doi: 10.17221/8453-CJAS

Comparison of models to estimate genetic parameters for scores of competitive sport horse events in the Czech Republic

A. Novotná, A. Sviťáková, J. Schmidová

Institute of Animal Science, Prague-Uhříněves, Czech Republic

ABSTRACT: The objective of the present study was to estimate genetic parameters and predict breeding values of sport horses in the Czech Republic using animal model variations. The data set for the evaluation was composed of edited records of show jumping competitions in the Czech Republic in years 1991–2013. Input data were not normally distributed; hence Blom transformation was used for the variable filtration. The Gibbs sampling algorithm was used for the genetic parameters estimation. Two models were examined. The first was a random regression model including the effect of a horse's experience in competition (expressed as the length of the horse's sporting career in days), fixed effects of sex, age, and event, and random effects of rider, permanent environment, and animal. The second model was a multi-trait model with fixed effects for sex, age, and event and random effects for rider, permanent environment, and animal. In this latter case, horse performance was classified as three traits. The first trait was jumping results from obstacle heights of 90–110 cm, the second of 120–135 cm, and the third of 140–155 cm. In the random regression model, heritability estimates ranged from 0.01 to 0.11; whereas in the multi-trait model, heritabilities were 0.07, 0.11, and 0.14 for the first, second, and third trait, respectively. Results indicate that both models could be used to predict breeding values of sport horses in the Czech Republic. The multi-trait model revealed that heritability estimates increased with the increasing height of obstacle. In the random regression model, breeding values differed according to a horse's experience in competition, allowing adjustment of the breeding value for the environmental effect of a past experience.

Keywords: show jumping; sport horses; breeding value; random regression model; multi-trait model

INTRODUCTION

In countries with advanced sport horse genetic improvement programs, breeding value prediction by the Best Linear Unbiased Prediction (BLUP) animal model is a commonly used procedure. Foreign breeding programs base breeding value prediction either upon a horse's performance tests or upon own sport results.

Show-jumping competition results can be modelled as a repeated measurement of a trait for which the phenotype of an animal is represented as a continuous function of time (Posta et al. 2010), which can be characterized by a trajectory with a theoretically infinite number of measurements. The random regression model has been used to evaluate performance of German Trotters (Bugislaus et al. 2006), Thoroughbred horses in Brazil (Buxadera and Da Mota 2008), and Hungarian Sporthorses (Posta et al. 2010). The multi-trait animal model was used by Janssens et al. (1997) in Belgium, Alridge et al. (2000) in Ireland, Viklund et al. (2011) in Sweden, and Luehrs-Behnke et al. (2002) in Germany.

Nowadays, genetic merit of the sport horse population in the Czech Republic lags considerably behind that of other European populations and has been heavily impacted by imported genetic material. Warmblood horse breeding in the Czech Republic is prominently directed to jumping performance. Evaluation is based solely on sport competition results, the only currently applicable informa-

---

Supported by the Ministry of Agriculture of the Czech Republic (Project No. MZeRO0714).

doi: 10.17221/8453-CJAS

tion for genetic evaluation of sport horses in the Czech Republic. This is so because the database of sport event results is extensive and expanding with new competitors without pronounced preselection while results from several competitions are mostly available for each horse.

Although there is a large and appropriate underlying database in the Czech Republic, routine evaluation of horse performances using the BLUP animal model has not existed so far. The only official system for sport horse evaluation (initiated in 1985) are the values based on phenotypic measurements (based upon number of penalty points or obtained percent) of show jumping, dressage, and eventing results. The basis of the method is a recalculation matrix that assigns values by taking into account the difficulty of each event. Based upon these results, sport horses are ranked using the Auxiliary penalty points system.

For stallions, the absolute sports value (a value of show jumping, dressage, eventing competitions) is determined. It is derived from Auxiliary penalty points as a weighted mean of the number of Auxiliary penalty points in all difficulties from all offspring of each individual stallion.

The objective of this study was to estimate genetic parameters and predict breeding values for show jumping performance of horses in the Czech Republic through examination of differences among analytic models and procedures.

MATERIAL AND METHODS

The dataset was created by editing the results of show jumping competitions in the Czech Republic between the years 1991–2013. The results were restricted to classic competitions evaluated according to the rules of the Czech Equestrian Federation. In total, 579 948 records from 20 126 horses were recorded in this time period.

Data. The input data for jumping performance, expressed as penalty points, were not normally distributed; therefore Blom normalizing transformation was applied using PROC RANK of SAS software (Statistical Analysis System, Version 9.1, 2005).

To estimate the genetic parameters, each trait was defined such that there would be a sufficient number of observations for a valid analysis. Riders with fewer than ten competitions and having ridden fewer than three horses were excluded from the database. Horses with fewer than five results, or horses ridden by only one rider, and competitions with less than three participants were also excluded. Also excluded were horses with adequate sport records but having fewer than four paternal half-sibs. Incomplete data (competitions with unknown horses and riders) were excluded from the data set as well. Only horses from 4 to 25 years of age were considered. After filtering, 282 437 results from 6209 horses were used in subsequent analyses. Variance components were estimated by the Gibbs sampling algorithm in GIBBS1f90 software (Version 1.37, 2002). The number of iterations was set to 120 000 with a burn-in of 40 000 iterations for all Gibbs calculations. Every twentieth iteration was sampled for an estimation of the posterior. Both models achieved stabilized results by the end of the chosen burn-in criterion.

Pedigree. Including four generations of ancestors, 74 261 horses were included in the pedigree file. Because there were 165 different breed type classes, it was not possible to create unknown groups of generations based purely on breed type due to small amount of data for each type.

Phantom groups were used in the pedigree in the most distant (fifth) generation of ancestors for the breed of horse (e.g. Czech Warmblood, Holsteiner horse, Thoroughbred, etc.), or they were pooled according to their ancestral relationships into the following breed types: Western Diluvial, Oriental – English, Oriental – Arabic, Oriental – other, Mongolian, and Nordic.

Czech Warmblood (the most common breed in the Czech Republic) was additionally split based on the birth year of the horse. In total, 20 unknown categories of ancestors were created, where the minimum number of individuals in each trait was 20. Both ancestors of an individual were known in 76.4% of cases, only one ancestor was known in 2.56% of cases, and in 21.04% of cases both sire and dam were unknown.

The models. Two models were chosen to estimate genetic parameters and predict breeding values for competitive scores of sport horses. The first model was a single-trait model with random regression (RR) and the second was a multi-trait model with quadratic regression (MT). The effect "horse's experience in competition", which was the length of the horse's sporting career in days (RR) or in years (MT), was incorporated in both of the models.

The RR model. The effect of each horse's experience in competition was expressed as days in the random regression analysis and was modelled with Legendre polynomials (LPs) from the first to the third order of fit.

$$Y_{ijklmno} = m + Sex_i + Age_j + Event_k + YearBreed_l \times fp_{YB} + Rider_m + PE_n \times fp_{PE} + Animal_o \times fp_A + e_{ijklmno}$$

where:

$Y_{ijklmno}$ = measurement variable of the competition result of the horse
m = population mean
Sex_i = fixed effect of the gender (three levels)
Age_j = fixed effect for nine age groups (4, 5, 6, 7, 8, 9, 10, 11–15, 16–25)
$Event_k$ = fixed effect of date and location of the event (27 525 levels)
$YearBreed_l$ = fixed effect of year of birth and breed of the horse (37 levels)
fp_{YB} = fixed regression coefficient of the $YearBreed_l$
$Rider_m$ = random effect of the rider (11 815 levels)
PE_n = random effect of the permanent environment of the horse, computed only for horses with competition results (20 126 levels)
fp_{PE} = random regression coefficient of the permanent environmental effect
$Animal_o$ = random effect of the additive genetic effect (74 261 levels)
fp_A = random regression coefficient of the additive genetic effect
$e_{ijklmno}$ = random error

Due to the small number of observations in some years, birth year of a horse was represented by seven categories (1971–1980, 1981–1985, 1986–1990, 1991–1995, 1996–2000, 2001–2005, 2006–2009). A total of 165 breeds were represented in the pedigree file, so compression into the following seven categories was deemed appropriate: Thoroughbred, Czech Warmblood, other Czech sport breeds (Slovak Warmblood, Kinsky horse, Moravian Warmblood), European breeds most influenced by Czech Warmblood (Hannoverian horse, Holsteiner horse, Oldenburger horse, Dutch Warmblood, Belgian Warmblood), ponies, coldblood, other warmblood horses.

Variability of traits in relation to age was fitted by functions (f) of orthogonal LPs:

$$f = p'b$$

where:

b = vector of regression coefficients
p = vector of parameters of the function (LPs)

Age standardization (as) was performed in LPs:

$$as = 2\,((x_i - x_{\min})/(x_{\max} - x_{\min})) - 1$$

where:

x_i = age at the day of event
$x_{\min}$ = minimum age (the youngest age at which a horse competed)
$x_{\max}$ = maximum age (the oldest age at which a horse competed)

After age standardization, the first four parameters (p) were (Pribyl et al. 2007):

$$p0 = 1$$
$$p1 = as\sqrt{3}$$
$$p2 = 0.5\,(3as2 - 1)\sqrt{5}$$
$$p3 = 0.5\,(5as3 - 3as)\sqrt{7}$$

The MT model. The multi-trait model was used when competitions were divided into three categories according to the height of obstacle (as a proxy for the level of difficulty). The first included jumping results with obstacles from 90 to 110 cm (TR1), the second from 120 to 135 cm (TR2), and the third from 140 to 155 cm (TR3). Numbers of observations for the estimation of genetic parameters and the prediction of breeding values were as follows: 183 848 and 379 708 for TR1, 88 170 and 176 317 for TR2, and 10 419 and 23 923 for TR3.

In this MT model, each horse's experience in competition is recorded throughout its entire jumping career. If a horse begins competing as four-year old, that is taken as its first year of experience. If another horse begins competing at seven years of age, it is its first year of experience and its eighth year of age is the second year of experience, etc.

$$Y_{ijklmno} = m + Sex_i + Age_j + Event_k + b_1fp + b_2fp^2 + Rider_m + PE_n + Animal_o + e_{ijklmno}$$

where:

$Y_{ijklmno}$ = measurement variable for the three competition result traits
m = population mean
Sex_i = fixed effect of gender (three levels)
Age_j = fixed effect for nine age groups (4, 5, 6, 7, 8, 9, 10, 11–15, 16–25)
$Event_k$ = fixed effect of date and location of the event (27 525 levels)
fp = the horse's competition experience in years
b_1 = linear regression coefficient
b_2 = quadratic regression coefficient
$Rider_m$ = random effect of the rider (11 815 levels)
PE_n = random permanent environmental effect of the horse (20 126 levels)
$Animal_o$ = random additive genetic effect (74 261 levels)
$e_{ijklmno}$ = random error

doi: 10.17221/8453-CJAS

Breeding values. Breeding values predicted using the RR model were used to create performance curves for specific stallions. Such curves depict change in predicted genetic merit across a horse's experience in competition, and the population average is subtracted from this value to put the breeding values at the same level.

Using the MT model, we obtained the relative breeding values (RBVs) as a reversion of the breeding values according to the following equation:

$$\mathrm{RBV} = \mathrm{m} + 1/((\mathrm{EBV}/\mathrm{SD}_{EBV}) \times \mathrm{SD}_{D})$$

where:

RBV = predicted relative breeding value on the common publication scale

m = mean of EBV equal to 100 for the population included in the BLUP method

EBV = breeding value on the original scale of the analyses

SD_{EBV} = genetic standard deviation of the EBV

SD_{D} = desired standard deviation fixed to 20 units

For the MT model, reliability of breeding value was calculated for all horses from the pedigree (74 261 animals) as well. This was based on the method described by Misztal and Wiggans (1988).

RESULTS AND DISCUSSION

Transformation of the measured variable. The basic statistical data on the jumping performance used for the computations were identical for both model equations (Table 1). Jumping performance expressed by penalty points was not normally distributed, due to high numbers of performances with zero, 4, 8 or 12 penalty points and to high numbers of unfinished or disqualified competitions (999 penalty points). During filtering, all records with more than 50 penalty points were merged and to exclude negative figures after Blom transformation, all values were increased by a constant (Table 2). Only Blom transformed penalty points were used for estimation of genetic parameters (Figure 1). Blom transformation is often used when data from competitions are available for breeding value prediction, as reported in the studies by Janssens et al. (1997), Reilly et al. (1998), Posta et al. (2009a,b), and Ricard and Legarra (2010).

Table 1. Numbers of observations in the BLUP method and after filtering for parameter estimation in the Gibbs method

	No. in BLUP	No. after filtering for Gibbs
Observations	579 948	282 437
Horses with sport records (PE)	20 126	6 209
Riders	11 815	5 201
Events	27 525	21 918
Horses in the pedigree	74 261	27 895

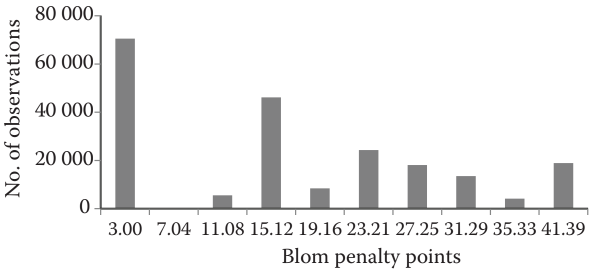

Figure 1. Distribution of the observations for Blom transformation of penalty points

Heritability. In the RR model, the length of a horse's experience in competition ranged from 1 to 6350 days, with a mean of 1159 days and standard deviation of 950 days. Heritability estimates ranged from 0.01 to 0.11, averaging 0.07. The highest value of 0.11 was by the length of experience of 1800 days. Proportion of variance attributed to the permanent environmental effect increased with length of a horse's competitive experience (Figure 2). From 4100 days onward (roughly 11 years of competitive participation), the curve quickly rose and had an inflated standard deviation, probably due to the small number of observations. Genetic variance at first increased, reaching the maximum at about 1800 days (i.e. 5 years of competition). Slightly higher heritabilities (0.03–0.30) were reported by Posta et al. (2010), who used a similar RR model but with regression based on the age of the horse.

Heritability estimates for the MT model were 0.07, 0.11, and 0.14 for TR1, TR2, and TR3, respectively (Table 3). The longest observed competitive career for a horse was 18 years (Figure 3). The increasing impact of the permanent environmental effect and the increasing level of difficulty of the competitions can be explained by the fact that all the horses competing at a high level of difficulty were well experienced and had extensive preparation and quality care, leading to reduced variation among the horses. Posta et al. (2010) also reported that the performance of show jumping horses later in their career was less influenced by

doi: 10.17221/8453-CJAS

Table 2. Characteristics of dependent variables for the estimation of genetic parameters

Dependent variable	*n*	Mean	Standard deviation	Minimum	Maximum
Penalty points	238 897	89.80	277.67	0	999.00
Blom penalty points	238 897	16.42	13.59	0.77	41.99

environmental effects than was the performance at the beginning of their career. The results of our work clearly show that heritability estimates increase with competition difficulty.

This increasing heritability may result because of the similar conditions among horses that are bred and trained for higher-level competitions; whereas horses that participate in lower-level competitions may not be as thoroughly trained, and their breeding and stabling conditions may have been less favourable. In consequence, systematic effects on the horse's performance would be less predictable in lower than in higher level competitions, leading to increased error variance and

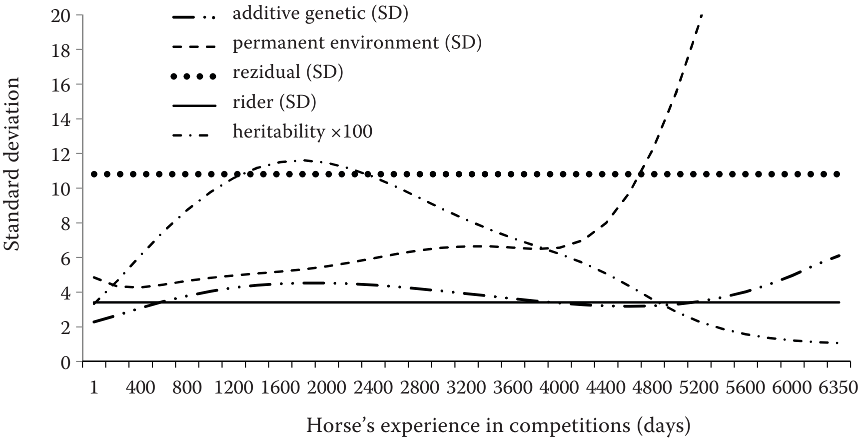

Figure 2. Variances and residual effects estimated with the random regression model

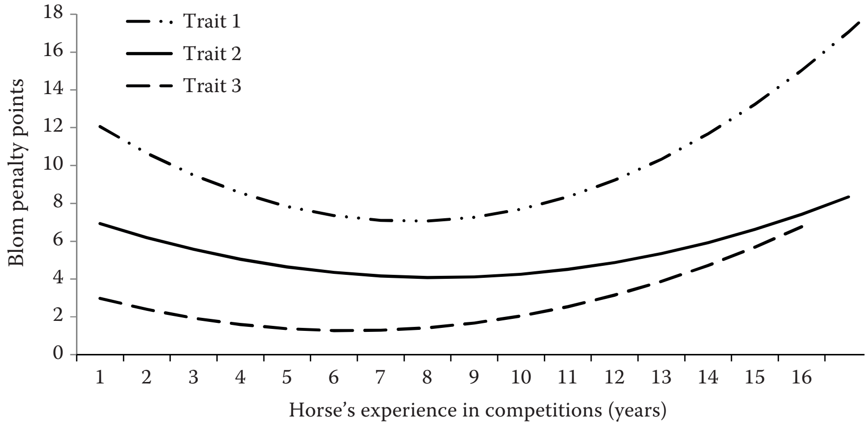

Figure 3. Blom penalty points for three traits reflecting height of obstacles and depending on a horse's experience in competition

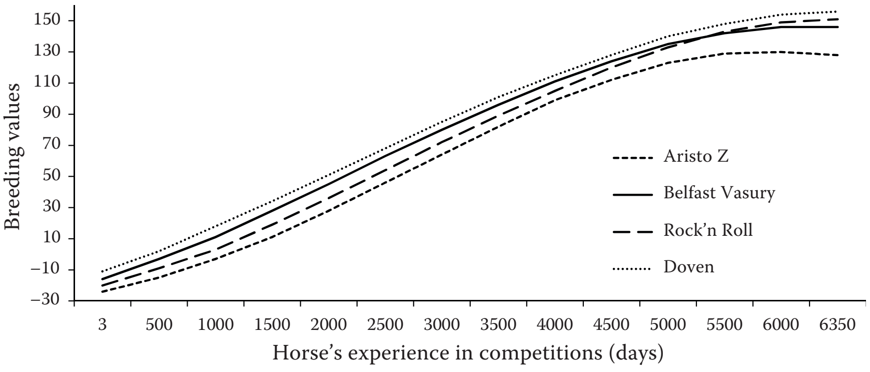

Figure 4. Breeding values for stallions estimated with the random regression model

Table 3. Percentages of total variance attributed to random effects in the analytical model, heritability estimates, and repeatabilities for sport horse jumping competition

Random effect	Random regression model (%)	Multitrait model (%)		
		trait 1	trait 2	trait 3
Rider	6.2	6.6	11.1	16.5
Permanent environment	23.8	10.1	11.0	15.5
Additive genetic effect	8.3	6.8	10.9	13.9
Residual	61.7	76.6	67.0	54.1
Total variance	100.0	100.0	100.0	100.0
Heritability	0.07	0.07	0.11	0.14
Repeatability	0.17	0.17	0.22	0.29

lower heritability estimates. In analyses also using the MT model, Alridge et al. (2011) and Viklund et al. (2011) also reported higher heritability estimates with the increasing difficulty of competition. Aldridge et al. (2000) reported that performance in low- and medium-level show jumping competitions could predict the genetic potential in high-level competitions; whereas Ducro (2011) concluded that performance at different levels of difficulty could be considered as different characteristics. Due to advantages of the MT model, it is worthwhile using it for genetic evaluation of sport horses in equestrian competitions. Table 4 shows the relative number of horses competing in the given trait according to the number of years of horse's experience. Horses having up to 10 years of competitive experience make up the largest proportion of the dataset, which we presume is due to cumulative injuries associated with longer competitive careers. As well, injuries are one of the possible reason for the retirement of a horse from its sports career.

Table 4. Distribution of animals in the multitrait model in the three traits by years of a horse's experience in competitions

Horse's experience in competitions (years)	Trait 1	Trait 2	Trait 3
	(%)		
1	23.1	5.6	1.6
2	19.4	14.6	3.4
3	15.3	16.2	10.2
4	11.8	15.5	18.4
5	8.9	13.1	17.9
6	6.4	10.8	14.2
7	4.7	7.8	11.2
8	3.5	5.7	9.0
9	2.6	4.1	6.0
10	1.7	2.7	3.9
11	1.1	1.7	1.7
12	0.6	1.1	1.2
13	0.4	0.6	0.6
14	0.3	0.3	0.3
15	0.1	0.2	0.3
16	0.1	0.1	0.1
17	0	0	0
18	0	0	0
Total	100	100	100

Correlations. Using the MT model, the correlation between the effect of age of a horse and the horse's experience was analyzed. If a high correlation between these effects was detected, then one could be excluded from the model. Correlations actually observed between these effects were between 0.17 and 0.41, so both of them were retained in the model.

Genetic correlations between traits from the MT model were high, 0.93 between TR1 and TR2, 0.95 between TR2 and TR3, and 0.81 that between TR1 and TR3.

Prediction of breeding values. The RR analyses provided results that can be used to create a

Table 5. Breeding values estimated using the random regression model for selected stallions in different days of a horse's experience in competitions

Days	Aristo Z	Belfast Vasury	Rock n'Roll	Doven
500	–15	–3	–9	2
1500	11	28	19	34
2000	28	45	36	51
3500	82	96	89	101
4500	112	124	120	128
6000	130	146	149	154

Table 6. Relative breeding values (RBV) and their reliability using the multitrait model for selected stallions

Name of stallion	Year of birth	Number of offspring	Trait 1		Trait 2		Trait 3	
			RBV	reliability	RBV	reliability	RBV	reliability
Aristo Z	2001	30	151	0.74	156	0.76	157	0.68
Belfast Vasury	1996	5	92	0.56	88	0.56	85	0.48
Rock n'Roll	1990	147	128	0.91	127	0.91	122	0.82
Doven	1997	8	55	0.66	54	0.64	55	0.55

weighted sum of a set of the continuous covariable, as described by Meyer (2005). Breeding values for the entire analyzed interval, estimated using the RR model, are shown for selected stallions in Figure 4. In this representation, the lower the curve for a stallion, the higher the predicted genetic merit. Curves for each stallion have a similar increasing trajectory, i.e. breeding values are negative early on, gradually increasing to high positive values. This may be due to the small number of observations from 4100 days (11 years) of competitive experience. Posta et al. (2010) used random regression on age of the horse and published curves of predicted breeding values at five advancing ages for the most popular stallions in Hungary at that time. They concluded that estimated breeding values for the stallions were relatively stable across time.

From the results of our work, it is not possible to confirm that the predicted breeding values of the stallions had stabilized. Differences in the breeding values at the beginning, during and at the end of a horse's competitive experience were observed (Table 5). This might be because long sport careers (6350 days) of only a few horses were examined; whereas only 2.1% of horses from the overall number in our database had 4100 or more days of competition, which may have caused the deviation of predicted breeding values.

Relative breeding values were calculated using only the MT model. The highest RBVs corresponded to the lowest Blom penalty points. The average reliability calculated for all horses in the pedigree (74 261 animals) were 0.22, 0.22, and 0.19 for TR1, TR2, and TR3, respectively with standard deviations of 0.22, 0.22, and 0.19. The minimum was 0 and the maximum was 0.94. The average reliability for a horse with a competitive record (20 126) was 0.46 with standard deviation of 0.11. Selected stallions with RBVs and reliabilities in all three traits are shown in Table 6.

CONCLUSION

Genetic parameters and breeding values for performance traits in sport horses were estimated by the BLUP-animal model method, and from the results we conclude that both model equations can be used. Each has advantages. In the RR model, we can more accurately account for differences among trainers than when using one overall indicator. This allows the option to select horses at a younger age. From the MT model, we conclude that predicted genetic merit of a horse can increase with the increasing level of difficulty in a competition. Hence the breeders should choose stallions based on TR3, because the aim of most sport horse breeding organizations is to produce quality horses for the highest levels of competition. Even though the heritability of show-jumping horse performance is low, breeding organizations in the Czech Republic should start evaluating breeding horses using the method BLUP-AM, as this is the only means to achieve genetic gain.

***Acknowledgement**.* This research work was done in close cooperation with the Association of the Czech and Slovak Warmblood Breeders. The authors express great appreciation to Dr. William D. Hohenboken for his helpful comments on the manuscript.

REFERENCES

Alridge L.I., Kelleher D.L, Reilly M., Brophy P.O. (2000): Estimation of the genetic correlation between performances at different levels of show jumping competitions in Ireland. Journal of Animal Breeding and Genetics, 117, 65–72.

Bugislaus A.E., Roehe R., Willms F., Kalm E. (2006): The use of random regression model to account for change in racing speed of German trotters with increasing age. Journal of Animal Breeding and Genetics, 123, 239–246.

doi: 10.17221/8453-CJAS

Buxadera A.M., Da Mota M.D.S. (2008): Variance components estimations for race performance of thoroughbred horses in Brazil by random regression model. Livestock Science, 117, 298–307.

Ducro B. (2011): Relevance of test information in horse breeding. PhD Diss. Wageningen, the Netherlands: Wageningen Univ. 170 p.

Janssens S., Geysen D., Vandepitte W. (1997): Genetic parameters for show jumping in Belgian sporthorses. In: Proc. 48th Annu. Mtg., European Association for Animal Production (EAAP), Vienna, Austria.

Luehrs-Behnke H., Roehe R., Kalm E. (2002): Genetic associations among traits of the new integrated breeding evaluation method used for selection of German warmblood horses. Veterinarija ir Zootechnika, 18, 90–93.

Meyer K. (2005): Random regression analyses using B-splines to model growth of Australian Angus cattle. Genetics Selection Evolution, 37, 473–500.

Misztal I., Wiggans G.R. (1988): Approximation of prediction error variance in large-scale animal models. Journal of Dairy Science, 71, 27–32.

Posta J., Komlosi I., Mihok S. (2009a): Breeding value estimation in the Hungarian Sport Horse population. Veterinary Journal, 181, 19–23.

Posta J., Mihok S., Markus S., Komlosi I. (2009b): Analysis of Hungarian sport horse show jumping results using different transformations and models. Archiv fur Tierzucht, 52, 451–458.

Posta J., Malovhr S., Mihok S, Komlosi I. (2010): Random regression model estimation of genetic parameters for show-jumping results of Hungarian Sporthorses. Journal of Animal Breeding and Genetics, 127, 280–288.

Pribyl J., Krejcova H., Pribylova J., Misztal I., Bohmanova J., Stipkova M. (2007): Trajectory of body weight of performance tested dual-purpose bulls. Czech Journal of Animal Science, 52, 315–324.

Reilly M., Foran M.K., Kelleher D.L., Flanagan M.J., Brophy P.O. (1998): Estimation of genetic value of show jumping horses from the ranking of all performances in all competitions. Journal of Animal Breeding and Genetics, 115, 17–25.

Ricard A., Legarra A. (2010): Validation of models for analysis of ranks in horse breeding evaluation. Genetics Selection Evolution, 42, 3.

Viklund A., Nasholm A., Strandberg E., Philipsson J. (2011): Genetic trends for performance of Swedish Warmblood horses. Livestock Science, 141, 113–122.

Received: 2014–10–28
Accepted after corrections: 2015–02–04

Corresponding Author

Ing. Alexandra Novotná, Institute of Animal Science, Přátelství 815, 104 01 Prague 10-Uhříněves, Czech Republic
Phone: +420 267 009 629, e-mail: novotna.alexandra@vuzv.cz