

# Relationship between lactation curve function and phenotypic variance in random regression Test Day models

N.P.P. Macciotta<sup>1</sup>, D. Vicario<sup>2</sup>, A. Cappio-Borlino<sup>1</sup>

<sup>1</sup>Dipartimento Scienze Zootecniche, Università di Sassari, Italy

<sup>2</sup>Associazione Nazionale Allevatori Pezzata Rossa Italiana, Udine, Italy

*Corresponding author:* Nicolò Pietro Paolo Macciotta. Dipartimento Scienze Zootecniche. Via E. De Nicola 9, 07100 Sassari, Italy – Tel: +39 079 229298 – Fax: +39 079 229302 – Email: macciott@uniss.it

**RIASSUNTO** – *Relazione tra modello della curva di lattazione e stima della varianza fenotipica nei modelli Test Day di regressione random.* I dati della produzione giornaliera di latte di 780 vacche primipare di razza pezzata Rossa Italiana sono stati analizzati con diversi modelli di regressione *random*. Il tipo di modello influenza notevolmente la stima della varianza fenotipica, con le funzioni polinomiali che mostrano forti distorsioni delle stime nella parte iniziale e finale della lattazione. Anche la forma della curva di lattazione influenza i risultati, con un miglioramento delle stime quando l'analisi viene condotta separatamente entro curve di forma omogenea (standard o atipica).

**KEY WORDS:** phenotypic variance, random regression models, lactation curve.

**INTRODUCTION** – In Random Regression models (RRM), the most updated version of Test Day (TD) models, the lactation curve is split into a fixed average curve and a random animal specific part (deviation from the average curve) (Schaeffer, 2004). The variance component of the RR coefficients determines the (co) variance function of each pair of days in milk (DIM) (Pool and Meuwissen, 2000). Very different patterns of variance functions have been reported in literature, and several authors pointed out a possible rule of the type of function chosen as RR sub-model and data structure (Kettunen *et al.*, 2000; Meyer, 1998). Aim of this work is to investigate some possible reasons for such results, in particular the effects of the mathematical function and of the possible occurrence of different shapes of lactation curve (regular and atypical).

**MATERIALS AND METHODS** – Data were 6284 TD records of milk yield belonging to 780 Italian Simmental heifers in 53 herds, extracted from the historical archive of the breed. Edits were on number of test per animal (>6), calving season (1, January-March), number of tests per herd (>70), type of test (A4 or A6). Average DIM at first TD was  $21 \pm 10$ . Data were analysed with the following RR phenotypic model:

$$y_{ij} = x'_{ij}\beta + \phi'_{ij(m)} k_{i(m)} + e_{ij} \quad [1]$$

where  $y_{ij}$  is TD milk yield  $j$  of animal  $i$ ,  $x'$  is the incidence row vector of fixed effects  $\beta$  (herd, calving year, lactation stage),  $\phi'_{ij(m)}$  is the row vector of DIM functions specific to the  $m$ -th sub-model chosen,  $k_{i(m)}$  is the vector of RR coefficients,  $e_{ij}$  is the random residual. The phenotypic variance function was estimated as:

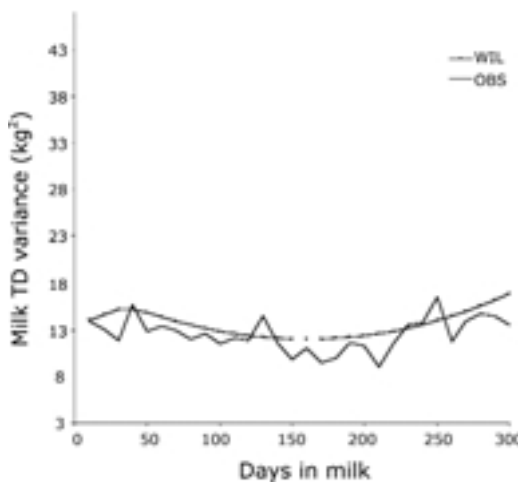
$$V = \phi'_{ijm} K \phi_{ijm} + \sigma_e^2 \quad [2]$$

where  $K$  is the estimated (co)variance matrix of random regression coefficients,  $\sigma_e^2$  is the residual variance. The  $m$  functions used as RR sub-models were: the Wilmink function (WIL), the Ali and Schaeffer (AS) five

parameter model, a fourth order orthogonal Legendre polynomial (LEG). The lactation curve shape (standard or atypical) was assessed on the basis of the sign of the WIL function parameters. Goodness of fit was evaluated by comparing variance patterns predicted by [2] and that observed between residuals when only fixed effects of model [1] were fitted.

**RESULTS AND CONCLUSIONS** – The relationship between mathematical function used as RR sub-model and estimated variance can be observed in figure 1, where patterns of the observed variance (OBS) and of those estimated with the different RRM are reported. The pattern of OBS variance is in agreement with a previous report for dairy cattle (Pool and Meuwissen, 2000). Both WIL and AS estimates roughly follow the OBS pattern even if with a tendency to overestimate for WIL and to under estimate for AS in the first part of lactation.

Figure 1. Phenotypic variance observed and estimated with different RR models in the whole data set.



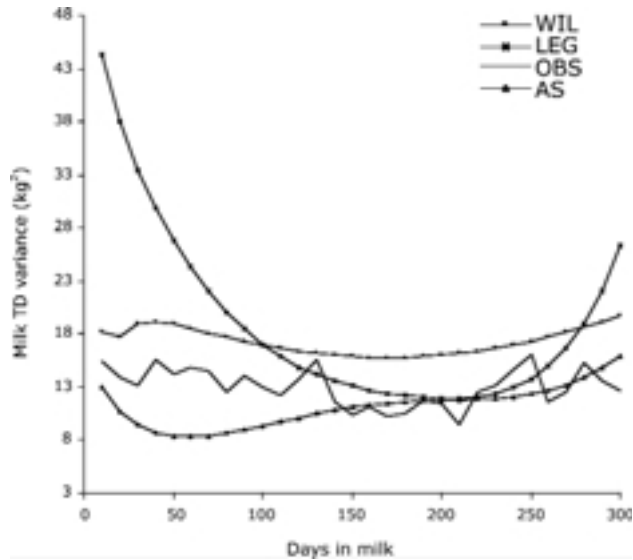
LEG estimates shows very high values at the ends of the lactation trajectory, especially in the first part. Actually several authors pointed out that data point at the edges of the lactation trajectory have a relatively large impact on the regression coefficient estimates when polynomials are used as the covariance function (Meyer, 1998; Robert-Graniè *et al.*, 2004). The effect of the presence of different types of curve shape (20% of atypical curves in the whole data set of this study) on RRM can be observed in table 1, where the **K** matrices estimated with the WIL function in the whole data set and the data set with only standard curve are reported.

Table 1. Estimated (Co)variance matrices of RR coefficients for the RRM including the WIL function ( $y = a + bt + ce^{kt}$ ) in the whole and the reduced (only standard shapes) data sets.

parameter	All curves			Standard curves		
	a	c	b	a	c	b
a	15.482			14.695		
c	-17.903	67.998		-11.316	34.148	
b	-0.041	0.038	0.0002	-0.040	0.028	0.0002

In the whole data set, the existence of a mixture of standard and atypical shapes results in an increased variance of the  $c$  parameter, the one that regulates the raise of milk yield to the lactation peak in the first part of the curve. On the basis of this result, an improvement of variance estimates can be obtained by developing the analysis separately within each shape of curve. This finding is confirmed by the comparison between figure 1 with figure 2, that reports the OBS and estimated variance for a reduced data set that included only curves with standard shape.

Figure 2. Phenotypic variances for a reduced data set of standard curves.



Results of the present study suggest a particular care when estimating phenotypic (co)variance by Random Regression TD models. The choice of the mathematical function and the occurrence of different shapes of the lactation curve may affect the estimates. These results should be checked in larger data set sets and, if confirmed, possible consequences on the estimation of genetic component of variance should be investigated.

**ACKNOWLEDGEMENTS** – Research supported by MIUR, grant PRIN 2003.

**REFERENCES** – **Kettunen**, A., Mantysaari, E., Poso, J., 2000. Estimation of genetic parameters for daily milk yield of primiparous Ayrshire cows by random regression test day models. *Livest. Prod. Sci.* 66: 251-261. **Meyer**, K., 1998. Estimating covariance function for longitudinal data using a random regression model. *Genet. Sel. Evol.* 30: 221-240. **Pool**, M.H., Meuwissen, T.H.E., 2000. reduction of number of parameters needed for a polynomial random regression test day model. *Livest. Prod. Sci.* 64: 133-145. **Robert-Graniè**, C., Foulley, J.L., Maza, E., Rupp, R., 2004. Statistical analysis of somatic cell scores via mixed model methodology for longitudinal data. *Anim. Res.* 53: 259-273. **Schaeffer**, L.R., 2004. Applications of random regression models in animal breeding. *Livest. Prod. Sci.* 86: 35-45.