

## Genetic Analysis of Milk Yield in Iranian Holstein Cattle by the Test Day Model

### Research Article

Y. Naderi<sup>1\*</sup>, N. Emam Jome Kashan<sup>1</sup>, R. Vaez Torshizi<sup>2</sup> and M. Amin Afshar<sup>1</sup>

<sup>1</sup> Department of Animal Science, Science and Research Branch, Islamic Azad University, Tehran, Iran

<sup>2</sup> Department of Animal Science, Faculty of Agriculture, Tarbiat Modares University, Tehran, Iran

Received on: 27 Nov 2014

Revised on: 15 Mar 2015

Accepted on: 31 Mar 2015

Online Published on: Dec 2015

\*Correspondence E-mail: [y.naderi@iau-astara.ac.ir](mailto:y.naderi@iau-astara.ac.ir)

© 2010 Copyright by Islamic Azad University, Rasht Branch, Rasht, Iran

Online version is available on: [www.ijas.ir](http://www.ijas.ir)

### ABSTRACT

Using monthly test day records the genetic parameters of Iranian Holstein cattle in first lactation were studied. Data of 277400 test-day milk records from 65320 cows and 2210 sires were analyzed by an animal random regression model using restricted maximum likelihood methodology. The model included herd-test-date, interaction between year-season of calving, days in milk (linear and quadratic) and dam age (linear and quadratic) as fixed effects and random regression coefficients for additive genetic and permanent environmental effects. The average of 305 days milk yield was 9760 ( $\pm 1324$ ) kilogram. Differences of milk yield among provinces were significant ( $P < 0.05$ ). The average of heritability estimates of milk was 0.50. The genetic correlations between adjacent test-day records were high and decreased with increase in interval between tests.

**KEY WORDS** legendre polynomials, milk yield, random regression.

### INTRODUCTION

The production level of milk yield is the most important factor in the economic efficiency of a dairy farm. The production level of a dairy cow is determined by the joint effects of genetic and environmental factors. The estimation of co-variance components is essential for the development of genetic evaluation systems based on test day (TD) yields. Genetic evaluation of dairy sires and cows has evolved immensely over the years. From the initial stages when simple dam-daughter comparisons were made, rapid advances in computer hardware improvements in computing algorithms have made possible to implement modern methods for analysis. The use of the test day model approach allows a more detailed statistical model to be developed, which accounts for environmental variation specific to individual test day yields and genetic effects associated with individual animals. It offers the opportunity to directly account for

short-term environmental factors specific to individual yields such as gestation period. The test day model also overcomes the need to predict 305-day yields or for projection of incomplete lactations (Hammami *et al.* 2008). The advantages of random regression test-day models over other approaches of evaluating test-day records include (Hammami *et al.* 2008) the following advantages:

1. A continuous treatment of observation over time and incorporation of heterogeneous (co)variances among measures along time (including days that were not sampled) with a potentially reduced number of parameters compared with the multiple trait approach (Lidauer *et al.* 2003).
2. Every record contributes information at the value of the control variable at which it is measured. Arbitrary or inappropriate corrections for the differences in the control variable are therefore unnecessary (Van der Werf, 1997).
3. With regards to estimation of variance components, random regression models facilitate parsimonious descriptions

of changing and potentially complex covariance structures, thereby utilizing the data more efficiently and generating breeding values of higher accuracies (Jamrozik and Schaeffer, 1997).

4. The random regression model also allows cows to be evaluated on the basis of any number of test day records during lactation. Related to this, as only 8 to 10 test day yields per cow per lactation may be collected, this could result in lower costs of recording (Schaeffer *et al.* 2000).

5. The random regression model for test day yields can account more precisely for environmental factors that could affect cows differently during lactation (Schaeffer and Dekkers, 1994).

6. Due to emphasis on more yield information, a random regression model results in top animals which are less related and hence results in reduced rates of inbreeding compared to lactation models (Mrode and Coffey, 2008).

Theoretically, any function can be used in random regression model as a basic function (Meyer, 2005). Legendre polynomials are the most common, because the correlations between parameters are lower than with other functions (Kirkpatrick *et al.* 1990). Orthogonal polynomials are able to model lactation curves for a range of covariance structures, but they also have undesirable properties (Misztal, 2006). The fit at the extremes of the trajectories may be poor especially for high orders of fit (Meyer, 2005) and there may be problems of convergence for large data sets. Cubic and quadratic orders of Legendre polynomials for additive genetic (AG) and permanent environmental (PE) effects, respectively, were selected (Kirkpatrick *et al.* 1990). Homogeneous residual variance was assumed throughout the lactation. The use of random regression (RR) model makes it possible to study differences in TD records over time, and to better understand the genetics of lactation (Swalve and Guo, 1999).

While being conceptually appealing, practical applications of random regression models in animal breeding have been plagued by problems associated with large numbers of parameters to be estimated, poor polynomial approximation and therefore the necessity of analyzing much larger sets of data, implausible estimates at the extremes values and associated high computational requirements (Meyer, 2005; Misztal, 2006).

Milk production is influenced by exactly the same environmental factors whether a test day model or lactation model is used in genetic evaluation. However, for a test day model, the stage of lactation is an important consideration, because of the curvilinear relationship that exists between the stage of lactation and milk production (Swalve and Guo, 1999). The test day models often use types of covariates or mathematical functions, in a regression, to account for stage of lactation.

The adoption of test day model over the lactation model replaced the use of herd-year-season (HYS) with herd-test-date (HTD). The HTD accounts for the effects of herd and the year and the season of production whereas HYS effect is commonly used to account for the effects of the individual herd, the year, and the season of calving and interactions among them. With a test day model, further effects that can be fitted in the analysis include age at calving, parity and pregnancy (Swalve and Guo, 1999).

The random regression test day model can account for many environmental factors that could affect cows differently during the lactation (Schaeffer and Dekkers, 1994). The lactation curve is split into two parts: a fixed part (average lactation curve) and a random animal specific part (deviation from the average curve).

The objective of the present study was to estimate the genetic parameters of milk yield in different test days (TD) of the first lactation of Iranian Holstein cows.

## MATERIALS AND METHODS

The data of milk in different test days of Holstein dairy cows were provided by the Animal Breeding Center of Iran. The study included TD records 5 to 300 days in milk collected from 2001 to 2014 of cows with at least one known parent and age at first calving between 18 to 45 months (Table 1). The data of cows with the first test day at least 60 days after parturition and TD intervals less than 15 days were discarded. The significant fixed effects on milk yield were determined using the GLM procedure of SAS (SAS, 2005). The genetic parameters of milk yield were estimated by restricted maximum likelihood using a random regression animal model (Meyer, 2007). A single-trait model with Legendre polynomials describing the trait curve in the population was used to fit data. The days in milk values transformed to the range of -1 to +1 using the formula 1 (Meyer, 2000):

$$d_t^* = -1 + 2(d_t + d_{min}) / (d_{max} - d_{min})$$

Where:

$d_{min}$  and  $d_{max}$ : minimum and maximum values for the days in milk. For the  $t$ -th standardized days in milk ( $d_t^*$ ), the  $k$ -th polynomials is given as (Meyer, 2000):

$$\phi(d_t^*)_k = \frac{1}{2^k} \sqrt{\frac{2k+1}{2}} \left( \sum_{m=0}^k (-1)^m \binom{k}{m} \binom{2k+1}{r} \right) (d_t^*)^{r-2m}$$

Where:

$k/2 = (k-1)/2$  if  $k$ : odd.

$m$ : index number needed to determine the  $k$ -th polynomial (Meyer, 2000).

The model equation for the analysis of data can be expressed as (Meyer, 2000):

$$Y_{ijk} = F_i + \sum_{k=0}^n \theta_{jtk} \beta_k + u_j + pe_j + e_{ijk}$$

Where:

$Y_{ijk}$ : test day milk yield record.

$F_i$ : fixed effects in the model (including herd-test date, year-season of calving, days in milk and age of dam as co-variate).

$\beta_k$ : fixed regression coefficients to describe the curve of trait for the whole population.

$u_j$  and  $pe_j$ : refer to the  $j$ -th animal additive genetic and permanent environmental effects, respectively.

$\theta_{jtk}$ :  $k$ -th Legendre polynomial for the standardized time  $t$  of milking (days in milk).

$n$ : order of Legendre polynomials for fitting traits curve for the population.

$e_{ijk}$ : residual effect describing the unknown effects on traits. The days in milk were partitioned into 10 equal periods of about 30 days and an independent residual variance structure was assumed. In matrix notation, the model is:

$$y = Xb + Qu + Zp + e$$

Where:

$y$ : vector of test day milk yield.

$b$ : vector of solution for fixed effects and regression coefficients of milk yield on days in milk.

$u$  and  $p$ : vectors of animal additive genetic and permanent environmental effects, respectively.

$X$ ,  $Q$  and  $Z$ : incidence matrices relating effects to the observations. The co-variance structure was assumed as:

$$\text{var} \begin{bmatrix} u \\ p \\ e \end{bmatrix} = \begin{bmatrix} G & 0 & 0 \\ 0 & P & 0 \\ 0 & 0 & R \end{bmatrix}$$

Where:

$G$ :  $A \sigma_a^2$ .

$P$ :  $I \sigma_p^2$ .

$R$ : diagonal matrix containing the 10 constant variances of residuals and  $A$  is the numerator relationship matrix among animals.

$\sigma_a^2$  and  $\sigma_p^2$ : variances of additive genetic and permanent environmental effects, respectively.

$I$ : denotes an identity matrix.

Genetic analysis was performed using the WOMBAT software with REML procedure for estimation of co-variance components (Meyer, 2007).

The variance components and genetic parameters were estimated using random regression model and restricted maximum likelihood method.

**Table 1** Numbers of records and means and standard deviations (SD) of milk yield

Test day records	277400
Mean of milk yield $\pm$ (SD) (kg)	32.0 $\pm$ 7.6
Cows with record	65320
Number of sires	2210
Number of herds	82
Average daughters per bull	30
Average test-day records per cow $\pm$ (SD)	5.0 $\pm$ 0.2
Number of animals in pedigree file	74860

The best model determined through the logarithm of maximum likelihood, the Akaike's information criterion (AIC) and Bayesian information criterion (BIC). The log-likelihood of a model given the data reflects the overall fit of the model and smaller values indicate poor fit. Because the likelihood tends to favor complex models with many parameters, the more conservative AIC and BIC has been suggested (Burnham and Anderson, 2002; Schaeffer and Jamrozik, 2008). The AIC and BIC penalizes models for the addition of parameters, and thus selects a model that fits well but has a minimum number of parameters (i.e. simplicity and parsimony) (Burnham and Anderson, 2002; Schaeffer and Jamrozik, 2008). All of the above criteria were considered and the model with the highest log-likelihood, lowest AIC and BIC, was considered to be the most appropriate. Third order Legendre polynomials was selected to describe the lactation curve of the animals with respect to the genetic and permanent environmental effects in the model (Table 2).

## RESULTS AND DISCUSSION

Genetic variance of milk production was lowest on the second to fourth test days, and increased from the fourth to the tenth test day (Figure 1). The trend of additive genetic variance of milk production was consistent with other studies (Druet *et al.* 2003).

The permanent environmental variance was higher in mid and the last two months of lactation (Figure 1). The trend of this effect was consistent with other studies (Mayeres *et al.* 2004; Zavandilova *et al.* 2005). The heterogeneity of the residual variance in days 5 to 305 of lactation was modeled using 10 intervals with equal lengths. The residual variance was almost constant across the lactation (Figure 1).

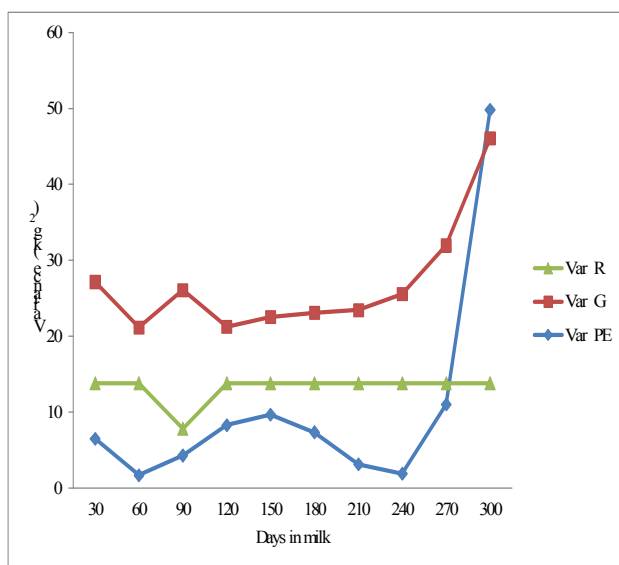
The estimated heritability of milk production decreased in our study from the 2<sup>nd</sup>-5<sup>th</sup> test day and was highest in the 8<sup>th</sup> test day (Figure 2). The trend of changes of heritability was consistent with other studies (Melo *et al.* 2007; Biassus *et al.* 2010).

**Table 2** The criteria used to select the best function for milk production (the best model is indicated in bold type)

Model	Parameters	BIC	AIC	Log <sub>L</sub>
1	3	3048.532	3040.702	-3037.702
2	7	3023.724	3005.460	-2998.460
<b>3</b>	<b>13</b>	<b>3032.485</b>	<b>2998.576</b>	<b>-2985.576</b>
4	21	3143.898	3089.138	-3068.138
5	31	3186.321	3105.508	-3074.508
6	43	3243.216	3131.152	-3088.152

BIC: Bayesian information criterion and AIC: Akaike's information criterion.

The heritability estimates of first lactation milk yield estimated by random regression models is reported from 0.14 (Strabel and Misztal, 1999) to 0.51 (Olori *et al.* 1999). Some authors reported higher heritabilities at the beginning and at the end of lactation (Olori *et al.* 1999; Melo *et al.* 2007).



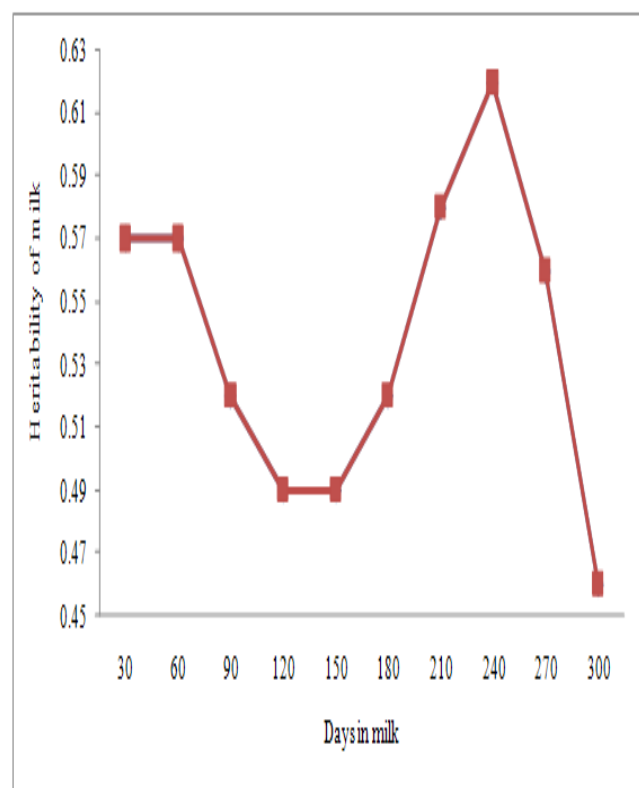
**Figure 1** Trajectory of genetic (G), permanent environmental (PE) and residual (R) variances in milk production by different test days

Genetic and phenotypic correlations between test day records were high, but decreased with increasing time differences between test days (Table 3). The trend of changes in genetic and phenotypic correlation was consistent with other studies (Druet *et al.* 2003).

The high genetic correlations between different test days show that these changes are caused by genetic factors and the impact of environmental factors between the test days is not high (Figure 3).

For the south African Holstein and Jersey cow populations, Mostert *et al.* (2006) reported that genetic correlations between test day milk yields of different lactations differed from one. This study led to the implementation of a fixed regression test day model, but recommended the use of random regression functions in the genetic evaluation of south African dairy cattle.

The attempt to improve the accuracy of estimated breeding values, reduce the generation interval and boost response to selection for dairy cattle and the quest to provide more comprehensive management information to dairy farmers are stimulating interest in advancing the conceptual framework of the TDM.



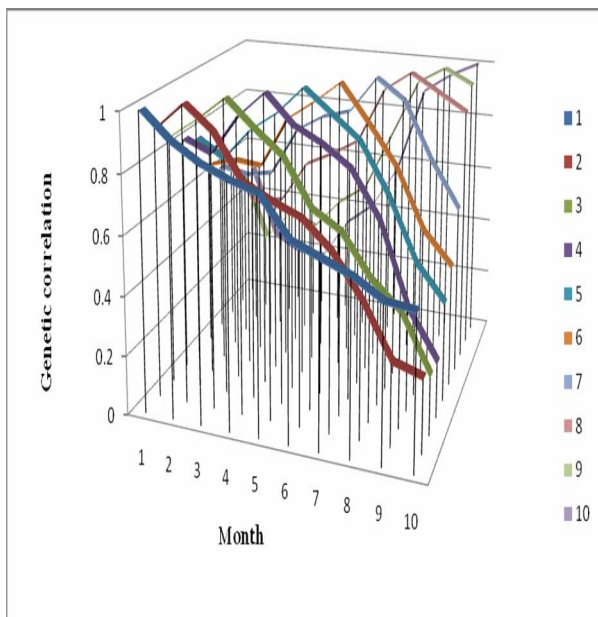
**Figure 2** Trajectory of estimated heritability for milk production in different test days

The random regression model approach probably yields the potential to realize these benefits for the Iranian Holstein dairy cattle genetic evaluation program. The future research should be focused on 1) defining the random regression model to be implemented 2) investigating the environmental effects to be included in the model and 3) estimating the covariance structure among observations and genetic parameters for traits to be included in the breeding program.

**Table 3** Additive genetic (below the diagonal) and phenotypic (above the diagonal) correlation coefficients for milk production in different test days

DIM	30	60	90	120	150	180	210	240	270	300
30	-	0.84	0.8	0.71	0.7	0.63	0.5	0.47	0.44	0.45
60	0.91	-	0.91	0.68	0.64	0.6	0.42	0.4	0.28	0.31
90	0.85	0.92	-	0.84	0.78	0.55	0.51	0.32	0.2	0.09
120	0.81	0.77	0.91	-	0.92	0.87	0.76	0.64	0.3	0.15
150	0.78	0.72	0.83	0.9	-	0.88	0.71	0.61	0.2	0.16
180	0.65	0.68	0.67	0.85	0.92	-	0.77	0.71	0.3	0.22
210	0.61	0.59	0.61	0.78	0.84	0.87	-	0.95	0.7	0.51
240	0.57	0.45	0.47	0.62	0.67	0.73	0.93	-	0.9	0.79
270	0.51	0.27	0.38	0.34	0.45	0.52	0.73	0.94	-	0.93
300	0.5	0.24	0.19	0.18	0.33	0.41	0.57	0.88	0.95	-

DIM: days in milk.

**Figure 3** Additive genetic correlation for milk production in different test days

These are the requisite steps towards adoption of a random regression model framework for analysis of dairy test day records. Test-day records of cows collected in early lactation and standardized for age and season are of potential value for early genetic evaluation of bulls and cows for 305-day yields.

## CONCLUSION

We propose that the Animal Breeding Center uses the random regression and multitrait models for breeding value prediction of waiting proof bulls. Comparison of the phenotypic performance and predicted breeding value of animals selected to the results of two models can provide the information to judge the validity of the different models.

## ACKNOWLEDGEMENT

The authors are grateful to the Animal Breeding Center of Iran for supplying the research data.

## REFERENCES

- Biassus I.D.O., Cobuci J.A., Costa J.A., Rorato P.R.N., Neto J.B. and Cardoso L.L. (2010). Persistence in milk, fat and protein production of primiparous Holstein cows by random regression models. *Brazilian J. Zootec.* **39**(12), 2617-2624.
- Burnham P.K. and Anderson R.D. (2002). Model Selection and Inference: A Practical Information Theoretic Approach. Springer Verlag, New York, NY.
- Druet T., Jaffrezic F. and Ducrocq V. (2003). Modeling of lactation curves and estimation of genetic parameters for first lactation test-day records of French Holstein cows. *J. Dairy Sci.* **86**, 2480-2490.
- Hammami H., Rezik B., Soyeurt H., Ben Gara A. and Gengler N. (2008). Genetic parameters for Tunisian Holsteins using a test-day random regression model. *J. Dairy Sci.* **91**, 2118-2126.
- Jamrozik J. and Schaeffer L. (1997). Estimates of genetic parameters for a test day model with random regression for yield traits of first lactation Holsteins. *J. Dairy Sci.* **80**, 762-770.
- Kirkpatrick M., Lofsvold D. and Bulmer M. (1990). Analysis of the inheritance, selection and evolution of growth trajectories. *Genetics.* **124**, 979-993.
- Lidauer M., Mantisaari E.A. and Strandén I. (2003). Comparison of test-day models for genetic evaluation of production traits in dairy cattle. *Livest. Prod. Sci.* **79**, 73-86.
- Mayeres P., Stool J., Boormann J., Reents R. and Gengler N. (2004). Prediction of daily milk, fat and protein production by a random regression test-day model. *J. Dairy Sci.* **87**, 1925-1933.
- Melo C., Packer I.U., Costa C.N. and Machado P.F. (2007). Genetic parameters for test day milk yields of first lactation Holstein cows by random regression models. *Animal.* **1**, 325-334.
- Meyer K. (2000). Random regressions to model phenotypic variation in monthly weights of Australian beef cows. *Livest. Prod. Sci.* **65**(1), 19-38.
- Meyer K. (2005). Advances in methodology for random regression analyses. *Australian J. Exp. Agric.* **45**, 847-858.
- Meyer K. (2007). WOMBAT - a tool for mixed model analysis in quantitative genetics by REML. *J. Zhejiang Univ. Sci.* **8**, 815-821.
- Misztal I. (2006). Properties of random regression models using linear splines. *J. Anim. Breed. Genet.* **123**, 74-80.
- Mostert B.E., Theron H.E., Kanfer F.H.J. and Van Marle-Koster E. (2006). Adjustment for heterogeneous variances and a calv-

- ing year effect in test-day models for national genetic evaluation of dairy cattle in South Africa. *South African J. Anim. Sci.* **36**, 165-174.
- Mrode R. and Coffey M. (2008). Understanding cow evaluations in univariate and multivariate animal and random regression models. *J. Dairy Sci.* **91**, 794-801.
- Olori V.E., Hill W.G., McGuirk B.J. and Brotherstone S. (1999). Estimating variance components for test day milk records by restricted maximum likelihood with a random regression animal model. *Livest. Prod. Sci.* **61**, 53-63.
- SAS Institute. (2005). SAS<sup>®</sup>/STAT Software, Release 9.1. SAS Institute, Inc., Cary, NC. USA.
- Schaeffer L.R. and Dekkers J.C.M. (1994). Random regressions in animal models for test-day production in dairy cattle. Pp. 443-446 in Proc. 5<sup>th</sup> World Congr. Genet. Appl. Livest. Prod., Guelph, Canada.
- Schaeffer L.R., Jamrozik J., Kistemaker G.J. and Van Doormaal B.J. (2000). Experience with a test-day model. *J. Dairy Sci.* **83**, 1135-1144.
- Schaeffer L.R. and Jamrozik J. (2008). Random regression models: a longitudinal perspective. *J. Anim. Breed. Genet.* **125**, 145-146.
- Strabel T. and Misztal I. (1999). Genetic parameters for first and second lactation milk yield of Polish black and white cattle with random regression test-day models. *J. Dairy Sci.* **82**, 2805-2810.
- Swalve H.H. and Guo Z. (1999). An illustration of lactation curves stratified by lactation yields within herd. *Arch. Tierz.* **42**, 515-525.
- Van der Werf M. (1997). Random Regressions in Animal Breeding. Course notes available at: <http://www.personal.une.edu.au/~jvanderw/CFcoursenotes.pdf>. Accessed Dec. 2009.
- Zavandilova L., Jamrozik J. and Schaeffer L.R. (2005). Genetic parameters for test day model with random regression for production traits of Czech Holstein cattle. *Czech. J. Anim. Sci.* **50**, 142-154.
-