

1    **56<sup>th</sup> Annual Meeting of the European Association for Animal Production (EAAP) June**  
2                               **5-8, 2005, Horse commission H1-26, Uppsala, Sweden**

3    **THE ESTIMATION OF BREEDING VALUES OF ENGLISH THOROUGHBREDS IN**  
4                               **THE CZECH REPUBLIC**

5  
6                               Sonia SVOBODOVA<sup>a</sup>, Christine BLOUIN<sup>b</sup> Bertrand LANGLOIS<sup>b\*</sup>,

7  
8                               <sup>a</sup>Faculty of Agronomy, Animal Breeding Institut, University of Agriculture and Forestry,  
9                               Gregor Mendel, 613 00 Brno, Czech Republic

10                              <sup>b</sup>Station de Génétique Quantitative et Appliquée, 78352 Jouy-en-Josas Cedex, France

11  
12                              \* Corresponding author: Bertrand.Langlois@jouy.inra.fr

13  
14    **Abstract** -The aim of this study was to estimate the breeding value of English Thoroughbreds  
15    in the Czech Republic using racing results from a 22-year period (1980-2001). The data  
16    includes the performance of two and three-year-old horses which raced in flat races at  
17    hippodromes in the Czech Republic. The racing results (30 203) were available corresponding  
18    to 6 333 horses descending from 762 sires and 2 836 dams. Different criteria were applied to  
19    analyse the performance: Log of earnings per race, a normalized ranking value, distance of  
20    the race when placed, earnings and number of starts for 2, 3, 2+3 year-old horses. After  
21    preliminary studies a year effect or a sex by year effect was finally retained. Variance  
22    component estimation using VCE software gave the following values for heritability  
23    (±standard errors): 0.14±0.01 and 0.16±0.01 for the Log of earnings per race and the ranking  
24    value. Repeatability was 0.31 and 0.35, respectively. The maternal environment component  
25    was evaluated as 0.02±0.004 for the Log of earnings per race and 0.03±0.004 for the ranking

value. We found, that the Log of earnings per race and the ranking value were two appropriate criteria when taking into account racing performance in selection for Thoroughbred in the Czech Republic. The genetic correlation of the two criteria was  $0.98 \pm 0.003$ . The heritability for the distance when placed was  $0.18 \pm 0.01$ . Genetic correlation of Log of earnings per race and distance was medium,  $0.38 \pm 0.05$  and of the same order,  $0.39 \pm 0.05$  for ranking value and distance. In the case where we used the Log of annual earnings and number of starts, the heritability were for Log of earnings:  $0.15 \pm 0.03$  for two-year-old,  $0.34 \pm 0.03$  for three-year-old and  $0.32 \pm 0.03$  for two and three-year-old career together and respectively,  $0.12 \pm 0.03$ ,  $0.21 \pm 0.03$  and  $0.20 \pm 0.02$  for number of starts. The genetic correlation between earnings and number of starts were respectively:  $0.26 \pm 0.14$ ,  $0.33 \pm 0.06$  and  $0.19 \pm 0.07$ . Restricted to horses earning money the two consecutive years, genetic correlation between number of starts for two and three year-olds was medium;  $0.35 \pm 0.05$  and between earnings for the same ages it was high  $0.80 \pm 0.04$ .

**Key words:** Thoroughbred / flat races / estimation of breeding value / BLUP-animal model / Czech Republic

## 1. INTRODUCTION

1. The history of the Czech turf dates back to 1839, when was established the Czech racing company for Bohemia and Moravia was established. Under the leadership of this company the first public race took place in Prague race on 9 October 1839, but the first official racing days were held on 12 and 15 October 1839. In 1906 the first hippodrome Velka Chuchle was opened on the outskirts of the capital city Prague. Velka Chuchle became the central place for gallop races. The first Czechoslovak Derby was held here in 1921 and was called "Prix du Czechoslovak Jockey Club". It was a sign of inclination of the new republic to France after World War I, when the Derby is called "Prix du Jockey Club". From 1922 Derby was called the Czechoslovak Derby and at present it is the Czech Derby. World War II had very bad effect on the turf in Czech country. In 1948 the Jockey Club was abolished and re-opened after 1989 [1]. At present, a number of important Czech and international races are held in

the Czech Republic. The Czech turf has its place in the breeding and selection of Thoroughbred in Europe.

The performance of Thoroughbred horses results from a long-term selection for maximum gallop speed and now for the ability to win in races [9]. The selection of Thoroughbred horses is implemented using a system of races acting as a performance test. Results of these races are used for a comparison of inter-generation and intra-generation performances of Thoroughbreds [10].

The criteria used to estimate the racing ability are timing, handicap weight, handicap lengths as for the performance rate [6, 11] and earnings [5, 8, 9]. The heritability of the criteria of handicap weight, performance rate and earnings ( $0.30 < h^2 < 0.40$ ) is substantially higher than that of the parameters derived from timing ( $h^2 < 0.20$ ) [6, 7]. Misar, Jiskrova, Pribyl [10] used for estimation of the breeding value the criterion of General handicap (Gh – weight in kg) and an index of performance based on earnings (IDP – earnings divided by the mean value of horses of the same year, age, sex category), because these performance characteristics have high or medium coefficients of heritability at the opposite of racing times which heritability and repeatability are low [4].

In the present study of the results of Czech races for breeding purposes, for the estimation of breeding value we used the criteria of earnings at different levels; per race, per year and for the career, rank at finish in the race, distance of the race when placed and number of starts for 2, 3 and 2+3 years old.

A number of propositions for routine estimations of breeding value of Thoroughbred in the Czech Republic were also explored.

## **2. MATERIAL AND METHODS**

### **2.1. Setting up the database**

The racing results and pedigree of Thoroughbred were collected from Year-books (1980-2001) and from Stud Books of Czech Republic. Results of two and three-year-old horses were used to set up the data base. The data consisted of 30 203 racing results of 6 333 thoroughbreds descending from 762 sires and 2 836 dams. The following information was taken into account: name of the horse year of birth, sex, number of races, year of race, distance, category of race. For each horse in each race we also recorded the rank at finish and the corresponding earnings. The breeder, the trainer, the rider and its category and the category of race were also recorded; included was also information about the pedigree of the horses over two generations.

From this basic information some other synthetic variables could also be calculated, e.g. the annual earnings for two or three-year-olds or career earnings. The same was done for number of starts.

## **2.2. Criteria for the measurement of performance**

To estimate the breeding value we used the following criteria: earnings per race, rank at finish, distance of the race when placed, earnings at 2, 3 and 2+3 years and the corresponding number of starts. For the calculations some transformations had to be done:

Earnings were normalized by a Log transformation (Log of earning). At the race level “non-real earnings” were calculated for non-placed horses by multiplying earnings by 0.5 for each increasing rank as done by Chico [4]. After normalization by a Log transformation of these “non-real earnings”, all the non-placed horses were equalized and received the same value, which is the mean of the Log of the “non-real earnings” of non-placed horses.

Rank was transformed using a Normal score, which can be found in statistical tables (e.g. the Normal standard deviation expectation of rank  $k$  out of  $N$  individuals). As for earnings non-ranked horses were equalized and received the same value.

We considered these performance criteria at three levels: The level of the race, the level of the year and the level of the career:

Level of race was considered for evaluation of genetic values of Log of earnings, rank and distance when placed.

Level of year was considered for annual earnings and number of starts for Two and Three-year-old horses.

Level of career (2+3) was followed in all cases, when the horses had performances for at least one year.

For the Normal scores, on the contrary of earnings, where the level of the race is taken into account by the amount of money distributed, no differences are made between the races because by construction the mean score for a race equals zero. We therefore introduced a pre-correction for the effect of the race as done by Belhajyahia et al. [2]. The score  $S_{ij}$  of horse  $i$  in race  $j$  is considered to be influenced by two effects, that of the race  $r_j$  and that of the horse  $h_i$ :

$$S_{ij} = h_i - r_j + e_{ij}$$

Where  $e_{ij}$  is a random residual and  $h_i$  and  $r_j$  are considered as fixed as it was implicitly supposed in the original performance rate [6, 11]. In contrast we treated the horse effect as random to take into account the degree of repeatability of the horse's performance. We therefore proposed this kind of pre-correction of data for the race level leading to the so called ranking value. Different runs coupled with variance component estimation (see further) were implemented until the stabilisation of the repeatability and consecutively of the ranking values  $R_{ij} = S_{ij} + r_j$  was achieved.  $R_{ij}$  acts as the measure of the performance of horse  $i$  in the race  $j$ .

### 2.3. Genetic analysis

In first analyses, fixed effects were studied using the GLM programme package SAS [12] with or without a random horse effect. For estimation of breeding value of performance of the

Thoroughbreds BLUP–Animal model was used. The following animal model was fitted by using VCE and PEST software [7]. As many assumed fixed effects were not estimable in reality, in our genetic analysis we only considered the effect of the year and sex. To avoid problems with interaction a Sex by Year effect was considered. Effect of age was not really estimable, because 2 and 3 year-olds are running separately. We prefer therefore to run separate analyses per age classes than a general one on the whole data fitting a model with an age effect. The effects of the animal additive genetic value and that of the specific environment to an animal and the effect of a common environment to the progenies of the same mare were considered as random effects. In some analysis, the maternal effect was not used for comparison of results with and without this effect. In some cases the equation for the specific environment to an animal due to the missing of the notion of repeatability of the performance (horses with one performance – one number of starts, one year of racing) also disappeared from the model.

The following equation provides the model for the most complex situation, and can sometimes be simplified:

$$\mathbf{y} = \mathbf{X} \mathbf{b} + \mathbf{Z} \mathbf{g} + \mathbf{W} \mathbf{m} + \mathbf{Z} \mathbf{p} + \mathbf{e}$$

Where:

y = vector of observations (Log of earnings, ranking value, distance when placed, number of starts)

b = vector of fixed effects (sex by year effect)

g = vector of additive genetic values (parentage)

m = vector of maternal environmental effects (used in some analysis)

p = vector of the specific environment to an animal (in the analysis with repeatability of performance)

1  $e$  = vector of errors

2  $X, Y, Z$  = the incidence matrices

3

4 The expectations of this linear model are:

5 
$$E \begin{bmatrix} y \\ g \\ m \\ p \\ e \end{bmatrix} = \begin{bmatrix} Xb \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

6

7 The variance covariance matrix is:

8 
$$V \begin{bmatrix} g \\ m \\ p \\ e \end{bmatrix} = \begin{bmatrix} A\mathbf{s}^2_g & 0 & 0 & 0 \\ 0 & I\mathbf{s}^2_m & 0 & 0 \\ 0 & 0 & I\mathbf{s}^2_p & 0 \\ 0 & 0 & 0 & I\mathbf{s}^2_e \end{bmatrix}$$

9

10  $s^2_g = h^2 s^2_y$

11  $s^2_m = \mu s^2_y$

12  $s^2_p = (r - \mu - h^2) s^2_y$

13  $s^2_e = (1 - r) s^2_y$

14

15 Where:

16  $A$  = relationship matrix;  $I$  = identity matrix;  $h^2$  = heritability;  $\mu$  = maternal environment

17 component of variance in %;  $r$  = repeatability

18

19

20

### **3. RESULTS**

#### **3.1. Estimation of fixed effects.**

##### **3.1.1 Effect of age, year, sex, category of race, category of rider.**

To the model we tried to assign the fixed effect of age, year, sex, category of race and category of rider. The GLM results of these fixed effects were found significant in models excluding the random effect of the horse. Including this random effect of the horse, most of these effects were not estimable, with the exception of sex and year. Most of these effects which were significant in simplified models could not be incorporated in the more complete genetic model, where they were not estimable, because they confounded with the effect of the horse.

##### **3.1.2. Effect of sex by year (BLUE) on annual earnings**

Effect of sex by year: This combination of both effects was significant in all cases ( $P < 0.0001$ ), involving yearly or career criteria. For criteria per race, year and sex effects were not significant, except the effect of year remaining alone for Log of earnings per race. In the other cases, the sex by year combination allowed to avoid any interactions resulting mainly in yearly changes in the policy of allocating money to males and females.

#### **3.2. Estimation of components of variance**

##### **3.2.1. Log of earning per race and ranking value (for all starters)**



Table 1 – Variance component estimation in % for Log of earnings per race and for the ranking value (all starting horses).

Traits	Log. of earnings per race	Ranking value
Residual	0.692(±0.05)	0.828(±0.001)
Maternal environment	<b>0.023(±0.04)</b>	0.655(±0.005)
		<b>0.028(±0.004)</b>
Permanent horse environment	0.141(±0.09)	0.982(±0.013)
		<b>0.155(±0.009)</b>
Additive genetic value = $h^2$ (diagonal)	<b>0.144(±0.010)</b>	<b>0.980(±0.003)</b>
= $r_g$ (above the diagonal)		<b>0.162(±0.011)</b>
Repeatability	<b>0.308(±0.014)</b>	<b>0.347(±0.014)</b>

The Table.1 shows the results of evaluation of genetic parameters for Log of earning per race and the corresponding ranking value for two and three-year-old horses. Maternal environmental effect, was lower than 3%. The genetic correlation of the two criteria was very high:  $0.98 \pm 0.003$ . The repeatability for Log of earning was  $0.31 \pm 0.02$  and for ranking value  $0.35 \pm 0.02$ . Heritability was estimated as  $0.14 \pm 0.01$  and  $0.16 \pm 0.01$  respectively.

**3.2.2. Log of earnings per race+ ranking value + distance when placed (placed horses only).**

Table 2 – Variance component estimation in % for Log of earnings per race, for the ranking value and for the distance when placed (restricted to horses placed).

Traits	Log. of earnings per race	Ranking value	Distance when placed
Residual	0.601(±0.007)	0.770(±0.003)	0.186(±0.006)
		0.677(±0.007)	0.008(±0.006)
			0.717(±0.007)
Permanent horse environment	0.207(±0.008)	0.888(±0.010)	-0.346(±0.064)
		0.149(±0.008)	-0.189(±0.066)
			0.100(±0.010)
Additive genetic value	<b>0.192(±0.008)</b>	<b>0.945(±0.008)</b>	<b>0.383(±0.049)</b>
= $h^2$ (diagonal)		<b>0.174(±0.008)</b>	<b>0.394(±0.050)</b>
= $r_g$ (above the diagonal)			<b>0.183(±0.012)</b>
Repeatability	<b>0.399( ±0.016)</b>	<b>0.323(±0.016)</b>	<b>0.283 (±0.022)</b>

The heritability of parameters for two and three-year-old horses was evaluated (Table 2) without the maternal environmental effect. The genetic correlations of Log of earning and distance when placed were medium and very similar. The genetic correlation of Log of earning and distance was  $0.38 \pm 0.05$  and  $0.39 \pm 0.05$  for ranking value and distance respectively. However, the genetic correlation of Log of earning and ranking value was still very high:  $0.95 \pm 0.01$ . The heritability of the distance when the horse is placed is medium:  $0.18 \pm 0.01$ . The repeatability  $0.28 \pm 0.02$  is low but allows breeding value evaluation. The selection of data and not taking into account the low maternal effect did not change the repeatability of earning  $0.40 \pm 0.02$  and ranking value  $0.32 \pm 0.02$  very much. However,

heritability of earning is slightly increased  $0.19 \pm 0.01$  and remained nearly the same for the ranking value  $0.17 \pm 0.01$ .

### 3.2.3 Log of earning per year or career and corresponding number of starts.

Table 3 – Variance component estimation in % for the Log of annual earnings and for the number of starts for two-year-olds.

Traits	Log. of annual earnings	Number of starts
Residual	$0.826(\pm 0.029)$	$0.487(\pm 0.019)$ $0.875(\pm 0.025)$
Maternal environment	<b><math>0.027(\pm 0.027)</math></b>	$-0.879(\pm 1.798)$ <b><math>0.007(\pm 0.007)</math></b>
Additive genetic value = $h^2$ (diagonal) = $r_g$ (above the diagonal)	<b><math>0.147(\pm 0.029)</math></b>	<b><math>0.258(\pm 0.144)</math></b> <b><math>0.118(\pm 0.027)</math></b>

The genetic parameters were estimated in group of 2, 3 and 2 + 3-year-old horses. Tables 3 to 6 show the results of these estimations. The genetic correlation between earnings and number of starts was low and not so well estimated. Standard errors of genetic correlations were very high, from 6% to 14%. The effect of maternal environment ranged between 2% and 4% for the Log of earnings and between 1 and 6% for the number of starts. It is not very important.

The genetic parameters were estimated in group of 2, 3 and 2 + 3-year-old horses. Tables 3 to 6 show the results of these estimations. The genetic correlation between earnings and number of starts was low and not so well estimated. Standard errors of genetic correlations were very high, from 6% to 14%. The effect of maternal environment ranged between 2% and 4% for the Log of earnings and between 1 and 6% for the number of starts. It is not very important.

In all cases, heritability of number of starts was lower than the heritability of the Log of earnings. Three-year-old horses achieved the highest results; the heritability of the Log of earning and the heritability of the number of starts was  $0.34 \pm 0.03$  and  $0.21 \pm 0.03$  respectively.

Table 4— Variance component estimation in % for the Log of annual earnings and for the number of starts for three-year-olds.

Traits	Log. of annual earnings	Number of starts
Residual	$0.637(\pm 0.029)$	$0.253(\pm 0.018)$
Maternal environment	<b><math>0.022(\pm 0.013)</math></b>	$0.744(\pm 0.024)$
Additive genetic value = $h^2$ (diagonal)	<b><math>0.341(\pm 0.033)</math></b>	<b><math>0.451(\pm 0.272)</math></b>
= $r_g$ (above the diagonal)		<b><math>0.044(\pm 0.013)</math></b>
		<b><math>0.329(\pm 0.057)</math></b>
		<b><math>0.212(\pm 0.026)</math></b>

#### 3.2.4. Log of earnings and number of starts for 2 and 3-year- olds (data restricted to horses without missing values).

Table 6 show the incidence of the selection of horses having earnings both at 2- and 3-years of age on the estimation of genetic parameters:

Heritability for the number of starts of the 2-year-olds ranged from  $0.12 \pm 0.03$  (Table 3) to  $0.10 \pm 0.02$ . For the number of starts of the 3-year- olds it ranged from  $0.22 \pm 0.03$  (Table 4) to  $0.23 \pm 0.02$ .

Heritability for the earnings of the 2-year-olds ranged from  $0.15 \pm 0.03$  (Table 3) to  $0.21 \pm 0.03$ .

For the earnings of the 3-year-olds it ranged from  $0.34 \pm 0.03$  (Table 4) to  $0.39 \pm 0.02$ ; not a very big variation at all. However, estimations of the genetic correlation between earnings and

1 0.42±0.05 (Table 6). The obtained genetic correlation between ages for number of starts and  
 2 earnings for 2 and 3-year-olds was respectively 0.34±0.05 and 0.80±0.04. They therefore  
 3 appear to be moderately reliable.

4 Table 5 – Variance component estimation in % for the Log of annual earnings and for the  
 5 total number of starts for 2+3-year-olds.

	Log. of career earnings	Total number of starts
Residual	0.641(±0.023)	0.480(±0.018)
Maternal environment	<b>0.040(±0.013)</b>	0.741(±0.019)
		<b>0.498(±0.151)</b>
Additive genetic value = $h^2$ (diagonal)	<b>0.319(±0.026)</b>	<b>0.058(±0.013)</b>
= $r_g$ (above the diagonal)		<b>0.191(±0.074)</b>
		<b>0.201(±0.022)</b>

6 Table 6 – Variance components estimation in% for the Log of annual earnings and for the  
 7 annual number for starts for two and three year-olds.

Traits	Annual number of starts		Log. of annual earnings	
	2-year-olds	3-year-olds	2-year-olds	3-year-olds
Residual	0.903(±0.017)	-0.034(±0.015)	0.527(±0.015)	0.227(±0.021)
		0.770(±0.021)	-0.158(±0.021)	0.230(±0.023)
			0.790 (±0.025)	0.253(±0.023)
				0.606(±0.023)
Additive genetic value	<b>0.097(±0.017)</b>	<b>0.347(±0.052)</b>	<b>0.123(±0.111)</b>	-0.193(±0.090)
= $h^2$ (diagonal)		<b>0.230(±0.021)</b>	0.183(±0.068)	<b>0.423(±0.055)</b>
= $r_g$ (above the diagonal)			<b>0.210(±0.025)</b>	<b>0.801(±0.038)</b>
				<b>0.394(±0.023)</b>

## 4. DISCUSSION

We tried to assign the effect of year, sex, age, category of race, category of rider to the model. Since we had only the results of 2-year-old horses, running separately from 3-year-olds for comparison it was impossible to estimate the effect of age for the same horse. This led to non significant results either for earnings per race and for the ranking value. This is logical for ranking value but expresses similar money allocation per race for 2 and 3-year-olds in the Czech racing programme. Other effects in combination with horse effects were not estimable due to too much confounding. The category of race, category of rider and ‘quality’ of the horse are confounded, best horses running in the best races ridden by the best jockeys and vice versa. This led to numerous empty cells not compatible with the number of levels to be estimated. These effects were therefore not estimable but evidently they do not pose a serious problem. To make a correction for the racing level would not be appropriate. To make a correction for category of rider would only be interesting when different categories of riders are competing in the same category of race, which is not often the case. The only available adjustment was therefore to correct for year and for sex.

Estimation of the variance components led to somewhat classical results: the maternal effect was in the range of 2-3% for Log of earning per race and ranking value. Heritability was between 15-20% for the Log of earning per race and the ranking value, the repeatability between 30-40%. These results did not differ very much from those obtained in Poland [13] and Germany [3] for similar criteria. Repeatability for the distance when placed was low, 28%, but the heritability -18% was comparable to the heritability of ranking value and earning per race. This confirm the general agreement about the heritability of the aptitude for the distance qualifying horses as sprinter, miler, classic or stayer and the first results obtained in

1 Australia [14]. Estimation of breeding values for distance when placed is therefore possible  
2 and could be an information interesting for breeders. The criteria of Log of earning and  
3 ranking value being very highly correlated (95-98%), the estimations of the genetic  
4 correlations of the distance when placed were respectively 38 and 39% with the earning per  
5 race and the ranking value.

6 Comparing genetic parameters of Tables 1 and 2 for the earning per race and the ranking  
7 value show that taking into account non-placed horses and maternal effect did not change the  
8 estimations of heritability and repeatability for the ranking value very much but led to a slight  
9 decrease of them for earnings per race. Taking in account non- placed horses as well as the  
10 maternal effect did not seem to change the genetic approach very much, particularly in the  
11 case of the ranking value, which appears more stable than earnings relative to these variation  
12 factors.

13 There were also no great differences between the criteria Log of earnings and ranking value.  
14 Both traits at the race level, showed sufficiently high genetic variation to allow breeding value  
15 estimation. Knowing the mean number of races in a horse's career, which is around 6, it can  
16 lead to an efficient selection on the racing ability. The genetic correlation presented clearly  
17 shows the similarity of the two criteria, so we have to choose one of them.

18 Since Logarithmic transformation allows to obtain a normally distributed underlying variable  
19 of performance only approximately, and because many subjective factors influence the  
20 management of earnings distribution in a practical racing programme we would recommend  
21 to give priority to ranking values where these problems do not exist.

22 At the year or career level we estimated the genetic parameters of Log of earnings and  
23 number of starts for 2-year-olds, 3-year-olds and 2 and 3-year-old horses together. The  
24 maternal effect was in the range of 1%-6%, heritability between 15-39% for the Log of  
25 earnings and between 12-23% for the number of starts. The genetic correlation of the two

criteria was in the range of 12-42%. Standard error of these estimations were sometimes high, between 5 and 14%. It is obvious that the performance of the 2-year-olds is genetically less informative than of the 3-year-olds. However, annual earnings are highly correlated (80%). For estimation of breeding value the early information on 2-year-olds is therefore to be taken in account. The estimation of two variables is not the best solution because they are difficulties in the estimation of genetic correlations due to the selection of data for 2 and 3-year-olds. The total earnings at the end of 3 years of life is as heritable (32%) as the annual earnings of the 3-year-olds (34%). To use this variable could be a good way of achieving a synthetic view. In this choice which is common by breeders on rough data, we can propose two breeding value estimation: an early one only on 2-year-olds' Log of annual earnings and a more synthetic one on 2+3 Log of career earnings. These estimations will bring progress, first concerning the Logarithmic transformation of earnings and second with the advantage of the "animal model method". for the optimisation of the use of the information coming from parentage. They shall therefore improve common practices of breeders of Thoroughbred in the Czech Republic.

But one can appreciate the difficulty with annual or career earnings to have a good evaluation of the effect of the number of starts: It depends partly of environmental factors and partly of the quality of the horse. This explains the low value of heritability (10-23%). The same earning can also be achieved with different numbers of starts: as an example a lot of starts at low level could earn the same money than few starts at high level. This will also induce non-linear relations between the number of starts and total earnings. These problems may explain the relatively weak genetic correlations obtained (19% for career earnings and number of starts).

Because the chance to have earnings depends also of the number of starts, in a thorough analysis of this kind of data we should consider different threshold of truncation according to



1 the number of starts. This is usually not performed. We would therefore recommend, if  
2 possible, using data at the race level which avoids all these difficulties. It is proposed here in a  
3 very simple manner with the ranking value. More sophisticated treatments may be proposed,  
4 but it is not certain that they would be useful.

5 An information system has to be established including the criteria of earnings, rank and  
6 distance to inform on phenotypic and breeding values of the Czech Thoroughbred. Phenotypic  
7 values could be given at the level of each race or at the level of the racing career according to  
8 the methodology presented here for the ranking value. Breeding values should be calculated  
9 each year using the BLUP–animal model methodology described. In addition to the  
10 evaluation on the ranking value an evaluation of the distance aptitude could be added. This  
11 procedure should be given priority over the more traditional one based on total rough  
12 earnings.

## 14 **5. CONCLUSION**

16 Our results show that earnings and ranking values are two appropriate criteria to select the  
17 English Thoroughbred for racing ability in Czech Republic. Due to their very high genetic  
18 correlation, more than 90%, we can recommend the choice of one of the two criteria.

19 Taking into account non-placed horses did not change the estimation of genetic parameters  
20 very much. The same is true for taking into account the maternal effect.

21 The addition of the parameter distance when placed appears to be a good step in the  
22 estimation of the breeding value. It provides information on the racing ability of  
23 Thoroughbred interesting for breeders.

24 We cannot really recommend the use of the criteria number of starts. Genetic correlation  
25 between number of starts and Log of earnings were low and the standard error too high. This

parameter was not well estimated. However, number of starts being in strong phenotypic relation with the annual or career earnings, it is important to find the optimal manner to take it into account when using these criteria.

The difficulty of properly taking into account the number of starts for annual or career earnings inclined us to prefer criteria at the race level where this problem does not exist. At this level the ranking criteria avoiding distribution problems and the subjectivity of earnings should be preferred.

## REFERENCES

[1] Anonym: Jockey Club Ceske republiky-vyrocní zpráva 1996-1997, Praha, JC CR, 1998, 20s.

[2] Belhajyahia, T., Blouin, Ch., Langlois, B., Harzalla, H.: Breeding evaluation of Arab horses from their racing results in Tunisia by a BLUP with an animal model approach, Anim. Res., 52, 2003; 481-488.

[3] Bugislaus, A.E., Roehe R., Uphaus H., Kalm E. Development of genetic models for estimation of racing performances in German Thoroughbred. Archiv für Tierzucht 47 (6) 505-516.

[4] Chico M.D.: Genetic analysis of thoroughbred racing performance in Spain, Ann. Zootech., 43, 1994; 393-397.

[5] Dusek, J.: Koeficienty dedivosti exteriéru a vykonnosti koni, Stud. Inform. UVTIZ, Živocisná výroba, 3, 1981; 88s.

[6] Gillespie R.H., A new way to evaluate race horses, Performance Rate. The Thoroughbred Record, 17 (1971), 961-977.

- [7] Groeneveld, E.: REML VCE V3.2, User's Guide, Institute of Animal Husbandry and Animal Ethology, Mariensee, D-31535 Neustadt, Germany, 1996; 52s.
- [8] Langlois, B.: Heritability of racing ability in thoroughbreds – a review. *Livest. Prod. Sci.*, 7, 1980; 591-605.
- [9] Langlois, B.: A consideration of the genetic aspects of some current practices in Thoroughbred horse breeding. *Ann. Zootech.* 1996, 45, 41-51.
- [10] Misar, D., Jiskrova, I., Pribyl, J. Estimation of breeding value of English Thoroughbred sires in the Czech Republic, *Czech J. Anim. Sci.*, 45, 2000; 201-208.
- [11] Robertson W., Explanation of Performances Rates. *The Thoroughbred Record* 17 (1971), 1132.
- [12] SAS/STAT User's Guide, Version 6, Fourth Edition, Volume 2, SAS Institute Inc., Cary, 1989; 846s.
- [13] Sobczynska M., Lukaszewicz M. Genetic parameters of racing merit of Thoroughbred horses in Poland. *J. Anim. Breed. Genet.* 121 (2004), 302-306
- [14] Williamson S. A., Beilharz, R.G. The inheritance of speed, stamina and other racing performance characters in the Australian thoroughbred. *J. Anim. Breed. Genet.* 115 (1998), 1-16.

## ACKNOWLEDGEMENTS

We are grateful to the French embassy in Praha for having facilitated the stay in France and to the support by the Ministry of Education, Youth and Sports of the Czech Republic (Project No. 432100001).