55th Annual Meeting of the EAAP, Bled, 2004, poster G4.65

Estimation of genetic parameters for milk production of Czech Spotted cattle

E.Němcová, L.Dědková and J. Wolf

Research Institute of Animal Production, P.O.Box 1, CZ 10401 Prague-Uhříněves, Czech Republic. E-Mail: nemcova.eva@vuzv.cz

Abstract. Genetic parameters for milk production traits of Czech Spotted cattle were estimated with the aim of their future use in genetic evaluation. About 150,000 test-day records from approximately 10,000 animals were taken from the official milk recording database, the year of test being in the range between 1995 and 2002. Model equations included the fixed effect of herd-test-day and fixed regression on the stage of lactation (within subgroups of cows) and the permanent environmental and animal effects modelled by random regression. Third order Legendre polynomials (with four coefficients) were used for all regressions. The models employed differed in the definition of the subgroups for fixed regression on stage of lactation, or (alternatively) in the definition of the residual effect. Gibbs sampling and REML were used for the parameter estimation. The heritabilities were found in the range from 0.18 to 0.25 in the first, from 0.19 to 0.27 in the second and from 0.16 to 0.38 in the third lactation. The genetic correlations between lactations varied from 0.76 to 0.82 between the first and second; from 0.63 to 0.68 between the first and third; and from 0.71 to 0.82 between the second and third lactations.

Introduction

During the recent years, the test-day model (TDM) has become widely used in the genetic evaluation of milk production in cattle (*Swalve, 1993; Swalve, 1995; Freeman, 1998; Meyer, 2001*). The advantages of TDMs compared with lactation-sum models involve the ability to simultaneously account for environmental effects of each test-day and to model lactation trajectories for specific genotypes. Furthermore, they provide the possibility of accessing persistency of milk production during lactation. The main disadvantages include increased computational requirements and more genetic parameters (*Freeman, 1998*). The confounding of genetic and environmental effects can lead to erroneous estimation of effects (*Meyer, 2001*). The maximization of genetic gain and minimization of the residual variance are proposed as criteria for the choice of the best model.

The implementation of TDMs requires accurate genetic parameters that describe the covariance structure of test-day production records within and across lactations (*De Roos et al., 2004*). Numerous researches have reported estimates of genetic parameters for test-day milk production traits using various models. The most popular methods have been random regression models where the covariances among random regression coefficients are estimated directly form data (*Jamrozik et al., 1998; Jakobsen et al., 2002; Samoré et al., 2002*) and indirect methods where production records from different lactation stages are analyzed in a multi-trait evaluation and a covariance function is fitted to the covariance matrices (*Lidauer et al., 2000; Gengler et al., 2001; Emmerling et al., 2002*).

The implementation of the test-day model into the genetic evaluation of cattle in the Czech Republic is important not only for increasing the accuracy of the genetic evaluation which will result in a higher genetic gain and a shorter generation interval but also for the maintenance of the comparability of the Czech cattle genetic evaluation with other countries, because the implementation of the TDM is a worldwide trend.

The aim of the present study was to estimate genetic parameters from a multiple-lactation random regression model for Czech Spotted cattle; these parameters are expected to be used in the future genetic evaluation.

Material and Methods

The primary dataset for the estimation of genetic parameters consisted of TD records from the milk recording database for Czech Spotted cattle which were collected from 1995 to 2002. The traits analysed were milk test-day yields between 6 to 305 days in milk (DIM) of the first, second and third lactation. A minimum of 8 TD records were required for each lactation and the interval between two consecutive TDs had to be between 20 and 75 days. Furthermore, it was demanded that not less than 4 cows were tested at each test day in each herd. After editing the datasets included TD records from 312,768 lactations of 201,481 cows, i.e. 1.5 lactations per cow on average. Two subsets (samples A and B with approximately 10,000 first lactations) were created by random using the primary dataset. The cow was required to have the first lactation; the second and third lactations were used if available. Dataset C included only cows with all three lactations.

The basic statistical characteristics were computed for all datasets (see Table 1). The subsets A and B did not differ from the primary dataset. Dataset C showed a higher milk yield in the third lactation probably due to a higher number of third lactations (selected cows). The total number of generations in the pedigree was 4. Phantom parent groups were not used.

Table 1: Basic statistical characteristics of the used datasets

	Primary dataset	Dataset (sample from primary dataset)		
		A	B	C
Number of animals, lactations and TDs				
Cows with observations	201 481	10 556	9 845	11 500
Animals in pedigree		50 275	47 729	44 996
No. of herds	3 303	145	145	155
No. of test-days	2 796 788	139 889	132 085	298 437
No. of lactations	312 768	15 769	14 655	32 399
Number and proportion (in parentheses) of lactations				
1st lactation	210 481 (67 %)	10 556 (67 %)	9 845 (67 %)	10 739 (33 %)
2nd lactation	74 865 (24 %)	3 900 (25 %)	3 640 (25 %)	10 739 (33 %)
3rd lactation	27 422 (9 %)	1 313 (8 %)	1 170 (8 %)	10 739 (33 %)
Average milk yield and its standard deviation (in parentheses)				
Milk (kg)	16.86 (6,00)	16.70 (5,80)	16.83 (5,75)	19.05 (7,09)
DIM (day)	148.52 (82.63)	148.91 (82.86)	147.91 (82.44)	149.02 (82.88)

Note: DIM – day in milk

Data sets A, B and C were analysed by a multiple-lactation model in which TD yields in the first, second and third lactations were considered as different traits. The model equations included the fixed effect of herd-test-day and fixed regression on the stage of lactation (within subgroups of cows). The permanent environmental and animal effects were modelled by random regression. The model was assumed to be the same for each parity and trait combination. The model equation was:

$$y_{nkitjl} = HTD_{ni} + \sum_{m=1}^{4} \beta_{nkm} z_{tm} + \sum_{m=1}^{4} a_{njm} z_{tm} + \sum_{m=1}^{4} p_{njm} z_{tm} + e_{nkitjl}$$, where:

y_{nkitjl}	is record *l* on cow *j* made on day *t* within herd-test day effect *i*, for a cow belonging to subclass *k* for age and season of calving for parity *n*,
HTD_{ni}	is the fixed herd-test day effect *i* in parity *n*,
β_{nkm}	are fixed regression coefficients specific to subclass *k* in parity *n*,
a_{njm}	are random regression coefficients specific to animal *j* in parity *n*,
p_{njm}	are random regression coefficients specific to the permanent environmental (PE) effect of cow *j* in parity *n*,
z_{tm}	are covariates associated with DIM, assumed to be the same for both fixed and random regressions,
e_{nkitjl}	is the residual effect.

Third order Legendre polynomials (with four coefficients) were used for both the fixed and random regressions on the scale from 6 to 305 DIM. Let $\mathbf{p_j}$ represent a 12x1 vector of random permanent environmental regression coefficients for cow *j* with the covariance matrix $\mathbf{P}$. The PE covariance matrix for all cows is then $\mathbf{I} \otimes \mathbf{P}$. Furthermore, let $\mathbf{a_j}$ be a 12x1 vector of the random regression coefficients for animal *j* with the covariance matrix $\mathbf{G}$. $\mathbf{A} \otimes \mathbf{G}$ is then the genetic covariance matrix for all animals with $\mathbf{A}$ being the additive genetic relationship matrix. Different residual variances were allowed for different lactations and time periods within a lactation. Residual effects on different DIM were assumed to be uncorrelated both within and between cows. The model used was a special case of *Jamrozik et al. (1998)* for the multiple-lactation, multiple trait situation.

Cows were assigned to the subclasses for fixed regression that were defined in two ways (see Table 2)**:**

Table 2: Definition of subclasses for fixed regression

Definition of SFR	No. of levels in each parity	Subclasses were create according to:
L	480	age at calving (only 1st lactation); season of calving; days open; previous calving interval (only for 2nd and 3rd lactations)
M	12	age at calving; season of calving

Note: SFR = subclass for fixed regression

In total, seven analyses were carried out. They differed in the subsets involved, definition of subclasses for fixed regression and definition of residual variance periods (see Table 3).

Table 3: Survey of analyses

Analysis	Sample	Method	Model
I	A	GS	definition of SFR: L, residuum in 4 periods
II	B	GS	same as analysis I.
III	A	GS	definition of SFR: M, residuum in 4 periods
IV	A	GS	definition of SFR: L, residuum in 9 periods
V	C	GS	same as analysis I. (all 3 lactations were required)
VI	A	REML	same as analysis I., (only 1st lactation)
VII	A	AI-REML	same as analysis I., (only 1st lactation)

Note: GS = Gibbs sampling
SFR = subclass for fixed regression

Bayesian estimation using Gibbs sampling was used to generate variances and covariances from their respective posterior distributions. Blocked samplings with multivariate normal and inverted Wishart distributions were used. For each trait, 55,000 samples were generated and 5,000 burn-in samples were discarded. Estimates were obtained as posterior means of 50,000 samples.

Results and discussion

Table 4: Estimates of genetic variances (GE), variances of permanent environmental effect (PE), residual variances (RE) and heritabilities for the 305-day milk yield in analyses I to VII

Analysis	Lactation	Variances			Gen. correlations		Heritability
		GE	PE	RE	2nd lact.	3rd lact.	
I	1	155433	393081	82722	0.8189	0.6333	0.2462
	2	268554	614429	111288		0.7109	0.2701
	3	441108	575827	131712			0.3840
II	1	163350	397149	83176	0.7557	0.6151	0.2538
	2	233983	629587	117239		0.7421	0.2386
	3	452448	600032	133688			0.3814
III	1	156754	410326	83454	0.8184	0.6702	0.2410
	2	253083	668054	112136		0.7096	0.2449
	3	467175	648811	133256			0.3739
IV	1	158945	407863	33689	0.8221	0.6765	0.2340
	2	272465	650078	44906		0.7495	0.2738
	3	449117	664662	54408			0.3394
V	1	119410	475705	86580	0.7637	0.6306	0.1752
	2	201270	763861	120282		0.8224	0.1854
	3	261696	966953	148744			0.1900
VI	1	136993	405137	1044			0.2522
VII	1	134223	407358	1044			0.2473

The estimated genetic parameters for the individual analyses are presented in Table 4 and Fig. 1 and 2. Higher genetic and permanent environmental variances occurred at the beginning of the lactation to 45th DIM where the first time period for residual variance ends (Fig. 3). In the course of the middle part of the lactation the curves for these variances were flat, whereas in the last third of the lactation the variances increased again, above all the permanent environmental variance. This rise was especially striking in the third lactation.

Fig 1: Heritabilities for milk yield in 305-day lactation in analyses I-V

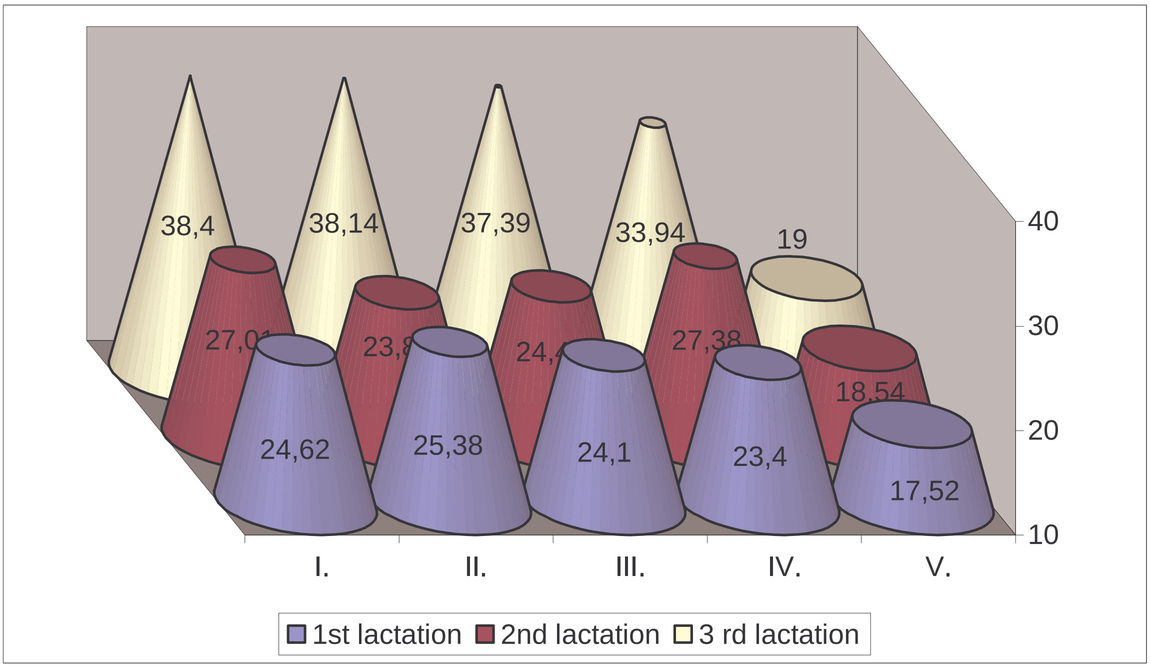
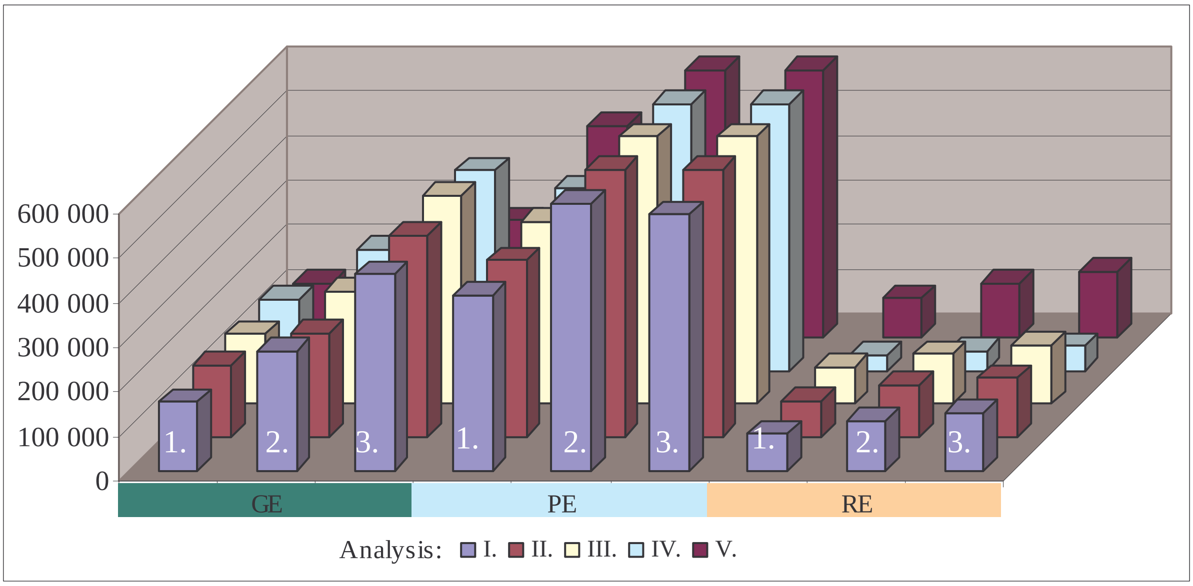


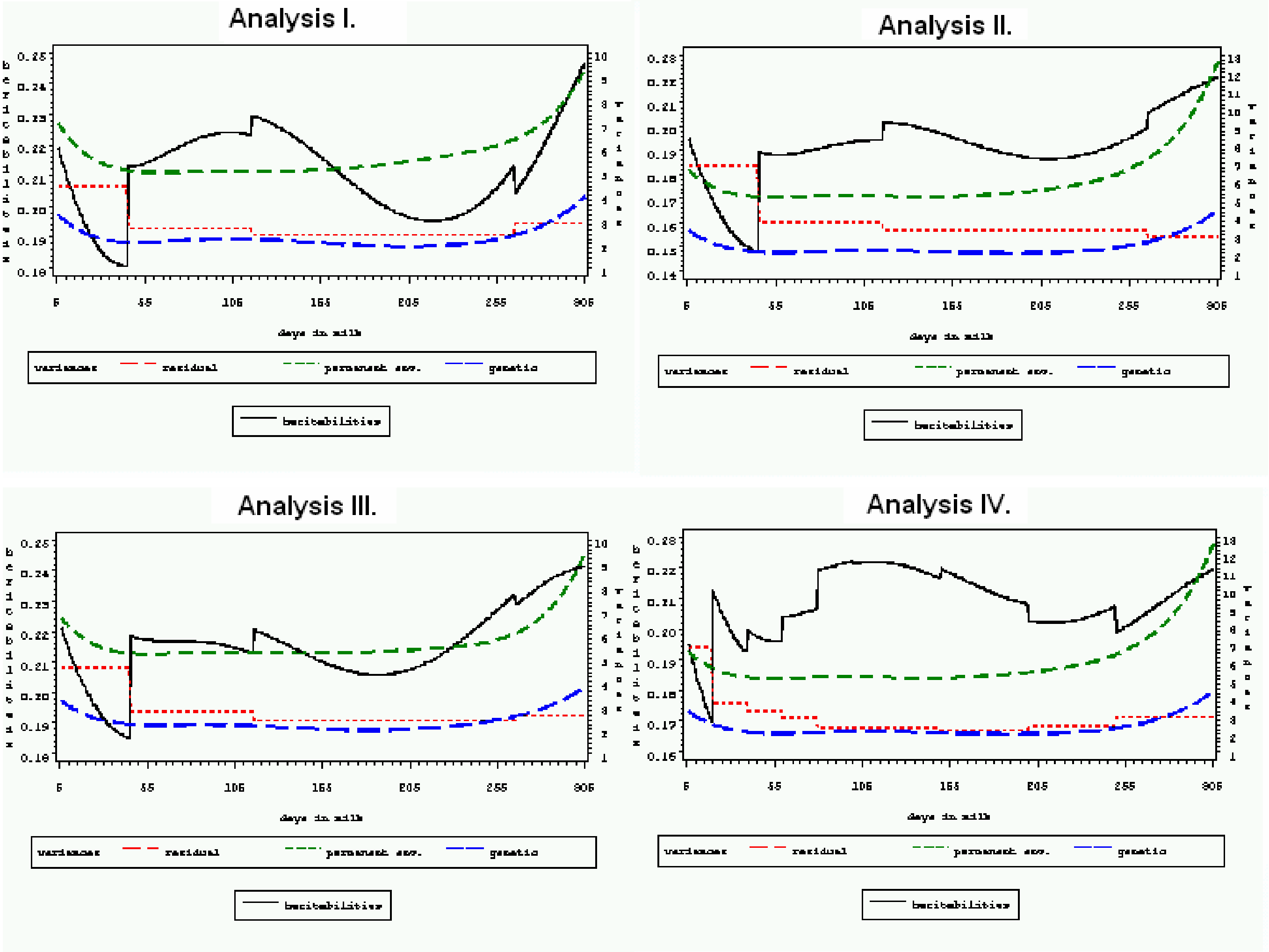
Fig 2: Estimates of genetic variances, variances of permanent environmental effect and residual variances for milk yield in 305-day lactation



The course of the residual variance in the first lactation was flatter than in the later lactations. This trend holds also for the other variances and heritability (see Fig. 4).

The genetic variances, variances of the permanent environmental effect and residual variances calculated for 305-day lactation increased with parity. But the variances of the permanent environmental effect and residual variances did not go up from the second to third lactation. The additive genetic variance was still rising. Therefore the heritabilities increased by about 10% in the third lactation.

Fig 3: Course of variances and heritability for milk yield in first lactations (results from analyses I, II, III and IV)



Analysis I.

Analysis II.

Analysis III.

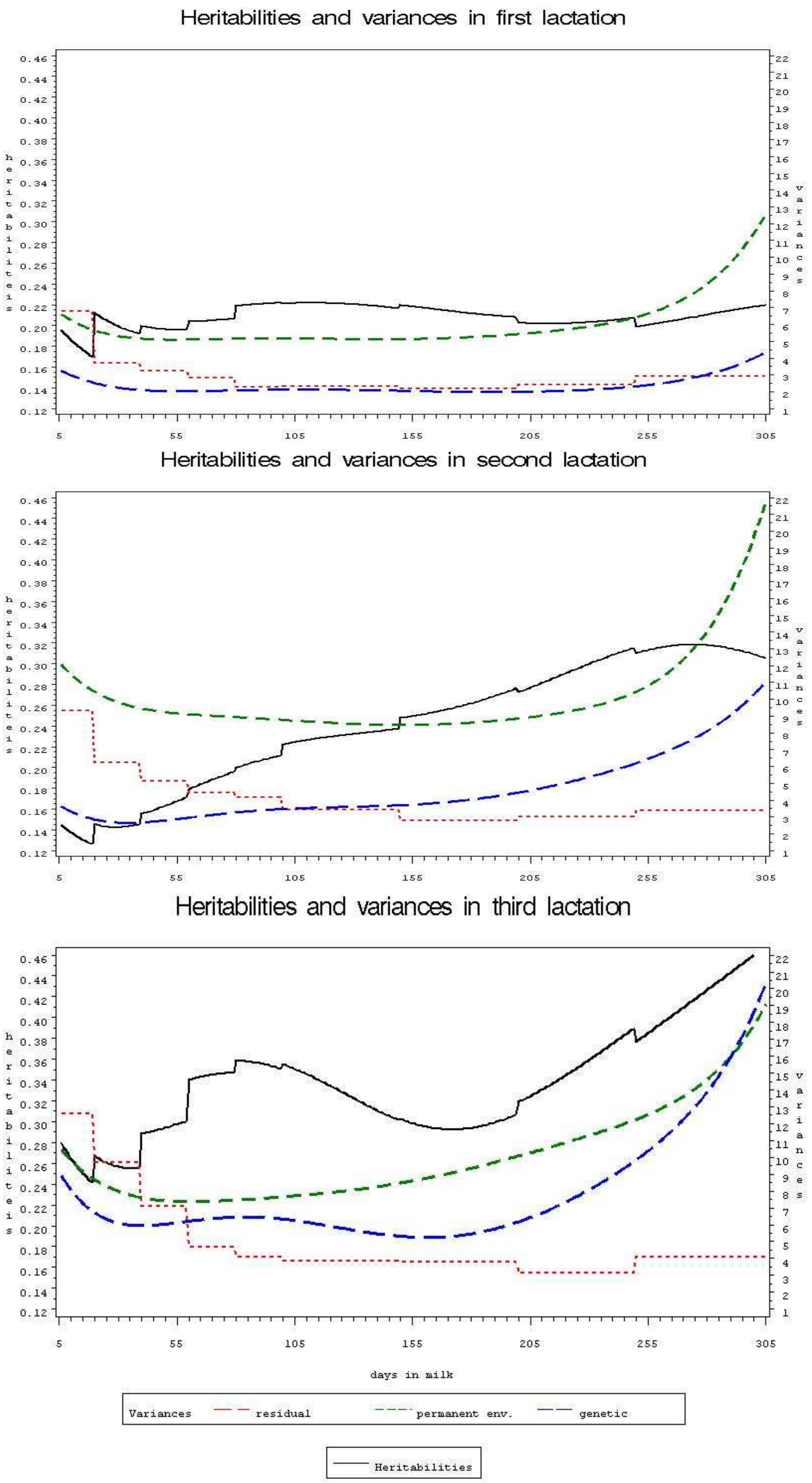
Analysis IV.

De Roos et al. (2004) mentioned that studies using random regression models often gave larger variances at the borders of lactation (e.g. *Van der Werf et al., 1998; Rekaya et al., 1999*). The increase is more apparent in the second and third lactation (*Jamrozik et al., 1997; De Roos et al., 2004*). A possible reason for this could be confounding between herd and genetic effects (*Jamrozik et al. 2001*, *Gengler et al. 2001). De Roos (2004)* found the decrease of genetic variances when the effect of the herd was modelled by random regression.

In analysis V (sample C), the genetic variance was almost constant and lower in the second and third lactations. On the contrary, the permanent environmental variances and residual variances were the highest from the all analyses. Consequently, the heritabilities were very low. The results were probably biased by selection. The creation of samples A and B was consistent with Interbull recommendations (*Interbull, 2001*) as well as with procedures used by other authors (e.g. *Swalve, 1995, De Roos et al., 2004*).

Comparing results from analyses I and II, i.e. for sets sampled by the same procedure, differences occurred mainly in the second lactation milk yield. The differences were caused by different proportions of variances.

Fig. 4: Variances in 1st, 2nd and 3rd lactations for milk production (results from analysis IV)



Heritabilities and variances in first lactation

Heritabilities and variances in second lactation

Heritabilities and variances in third lactation

The highest genetic correlations were found between the first and second lactation and it was out of accord with the results found by *Dědková and Wolf (2001)* in Holstein cattle in Czech Republic. The estimated genetic correlations from analysis V were consistent with most published results. *De Roos (2004)* states that random regression models often provide lower genetic correlations between lactation in comparison with multi-trait models. The low genetic correlation found in this study could be caused by the low number of records in the third lactation. This assumption is supported by results from analysis V.

The higher number of the time periods of residual variances in analysis IV caused a decrease of the total variance. The proportions of particular variances and the heritabilities remained the same as in the other analyses. This approach will be the bases for the construction of covariance functions for genetic parameter estimates that are now frequently used procedures besides of the direct estimation method (*Liu et al., 2000; Misztal et al., 2000; De Roos et al., 2004*). The estimates from REML as well as AI-REML corresponded to the results of the Gibbs sampling method. *Reents (1995)* and *Misztal et al. (2000)* dealt with comparison of these methods.

Conclusions

The presented estimates of genetic parameters for milk yield of the Czech Spotted population correspond mostly to published results for dairy cattle. The results were more influenced by the change in model than by the data sample.

Acknowledgments

Thanks are due to Prof. I. Misztal (University of Georgia, USA) and Dr. J. Jamrozik (CGIL, University of Guelph, Canada) for making available program packages. The authors acknowledge the financial support from the project NAZV QC 1235 of the Ministry of Agriculture of the Czech Republic.

References

- ***De Roos, A.P.W, et al. (2004):*** Random herd curves in a test-day model for milk, fat, and protein production of dairy cattle in The Netherlands. J. Dairy Sci., 87, 2693 – 2701.
- ***Dědková, L. et Wolf, J. (2001):*** Estimation of genetic parameters for milk production traits in Czech dairy cattle populations. Czech J. Anim. Sci., 46, 292-301.
- ***Emmerling, R. et al. (2002):*** Multiple lactation random regression test-day model for Simmental and Brown Swiss in Germany and Austria. Interbull. Bull., 29, 111 – 117.
- ***Freeman, A. E. (1998):*** Dairy cattle breeding. Proc. 6th World Congr. Genet. Appl. Livest. Prod., Armidale, Australia, Vol. 23, 293 - 294.
- ***Gengler, N. et al. (2001):*** Indirect estimation of (co)variance functions for test-day yields dutiny first and second lactation in the United States. J. Dairy Sci. 84 (http://www.adsa.org/jds/jds9296E/jds9296.html).
- ***Interbull (2001):*** Interbull guidelines for national & international genetic evaluation systems in dairy cattle with focus on production traits. Interbull Bull., 28.
- ***Jakobsen et al.(2002):*** Genetic parameters for milk production and persistency for Danish Holsteins estimated in random regression models using REML. J. Dairy Sci., 85, 1607 – 1616.
- ***Jamrozik, J. et al. (1997):*** Multiple trait random regression test day model for production traits. Intebrull Bull., 16, 43 – 47.

- ***Jamrozik, J. et al. (1998):*** Genetic parameters for production traits and somatic cell score of Canadian Holsteins with multiple trait random regression model. Proc 6th World Congr. Genet. Appl. Livest. Prod., Armidale, Australia, Vol. 23, 303 – 306.
- ***Jamrozik, J. et al. (2001):*** Bayesian estimation of genetic parameters for test day records in dairy cattle using linear hierarchical models. Livest. Prod. Sci., 71, 223 – 240.
- ***Lidauer, M. et al. (2000):*** Multiple-trait random regression model for all lactation. Interbull Bull., 25, 81 -86.
- ***Liu, Z. et al. (2000):*** Estimating parameters of a random regression test day model for first three lactation milk production traits using the covariance function approach. Interbull Bull., 25, 74 – 80.
- ***Meyer, K. (2001):*** Estimating genetic covariance functions assuming a parametric correlation structure for environmental effects. Genet. Sel. Evol., 33, 557 – 585.
- ***Misztal, I. et al. (2000):*** Strategies for estimating the parameters needed for different test-day models. J. Dairy Sci., 83, 1125 – 1134.
- ***Reents, R. et al. (1995):*** Estimation of genetic parameters for test day records of somatic cell score. J. Dairy Sci., 78, 2847 – 2857.
- ***Rekaya, R. et al. (1999):*** Use of test day yields for the genetic evaluation of production traits in Holstein-Friesian cattle. Livest. Prod. Sci., 57, 203 – 217.
- ***Samoré, A.B.P. et al. (2002):*** Genetic parameters for production traits and somatic cell scores estimated with a multiple trait random regression model in Italia Holsteins. Proc. 7th World congr. Genet. Appl. Livest. Prod., Montpellier, France, 29, 63 -66.
- ***Swalve, H.H. (1993):*** Genetic parameters for test-day models. Aarmus, 19. – 23. 8. 1993 (Interbull Bulletin No 8).
- ***Swalve, H.H. (1995):*** Test day models in the analysis of dairy production data – a review. Arch. Tierz., 38, 591 – 612.
- ***Van der Werf, J.H.J. et al. (1998):*** The use of covariance functions and random regressions for genetic evaluation of milk production based on test day records. J. Dairy Sci., 81, 3300 – 3308.