Anim. Res. 54 (2005) 499–509
© INRA, EDP Sciences, 2005
DOI: 10.1051/animres:2005035

Original article

Estimation of genetic parameters of Thoroughbred racing performance in the Czech Republic

Sonia SVOBODOVA[a], Christine BLOUIN[b], Bertrand LANGLOIS[b]*

[a] Faculty of Agronomy, Animal Breeding Institute, University of Agriculture and Forestry, Gregor Mendel, 613 00 Brno, Czech Republic
[b] Station de Génétique Quantitative et Appliquée, 78352 Jouy-en-Josas Cedex, France

(Received 16 July 2004; accepted 24 June 2005)

Abstract – The aim of this study was to estimate the breeding value of English Thoroughbreds in the Czech Republic using racing results from a 22-year period (1980–2001). The data include the performance of two and three-year-old horses which raced in flat races at hippodromes in the Czech Republic. The racing results (30 203) were available corresponding to 6 333 horses descending from 762 sires and 2 836 dams. Different criteria were applied in order to analyse the performance: Log of earnings per race, a normalized ranking value, distance of the race when placed, earnings and number of starts for 2, 3, 2+3 year-old horses. After preliminary studies, a year effect or a sex by year effect was finally retained. Variance component estimation using VCE software gave the following values for heritability (±standard errors): 0.14 ± 0.01 and 0.16 ± 0.01 for the Log of earnings per race and the ranking value. Repeatability was 0.31 and 0.35, respectively. The maternal environment component 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 Thoroughbreds in the Czech Republic. The genetic correlation of the two criteria was 0.98 ± 0.003. The heritability for the distance when placed was 0.18 ± 0.01. The genetic correlation of the Log of earnings per race and distance was medium, 0.38 ± 0.05 and of the same order, 0.39 ± 0.05 for ranking value and distance. In the case where we used the Log of annual earnings and the number of starts, the heritabilities were for the Log of earnings: 0.15 ± 0.03 for two-year-olds, 0.34 ± 0.03 for three-year-olds and 0.32 ± 0.03 for two and three-year-old careers together and respectively, 0.12 ± 0.03, 0.21 ± 0.03 and 0.20 ± 0.02 for number of starts. The genetic correlations between the earnings and the number of starts were respectively: 0.26 ± 0.14, 0.33 ± 0.06 and 0.19 ± 0.07. The genetic correlation between the number of starts for two and three year-olds which was restricted to horses earning money for two consecutive years, was average: 0.35 ± 0.05 and between earnings for the same ages was high 0.80 ± 0.04.

Thoroughbred / flat races / estimation of breeding value / BLUP – animal model / Czech Republic

Résumé – **Estimation de la valeur génétique du Pur Sang Anglais en République Tchèque.** L'objectif de cette étude était d'estimer la valeur génétique des Pur Sang de la République Tchèque

* Corresponding author: Bertrand.Langlois@dga.jouy.inra.fr

à l'aide des résultats de course sur une période de 22 ans (1980–2001). Notre fichier comprend les performances à 2 et 3 ans des chevaux ayant couru en plat sur les hippodromes Tchèques soit 30 203 résultats de courses de 6 333 chevaux issus de 762 étalons et 2 836 poulinières. La performance a été appréhendée par différents critères, le Log du gain par course, une valeur calculée à partir du rang d'arrivée, la distance de course si le cheval est placé, le gain et le nombre de départs à 2, 3 et 2+3 ans. Après quelques analyses préliminaires un effet année ou un effet sexe par année a finalement été retenu. L'estimation des composantes de la variance par l'utilisation du programme VCE a donné les valeurs suivantes pour l'héritabilité (± écart-type) 0,14 ± 0,01 et 0,16 ± 0,01 pour le Log du gain par course et la valeur issue du classement. La répétabilité a été de 0,31 et 0,35 respectivement. La composante environnement maternel a été évaluée à 0,02 ± 0,004 pour le Log du gain par course et à 0,03 ± 0,004 pour la valeur issue du classement. Ces deux critères constituent deux moyens appropriés d'appréhender les performances pour la sélection des Pur sang en République Tchèque. La corrélation génétique des deux critères a été de 0,98 ± 0,03. Pour les chevaux placés uniquement, l'héritabilité de la distance de course a été estimée à 0,18 ± 0,01, sa corrélation génétique avec le Log du gain par course a été trouvée moyenne 0,38 ± 0,05 et du même ordre pour la valeur issue du classement 0,39 ± 0,05. Dans le cas ou l'on utilise le Log du gain et le nombre de départs annuels, l'héritabilité a été pour le Log du gain 0,15 ± 0,03 pour les 2 ans, 0,34 ± 0,03 pour les 3 ans, 0,32 ± 0,03 pour les 2 et 3 ans ensemble et respectivement 0,12 ± 0,03, 0,21 ± 0,03 et 0,20 ± 0,02 pour le nombre de départs. La corrélation génétique entre gains et nombre de départs a été respectivement 0,26 ± 0,14, 0,33 ± 0,06 et 0,19 ± 0,07. En se limitant aux chevaux ayant gagné deux années consécutives, la corrélation génétique entre le nombre de départs à 2 et 3 ans a été moyenne : 0,35 ± 0,05 et entre gains aux mêmes âges elle a été forte : 0,80 ± 0,04.

Pur Sang / courses plates / estimation de la valeur génétique / BLUP – modèle animal / République Tchèque

1. INTRODUCTION

The history of the Czech turf dates back to 1839, when the Czech racing company for Bohemia and Moravia was established. Under the leadership of this company, the first public race took place in Prague 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 of Prague. Velka Chuchle became the central place for gallop races. The first Czechoslovak Derby was held here in 1921 and was called the "Prix du Czechoslovak Jockey Club". It was a sign of inclination of the new republic to France after World War I, when the Derby was called the "Prix du Jockey Club". From 1922 the Derby was called the Czechoslovak Derby and at present it is the Czech Derby. World War II had a very bad effect on the turf in the 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 Thoroughbreds 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. The 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 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) for the estimation of the breeding value, because these performance characteristics have high or medium coefficients of heritability in the contrary to racing times whose 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 year olds.

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

2. MATERIALS AND METHODS

2.1. Setting up the database

The racing results and pedigree of the Thoroughbreds were collected from Yearbooks (1980–2001) and from Stud Books of the Czech Republic. The 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: the 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; information about the pedigree of the horses over two generations was also included.

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 the number of starts.

2.2. Criteria for the measurement of performance

In order 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 the 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 the 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. The level of race was considered for the evaluation of genetic values of the Log of earnings, rank and distance when placed. The level of year was considered for annual earnings and number of starts for two and three-year-old horses. The 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 to 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 the first analyses, the fixed effects were studied using the GLM programme package SAS [12] with or without a random horse effect. For the estimation of the breeding value of performance of the Thoroughbreds, the BLUP – animal model was used. The following animal model was fitted by using VCE and PEST software [7]. Since many assumed fixed effects were not estimable in reality, in our genetic analysis we only considered the effect of the year and sex. In order to avoid problems with interaction, a Sex by Year effect was considered. The 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:

$$y = X\,b + Z\,g + W\,m + Z\,p + e$$

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); e = vector of errors; and X, Y, Z = the incidence matrices.

The expectations of this linear model are:

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

The variance covariance matrix is:

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

$\sigma^2_g = h^2\sigma^2_y$
$\sigma^2_m = \mu\,\sigma^2_y$
$\sigma^2_p = (r - \mu - h^2)\,\sigma^2_y$
$\sigma^2_e = (1 - r)\,\sigma^2_y$

where: A = relationship matrix; I = identity matrix; h^2 = heritability; μ = maternal environment component of variance in %; r = repeatability.

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

Traits	Log of earnings per race	Ranking value
Number of observations	28 051	28 051
Minimum	2.18	–1.74
Maximum	13.85	2.55
Average	6.76	0.06
Standard deviation	1.95	0.83
Number of equations	37 948	37 948
d.f. horses	10 423	10 423
d.f. maternal environment	2 669	2 669
d.f. permanent horse environment	5 872	5 872
d.f. fixed effects (years)	20	–
Variance component ratios (± standard-deviation) in %		
Residual	0.692(±0.05)	0.828(±0.001)
		0.655(±0.005)
Maternal environment	**0.023(±0.04)**	0.982(±0.013)
		0.028(±0.004)
Permanent horse environment	0.141(±0.09)	0.962(±0.004)
		0.155(±0.009)
Additive genetic value = h^2 (diagonal) = r_g (above the diagonal)	**0.144(±0.010)**	**0.980(±0.003)**
		0.162(±0.011)
Repeatability	**0.308(±0.014)**	**0.347(±0.014)**

3. RESULTS

3.1. Estimation of fixed effects

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

In 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 significant in models excluding the random effect of the horse. Most of these effects including this random effect of the horse, 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 the criteria per race, year and sex effects were not significant, except the effect of year remaining alone for the 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 earnings per race and ranking value (for all starters)

Table I 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. The maternal environmental effect was lower than 3%. The genetic correlation of the two criteria was very high: 0.98 ± 0.003. The

Table II. Variance component estimation 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
Number of observations	13 302	13 302	13 302
Minimum	5.30	–1.53	800
Maximum	13.85	2.55	3210
Average	8.48	0.78	1 543
Standard deviation	1.08	0.60	427
Number of equations	37 793	37 793	37 793
d.f. horses	8 329	8 329	8 329
d.f. permanent horse environment	4 262	4 262	4 262
d.f. fixed effects (years)	20	–	–
Variance component ratios (± standard-deviation) in %			
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	**0.192(±0.008)**	**0.945(±0.008)**	**0.383(±0.049)**
= h^2 (diagonal)		**0.174(±0.008)**	**0.394(±0.050)**
= r_g (above the diagonal)			**0.183(±0.012)**
Repeatability	**0.399(±0.016)**	**0.323(±0.016)**	**0.283 (±0.022)**

repeatability for Log of earning was 0.31 ± 0.02 and for ranking value was 0.35 ± 0.02. Heritability was estimated as 0.14 ± 0.01 and 0.16 ± 0.01, respectively.

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

The heritability of parameters for two and three-year-old horses was evaluated (Tab. II) 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 ± 0.05 and 0.39 ± 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 ± 0.01.The heritability of the distance when the horse is placed is medium: 0.18 ± 0.01. The repeatability 0.28 ± 0.02 is low but allows breeding value evaluation. The selection of data not taking into account the low maternal effect did not change the repeatability of earning 0.40 ± 0.02 and ranking value 0.32 ± 0.02 very much. However, the heritability of earning was slightly increased 0.19 ± 0.01 and remained nearly the same for the ranking value 0.17 ± 0.01.

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

The genetic parameters were estimated in groups of 2, 3 and 2 + 3-year-old horses. Tables III to VI show the results of these estimations. The genetic correlation between earnings and the 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 was not very important.

Table III. Variance component estimation for the Log of annual earnings and for the number of starts for two-year-olds.

Traits	Log of annual earnings	Number of starts
Number of observations	2 579	2 579
Minimum	6.40	1
Maximum	12.32	8
Average	9.08	**3.1**
Standard deviation	1.11	1.4
Number of equations	15 020	15 020
d.f. horses	5 942	5 942
d.f. maternal environment	1 526	1 526
d.f. fixed effects (sex × years)	42	42
Variance component ratios (± standard-deviation) in %		
Residual	0.826(±0.029)	0.487(±0.019)
		0.875(±0.025)
Maternal environment	**0.027(±0.027)**	–0.879(±1.798)
		0.007(±0.007)
Additive genetic value = h^2 (diagonal) = r_g	**0.147(±0.029)**	**0.258(±0.144)**
(above the diagonal)		**0.118(±0.027)**

Table IV. Variance component estimation for the Log of annual earnings and for the number of starts for three-year-olds.

Traits	Log of annual earnings	Number of starts
Number of observations	3 527	3 527
Minimum	6.00	1
Maximum	14.43	15
Average	9.25	**4.2**
Standard deviation	1.30	2.2
Number of equations	18 616	18 616
d.f. horses	7 364	7 364
d.f. maternal environment	1 902	1 902
d.f. fixed effects (sex × years)	42	42
Variance component ratios (± standard-deviation) in %		
Residual	0.637(±0.029)	0.253(±0.018)
		0.744(±0.024)
Maternal environment	**0.022(±0.013)**	0.451(±0.272)
		0.044(±0.013)
Additive genetic value = h^2 (diagonal) = r_g	**0.341(±0.033)**	**0.329(±0.057)**
(above the diagonal)		**0.212(±0.026)**

In all cases, the heritability of the 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 earnings and the heritability of the number of starts were 0.34 ± 0.03 and 0.21 ± 0.03 respectively.

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

Table VI shows the incidence of the selection of horses having earnings both at

Table V. Variance component estimation for the Log of annual earnings and for the total number of starts for 2+3-year-olds.

Traits	Log of career earnings	Total number of starts
Number of observations	4 358	4 358
Minimum	6.00	1
Maximum	14.43	18
Average	9.49	**5.8**
Standard deviation	1.29	2.9
Number of equations	21 506	21 506
d.f. horses	8 525	8 525
d.f. maternal environment	2 186	2 186
d.f. fixed effects (sex × years)	42	42
Variance component ratios (± standard deviation) in %		
Residual	0.641(±0.023)	0.480(±0.018)
		0.741(±0.019)
Maternal environment	**0.040(±0.013)**	0.498(±0.151)
		0.058(±0.013)
Additive genetic value = h^2 (diagonal) =	**0.319(±0.026)**	**0.191(±0.074)**
r_g (above the diagonal)		**0.201(±0.022)**

Table VI. Variance component estimation for the Log of annual earnings and for the 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
Number of observations	3 426	4 031	2 579	3 527
Minimum	1	1	6.40	6.00
Maximum	8	15	12.32	14.43
Average	**2.8**	**3.9**	9.08	9.25
Standard deviation	1.5	2.2	1.11	1.30
Number of equations	24 268	24 268	24 268	24 268
d.f. horses	8 525	8 525	8 525	8 525
d.f. fixed effects (sex × years)	42	42	42	42
Variance component ratios (± standard deviation) in %				
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	**0.097(±0.017)**	**0.347(±0.052)**	**0.123(±0.111)**	–0.193(±0.090)
= h^2 (diagonal)		**0.230(±0.021)**	0.183(±0.068)	**0.423(±0.055)**
= r_g (above the diagonal)			**0.210(±0.025)**	**0.801(±0.038)**
				0.394(±0.023)

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 ± 0.03 (Tab. III) to 0.10 ± 0.02. For the number of starts of the 3-year-olds, it ranged from 0.22 ± 0.03 (Tab. IV) to 0.23 ± 0.02.

The heritability for the earnings of the 2-year-olds ranged from 0.15 ± 0.03 (Tab. III) to 0.21 ± 03.

For the earnings of the 3-year-olds it ranged from 0.34 ± 0.03 (Tab. IV) to 0.39 ± 0.02; this was not a very big variation at all. However, the estimations of the genetic correlation between earnings and number of starts at the same age were 0.26 ± 0.14 for 2-year-olds in one case (Tab. III) and 0.12 ± 0.11 in the other (Tab. VI) and respectively for 3-year-olds 0.33 ± 0.06 (Tab. IV) and 0.42 ± 0.05 (Tab. VI). The genetic correlations obtained between ages for the number of starts and earnings for 2 and 3-year-olds were respectively 0.34 ± 0.05 and 0.80 ± 0.04. They therefore appear to be moderately reliable.

4. DISCUSSION

We tried to assign the effect of year, sex, age, category of race, and category of rider to the model. Since we only had 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 the ranking value but expresses similar money allocation per race for 2 and 3-year-olds in the Czech racing program. Other effects in combination with horse effects were not estimable due to too much confounding. The categories of race, of rider and “quality” of the horse were confounded, with the 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 the category of the 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 sex.

The estimation of the variance components led to somewhat classical results: the maternal effect was in the range of 2–3% for Log of earnings per race and ranking value. Heritability was between 15–20% for the Log of earnings 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. The repeatability for the distance when placed was low, 28%, but the heritability, 18% was comparable to the heritability of the ranking value and earnings per race. This confirmed the general agreement about the heritability of the aptitude for the distance qualifying horses as sprinters, milers, classics or stayers and the first results obtained in Australia [14]. The estimation of breeding values for distance when placed is therefore possible and could be an information interesting for breeders. The criteria of Log of earnings and ranking value being very highly correlated (95–98%), the estimations of the genetic correlations of the distance when placed were respectively 38 and 39% with the earnings per race and the ranking value.

Comparing genetic parameters of Tables I and II for the earnings per race and the ranking value show that taking into account non-placed horses and maternal effect did not change the estimations of heritability and repeatability for the ranking value very much but led to a slight decrease of them for earnings per race. Taking into account none-placed horses as well as the maternal effect did not seem to change the genetic approach very much, particularly in the case of the ranking value, which appears more stable than the earnings relative to these variation factors.

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

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

At the year or career level, we estimated the genetic parameters of Log of earnings and number of starts for 2-year-olds, 3-year-olds and 2 and 3-year-old horses together. The maternal effect was in the range of 1–6%, heritability between 15–39% for the Log of earnings and between 12–23% for the number of starts. The genetic correlation of the two criteria was in the range of 12–42%. The standard error of these estimations was sometimes high, between 5 and 14%. It is obvious that the performance of the 2-year-olds was genetically less informative than that of the 3-year-olds. However, annual earnings were highly correlated (80%). For the estimation of breeding value, the early information on 2-year-olds therefore has to be taken into account. The estimation of two variables was not the best solution because there were 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 were as heritable (32%) as the annual earnings of the 3-year-olds (34%). The use of this variable could be a good way of achieving a synthetic view. In this choice which is common by breeders on raw data, we can propose two breeding value estimations: an early one only on the 2-year-olds' Log of annual earnings and a more synthetic one on the 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 Thoroughbreds in the Czech Republic.

But one can appreciate the difficulty with annual or career earnings in order to obtain a good evaluation of the effect of the number of starts: it depends partly on environmental factors and partly on 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 a low level could earn the same money as a few starts at a 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 also depends on the number of starts, in a thorough analysis of this kind of data we should consider a different threshold of truncation according to the number of starts. This is usually not performed. We therefore recommend, if possible, using data at the race level which avoids all these difficulties. This is proposed here in a very simple manner with the ranking value. More sophisticated treatments may be proposed, but it is not certain that they would be useful.

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

5. CONCLUSION

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

Taking into account none-placed horses did not change the estimation of genetic parameters very much. The same was true when considering the maternal effect.

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

We cannot really recommend the use of the criteria, number of starts. Genetic correlation between number of starts and Log of earnings were low and the standard error too high. This parameter was not well estimated. However, since the number of starts is in a 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.

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).

REFERENCES

[1] Anonym: Jockey Club Ceske republiky-vyrocni zprava 1996–1997, Praha, JC CR, 1998, 20 p.

[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 (2004) 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 exterieru a vykonnosti koni, Stud. Inform. UVTIZ, Zivocisna vyroba, 3, 1981, 88 p.

[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, 31535 Neustadt, Germany, 1996, 52 p.

[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. 45 (1996) 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, 4th ed., Vol. 2, SAS Institute Inc., Cary, 1989, 846 p.

[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.