Journal of Central European Agriculture, 2014, 15(4), p.51-63 DOI: 10.5513/JCEA01/15.4.1507

Methods for calculating the breeding values of missing traits and their comparison

Metody pro dopočítání plemenné hodnoty pro chybějící znak v selekčním indexu a jejich porovnání

Lenka KRPÁLKOVÁ[1*], Josef PŘIBYL[1], Luboš VOSTRÝ[1], Mojmír VACEK[2] and Luděk STÁDNÍK[2]

[1]Institute of Animal Science, Prague-Uhříněves, Czech Republic
[2]Czech University of Life Sciences Prague, Faculty of Agrobiology, Food and Natural Resources, Department of Animal Husbandry, Prague, Czech Republic

* correspondence

Lenka Krpálková; Institute of Animal Science; Department of Cattle Breeding; Přátelství 815; 104 00 Prague 10 – Uhříněves; Czech Republic; phone number: + 420 267 009 529; e-mail: lenka.krpalkova@centrum.cz

Abstract

The objective of this study was to calculate the breeding values (BVs) of traits missing in a selection index. Different traits can be evaluated within the breeding programs of given countries. The BV of a trait can be calculated based on genetic correlations with other traits. Similarly, the BV of a missing trait can be calculated for imported bulls. Two methods of calculation were used. Method A was based on a regression of BVs. Method B was based on performing a de-regression of BVs and their retroactive calculation. Both of these methods were tested using a Czech and a Canadian database of BVs for Holstein bulls. The Czech database of Holstein bulls contained 766 bulls and the Canadian database 851. Two calculations were performed for bulls with low reliability of estimated BVs, the first calculation with their genetic correlation matrix and the second with a genetic correlation matrix created from a set of bulls with high reliability of BVs. These newly calculated BVs (CBVs) were then compared with the national BVs (NBVs) using correlation coefficients. The highest correlations were achieved with high reliability bulls when all traits were included into the calculation (34 evaluated traits). The correlations of these bulls averaged 0.82, with an average standard deviation of 0.19. The lowest correlations were found when low reliability bulls were included and the genetic correlation matrix from the high reliability bulls was applied. That average correlation was 0.74 and standard deviation 0.25. When only 15 traits were evaluated in the model, the average correlation for all sets was 0.68 with standard deviation of 0.28. These results show that calculating the BV of a missing trait is possible using both methods. Method B was slightly more accurate in its prediction.

Key words: breeding value, de-regression, genetic correlation, Holstein cattle, missing traits

Abstrakt

Cílem této studie bylo dopočítat plemenné hodnoty (PH) býků, které chybí v selekčním indexu. V rámci šlechtitelských programů v jednotlivých zemích mohou být hodnoceny různé znaky. PH znaků lze vypočítat na základě genetických korelací s jinými znaky. Podobně lze PH chybějícího znaku vypočítat i pro importované býky. Byly použity dvě metody výpočtu. Metoda A byla založena na regresi PH. Metoda B byla založena na odregresování PH a na následném zpětném výpočtu chybějící PH pomocí korelací. Obě metody byly testovány na české a kanadské databázi PH holštýnských býků. Česká databáze holštýnských býků obsahovala 766 býků a kanadská databáze 851. U býků s nízkou spolehlivostí odhadovaných PH byly provedeny dva výpočty, první výpočet na základě jejich vlastní genetické korelační matice a druhý pomocí genetické korelační matice vytvořené podle souboru býků s vysokou spolehlivostí PH. Tyto nově vypočtené PH (CBVs) pak byly porovnávány s původními národními PH (NBVs) pomocí korelačních koeficientů. Nejvyšší korelace byly dosaženy u býků s vysokou spolehlivostí odhadu, kde byly zahrnuty všechny znaky ve výpočtu (34 znaků). Korelace u těchto býků činili v průměru 0,82 a průměrná směrodatná odchylka byla 0,19. Nejnižší korelace byly zjištěny u býků s nízkou spolehlivostí odhadu při použití genetické korelační matice s vysokou spolehlivosti odhadu. Průměr korelace byl 0,74 a směrodatná odchylka 0,25. Při hodnocení pouze 15 vlastností v modelu byla průměrná korelace u všech souborů 0,68 se směrodatnou odchylkou 0,28. Výsledky ukazují, že výpočet PH chybějícího znaku je možné při použití obou metod. Metoda B byla v předpovědi nepatrně přesnější než metoda A.

Klíčová slova: genetická korelace, holštýnský skot, chybějící znak, odregresování, plemenná hodnota

Detailed Abstract

V každé zemi se u holštýnského plemene hodnotí různé vlastnosti v závislosti na šlechtitelských programech daných zemí. Cílem této práce bylo najít vhodnou metodu pro dopočítání chybějící vlastnosti (plemenné hodnoty), která se nehodnotí v našem šlechtitelském programu holštýnských býků, ale hodnotí se v jiné zemi. Dopočítaná chybějící plemenná hodnota může být následně použita v domácím selekčním indexu holštýnských býků.

Hodnoceny byly dvě metody. Metoda A byla založená na regresi plemenných hodnot bez odregresování. Metoda B byla založená na odregresování plemenných hodnot a na zpětném výpočtu plemenných hodnot. Obě tyto metody byly testovány na české a kanadské databázi plemenných hodnot holštýnských býků. Tyto databáze byly rozděleny na býky s vysokou spolehlivostí odhadu a s nízkou spolehlivostí odhadu plemenných hodnot. U býků s nízkou spolehlivostí odhadu byly provedeny dva výpočty a to s jejich vlastní korelační maticí a druhý výpočet s korelační maticí stanovenou u býků s vysokou spolehlivostí. Do výpočtu se zahrnovaly všechny vlastnosti (plemenné hodnoty užitkovosti, znaků zevnějšku, plodnosti, dlouhověkosti a obtížnosti porodů), ale též byly provedeny výpočty pouze při polovičním počtu zahrnutých vlastností (plemenných hodnot). Uskutečnilo se 6 výpočtů pro každou zemi oběma metodami, tzn. celkem 24 výpočtů. U všech výpočtu se jednotlivé plemenné hodnoty býků dopočítaly jako chybějící. Následně se tyto nově dopočítané plemenné hodnoty porovnaly se skutečnými plemennými hodnotami pomocí Pearsonových korelačních koeficientů ve statistickém programu SAS 9.2.

U obou metod jsme získaly velmi podobné výsledky. Metoda B byla v předpovědi nepatrně lepší než metoda A. Nejvyšší korelace byly dosaženy u býků s vysokou spolehlivostí při zahrnutí všech vlastností. Průměrné korelace těchto býků u kanadského a českého souboru byly 0,82 s průměrnou směrodatnou odchylkou 0,19. Nejnižších korelací bylo dosaženo při zahrnutí býků s nízkou spolehlivostí za použití korelační matice býků s vysokou spolehlivostí. Průměrná korelace byla 0,74 se směrodatnou odchylkou 0,25. Při hodnocení pouze 15 vlastností v modelu byla průměrná korelace u všech souborů 0,68 se směrodatnou odchylkou 0,28.

Z výsledků vyplývá, že na základě korelací mezi plemennými hodnotami je možné dopočítat chybějící plemennou hodnotu. Korelace mezi skutečnými a dopočítanými plemennými hodnotami vykazují vysoké hodnoty.

Introduction

Genetic parameters are crucial for those traits included as selection criteria into a breeding program, as are the possibility and cost of monitoring and evaluating such traits, their economic value, and their relationship to the overall economics of the breed. Missing traits which are not included into the selection of parents could be the cause of only suboptimal genetic progress in the next generation, but that progress depends also on the accuracy of the estimated breeding value (BV) (Mark et al., 2007). Use of a regression equation is among the possible approaches for calculating the BVs of missing traits (Powell, 1988). Regression models for calculating the BVs of missing traits are also used in the international evaluation of bulls by the multiple across country evaluations (MACE) method (Mark et al., 2007). Jorjani and Fikse (2004) compared the methods based on regression equations and the so-called simple-MACE method. Their analysis showed that the multi-trait regression equations were always better in predicting than were single-trait regression equations. The simple-MACE method was better in prediction than was the regression equation when genetic correlations between traits were low. Schaeffer et al. (1996) included among those factors affecting the accuracy of newly CBVs in the MACE method the genetic trends within countries, evaluations of imported bulls, year of birth for those bulls included into the analysis, and estimates of genetic correlations between countries. Loberg et al. (2009) estimated BVs for bulls without national BVs using correlations with evaluated bulls in other countries. The accuracy of the CBVs ranged from 0.92 to 0.97, and loss of genetic gain was 2.82% to 11.9%. In the present study, it was hypothesised that the possibility of CBVs for some important traits within and also between countries is a suitable and simple method for obtaining the required BVs with high accuracy of prediction. The objective of this work was to find method for estimate the BVs of traits which are not evaluated within the breeding programs of some countries but which are evaluated in others without necessarily testing of animals in the population.

Materials and methods

The data used in calculating the BVs of the missing traits was comprised of national BVs of Holstein bulls from the Czech Republic and Canada. The database of Czech national BVs was provided by Plemdat Corporation, s.r.o. (www.plemdat.cz), and the database of Canadian national BVs was provided from the Canadian Dairy Network (www.cdn.ca). Two methods were used to calculate BVs of the missing traits. The calculations were performed using IML/SAS 9.1 (SAS, 2004). Both national databases contained BVs with their reported reliabilities for 34 evaluated traits included in the calculation. These databases were divided into bulls with low versus high reliability of NBVs (Table 1). Bulls in the calculation sets had BVs for all evaluated traits and their reliability of NBVs was within a predefined range (Table 1). Other selection criteria were not considered. In the end, the Czech database's set of Holstein bulls included 766 animals and the Canadian database 851. The Canadian database contained BVs only for production and type traits. Most of the traits coincided with those in the Czech database. Traits 1–26 were the same for the Czech and Canadian sets of bulls, while traits 27–34 (Canadian) and 27*–34* (Czech) differed between the 2 sets. Traits 27–34 in the Canadian set were not evaluated either in the Czech database of bulls or in the Czech breeding program for Holstein cattle.

Table 1. Sets of bulls divided by reliability (r^2) of national breeding values (NBVs)[1]
Tabulka 1. Soubory býků rozdělené podle spolehlivosti (r^2) národních plemenných hodnot (NBVs)[1]

Database	r^2	n[4]	Range of r^2 (BVs No. 1–5)	Average of r^2 (BVs No. 1–5)	Range of r^2 (BVs No. 6–34)	Average of r^2 (BVs No. 6–34)
Czech	High[2]	382	> 0.88	0.97	> 0.83	0.90
	Low[3]	384	≤ 0.88	0.81	≤ 0.83	0.63
Canadian	High[2]	507	> 0.93	0.99	> 0.88	0.97
	Low[3]	344	≤ 0.93	0.85	≤ 0.88	0.75

[1] Breeding values (BVs) are for traits 1–5 and 6–34 as shown in Table 2.
[2] Sets of bulls with high reliability of NBV.
[3] Sets of bulls with low reliability of NBV.
[4] n = number of bulls in set.

Method A

The modified equation 1 was used to calculate the CBVs of those traits treated as missing. This calculation is used in the MACE method (Mark et al., 2007 and Schaeffer, 2001). The CBVs of the missing traits were converted using the genetic variance and covariance between the traits.

$$\mathbf{BV}_{\mathrm{i}} = \mathbf{G}_{\mathrm{i}} \cdot \mathbf{G}^{-1} \cdot \mathbf{BV}_{\mathrm{n}} \ , \qquad [1]$$

where BV$_i$ = CBV of the missing trait, G$_i$ = vector of the genetic covariance between the missing and evaluated traits, G^{-1} = genetic variance-covariance matrix between the evaluated traits, and BV$_n$ = vector of the NBVs of the known traits. The number of values in the vector BV$_n$ was always without the CBV of the trait considered as missing, i.e., BV$_{n(n-1;1)}$. The vector G$_i$ contained only covariance of the missing trait with the available evaluated traits without the variance of missing traits, i.e., G$_{i(1;\ n-1)}$. The elements of variance-covariance matrix G were determined using formulas 2 and 3. The matrix G contains variance and covariance only for traits considered as known, i.e., G$_{(n-1;\ n-1)}$.

$$\boldsymbol{\sigma}_{\mathrm{g}}{}^{2} = \boldsymbol{\sigma}_{\mathrm{BV}}{}^{2} / \mathbf{r}^{\mathbf{2}} \ , \qquad [2]$$

where σ_g^2 = genetic variance, σ_{BV}^2 = variance of BV, and r^2 = reliability of NBV. Standard deviations (σ_g) of the traits were calculated from genetic variance (σ_g^2), then were used to calculate the genetic covariance ($\mathbf{cov}_{(g_1,g_2)}$):

$$\mathbf{q}_{(\mathrm{BV}_1,\,\mathrm{BV}_2)} = \frac{\mathbf{cov}_{(\mathrm{g}_1,\mathrm{g}_2)}}{\boldsymbol{\sigma}_{\mathrm{g}_1} \cdot \boldsymbol{\sigma}_{\mathrm{g}_2}} \ . \qquad [3]$$

Genetic correlations ($\mathbf{q}_{(BV_1,\,BV_2)}$) between the traits were calculated by the CORR procedure in SAS 9.1 (SAS, 2004).

Method B

In a first step, the BVs were de-regressed and a vector with the daughters' yield deviations (DYD) values was obtained (Liu et al., 2004). Subsequently, the multi-traits model was calculated for each individual sire (best linear unbiased prediction method). De-regression of NBVs used the following equation:

$$\mathbf{DYD} = \left[\mathbf{Z}'\mathbf{R}^{-1}\right]^{-1} \cdot \left[\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}' + \mathbf{H}^{-1}\right] \cdot \left[\mathbf{BV}\right] \ , \qquad [4]$$

where BV = vector of known BVs of the traits, DYD = vector of de-regressed BVs, matrix Z = diagonal identity matrix connecting the BVs with individuals, matrix R = diagonal matrix with diagonal elements equal to σ_e^2/EDC and related to the relevant BV and sire (j) (Formula 5).

$$\mathbf{R}^{-1} = \mathbf{I} \otimes \frac{\boldsymbol{\sigma}_{\mathrm{e}}{}^{2}}{\mathbf{EDC}_{\mathrm{j}}} \ , \qquad [5]$$

where σ_e^2 = the residual variances, I = an identity matrix, and EDC = effective daughter contributions

$$\mathbf{H} = \mathbf{G} \otimes \mathbf{A}^{-1} , \qquad [6]$$

and where matrix H = matrix calculated with variance-covariance matrix G and additive genetic relationships matrix A.

EDC in the matrix R determines in the model the weight of the sire based on the number of daughters for the estimated BV (Liu et al., 2004). In this case, EDC is unknown if it is evaluated from the reliability coefficient of NBV and heritability coefficient of the trait (formulas 7 and 8). After this step, one average daughters' deviation adjusted for systematic environmental effects is obtained for each bull and each trait. EDC is associated with this deviation.

$$\mathbf{EDC} = \left(\frac{\mathbf{r}^2}{1-\mathbf{r}^2}\right)\cdot\mathbf{k}\,, \qquad [7] \qquad\qquad \mathbf{k} = \left(\frac{4-\mathbf{h}^2}{\mathbf{h}^2}\right), \qquad [8]$$

where r^2 = reliability coefficient of NBV for a sire and a trait, and h^2 = heritability coefficient of a trait.

The heritability coefficients (h^2) for Canadian and Czech estimates were taken from Přibyl et al. (2004), Sewalem et al. (2010) and Bastin et al. (2010). The elements of variance-covariance matrix G were determined in the same way as in method A (formulas 2 and 3). Only variances for de-regression of the BVs were used in the matrix G. Matrix A is in this case the identity matrix and therefore de-regressed BVs could be determined iteratively and separately only for each individual sire, and not simultaneously for many sires. Therefore, the calculation is more simple and it follows from this fact that H = G (Formula 6). The reliability of the CBVs was computed using the modified formula (Schaeffer, 2010):

$$\mathbf{r}^2 = \mathbf{a}_{\mathrm{ii}} - \left(\frac{\mathbf{c}_{\mathrm{ii}}}{\mathbf{h}_{\mathrm{ii}}}\right), \qquad [9]$$

where a_{ii} = diagonal element of the relationship matrix A (in the case of identity matrix: $a_{ii} = 1$), c_{ii} = diagonal element of the matrix C (left side of the inversion equations), and h_{ii} = diagonal element of the genetic covariance matrix.

Each trait in the 4 sets of bulls (Table 1) was calculated as missing, and then this new CBV was compared with the NBV using a correlation coefficient. All BVs of the traits in the calculations were converted by the correlation matrix for their own countries using both methods A and B. In the case of bulls with low reliability of NBVs, the calculation was also performed with correlation matrix r_1 made up of bulls with high reliability of NBVs. r_1 is Czech or Canadian genetic correlation matrix calculated using bulls with high reliability of NBVs and r_2 is Czech or Canadian genetic correlation matrix calculated using bulls with low reliability of NBVs (Table 2). The hypothesis of higher accuracy in the correlation matrix determined in bulls with high reliability of NBVs was assumed. Six calculations were performed for each country and using both methods, i.e., 24 calculations in total.

Results

Analysis and Comparison of Genetic Correlations of Traits in One Country

Genetic correlation matrices r_1 and r_2 differed depending on the reliability of OBVs in the Czech and Canadian sets of bulls. Correlation coefficients between production traits (traits 1–5) for bulls with high and low reliability of OBVs were almost the same in the two countries and sets. The opposite was true for the genetic correlation between the traits 32* (birthing difficulty [maternal effect]) and 34* (longevity) or the genetic correlation between the traits 33* (birthing difficulty [direct effect]) and 34* (longevity) in the Czech set. The difference in correlation matrix was approximately 0.5 for all those traits. In the Canadian set, the genetic correlation between traits 1 (milk [kg]) and 23 (bone quality), 1 (milk [kg]) and 28 (dairy strength) or 1 (milk [kg]) and 31 (udder texture) differed notably when the reliability of the OBVs was high versus when it was low. The difference in correlation matrix was approximately 0.25 for all those traits mentioned above. The differences in genetic correlations within a country depend on the chosen set of bulls and its parameters. The reliability of OBVs and their variability impact on the resulting genetic correlation between traits. The more different are the reliability of OBVs and their variability, the greater are the differences in the genetic correlation matrices.

Analysis and Comparison of Genetic Correlations of the Same Traits between Countries

Differences in the genetic correlation coefficients between countries are not the same for all traits and depend on the facts mentioned above (e.g., number of bulls in the set, reliability of their OBVs). Differences in genetic correlation matrices also depend on the description and evaluation of traits and those traits' levels and variability within the breed in the countries.

Correlations between NBVs and CBVs

Methods A and B performed almost the same in predicting the CBVs of the missing traits, as the mean correlations and their standard deviations (shown in Table 2) calculated when only this variable was changed did not vary substantially.

Table 2. Average correlations between national breeding values (NBVs) and converted breeding values (CBVs) in calculations with 34 and 15 traits
Tabulka 2. Průměrné korelace mezi národními plemennými hodnotami (NBVs) a nově dopočítanými plemennými hodnotami (CBVs) se zahrnutými 34 a15 znaky ve výpočtu

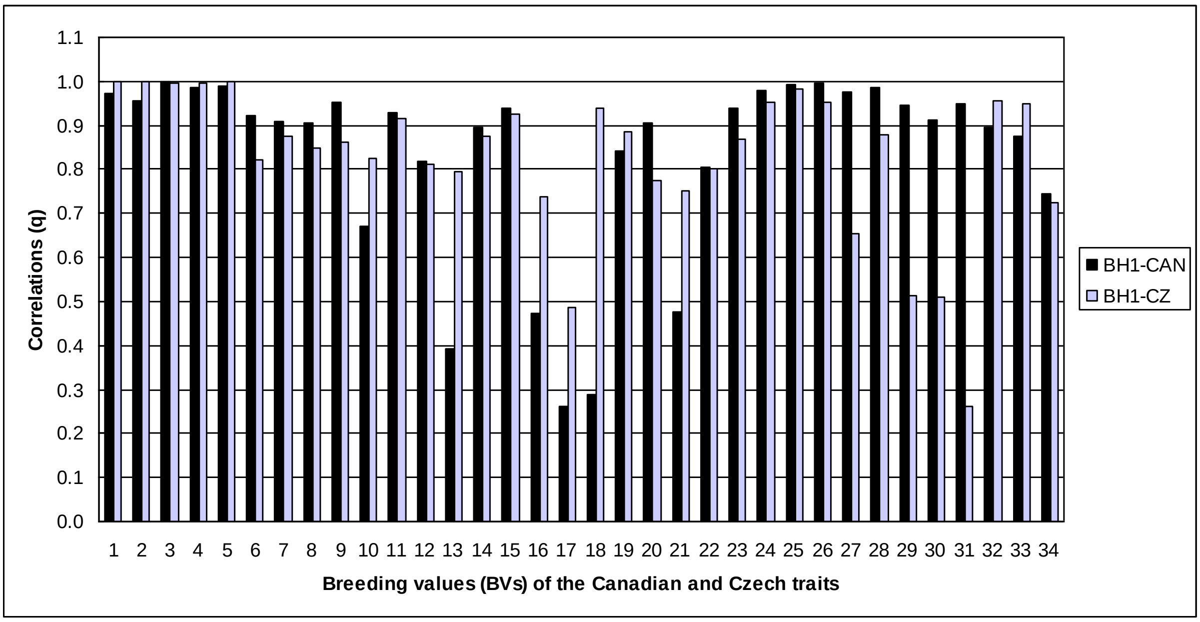
Canadian set of bulls			Czech set of bulls		
Calculation[1]	Average of correlation	Standard deviation of correlation	Calculation[1]	Average of correlation	Standard deviation of correlation
AH1$^{34)}$	0.79	0.21	AH1$^{34)}$	0.83	0.17
AL2$^{34)}$	0.81	0.20	AL2$^{34)}$	0.78	0.18
AL1$^{34)}$	0.76	0.27	AL1$^{34)}$	0.71	0.24
BH1$^{34)}$	0.83	0.21	BH1$^{34)}$	0.83	0.17
BL2$^{34)}$	0.81	0.19	BL2$^{34)}$	0.77	0.18
BL1$^{34)}$	0.77	0.25	BL1$^{34)}$	0.71	0.23

AH1 15)	0.67	0.31	AH1 15)	0.74	0.21
AL2 15)	0.64	0.33	AL2 15)	0.70	0.23
AL1 15)	0.64	0.33	AL1 15)	0.66	0.27
BH1 15)	0.69	0.30	BH1 15)	0.74	0.21
BL2 15)	0.69	0.28	BL2 15)	0.70	0.23
BL1 15)	0.65	0.33	BL1 15)	0.67	0.29

[1] Each of three letters in the calculation designation has a meaning as follows: first position – A = method A, B = method B; second position – L = low reliability of NBVs, H = high reliability of NBVs; third position – 1 = genetic correlation matrix r_1 of the relevant country used, 2 = genetic correlation matrix r_2 of the relevant country used, 34), 15) = average correlations between national breeding values (NBVs) and converted breeding values (CBVs) in calculations with 34 or 15 traits.

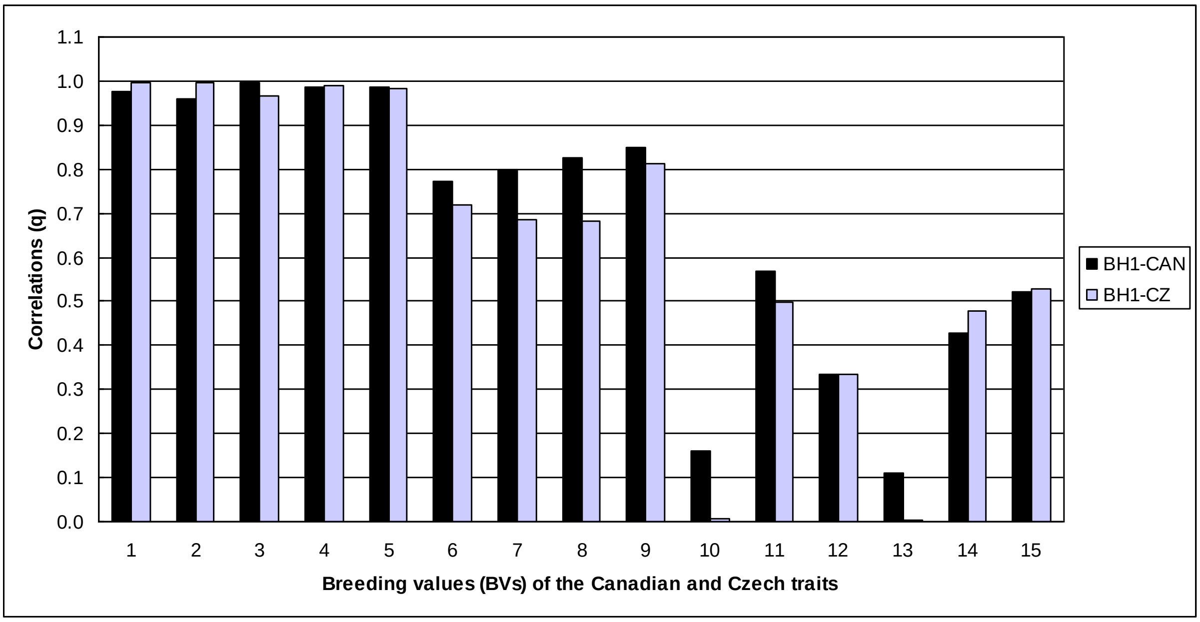
The best predictions of the CBVs for those traits considered as missing were found in calculations for bulls with high reliability of NBVs (calculations AH1 and BH1). CBVs' accuracies were lowest when calculated in the sets of bulls with low reliability of NBVs and using correlation matrix r_1 consisting of bulls with high reliability of NBVs (calculations AL1 and BL1). These general conclusions are evident in all the resulting Figures 1 and 2 and Tables 1.

Figure 1. Correlations between national breeding values (NBVs) and converted breeding values (CBVs) of 34 traits[1] used in calculations for the Canadian and Czech set of Holstein bulls[2]
Graf 1. Korelace mezi národními plemennými hodnotami (NBVs) a nově dopočítanými plemennými hodnotami (CBVs) se zahrnutými 34 znaky ve výpočtu v souboru kanadských a českých holštýnských býků



[1] Breeding values of traits 1–34 on the x-axis are: 1- Milk (kg), 2- Protein (kg), 3- Protein (%), 4- Fat (kg), 5- Fat (%), 6- Body frame, 7- Chest width, 8- Body depth, 9- Angularity, 10- Rump angle, 11- Rump width, 12- Rear leg rear view, 13- Rear legs set – side view, 14- Foot angle, 15- Fore udder attachment, 16- Front teat placement, 17- Teat length, 18- Udder depth, 19- Rear attachment height, 20- Suspensory ligament, 21- Rear teat placement, 22- Rear attachment width, 23- Bone quality, 24- Feet & legs, 25- Mammary system, 26- Conformation; Traits (No. 27–34) evaluated in Canadian set of bulls: 27- Rump, 28- Dairy strength, 29- Pin setting, 30- Loin strength, 31- Udder texture, 32- Heel depth, 33- Rear legs set, 34- Height at front end; Traits (No. 27–34) evaluated in Czech set of bulls: 27- Locomotion, 28- Body condition, 29- Somatic cell score, 30- Fertility of daughters, 31- Bull fertility, 32- Birthing difficulty (maternal effect), 33- Birthing difficulty (direct effect), 34- Longevity
[1] In the designation BH1 of the calculations for Canadian (CAN) and Czech (CZ) bulls, B represents method B, H stands for high reliability of NBVs, and 1 indicates that the genetic correlation matrix r_1 of the relevant country was used. P-values of correlations between all NBVs and CBVs were significant ($P < 0.01$).

Figure 2. Correlations between national breeding values (NBVs) and converted breeding values (CBVs) of 15 traits[1] used in the calculations for the Canadian and Czech
Graf 2. Korelace mezi národními plemennými hodnotami (NBVs) a nově dopočítanými plemennými hodnotami (CBVs) se zahrnutými 15 znaky ve výpočtu v souboru kanadských a českých holštýnských býků



[1] Breeding values of traits 1–15 on the x-axis are: 1- Milk (kg), 2- Protein (kg), 3- Protein (%), 4- Fat (kg), 5- Fat (%), 6- Body frame, 7- Chest width, 8- Body depth, 9- Angularity, 10- Rump angle, 11- Rump width, 12- Rear leg rear view, 13- Rear legs set – side view, 14- Foot angle, 15- Fore udder attachment

[2] In the designation BH1 of the calculations for Canadian (CAN) and Czech (CZ) bulls, B represents method B, H stands for high reliability of NBVs, and 1 indicates that the genetic correlation matrix r_1 of the relevant country was used. P-values of correlations between all NBVs and CBVs were significant ($P < 0.01$).

The reasons for low correlations between NBVs and the new CBVs were primarily the following: low genetic correlation between the traits; use of a different genetic correlation matrix, r_1, in bulls with low reliability of NBVs (calculations AL1 and BL1); and a smaller amount of information (calculations with the 15 traits). These facts complemented one another and significantly influenced the accuracy of the CBVs and conclusiveness of the correlation coefficients (P–value).

When comparing the calculations of AH1 and AL2 in the Canadian set of bulls (Table 2), the highest correlations between the NBVs and CBVs were found in bulls with low reliability of NBVs. The difference in the average correlation between the AH1 and AL1 calculations was only 0.02. This difference reflects higher correlation coefficients in the correlation matrix r_2 for some traits, e. g., traits 1 (kg milk) and 2 (kg protein). These traits had higher correlation coefficients with the other traits (on average, the correlation coefficients were higher by about 0.1) in the correlation matrix r_2.

Figures 1 and 2 present the accuracy of calculating the CBVs based on genetic correlations between traits for each of the included traits. Shown there are only those calculations using method B, the sets of bulls with high reliability of NBVs, and correlation matrix r_1 consisting of bulls with high reliability of NBVs. The most accurate predictions of CBVs were obtained in this case (i.e. for calculations designated BH1). Predictions for some traits differed from the aforementioned conclusions. The best prediction of the CBVs for traits 1, 2, 17, 18, and 21 (kg milk, kg protein, teat length, udder depth, and rear teat placement) in the Canadian set was achieved by the AL2 and BL2 calculations when including all 34 traits. Traits 1 and 2 (kg milk, kg protein) were discussed above. Traits 17 and 18 (teat length, udder depth) had slightly higher correlation coefficients in the genetic correlation matrix r_2 than in the matrix r_1 and genetic correlations with the other traits were low. Similar reasons related also to the AL2 and BL2 calculations for trait 31* (bull fertility) in the Czech set (i.e., very low genetic correlation with the other traits and higher correlations in the genetic correlation matrix r_2). The reason for low correlations between the NBVs and CBVs in traits 17 and 18 in the Canadian set and trait 31* (bull fertility) in the Czech set (Figure 1) lies in very low genetic correlation with other traits (varying within the range ± 0.2).

Figure 2 and Table 2 show findings when only the first 15 traits were evaluated. The general accuracy of the CBVs declined significantly when the smaller number of traits was included. This was not the case, however, for the production traits, 1–5 (kg milk, kg protein, % protein, kg fat, % fat), the correlations for which remained unchanged. The reason lies in the high genetic correlations between these production traits, as shown also in Table 2. The correlations between the NBVs and CBVs in the selected nonproduction traits decreased significantly. For example, traits 10 (rump angle) and 13 (rear legs set – side view) in Figure 2 for the sets from both countries had lower numbers of high genetically correlated traits than did the other traits.

Variability, Averages and Reliability of NBVs and CBVs

The higher the correlations between the NBVs and the CBVs, the more similar are the means and standard deviations (and conversely). However, this general rule has many exceptions. In method B, the reliability of the CBVs was calculated using Formula 9. For sets of bulls with high reliability of NBVs (Table 1) the reliability of the CBVs for the traits ranged from 0.97 to 0.99 (BH1 calculations).

Discussion

In calculating CBVs, the number of traits included is important, but even more important are the levels of the genetic correlations of other traits with the calculated trait. It is similarly possible to calculate missing selection indices based on correlations. Jorjani and Fikse (2004) used three highly correlated USA composite traits (net merit, cheese merit and fluid merit) to calculate the Swedish Total Merit Index (TMI). The correlations between Swedish TMI and the USA composite traits increased on average by 0.77 after adjustment for reliabilities. The importance of high correlations between traits in the genetic correlation matrices is shown in Figure 1 for traits 17 and 18 (teat length, udder depth) in the Czech set or for trait 31 (udder texture) in the Canadian set. In those calculations including all 34 traits, however, the CBVs were not calculated with sufficient accuracy because of their low genetic correlations with other traits. Another example can be seen in traits 1–5 (kg milk, kg protein, % protein, kg fat, % fat). These CBVs are predicted very accurately regardless of the total number of traits included (Figures 1, 2). The reason lies in the high genetic correlations among production traits. More correlated traits must be brought into calculation of the CBVs in cases of low correlation between the missing trait and other traits. Zavadilová et al. (2011) stated the importance of sufficiently high correlation between selected BVs when considering to estimate the BV for a trait that is itself difficult to measure for selection purposes.

It is mainly the correlation matrix that affects the accuracy of the CBV. Lower correlations in the genetic correlation matrix are reflected in lower accuracy of the CBV. Toghiany (2012) stated that the genetic parameters of a breed, and also the parameters in the genetic correlation matrix, are changing over time. Subsequently, Hossein-Zadeh (2012) determined the traits variations in accordance with changes of genetic base.

To calculate the BV of a missing trait, only traits can be used which have similar genetic correlations in both countries and also high correlations with the missing trait. Jorjani et al. (2005) examined the effects of using bulls across countries on estimates of genetic correlations to account for genotype–environment interaction. The accuracy of the genetic correlation matrix is affected by the factors mentioned above (e.g., number of bulls included, reliability of NBVs, variability in reliability of NBVs). The selected set of bulls should characterize the population in the best way possible. Powell (1988) recommended using BV with reliability R ≥ 0.75. Jorjani and Fikse (2004), Powell (1988) and Wilmink et al. (1986) are in agreement that high genetic correlations between traits and high reliability of BVs estimation are important in calculating the BV of the missing trait. Therefore, we can conclude that using higher reliability of NBVs (i.e. higher accuracy of the NBVs) as correlated traits is associated with greater accuracy of the new calculated CBV of missing traits.

Acknowledgments

We are grateful to Gale A. Kirking of English Editorial Services for his help with the final manuscript. The authors also thank Plemdat Corporation, s.r.o. for providing data. This research was supported by „S grant of MSMT CR" and by projects MZERO0174 from the Ministry of Agriculture, all of the Czech Republic.

References

Bastin, C., Loker, S., Gengler, N., Sekalem, A., Miglior, F., (2010) Genetic relationship between body condition score and reproduction traits in Canadian Holstein and Ayrshire first-parity cows. J Dairy Sci 93 (5), 2215–2228. DOI: 10.3168/jds.2009-2720.

Jorjani, H., Emanuelson, U., Fikse, W. F., (2005) Data subsetting strategies for estimation of across-country genetic correlations. J Dairy Sci 88, 1214–1224. DOI: 10.3168/jds.S0022-0302(05)72788-0.

Jorjani, H., Fikse, F., (2004) Conversion equations revisited: single- and multiple-trait conversion equations and comparison with simple-MACE. Interbull Bull No. 32, 25–30.

Liu, Z., Reinhardt, A., Bünger, A., Reents, R., (2004) Derivation and calculation of approximate reliabilities and daughter yield-deviation of a random regression test-day model for genetic evaluation of dairy cattle. J Dairy Sci 87, 1896–1907. DOI: 10.3168/jds.S0022-0302(04)73348-2.

Loberg, A., Jorjani, H., Fikse, W. F., (2009) Prospects of performing multiple-country comparison of dairy sires for countries not participating in Interbull international genetic evaluations. S Afr J Anim Sci 39, 86–89.

Mark, T., Fikse, W. F., Sullivan, P. G., VanRaden, P. M., (2007) Prediction of genetic correlations and international breeding values for missing traits. J Dairy Sci 90, 4805–4813. DOI: 10.3168/jds.2007-0248.

Hossein-Zadeh, N. G., (2012) Genetic parameters and trends for lactation length in the first three lactations of Holstein cows. Arch Tierz 55, 533–539.

Přibyl, J., Šafus, P., Štípková, M., Stádník, L., Čermák, V., (2004) Selection index for bulls of Holstein cattle in the Czech Republic. Czech J Anim Sci 49, 244–256.

Powell, R. L., (1988) Comparison of procedures to convert sire evaluations between countries. J Dairy Sci 71, 1609. DOI: 10.3168/jds.S0022-0302(88)79724-6.

SAS Institute Inc (2004) SAS/STAT® 9.1. User’s Guide. SAS Institute Inc., Cary, NC.

Sewalem, A., Kistemaker, G. J., Miglior, F., (2010) Relationship between female fertility and production traits in Canadian Holsteins. J Dairy Sci 93(9), 4427- 4434. DOI: 10.3168/jds.2009-2915.

Schaeffer, R. L., (2010) Linear Models and Animal Breeding (materials for a June 2010 training course in Norway). Centre for Genetic Improvement of Livestock Department of Animal and Poultry Science. University of Guelph, pp. 314, http://www.aps.uoguelph.ca/~lrs/ABModels/DATA/EuropeNotes.pdf.

Schaeffer, R. L., (2001) Multiple trait international bull comparisons. Livest Prod Sci 69, 145–153. DOI: 10.1016/S0301-6226(00)00255-4.

Schaeffer, R. L., Reents, R., Jamrozik, J., (1996) Factors influencing international comparisons of dairy sires. J Dairy Sci 79, 1108–1116. DOI: 10.3168/jds.S0022-0302(96)76464-0.

Toghiany, S., (2012) Genetic relationships between production traits and reproductive performance in Holstein dairy cows. Arch Tierz 55, 458-468.

Wilmink, J. B. M., Mcijaing, A., Engel, B., (1986) Conversion of breeding values for milk from foreign populations. Livest Prod Sci 14, 223–229. DOI: 10.1016/0301-6226(86)90081-3.

Zavadilová, L., Wolf, J., Štípková, M., Němcová, E., Jamrozik, J., (2011) Genetic parameters for somatic cell score in the first three lactations of Czech Holstein and Fleckvieh breeds using a random regression model. Czech J Anim Sci 56, 251–260.