Random regression models for genetic evaluation of performance of the Hungarian show-jumping horse population

Anita Rudiné Mezei – János Posta – Sándor Mihók
University of Debrecen Faculty of Agricultural and Food Sciences and Environmental Management,
Department of Animal Breeding, Debrecen
mezeia@agr.unideb.hu

SUMMARY

The aim of the study was to estimate genetic parameters for show-jumping competition performance using random regression model. Show-jumping competition results collected between 1996 and 2009 were analyzed. The database contained 272 951 starts of 8020 horses. Identity number and gender of the horse, rider, competition date, the level of the competition and placing were recorded in the database. Competition levels were categorized into five groups. Weighted – competition level used – square root transformed placing was used to measure performance of horses. The random regression model included fixed effects for gender, year and place of competition, and random effects for rider, animal and permanent environment.

Later performance of show-jumping horses measured with weighted square root ranks is less influenced by rider and permanent environmental effects than performance at the beginning of a horse's sporting career. Heritability increased continuously from 6.3 years of age (2296 age in days), values were in the range of 0.07 and 0.37. Higher heritability was found in later ages. Weak genetic and phenotypic correlation was found between the early 4–5–6 years of age and older (7, 8, 8+) age classes. From 8.5 years of age (3132 days old) there were strong genetic and phenotypic correlations between neighboring age groups. For the same age classes moderate and strong genetic and phenotypic correlation was found. Genetic correlation between 13.5 years of age and older horses was very strong.

***Keywords:** random regression, sporthorse, performance measurement*

ÖSSZEFOGLALÁS

A tanulmány célja genetikai paraméterek becslése volt sportlovak ugrósportban nyújtott teljesítményei alapján random regressziós modellel. A vizsgálat anyagát az 1996 és 2009 közötti díjugratás szakági eredmények jelentették. Az adatbázis tartalmazta 8020 ló 272 951 startját, a lovak azonosítóját, ivarát, lovasát, a verseny pontos dátumát, nehézségi szintjét, helyszínét és a helyezést. A teljesítmény méréséhez a verseny nehézségi szintjével súlyozott négyzetgyök függvénnyel transzformált helyezéseket használtuk fel. A nehézségi kategóriák öt csoportba voltak besorolva. A random regressziós modellben fix hatásként vettük figyelembe a ló ivarát, a verseny évét, helyszínét, véletlen hatásként szerepelt a lovas, a ló és az állandó környezeti hatás a modellben.

A lovak teljesítményét az életkor előrehaladtával egyre kisebb mértékben befolyásolta a lovas és az állandó környezet hatása. A becsült örökölhetőségi értékek 0,07 és 0,37 közöttiek voltak. Az életkor előrehaladtával 6,3 éves kortól kezdve a teljesítmény örökölhetősége a vizsgált mérőszámra folyamatosan nőtt. A fiatal 4–5–6 éves lovak és az idősebb (7, 8 éves, 8 évesnél idősebb) lovak teljesítménye között alacsony genetikai és fenotípusos korrelációkat becsültünk. A lovak 8,5 éves életkor felett szoros genetikai és fenotípusos korrelációkat becsültünk a szomszédos korcsoportok között. A genetikai korreláció a 13,5 éves és idősebb lovak között nagyon szoros volt.

***Kulcsszavak:** díjugratás, sportló, teljesítmény mérőszám*

INTRODUCTION

Sporthorse breeder's interest is focused on changes in performance within individual horses over their career. Changes can be observed not only in means, but there can be heterogeneity in variances and covariances at different ages. The way how to investigate changes in variances and covariances is by defining appropriate covariance functions (Kirkpatrick et al., 1990). Covariance functions are continuous functions that should make possible to estimate covariance within animals between measures at any two age t_i and t_j, and variances between animals at any age t_i.

To estimate covariance functions an alternative is using random coefficient regression model. Thus the regression coefficients are assumed to be random variables and there is a covariance structure among regression coefficients. These can be used to estimate covariance functions, i.e. variances and covariances between values for different ages.

Random regression models (RRM) are intended for use on longitudinal data or 'repeated' records situations where observations for a trait are collected several times during the course of an animal's life (Hill and Brotherstone, 1999).

From the point of view of animal breeding, interest lies in the genetic parameters that describe how traits change in time, because they reflect to what extent and how genetic changes in performance patterns over time can be achieved by selection (Huisman et al., 2002), thus providing greater information for breeders to select those horses which should be chosen as sires for the following generation.

During the last 15 years this type of longitudinal data has been analyzed by random regression models on many economical traits at different animal breeding scenarios (Schaeffer, 2004).

Random regression model have become a popular choice for modelling traits, which are measured repeatedly per individual, but change gradually and continually with time, because of improved modelling of variances and genetic parameters (Meyer, 2004). RRM provides higher volume of information than the repeatability animal model (Posta et al., 2010). The use of random regression model is highly recommended as a very useful tool, because it allows us to estimate the animal's breeding value through all the stages of its trajectory in racing competitions.

In the last decade some studies used this methodology in horses (Kaps et al., 2004; Bugislaus et al., 2006; Menendez-Buxadera and Mota, 2008; Posta et al., 2010; Gómez, 2011).

Kaps et al. (2004) investigated changes in variances and covariances of withers height of Lipizzan horses from birth to 36 months of age. Bugislaus et al. (2006) analysed the trajectory of racing time per kilometre by age (6 classes: 2, 3, 4, 5, 6 and over 6 year olds) in German Trotter Horses; whereas Menendez-Buxadera and Mota (2008) analysed the complete race time (in seconds) by distance (7 classes: 1000–1600 m) in Brazilian Thoroughbreds. Posta et al. (2010) estimated genetic parameters for show-jumping performance using shifted Blom score and difference between fence height and fault point as performance measurement variable. Gómez et al. (2011) have analyzed the changes in the trajectory of accumulated earnings for Spanish Trotter Horses during their sporting life.

The aim of the study was to estimate genetic parameters for performance of Hungarian show jumping horses during their career using random regression model.

MATERIAL AND METHODS

Show-jumping competition results collected between 1996 and 2009 were analyzed. The data used in this study were obtained from the Hungarian Equestrian Federation. The final dataset contained in total 272 951 competition records on 8020 individual horses after data screening, results were gathered from Hungary and other European countries. Identity number and gender of the horse, rider, year, level and location of the competition and placing were recorded in the dataset. Information about pedigree of horses were gathered and set up with help of the National Horse Breeder Information System. The pedigree file contained 40 142 animals four generation back.

Competitions were categorized into five groups based on their difficulty level. For the evaluation of show-jumping performance weighted square root ranks was used. Analysis before showed that the best goodness-of-fit value were in case of weighted square root (Rudiné et al., 2013). Sporthorses were divided into 6 age groups, 4, 5, 6, 7, 8 and more than 8 years old.

To estimate variances and covariances of regression coefficients Legendre polynomials were used (Kirkpatrick et al., 1990). The random regression coefficients for additive genetic and permanent environmental effects were fitted with first to fifth (n=1, 2, 3, 4, 5) orders of Legendre polynomial for the trait. The general form of a Legendre polynomial is:

$$\phi_n(t) = \frac{1}{2^n}\sqrt{\frac{2n+1}{2}}\sum_{k=0}^{\left[\frac{n}{2}\right]}(-1)^k\binom{n}{k}\binom{2n-2k}{n}q_t^{n-2k}$$

$$q_t = -1 + 2\frac{t - t_{min}}{t_{max} - t_{min}}$$

where n is the degree of the Legendre polynomial, k=0,..,n; q_t is the standardized time converted to a number between -1 and 1.

To estimate the trajectory of the variance components the following random regression model was used applying the VCE-6 program developed by Kovac and Groeneveld (2003):

$$Y_{jklmno}(t) = Gender_j + Year_k + Place_l + Rider_m + \sum_{n=1}^{p} t^n + \sum_{n=1}^{p}\alpha_o\phi_n(q_t) + \sum_{n=1}^{p}\gamma_o\phi_n(q_t) + e_{jklmno}$$

where:

Y_{jklmno} = a vector of observations for the particular trait of a specific horse,
$Gender_j$ = fixed effect of gender (3 levels),
$Year_k$ = fixed effect of year of competition (14 levels),
$Place_l$ = fixed effect of place of competition (443 levels),
$Rider_m$ = random effects of the rider (3352 levels),
t = age in days,
$Ø_n$ = the Legendre polynomial,
α_o = random regression coefficients of additive genetic effect,
γ_o = random regression coefficients of permanent environmental effect,
e_{jklmno} = random residual effect.

The regression coefficients are assumed to be random variables. Regressions were split additive genetic effect regression and permanent environment regression due to a horse. Gender of the horse, year, place of the competition were defined as fixed effect.

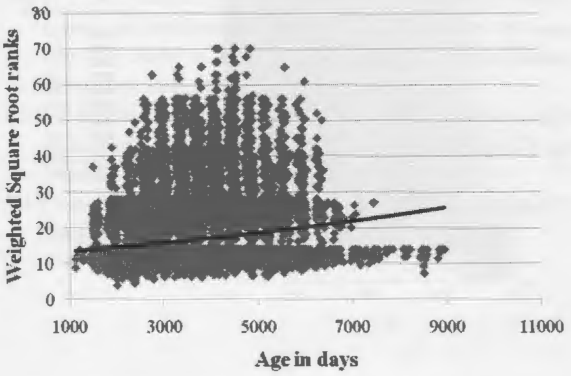
The eigenfunctions and eigenvalues were calculated from the (co)variance components of the random regression coefficients with SAS PROC IML procedure (SAS Institute Inc., 1999).

RESULTS

All fixed effects were statistically significant in the model as determined by SAS PROC GLM (SAS Institute Inc., 1999).

Figure 1 shows the distribution of mean weighted square root ranks on age in days. Mean performance increased with increasing age. Younger 4, 5, and 6 years old horses (1090–2554 days old) had weaker performance than older competitors. Better performance, talented horses are revealed from 7 years of age. The largest values, consequently the best performance was reached between 11 and 13 years of age (4100–4900 ages in days).

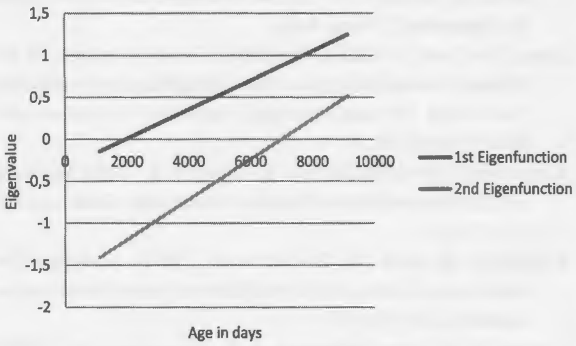
Figure 1: **Relationship between weighted square root ranks and age**



After examining the estimated variances and variance proportions for all models, the best fitting random regression model was the model with a polynomial of first order of fit (LP 1) for all random effects.

Estimates of the first eigenfunction for weighted square root ranks were positive from age 6.08 (2219 age in days) *(Figure 2)* indicating that selection on the sport performance traits could be successful from 2219 days old. In other words selection on this trait at any age from 6.08 should change show-jumping performance in the same direction. For weighted square root ranks the first eigenvalue explained 99.56% of the genetic variation, and the second eigenfunction was responsible for only 0.44% of the total genetic variation for the first ordered Legendre polynomial. Variances explained by eigenfunctions of second order Legendre polynomials were lower, so the first order LP 1 was the best fitting.

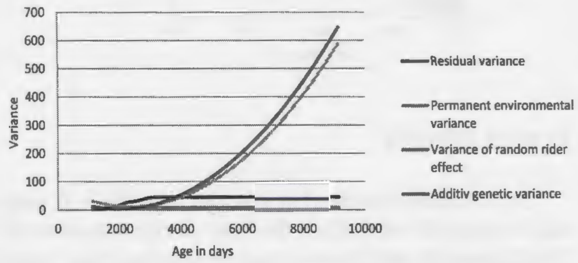
Figure 2: **Eigenfunctions of weighted square root ranks for the random animal effect estimated with random regression model**



The low values of the second eigenfunctions suggest that it is sufficient to focus only on the first eigenfunction. Eigenfunctions, estimated from the eigenvectors of the genetic (co)variance matrix (Kirkpatrick et al., 1990), provide an insight into the effects of selection across the growth trajectory. Eigenvalues indicate the amount of variance explained by its associated eigenfunction (Kingsolver et al., 2001).

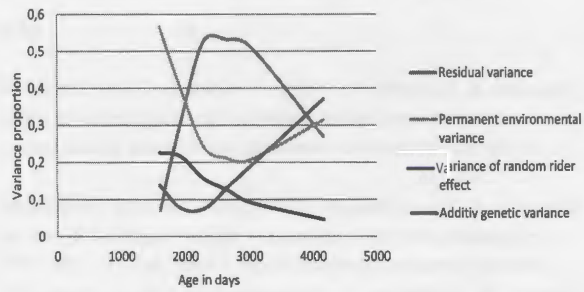
The variance of random animal effect (additive genetic variance) decreased in the early age *(Figure 3)*. The lowest genetic variance was estimated in the age of between 1854 and 2219 days (5–6 years old horses). In later age the genetic variance continuously increased. Homogeneous variance 7.63 was estimated for the random effect of rider. Variance of permanent environmental effect continuously increased. Variance of residual effect was estimated within age groups, it increased up to 9 years of age. From 9 years old (3314 age in days) residual variance 44.98 was constant.

Figure 3: **Variances for random and residual effects estimated with random regression model (LP 1)**



The heritability values were in the range of 0.07 and 0.37 *(Figure 4)*, it increased continuously after 6.3 years of age (2296 age in days). Variance proportion of permanent environmental effect decreased continuously up to 8.3 years of age than it increased slightly over time. Proportion of residual variance increased up to 6 years of age and afterwards its effect was lower and lower over time. Proportion of variance for random effect of rider continuously decreased.

Figure 4: **Variance proportions for random and residual effects estimated with random regression model (LP 1)**

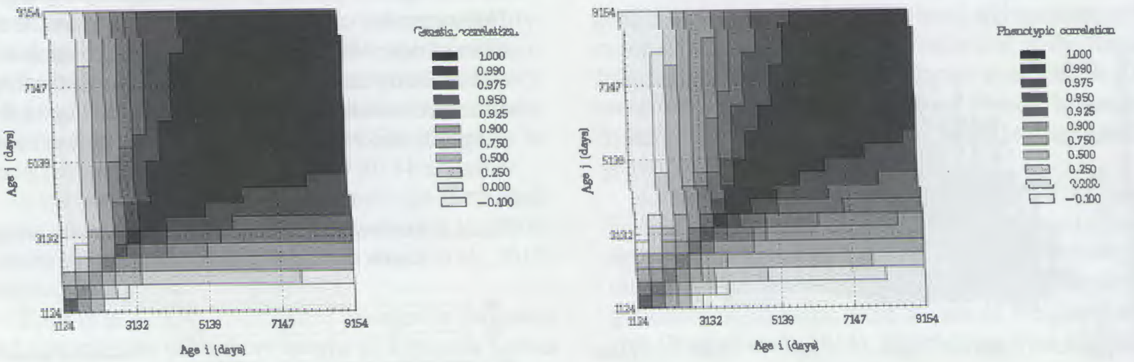


Heritabilities based on square root of placing were similar to those of Huizinga and van der Meij (1989) h^2=0.10–0.20; Koenen et al. (1995) h^2=0.17; Hassenstein et al. (1998) h^2=0.07–0.11; Luehrs Behnke et al. (2002) h^2=0.11; Jaitner and Reinhardt (2003) h^2=0.10; Sobczynska and Lukaszewicz (2004) h^2=0.15; and Viklund et al. (2011) h^2=0.11. Posta et al. (2010) estimated 0.037 and 0.385 heritabilities of shifted Blom score.

From 8.5 years of age (3132 days old) there were strong genetic and phenotypic correlations between neighboring age groups *(Figure 5)*.

For the same age classes moderate and strong genetic and phenotypic correlation was found. Weak genetic and phenotypic correlation was found between the early 4–5–6 years of age and older (7, 8, 8+) age classes. Genetic correlation between 13.5 years of age and older horses was very strong.

Figure 5: **Genetic and phenotypic correlations between performances in different age classes**



CONCLUSIONS

Performance reached at younger age (4–5–6 years old horses) do not reflect the later performance at all. Weak genetic and phenotypic correlation was found between the early 4–5–6 years of age and older (7, 8, 8+) age classes. From 6 years of age selection could be successful on weighted square root ranks. Later performance of show-jumping horses measured with weighted square root ranks is less influenced by rider and permanent environmental effects than performance at the beginning of a horse's sporting career. Heritabilities increased continuously from 6.3 years of age (2296 age in days), values were in the range of 0.07 and 0.37. Higher heritabilities were found in later ages. From 8.5 years of age (3132 days old) there were strong genetic and phenotypic correlations between neighboring age groups. For the same age classes moderate and strong genetic and phenotypic correlation was found. Genetic correlation between 13.5 years of age and older horses was very strong.

ACKNOWLEDGEMENTS

This work was made possible by the financial support of the „OTKA-PD83885" research projects. This work supported by the TÁMOP-4.2.2/B-10/1-2010-0024 project. The project is co-financed by the European Union and the European Social Fund. The Association of Hungarian Horse Breeders and Horse Organization, and Discipline of Show-jumping of The Hungarian Equestrian Federation are gratefully acknowledged for providing the data set for the study.

REFERENCES

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. J. Anim. Breed. Genet. 123: 239–246.

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

Gómez, M. D.–Molina, A.–Menendez-Buxadera, A.–Valera, M. (2011): Estimation of genetic parameters for the annual earnings at different race distances in young and adult Trotter Horse using a Random Regression Model. Livest. Sci. 137: 87–94.

Hassenstein, C.–Roehe, R.–Kalm, E. (1998): Estimation of genetic parameters of German sport horses accounting for competition in the statistical model. Proc. of the 6th World Congr. Genet. Appl. Livest. Prod. Armidale. Australia. 11–16. January. 24: 436–439.

Hill, W. G.–Brotherstone, S. (1999): Advances in methodology for utilizing sequential records. British Soc. of Animal Science. 24: 55–61.

Huisman, A. E.–Veerkamp, R. F.–van Arendonk, J. A. (2002): Genetic parameters for various random regression models to describe the weight data of pigs. J. Anim. Sci. 80: 575–582.

Huizinga, H. A.–van der Meij, G. J. W. (1989): Estimated parameters of performance in jumping and dressage competitions of the Dutch Warmblood horse. Liv. Prod. Sci. 21: 333–345.

Jaitner, J.–Reinhardt, F. (2003): National genetic evaluation for horses is Germany. 54th Annual Meeting of the EAAP. August 31–September 3. Roma. Italy.

Kaps, M.–Curik, I.–Baban, M. (2004): Genetic Analysis of the Withers Growth In Lipizzan Horse Using Random Regression. Proc. of the 55th Ann. Meeting of the EAAP. 5–9 September. Bled. Slovenia. H2.4.

Kingsolver, J. G.–Gomulkiewicz, R.–Carter, P. A. (2001) Variation, selection and evolution of function-valued traits. Genetica. 112–113: 87–104.

Kirkpatrick, M.–Lofsvold, D.–Bulmer, M. (1990): Analysis of the inheritance, selection and evolution of growth trajectories. Genetics. 124: 979–993.

Koenen, E. P. C.–van Veldhuizen, A. E.–Brascamp, E. W. (1995): Genetic parameters of linear scored conformation traits and their relation to dressage and show-jumping performance in the Dutch Warmblood riding horse population. Liv. Prod. Sci. 43: 85–94.

Kovac, M.–Groeneveld, E. (2003): VCE-5 User's Guide and Reference Manual Version 5.1. Institute of Animal Science Federal Agricultural Research Center (FAL). Neustadt. Germany.

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. 40: 90–93.

Meyer, K. (2005): Estimates of genetic covariance functions for growth of Angus cattle. J. Anim. Breed. Genet. 122: 73–85.

Posta, J.–Malovhr, S.–Mihók, S.–Komlósi, I. (2010): Random regression model estimation of genetic parameters for show-jumping results of Hungarian Sporthorses. J. Anim. Breed. Genet. 127: 280–288.

Rudiné Mezei, A.–Posta, J.–Mihók, S. (2013): Evaluation of Hungarian show-jumping results using different measurement variables. Acta Agraria Debreceniensis. 53: 81–85.

SAS Institute Inc. (1999): SAS / STAT Software Release 8.2. SPS. Cary. NC. USA.

Schaeffer, L. R. (2004): Application of random regression models in animal breeding. Livest. Prod. Sci. 86: 35–45.

Sobczynska, M.–Lukaszewicz, M. (2004): Genetic parameters of racing merit of thoroughbred horses in Poland. J. Anim. Breed. Genet. 121: 302–306.

Viklund, A.–Furre, S.–Philipsson, J.–Vangen, O. (2011): Nordic Interstallion – Competition data. Workshop on "Linear scoring in evaluation of sport horses – experiences of current practices and potential developments". 15–16 December. Stockholm. Sweden.