift variance and covariance between the management units (Sorensen and Kennedy, 1983). A simple measure resulting from Ā is to compute the genetic drift variance between management units that is given by:

$$\mathrm{GDV}_{ij} = \bar{A}_{ii} + \bar{A}_{jj} - 2 \cdot \bar{A}_{ij},$$

where:

$\bar{A}_{ii}$ (and/or $\bar{A}_{jj}$) = the diagonal element of matrix Ā

$\bar{A}_{ij}$ = the off-diagonal element of matrix Ā.

Hanocq et al. (1996) investigated the simulation of changes in genetic gain and reliability of breeding value estimation in relation to the connectedness of subpopulations through sires. If the populations were not connected, it was not possible to determine genetic differences between them. Clement et al. (2001) stated that the missing connectedness between herds distorted the estimation of breeding value because the effect of genotype and the effect of environment could not be distinguished. Simm et al. (2001) accentuated that the use of reference sires allowed across-herd genetic evaluation, higher intensity of selection and higher rate of genetic gain. Lewis et al. (1999) reported that the stability of mutual connectedness of herds influenced the reliability of selection decisions. Lewis and Simm

(2000) also stated that the connectedness of herds increased the reliability of breeding value estimation and provided a higher number of candidates for selection. Johnson et al. (2002) optimised by stochastic simulation the sire referencing scheme in Texel breed in New Zealand. They drew a conclusion that the highest selection gain could be achieved by the use of reference sires for insemination of the best breeding ewes in the flock. The use of reference sires increased the reliability of breeding value estimation. For economic reasons artificial insemination was the most suitable method for the use of reference sires. Kennedy and Trus (1993) reported that in the estimation of breeding value by animal model the connectedness between herds is reflected in the relationship matrix (A). The impact of data structure on the reliability of breeding value estimation was examined by Tosh and Wilton (1994); Přibyl and Přibylová (2001). The influence of sire referencing scheme on the reliability of breeding value estimation in beef cattle was investigated in various categories of animals in herds by Veselá et al. (2004).

The objective of the present paper is to evaluate the influence of reference sires on the estimation of breeding value and its reliability.

MATERIAL AND METHOD

The basic set for simulation was selected from the database of performance testing of beef breeds of cattle in the Czech Republic. Charolais was a used breed, and live weight at 120 days of age was an evaluated trait. The set comprised individuals with performance records from the years 1999–2003. Nine herds (76 HYS) of medium size were selected from the national database for subsequent simulations. There were 666 individuals with own performance, 380 dams, 52 sires, and the total number of all individuals including three generations of ancestors in their pedigree was 1 793.

Table 1 shows statistical characteristics of the basic set selected for simulations. The measured performance was from 71 kg to 266 kg, the average value being 173.40 kg.

Table 1. Characteristics of the basic set (recorded trait – live weight in kg at 120 days of age)

n	Mean	s	Min.	Max.
666	173.40	31.51	71	266

The first step of our study was to evaluate the degree of connectedness of the set.

For graphical representation we used the method of the multiplication of matrices $X′Z$, where X and Z are incidence matrices for management units (herd) and sire, proposed by Searle (1987).

The degree of connectedness was determined by means of genetic drift variance (GDV) that was calculated as the mean of GDV between the particular HYS (Kennedy and Trus, 1993).

Breeding value was calculated by a single-trait animal model according to the model equation that is used for routine estimation of breeding values from a field test of beef breeds in the Czech Republic (Přibyl et al., 2003):

$$y_{ijkl} = \mu + S_i + Po_j + V_k + gp_{ijkl} + gm_{ijkl} + pe_{ijkl} + e_{ijkl} \quad (1)$$

where:

y_{ijkl} = measured performance (weight at 120 days of age)
S_i = group of contemporaries (herd – year – season)
Po_j = calf sex (young bulls, heifers/singles, twins)
V_k = dam's age at calving (less than three years, four, five to seven years old, more than seven years)
gp_{ijkl} = an individual, breeding value for direct effect
gm_{ijkl} = an individual, dam's breeding value for maternal effect; direct and maternal effect are correlated with each other with the joint relationship matrix
pe_{ijkl} = permanent maternal environment
e_{ijkl} = random deviation

Effects of an individual (direct and maternal ones) and permanent maternal environment were used as random effects.

Population-genetic parameters that were substituted in the computation of breeding values are shown in Table 2. The values were taken over from literature (Přibyl et al., 2003).

Table 2. Population-genetic parameters

Standard deviations substituted to the computation (in kg)	
σ_{GD}	12.2
σ_{GM}	9.75
σ_{PE}	6.24
σ_e	14.13
Correlation between direct and maternal effect	
r_G	–0.18

σ_{GD} = standard deviation for direct genetic effect; σ_{GM} = standard deviation for maternal genetic effect; σ_{PE} = standard deviation for permanent maternal environment; σ_e = standard deviation for random uncontrolled environment; r_G = genetic correlation between direct and maternal effect

Breeding value was computed by a system of matrices in the SAS Software.

Reliability of the estimation of breeding value was computed from diagonal elements of the inverse of the left side of the matrix system (Henderson, 1984; Mrode, 1996):

$$r_i^2 = 1 - C_{ii} \frac{\sigma_e^2}{\sigma_G^2} \quad (2)$$

where:

r_i^2 = reliability of breeding value estimation for direct or maternal effect
C_{ii} = diagonal element of the left side of the matrix system after the inverse for an individual i ($C_{ii} = PEV/\sigma_e^2$)
σ_e = standard deviation for random uncontrolled environment
σ_G = standard deviation for direct (or maternal) genetic effect

In the next step we simulated a single use of insemination simulation in a 5-year interval within one generation and we generated the performance of calves added to the basic set. Every year 1 to 5 calves were added to the particular herds.

The generating of calves was divided into 3 variants. In variant 1 calves whose sires were unidentified were added. So the number of animals in the set increased but the added calves were not connected through the sire. In variant 2 the added calves were connected through 1–5 sires in each year of the birth of calves in the basic set (only one calf after the sire into each herd). In variant 3 all added calves were connected only through one sire each year (independently of the number of added calves). Each year different sires were generated, which corresponds to routine methods of breeding. The total number of generated calves was 45 after the addition of 1 calf, and 225 after the addition of 5 calves. The generating of calves avoids inbreeding.

The performance of added calves was calculated according to equation 1, where the values calculated from the basic set were substituted as fixed effects. The level of the effect of sex and dam's age was selected randomly on the basis of the representation of the particular levels of these effects in the basic set.

Random effects were generated by means of random numbers using the Monte Carlo method (Brascamp, 1978; Gibson, 1995) according to these equations:

$$g_{p_i} = 0.5 \times a_{i1} \times \sigma_{GD} + 0.5 \times a_{i2} \times \sigma_{GD} + a_{i3} \times \sqrt{0.5} \times \sigma_{GD} \quad (3)$$

$$g_{m_i} = a_{i2} \times r_G \times \sigma_{GM} + a_{i4} \times \sqrt{1 - r_G^2} \times \sigma_{GM} \qquad (4)$$

$$pe_i = a_{i5} \times \sigma_{PE} \qquad (5)$$

$$e_i = a_{i6} \times \sigma_e \qquad (6)$$

where:

g_{p_i} = breeding value of the *i*-th individual for direct effect (the sum of breeding values for direct effect of sire and dam and Mendelian sampling)

g_{m_i} = breeding value of the *i*-th individual's dam for maternal effect

pe_i = permanent environment of the *i*-th individual's dam

e_i = random uncontrolled environment of the *i*-th individual

σ_{GD} = standard deviation for direct genetic effect

σ_{GM} = standard deviation for maternal genetic effect

σ_{PE} = standard deviation for permanent maternal environment

σ_e = standard deviation for random uncontrolled environment

r_G = genetic correlation between direct and maternal effect

$a_{i1}, a_{i2}, a_{i3}, a_{i4}, a_{i5}, a_{i6}$ = random numbers from normal distribution N(0, 1) for the *i*-th individual

Each variant was generated ten times for fixed effects, and within these ten generations again ten times for random effects. In total, 100 repeated generations for each variant were performed. The generation of random numbers was done in the SAS programme.

Breeding value and its reliability were computed for each generation and each variant.

We also examined deviations of breeding values from the variant with the highest reliability of breeding value estimation that were computed from the values standardised per mean 0 and standard deviation 1. Mean square error was computed

on the basis of estimated deviations and PEV (and/ or reliability of breeding value estimation) (Van Vleck, 1987).

$$MSE = (Bias)^2 + PEV \qquad (7)$$

Correlations of breeding values of an offspring with mid-parent breeding values were computed for each variant.

RESULTS AND DISCUSSION

Connectedness of the set

Graphical representation of herd connectedness in the basic set through the particular sires is shown in Table 3. Numbers in the table represent the number of the offspring of a concrete sire in a concrete herd. This simple graphical method to represent the connectedness of the set was proposed by Searle (1987). This method provides information whether the herds are or are not connected through sires, but it cannot be used for a quantitative expression of the degree of connectedness. The table shows that the connectedness of herds is very low. Sires S1 to S48 contributed to the connectedness of herds H1 to H7. Herds H8 and H9 remain isolated from the aspect of genetic connectedness. A similar situation may be observed in the whole population. The aim of the use of reference sires was to improve herd connectedness.

Figure 1 shows average genetic drift variance (GDV = 0.71) in the original set and in simulated sets according to the particular variants. Lower GDV values indicate higher connectedness of the set (Kennedy and Trus, 1993). At the same time, the higher the number of simulated added calves, the
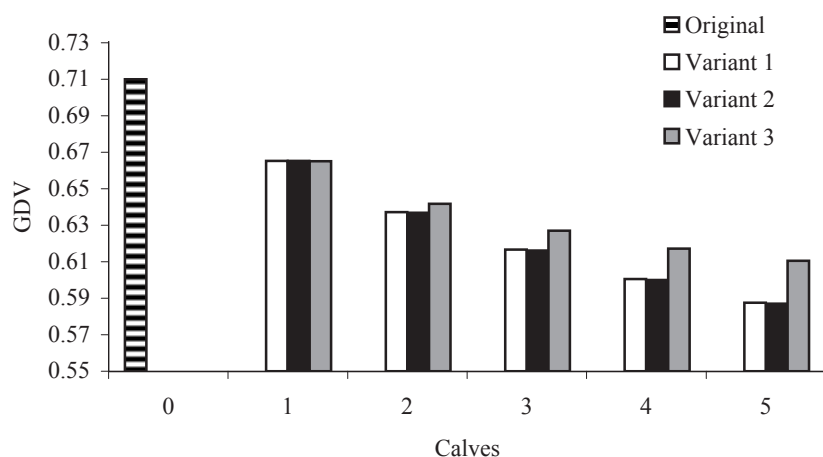

Figure 1. The degree of connectedness of the set expressed by means of genetic drift variance in relation to the number of added calves

Table 3. Graphical representation of herd (H1–H9) connectedness across sires (S1–S52)

	H1	H2	H3	H4	H5	H6	H7	H8	H9		H1	H2	H3	H4	H5	H6	H7	H8	H9
S1	1	0	0	0	0	0	0	0	0	S29	0	0	5	0	0	0	5	0	0
S2	3	0	0	0	0	0	0	0	0	S30	0	0	1	0	2	0	0	0	0
S3	2	0	0	0	0	0	0	0	0	S31	0	0	4	0	2	0	0	0	0
S4	33	0	0	0	0	0	0	0	0	S32	0	0	1	1	5	5	0	0	0
S5	1	0	0	0	0	0	0	0	0	S33	0	0	0	69	0	0	0	0	0
S6	1	0	0	0	0	0	0	0	0	S34	0	0	0	0	18	0	0	0	0
S7	1	0	0	0	0	0	0	0	0	S35	0	0	0	0	16	0	0	0	0
S8	16	0	0	0	0	0	0	0	0	S36	0	0	0	0	0	1	0	0	0
S9	2	0	0	0	0	0	0	0	0	S37	0	0	0	0	0	46	0	0	0
S10	2	2	1	0	0	0	3	0	0	S38	0	0	0	0	0	1	0	0	0
S11	0	68	2	0	0	0	0	0	0	S39	0	0	0	0	0	2	0	0	0
S12	0	2	0	0	0	0	0	0	0	S40	0	0	0	0	0	0	1	0	0
S13	0	6	0	0	0	0	0	0	0	S41	0	0	0	0	0	0	1	0	0
S14	0	2	0	0	0	0	0	0	0	S42	0	0	0	0	0	0	1	0	0
S15	5	1	2	0	0	0	0	0	0	S43	0	0	0	0	0	0	8	0	0
S16	1	0	1	0	21	12	0	0	0	S44	0	0	0	0	0	0	1	0	0
S17	4	0	5	0	3	0	0	0	0	S45	0	0	0	0	0	0	1	0	0
S18	1	0	2	0	0	0	0	0	0	S46	0	0	0	0	0	0	2	0	0
S19	0	0	1	0	0	0	0	0	0	S47	0	0	0	0	0	0	12	0	0
S20	0	0	17	0	0	0	0	0	0	S48	0	0	0	0	0	0	32	0	0
S21	0	0	13	0	0	0	0	0	0	S49	0	0	0	0	0	0	0	33	0
S22	0	0	12	0	0	0	0	0	0	S50	0	0	0	0	0	0	0	34	0
S23	0	0	4	0	0	0	0	0	0	S51	0	0	0	0	0	0	0	11	0
S24	0	0	1	0	0	15	0	0	0	S52	0	0	0	0	0	0	0	0	65
S25	0	0	4	0	0	0	0	0	0										
S26	0	0	2	0	0	0	0	0	0										
S27	0	0	2	0	0	0	0	0	0										
S28	0	0	3	0	0	0	0	0	0										

Graphical representation

H1–H9 = herd 1–9; S1–S52 = sire 1–52

Number of the offspring of a concrete sire in a concrete herd

lower the GDV values, which increased the effective number of sires' offspring in the set. Kennedy and Trus (1993) reported that an increase in the relationship within HYS would cause an increase in PEV (lower r^2) while an increase in the relationship between HYS would decrease PEV (higher r^2). So GDV is influenced by an increase in the relationship not only between HYS (decrease in GDV) but also within HYS (increase in GDV). In our case, the highest values of GDV were reached in simulated variant 3, in which the relationship increased not only between HYS but also within HYS. The lowest values of GDV were in variant 2, where the relationship increased only between HYS (simulated calves after a reference sire did not have any half-sibs in HYS). From the aspect of the degree of connected-

Table 4. Reliability of breeding value estimation for direct effect in the basic set

	n	Mean	*s*	Min.	Max.
Whole set	1 793	0.1107	0.1057	0.0000	0.5560
Calves with own performance	666	0.2244	0.0547	0.0494	0.3458
Sires of calves with own performance	52	0.2382	0.1328	0.0348	0.5560
Dams of calves with own performance	380	0.0836	0.0606	0.0004	0.3408

Table 5. Reliability of breeding value estimation for maternal effect in the basic set

	n	Mean	*s*	Min.	Max.
Whole set	1 793	0.0938	0.0828	0.0000	0.3899
Calves with own performance	666	0.0827	0.0416	0.0115	0.2897
Sires of calves with own performance	52	0.1187	0.1110	0.0028	0.3899
Dams of calves with own performance	380	0.2216	0.0599	0.0033	0.3669

ness the use of a higher number of reference sires with a lower number of offspring in herds seems the most suitable.

Reliability of breeding value estimation

The reliability of breeding value estimation for direct effect in the basic set (Table 4) ranged from $r^2 = 0$ to $r^2 = 0.556$ with the mean $r^2 = 0.1107$. The reliability of breeding value estimation for maternal effect in the basic set (Table 5) was in the range from $r^2 = 0$ to $r^2 = 0.3899$ with the mean $r^2 = 0.0938$.

A single-trait animal model was used because of computation simplicity. In the routine estimation of breeding value a multi-trait model is applied and the reliability of breeding values is therefore higher (Přibyl et al., 2003).

Table 4 shows the mean values of reliability of breeding value estimation for direct effect computed in the basic set. The whole set was divided into three parts (calves with own performance, sires of calves with own performance, dams of calves with own performance). The highest average reliability was found out in sires of calves with own performance ($r^2 = 0.2382$), somewhat lower was the reliability in calves with own performance

($r^2 = 0.2244$) and dams of calves with own performance had the lowest reliability ($r^2 = 0.0836$). The highest amount of information for the estimation of breeding value was available for sires of calves with own performance and calves with own performance. Only a limited amount of information on the performance of their offspring was available for dams of calves with own performance, therefore their reliability of breeding value estimation was the lowest.

Presented coefficients of reliability of breeding values were calculated from the performance of one generation. For this reason parents are evaluated mainly according to their progeny. More generations are evaluated in routine and parents have their breeding values estimated also according to own performance. Resulting reliability of parents' breeding values is the combination of own and progeny performance. Therefore the reliability of breeding values of parents will be higher than that of calves with own performance in practice.

Table 5 shows the mean values of the reliability of breeding value estimation for maternal effect. The reliability of breeding value estimation is markedly higher for dams of calves with own performance ($r^2 = 0.2216$) where information on the performance of their offspring is available.

Table 6 presents estimated coefficients of the reliability of breeding values for direct and maternal
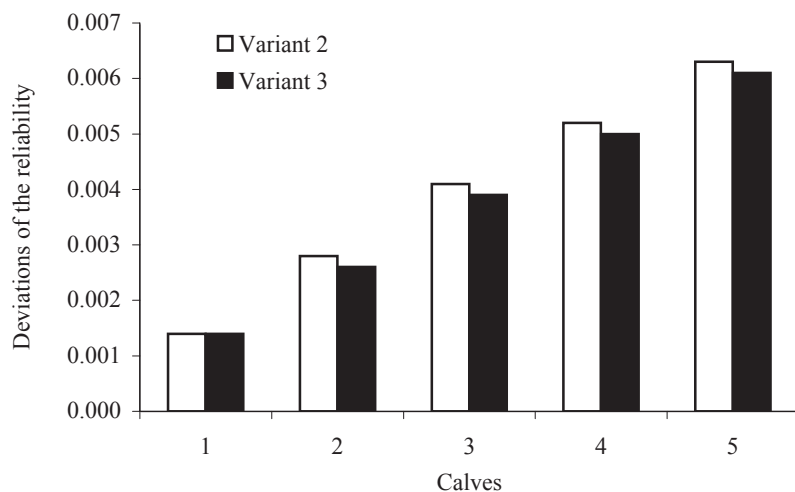

Figure 2. Deviations of the reliability of breeding value estimation for direct effect from variant 1 (whole set) in relation to the number of added calves
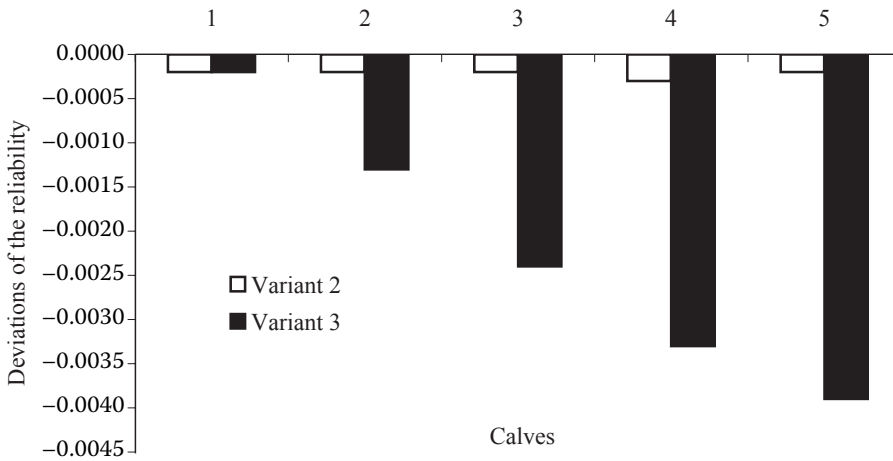
Figure 3. Deviations of the reliability of breeding value estimation for direct effect from variant 1 (basic set – calves with own performance) in relation to the number of added calves

effect for different variants of simulation in the whole set and in selected subsets. The size of the set increased with the number of simulated and added calves, and it also led to an increase in the coefficient of reliability. This is the reason why the sets with the identical number of generated calves are always compared with each other in the following analyses.

Figure 2 illustrates the deviation of the reliability of breeding value estimation in variant 2 and 3 from simulated unconnected variant 1 for direct effect in the whole set (basic set + generated animals). Obviously, the increase in reliability was almost identical in variant 2 and 3 compared to variant 1, where the simulated calves were not connected through the sire. Figure 2 also documents the slightly lower reliability of breeding value in variant 3 than in variant 2. It is explained by the lower effective number of sires' offspring (Van Vleck, 1987) than in variant 3. These results

confirm the results obtained by GDV computation (Figure 1).

Results will be different if we evaluate only calves with own performance from the basic set (Figure 3). The reliability of breeding value in connected variants 2 and 3 was slightly lower compared to unconnected variant 1, which can be explained again by a decrease in the effective number of offspring. Variations in the reliability of breeding value estimation for maternal effect documented the same tendency.

The absolute value of the reliability of breeding value estimation increased with the increasing number of individuals included in the evaluation. Figure 3 shows that the reliability of breeding values for individuals that were not directly connected through the sire (basic set) was more markedly influenced by the effective number of sires' offspring than by the own higher connectedness of herds. The addition of unrelated individuals into the basic
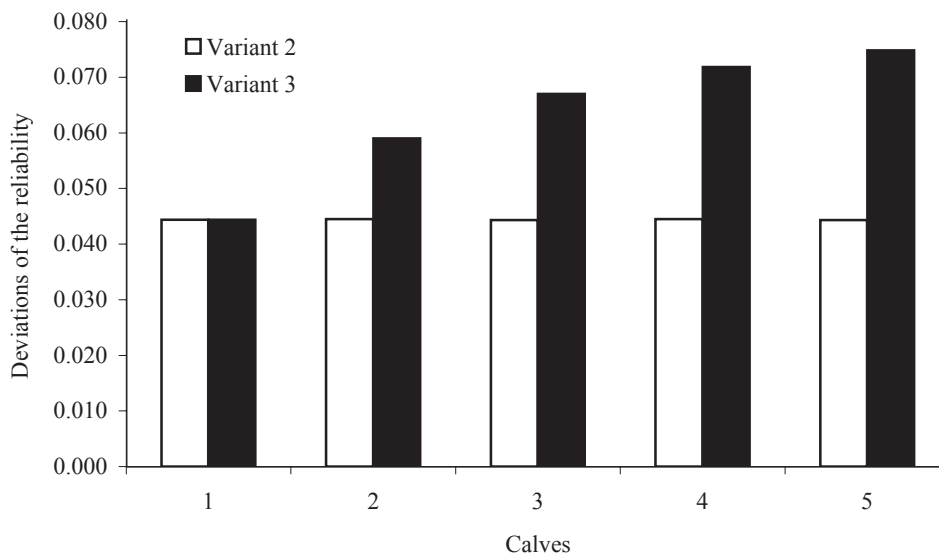

Figure 4. Deviations of the reliability of breeding value estimation for direct effect from variant 1 (added calves) in relation to the number of added calves

243

Table 6. Variations in the coefficient of breeding value reliability in different variants of simulation in the whole set and in selected subsets

Number of added calves		Reliability of breeding values for direct effect					Reliability of breeding values for maternal effect				
		1	2	3	4	5	1	2	3	4	5
Whole set	variant 1	0.1212	0.1278	0.1324	0.1360	0.1387	0.0950	0.0956	0.0960	0.0962	0.0962
	variant 2	0.1226	0.1306	0.1365	0.1412	0.1450	0.0950	0.0956	0.0959	0.0960	0.0961
	variant 3	0.1226	0.1304	0.1363	0.1410	0.1448	0.0950	0.0959	0.0967	0.0973	0.0978
Calves with own performance	variant 1	0.2456	0.2590	0.2686	0.2760	0.2816	0.0850	0.0866	0.0877	0.0886	0.0893
	variant 2	0.2454	0.2588	0.2684	0.2757	0.2814	0.0850	0.0866	0.0877	0.0886	0.0894
	variant 3	0.2454	0.2577	0.2662	0.2727	0.2777	0.0850	0.0865	0.0875	0.0884	0.0890
Sires of calves with own performance	variant 1	0.2622	0.2779	0.2894	0.2984	0.3055	0.1223	0.1247	0.1265	0.1279	0.1291
	variant 2	0.2619	0.2776	0.2891	0.2981	0.3052	0.1223	0.1248	0.1266	0.1280	0.1292
	variant 3	0.2619	0.2763	0.2865	0.2944	0.3006	0.1223	0.1246	0.1262	0.1276	0.1286
Dams of calves with own performance	variant 1	0.0891	0.0925	0.0949	0.0967	0.0981	0.2254	0.2279	0.2297	0.2311	0.2322
	variant 2	0.0890	0.0924	0.0947	0.0965	0.0980	0.2254	0.2280	0.2298	0.2312	0.2323
	variant 3	0.0890	0.0920	0.0940	0.0956	0.0968	0.2254	0.2278	0.2295	0.2308	0.2318
Added calves	variant 1	0.2407	0.2459	0.2499	0.2527	0.2550	0.0135	0.0138	0.0140	0.0141	0.0143
	variant 2	0.2851	0.2904	0.2942	0.2972	0.2993	0.0184	0.0188	0.0192	0.0194	0.0196
	variant 3	0.2851	0.3049	0.3169	0.3245	0.3298	0.0184	0.0218	0.0237	0.0249	0.0257
Dams of added calves	variant 1	0.0354	0.0362	0.0368	0.0372	0.0375	0.1406	0.1436	0.1459	0.1476	0.1489
	variant 2	0.0363	0.0372	0.0378	0.0383	0.0386	0.1441	0.1474	0.1498	0.1517	0.1530
	variant 3	0.0363	0.0380	0.0390	0.0397	0.0402	0.1441	0.1505	0.1546	0.1573	0.1594
Sires of added calves	variant 2	0.3516	0.3567	0.3605	0.3634	0.3655	0.0114	0.0116	0.0117	0.0118	0.0118
	variant 3	0.3516	0.4865	0.5606	0.6068	0.6385	0.0114	0.0158	0.0182	0.0197	0.0207

Czech J. Anim. Sci., 52, 2007 (8): 236–248

Original Paper

Table 7. Breeding value for direct effect in the basic set

	n	Mean	*s*	Min.	Max.
Calves with own performance	666	−0.20	7.28	−28.23	19.73
Sires of calves with own performance	52	0.49	6.72	−21.74	16.94
Dams of calves with own performance	380	−0.32	4.34	−18.72	10.26

Table 8. Breeding value for maternal effect in the basic set

	n	Mean	*s*	Min.	Max.
Calves with own performance	666	0.03	3.32	−11.48	9.59
Sires of calves with own performance	52	0.25	4.30	−11.86	13.40
Dams of calves with own performance	380	0.36	5.61	−18.26	16.33

set (variant 1) leads to a more marked increase in the effective number of offspring than the addition of individuals into the basic set that are mutually related through the sire.

On the contrary, in added individuals the resultant reliability of breeding value estimation was necessarily influenced substantially by an increase in the connectedness of herds through the sire and by an increase in the number of half-sibs (Figure 4). The highest reliability of breeding value estimation for direct effect was determined in calves with four half-sibs through the sire in the herd (variant 3 with five added calves). Their value of reliability coefficient was 0.33, i.e. much higher than in calves of the basic set. Similarly, the sires of added calves have the markedly higher reliability of breeding value estimation for direct effect than the sires of calves of the basic set.

Deviation in breeding values estimation

Another consequence of the use of reference sires we observed was the fluctuation of estimated breeding values. As the simulation was done on data from performance testing, we do not know the real breeding values of animals. The fluctuation of breeding values may be assessed only in relative terms on the basis of estimated values.

Tables 7 and 8 show the mean values of breeding value estimation for direct or maternal effect computed in the basic set for the particular subsets.

The correlations of breeding values of the basic set estimated for the particular variants were high (0.99 and more).

We examined deviations of breeding values from the variant with highest reliability of breeding value estimation – variant 2 with 5 added calves. Breeding
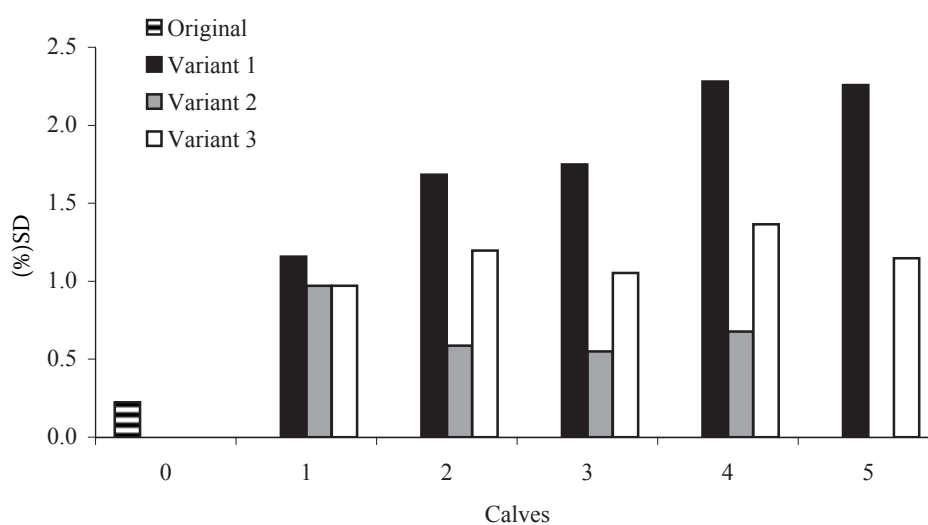

Figure 5. Deviations of standardised breeding values for direct effect from the variant with the highest reliability of breeding value estimation (variant 2 with 5 added calves) in the set of calves with own performance related to the standard deviation of breeding value for direct effect
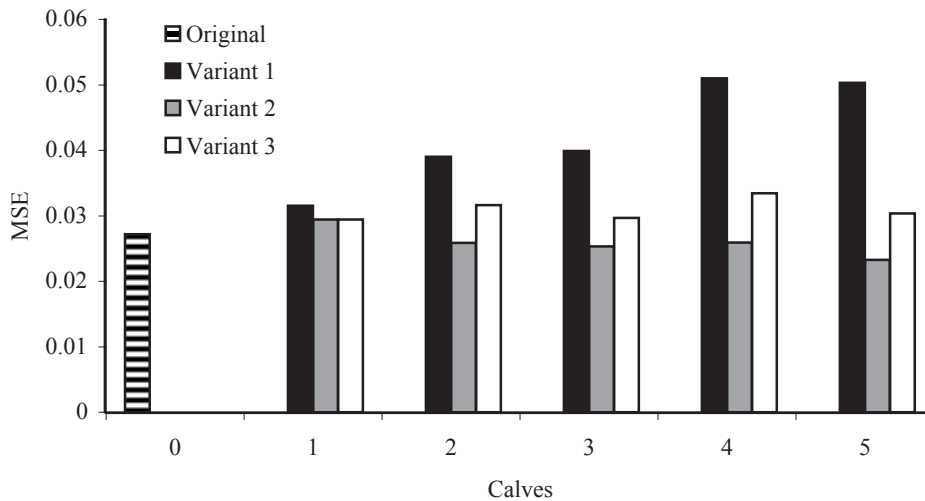
Figure 6. Estimated MSE for the particular variants (inclusion of deviation from variant 2 with 5 added calves)

values were standardized per mean 0 and standard deviation 1. Figure 5 illustrates the mean absolute deviations of standardised breeding values for direct effect of the particular variants from variant 2 with five added calves in the set of calves with own performance. Each value was calculated as the mean of 100 repeated generations, which should eliminate random fluctuations to a large extent. The largest deviations were in variant 1. The deviations ranged from 1.15 to 2.28%. Deviations increased with the number of added calves. Deviations of variant 3 were lower (0.97–1.36%). The deviations of breeding value estimation for maternal effect had a similar pattern like the deviations of breeding value estimation for direct effect. The addition of unrelated calves (variant 1) relatively decreased the connectedness of the whole set, which led to a distortion of the estimation of breeding values, and

consequently, to a larger fluctuation of estimated breeding values.

In Figure 6 the deviations of breeding values and reliability of breeding value estimation in the set of calves with own performance are summarised by means of MSE estimation (Van Vleck, 1987). As in our case the data are derived from a real set, we do not know the real breeding values, and so the deviations of breeding values and also MSE are related to the variant with the highest reliability of breeding value (variant 2 with five added calves) only in relative terms. The deviation of breeding values and MSE may be caused by many influences. In our case, when we use the same model equation for all variants and random variations in breeding values of generated added calves are minimised by a multiple replication, MSE may be induced mainly by a change in the structure of the set. The highest MSE was
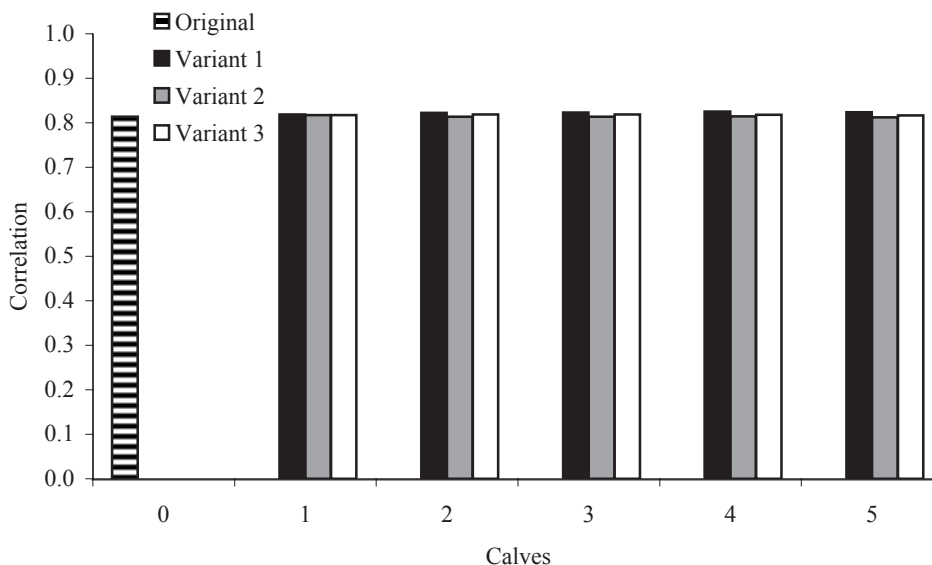

Figure 7. Correlations of the mid-parent breeding values and breeding values of their offspring
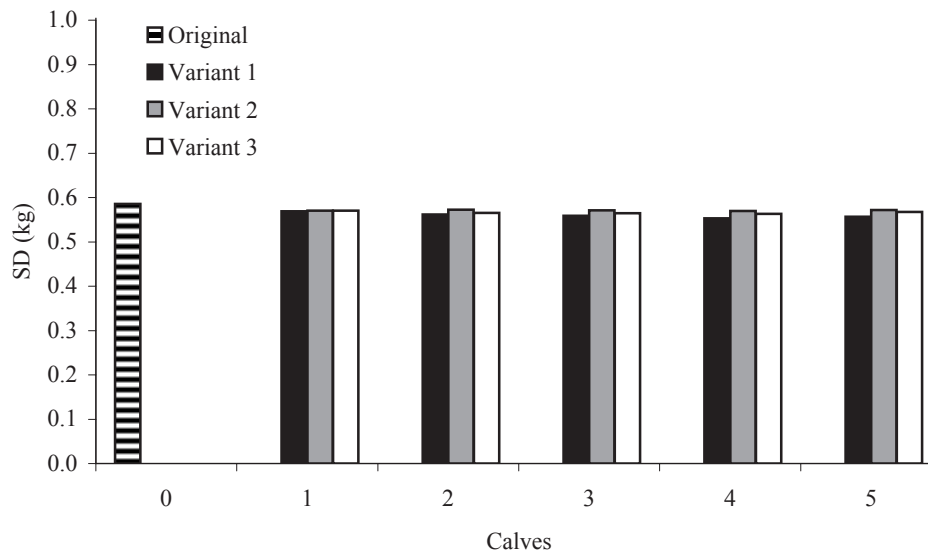
Figure 8. Standard deviations of Mendelian sampling

determined in variant 1, when only the number of individuals increased, without any connectedness.

Figure 7 illustrates the correlations of mid-parent breeding values with the breeding value of offspring in the set of calves with own performance as an alternative expression of standard deviations of Mendelian sampling that are represented in Figure 8 (Miglior and Van Doormaal, 2000). The breeding value of an offspring should approximate the mid-parent breeding value as much as possible. Differences between the expected and estimated breeding value are caused by Mendelian sampling with variance ($0.5 \times \sigma^2_G$) and by the error of breeding value estimation. The graphs do not document any significant differences between the variants. The correlations ranged from the value (0.81) for the original set to the value (0.82) for variant 1 with the addition of five calves. These values indicate that the model used in practice for the estimation of breeding value of a field test makes it possible to predict the breeding value of an offspring from the parents' breeding value with reliability approximately 0.66.

In our study we simulated a single use of reference sires within a five-year period of one generation. So reference sires connected the particular herds in one generation only, not in the succession of several generations. On the contrary, the majority of the papers dealing with this subject simulated insemination by reference sires in the succession of several generations. Hanocq et al. (1996) simulated insemination across 7 generations. The reliability of breeding value estimation increased as a result of high connectedness through reference sires in the successive generations.

CONCLUSION

A single use of reference sires in one generation aimed at the connectedness of herds led to a slight increase in the average reliability of breeding value from 0.121 to 0.145 (by 2.4%) in the whole set. This increase from 0.241 to 0.33 (by 8.9%) was recorded only in animals that were directly connected through reference sires. Higher reliability was mainly a result of the increased number of contemporaries. An indirect consequence of the increased connectedness of the set and higher use of reference sires in insemination was a decrease in the effective number of sires' offspring. This is the reason why an increase in the connectedness of the set influenced the reliability of breeding value of the whole set less than we would expect. The use of reference sires had an explicit influence on the fluctuation of breeding values. Complete evaluation of the impact of reference sires on estimated breeding value could be done only if we knew real breeding values of animals in the simulated set.

The use of a higher number of reference sires in insemination with only one offspring in each of the connected herds appeared as the most suitable approach from the aspect of the connectedness of the set and reliability of breeding value estimation.

The use of insemination to make herds connected only in one generation is insufficient. A marked influence of herd connectedness on the estimation of breeding value and its reliability, and on the result of breeding, is to be expected after reference sires have been used in several successive generations.

REFERENCES

Brascamp E.W. (1978): Methods of economic optimization of animal breeding plans. Research Institute for Animal Husbandry Schoonoord. Zeist, Raport B-134, 117 pp.

Clement V., Bibe B., Verrier E., Elsen J.M., Manfredi E., Bouix J., Hanocq E. (2001): Simulation of model adequacy and data structure on the estimation of genetic parameters for traits with direct and maternal effect. Genet. Sel. Evol., 33, 369–395.

Gibson J.P. (1995): An introduction to the design and economics of animal breeding strategies. University of Guelph, Canada, 224 pp.

Hanocq E., Boichard D., Foulley J.L. (1996): A simulation study of the effect of connectedness on genetic trend. Genet. Sel. Evol., 28, 67–82.

Henderson C.R. (1984): Application of linear models in animal breeding. University of Guelph, Press, 462 pp.

Johnson P.L., Sherlock R.G., Lopez-Villalobos N., Garrick D.J., Blair H.T. (2002): Optimisation of breeding scheme design for the New Zealand texel sire referencing scheme. 7th WCGALP Montpellier, Communication, 02–08.

Kennedy B.W., Trus D. (1993): Considerations on genetic connectedness between management units under animal model. J. Anim. Sci., 71, 2341–2352.

Lewis R.M., Simm G. (2000): Selection in sire referencing scheme in sheep. Livest. Prod. Sci., 67, 129–141.

Lewis R.M., Crump R.E., Simm G., Thomson R. (1999): Assessing connectedness in across-flock genetic evaluation. In: Proc. Brit. Society of Animal Science, 121 pp.

Meyer K. (1992): Variance components due to direct and maternal effects for growth traits of Australian beef cattle. Livest. Prod. Sci., 31, 179–204.

Meyer K. (2001): Estimates of direct and maternal covariance functions for growth of Australian beef calves from birth to weaning, Genet. Sel. Evol., 33, 487–514.

Miglior F., Van Doormaal B.J. (2000): Impact of Second Country Proofs on Genetic Evaluations. Interbull Bulletin, 25, 57–59.

Mrode R.A. (1996): Linear models for the prediction of animal breeding value. CAB International, Wallingford, 187 pp.

Přibyl J., Přibylová J. (2001): Expected reliability of cow's breeding value for average life-time milk production. Czech J. Anim. Sci., 46, 523–532.

Přibyl J., Misztal I., Přibylová J., Šeba K. (2003): Multiple-breed, multiple-traits evaluation of beef cattle in the Czech Republic. Czech J. Anim. Sci., 48, 519–532.

Přibylová J., Vostrý L., Veselá Z., Přibyl J., Bohmanová J. (2004): Breeding value for own growth of beef bulls in performance-test station. Anim. Sci. Paper Rep., 22, 97–103.

Robinson O.W. (1981): The influence of maternal effects on the efficiency of selection: a review. Livest. Prod. Sci., 8, 212–137.

Searle S.R. (1987): Linear model for unbalanced data. John Wiley and Sons, New York. 560 pp.

Simm G., Lewis R.M., Collins J.E., Nieuwhof G.J. (2001): Use of sire referencing scheme to select for improved carcass composition in sheep. J. Anim. Sci., 79, 255–259.

Sorensen D.A., Kennedy B.W. (1983): The use of the relationship matrix to account for genetic drift variance in the analysis of genetic experiments. Theor. Appl. Genet., 66, 217.

Šafus P., Přibyl J., Veselá Z., Vostrý L., Štípková M., Stádník L. (2006): Selection indexes for bulls of beef cattle. Czech J. Anim. Sci., 51, 285–298.

Tosh J.J., Wilton J.W. (1994): Effect of data structure on variance of prediction error and accuracy of genetic evaluation. J. Anim. Sci., 72, 2568–2577.

Van Vleck L.D. (1987): Contemporary groups for genetic evaluations. J. Dairy Sci., 70, 2456–2464.

Veselá Z., Přibyl J., Vostrý L., Štolc L. (2004): Effect of sire referencing scheme on reliability of the breeding value estimation in beef cattle. Anim. Sci. Paper Rep., 22, 151–157.

Veselá Z., Přibyl J., Šafus P., Vostrý L., Šeba K., Štolc L. (2005): Breeding value for type traits in beef cattle in the Czech Republic. Czech J. Anim. Sci., 50, 385–393.

Corresponding Author

Ing. Zdeňka Veselá, Institute of Animal Science, Přátelství 815, P.O. Box 1, 104 01 Prague 114-Uhříněves, Czech Republic
Tel. +420 267 009 664, e-mail: vesela.zdena@vuzv.cz