Estimation of genetic parameters for first-lactation test-day milk yield in Holstein Friesian cows fitting random regression models

Meseret S.¹, Tamir B.¹ and Negussie E.²*

¹Department of Animal Production Studies, College of Veterinary Medicine and Agriculture, Addis Ababa University, Ethiopia.
²Natural Resources Institute Finland (LUKE), Biometrical Genetics, Finland.
³Department of Animal and Range Sciences, Wolaita Sodo University, Wolaita Sodo, Ethiopia.

Received 14 March, 2015; Accepted 16 June, 2015

In this study, the effects of varying order of Legendre polynomials (LP) for permanent environmental (Pe) variance structure on the estimates of genetic parameters for first-lactation milk yield were evaluated fitting random regression (RR) test-day animal models. The total data set included 6850 test-day milk yield records from 800 first-lactation Holstein Friesian cows that calved between 1997 and 2013 and the pedigree file with total of 1779 animals. Four different random regression models (RR1, RR2, RR3 and RR4) all with second order LP for additive genetic effects but with varying order (intercept, 1st, 2nd and 3rd, respectively) of LP for modelling the Pe variances structure were tested for estimation of variance components and corresponding genetic parameters for milk yield. Variance components were estimated by average information restricted maximum likelihood method. The performances of competing RR models in the estimation of variance components were compared using estimates of log-likelihoods and the size of residual variances. Results showed that the estimates of log-likelihoods were higher and residual variances were lower for models that fitted second (RR3) and third order (RR4) LP for Pe effects. Heritability ($h^2$) and genetic correlations from RR3 and RR4 models ranged from 0.13 to 0.29 and 0.45 to 0.98, respectively. Models with lower order fits (RR1 and RR2) with either a constant or a linear term for Pe resulted in oscillatory trend for variance components and highly erratic $h^2$ estimates ranged from 0.18 to 0.52. Genetic correlations from these models were also implausible biologically indicating that models with lower order fits for the Pe effects were not robust enough to accurately model the variance structure at different stages of lactation. It is therefore suggested that at least a second or higher order polynomial fits are needed to model the Pe variance structure for the accurate estimation of genetic parameters for milk yield in first-lactation Ethiopian Holsteins.

Key words: Dairy cows, milk yield, random regression model, genetic parameters, test-day.

INTRODUCTION

Milk yield is one of the most important traits in dairy cattle. Genetic improvement in livestock is a particular cost effective technology, producing permanent and cumulative changes in performance (Wall et al., 2010). Improving milk yield through genetic selection is therefore a sustainable means of increasing farm profitability. Today, test-day models have been used for the analysis of longitudinal data or repeated records of production
traits worldwide in the evaluation of the genetic merit of dairy cows (Jensen, 2001; Intebull, 2012). Random regression model (RRM) is one of the test-day models used widely for genetic analysis of production traits in dairy cattle. The use of test-day records fitting RRM has several benefits including ability to account for environmental factors that affect cows at different stages of lactation. The use of test-day records fitting RRM has several benefits including ability to capture genetic and permanent environmental variations using Legendre polynomials functions particularly for dairy cattle data from tropical environments. Therefore, the objectives of this study were to evaluate varying order of Legendre polynomials for modelling additive genetic and permanent environmental effects and to estimate genetic parameters for test-day milk yield based on data from Holstein-Friesian cows in Ethiopia.

**MATERIALS AND METHODS**

**Data**

In this study, the data set contained 6,850 test-day milk yield records from 800 first-lactation Holstein-Friesian cows which were daughters of 149 sires that calved between 1997 and 2013 and belonging to two herds. The pedigree file had 1779 animals of which 800 animals had their own records. The data were obtained from the Ethiopian national dairy cattle milk recording database. Only test-day records between 5 and 305 days in milk (DIM) were used for the estimation of variance components. In addition, records of cows with age at calving less than 20 months or greater than 54 months were excluded. Age at calving was grouped into five classes (in months). These included cows less than 27, 28 to 33, 34 to 39, 40 to 45 and above 45 months of age at first calving. The data was edited to ensure that cows have at least five test-day records and the calving months were divided into three seasons (October-February, March-May and June-September).

**Models**

Four different RRMs were identified for the estimation of variance components and genetic parameters for test-day records of first-lactation Holstein-Friesian cows. The RR1 model was a RRM where second order Legendre polynomials was used for modelling the additive genetic variance and only the intercept term was used to model common permanent environment associated with all test-day yields of a cow.

\[ y_{ijkmn} = h_i + \sum_{r=1}^{4} b_{r} \phi(d_r) + a_g k + s y_l + h t m_m + p_r + \sum_{r=1}^{2} a_{nr} \phi(d_r) + \epsilon_{ijkmn} \]

The RR2, RR3 and RR4 were RRMs where Legendre polynomials of second order were used for additive genetic effects whilst first, second and third order was used for modeling the permanent environmental effects, respectively. The models for RR2, RR3 and RR4 can be described as follows:

**RR2**

\[ y_{ijkmn} = h_i + \sum_{r=1}^{4} b_{r} \phi(d_r) + a_g k + s y_l + h t m_m + \sum_{r=1}^{1} p_r \phi(d_r) + \sum_{r=1}^{2} a_{nr} \phi(d_r) + \epsilon_{ijkmn} \]

**RR3**

\[ y_{ijkmn} = h_i + \sum_{r=1}^{4} b_{r} \phi(d_r) + a_g k + s y_l + h t m_m + \sum_{r=1}^{1} p_r \phi(d_r) + \sum_{r=1}^{2} a_{nr} \phi(d_r) + \epsilon_{ijkmn} \]

*Corresponding author. E-mail: enyew.negussie@luke.fi.*

Abbreviations: LP, Legendre polynomials; Pe, permanent environmental; RR, random regression; h², heritability.

Author(s) agree that this article remain permanently open access under the terms of the Creative Commons Attribution License 4.0 International License.
calculated by dividing the estimated genetic covariance between days \( d_i \) and \( d_j \) by the square root of the genetic variances of the days \( d_i \) and \( d_j \). 

\[
\sigma_{\phi(d)} = \frac{\phi(d_i) G \phi(d_j)}{\sqrt{\phi(d_i) G \phi(d_i) \phi(d_j) G \phi(d_j)}},
\]

Genetic correlations: The genetic correlation between two days in lactation \( d_i \) and \( d_j \) was calculated by dividing the additive genetic covariance between days \( d_i \) and \( d_j \) by the product of the square root of the genetic variances of the days \( d_i \) and \( d_j \). 

\[
r_{\phi(d_i), \phi(d_j)} = \frac{\phi(d_i) G \phi(d_j)}{\sqrt{\phi(d_i) G \phi(d_i) \phi(d_j) G \phi(d_j)}},
\]

RESULTS AND DISCUSSION

Overall mean

Phenotypic means, standard deviations and coefficient of variation for test-day milk yield are given in Table 1. The overall average test-day milk yield was 11.1 kg. Milk yield peaked at around 35 DIM and then declined as the lactation progressed. The peak milk yield was reached earlier at about 3 to 4 weeks in early lactation which subsequently decreased as lactation progressed.

Comparison of model performances

Different models with the same fixed and random additive genetic effects but fitting either intercept, first, second or third order of Legendre polynomials for permanent environment were compared based on estimates of log-likelihoods and residual variances (Table 2). The result showed that with the increase in the order of Legendre polynomials for modelling the permanent environmental
effects there was a decrease in the residual variances and an increase in the log-likelihood estimates. The RR4 model fitting third order of Legendre polynomials for the permanent environmental effects had the lowest residual variances and highest log-likelihood estimates. This is the most preferred model as compared to the others and the least performing model was RR1 that considered a common permanent environmental effect for all test-day milk yields across the lactation trajectory. Similarly, Rekaya et al. (1999) and Cobuci et al. (2005) worked on first-lactation Holstein cows fitting a common permanent environmental effect that resulted in very large heritability estimates and negative genetic correlations between test-day records which they concluded as one of the least performing model for estimation of genetic parameters. In addition, Liu et al. (2000) and Torshizi et al. (2013) comparing different RRM s indicated that fitting higher order of random regression coefficients for the permanent environment than to the additive genetic was selected as the best for modelling milk yield in first-lactation dairy cows.

**Estimates of variance components**

Estimates of additive genetic and permanent environmental variances for milk yield during first-lactation are in Figure 1 and 2. The trend of additive genetic variances from RR1 and RR2 models were not very clear and it looks that they were slightly over estimated. Estimates of additive genetic variances from the simplest RR1 model ranged between 1.92 to 6.27 kg² and it was much higher than those from the other models. Marked decrease in the course of the genetic variance curve was observed for the first 30 days of lactation, suggesting that the models were less robust to describe the genetic variance in milk yield during the early lactation period (Cobuci et al., 2005). Rekaya et al. (1999) and Cobuci et al. (2005) have also reported that additive genetic variances estimated by the RRM when the permanent environmental effect was fitted only the intercept term were greater than that estimated by a model with equal random regression coefficients for both additive genetic and permanent environmental effects. The genetic variances estimated from RR2 model ranged from 1.20 to 2.84 kg² and this was also higher than the estimates from RR3 (1.21 to 1.74 kg²) and RR4 (1.21 to 1.78 kg²) models.

On the other hand, the estimates from the RR3 and RR4 models followed the same trend with slight differences observed in the magnitude of additive genetic variances. The additive genetic variances from RR3 and RR4 models were highest in mid lactation (between 185 to 245 DIM) and estimates were generally lower during the early stages of lactation. A much more similar additive genetic variance trend has also been reported by Torshizi et al. (2013). This is, however, on the contrary to Cobuci et al. (2005) who reported a higher additive genetic variance towards the beginning and end of lactation. The highest estimates of permanent environmental variances were observed in RR4 model at the beginning of lactation. The estimates from RR3 followed closely the estimates from RR4. On the other hand, the permanent environmental variance estimates from RR1 model were higher than all the other models along the entire course of lactation except the early lactation.

**Table 2.** Estimates of log-likelihoods and residual variances for RR1, RR2, RR3 and RR4 models.

| Random regression models | Log-likelihood | Residual variances |
|-------------------------|----------------|--------------------|
| RR1                     | -8657.49       | 2.38               |
| RR2                     | -8634.85       | 2.36               |
| RR3                     | -8618.49       | 2.33               |
| RR4                     | -8537.67       | 2.05               |

**Figure 1.** Estimates of additive genetic variances for RR1, RR2, RR3 and RR4 models for first-lactation Holstein Friesian cows across different days in milk.

**Figure 2.** Estimates of permanent environmental variances for RR1, RR2, RR3 and RR4 models for test-day milk yield for first-lactation Holstein Friesian cows across different days in milk.
Table 3. Estimates of heritability for milk yield at selected days in milk from RR1, RR2, RR3 and RR4 models in first-lactation Holstein Friesian cows.

| DIM | RR1 | RR2 | RR3 | RR4 |
|-----|-----|-----|-----|-----|
| 5   | 0.52| 0.29| 0.17| 0.13|
| 35  | 0.40| 0.19| 0.17| 0.17|
| 65  | 0.32| 0.18| 0.19| 0.19|
| 95  | 0.30| 0.24| 0.21| 0.20|
| 125 | 0.30| 0.31| 0.23| 0.23|
| 155 | 0.30| 0.36| 0.25| 0.26|
| 185 | 0.28| 0.37| 0.27| 0.27|
| 215 | 0.26| 0.33| 0.29| 0.28|
| 245 | 0.25| 0.28| 0.29| 0.28|
| 275 | 0.27| 0.25| 0.26| 0.28|
| 305 | 0.37| 0.30| 0.20| 0.21|

1DIM, days in milk.

Figure 3. Heritability estimates of test-day milk yield at different stages of lactation from RR1, RR2, RR3 and RR4 models in first-lactation Holstein Friesian cows.

Estimates of genetic parameters

Heritability

Estimates of heritabilities for the four RRM s are in Table 3 and Figure 3. The estimates of daily heritabilities showed that varying the combinations of order of Legendre polynomials for modelling additive genetic and permanent environmental clearly influenced the resulting parameter estimates. The inclusion of either the intercept term (RR1) or first order Legendre polynomials (RR2) for modelling the permanent environment variation resulted in wiggly heritability estimates along the trajectory with the highest estimates at the beginning and end of lactation. Similarly, Jamrozik and Schaeffer (1997) reported that daily heritability estimates with higher values at both beginning and end of lactations (border or wave effect) was the features of a RRM, in which the permanent environmental effect was constant during lactation. Similar results have also been reported by Rekaya et al. (1999) and Cobuci et al. (2005) for Holstein Friesian cattle population. In general, the assumption of constant permanent environment and modelling permanent environment with order of Legendre polynomials less than additive genetic effects across lactation did not permit the differentiation between variance estimates correctly. Therefore, genetic variances were overestimated and resulted in heritability estimates that are higher than those from the other models.

On the other hand, heritabilities estimated from the RR3 and RR4 models had a similar trend with slightly increasing heritability estimates followed by a slight decline towards the end of lactation which is quite expected. In these models heritability estimates along the lactation trajectory had a similar shape to the genetic variation but were less extreme at the beginning and end of the trajectory because of higher permanent environmental variances. The heritability estimates from RR3 and RR4 models were consistent and more plausible biologically. Various studies have also reported similar trends for the heritability curve with lower values at the beginning and an increase across lactation followed by slight decline towards the end of lactation (Liu et al., 2000; Druet et al., 2003; Negussie et al., 2008, Abdullahpour et al., 2010; Cobuci et al., 2011). On the contrary, Strabel et al. (2003) working in polish black and white cows reported that models with different order polynomials for additive genetic and permanent environmental effects resulted undesired shape of heritability curves.

Genetic correlations

Estimated genetic correlations between milk yields at different test-days from the different RRM s are given in Figure 4. In the RR1 model, genetic correlations ranged from -0.06 to 0.96 and in RR2 from -0.02 to 0.90. These estimates from the RR1 and RR2 seems not reasonable biologically and are in contrast to literature reports by Lidauer et al. (2003), El Faro et al. (2008), Bignardi et al. (2011) and Cobuci et al. (2011). On the other hand, the model with equal order (RR3) or higher order (RR4) of Legendre polynomials for fitting permanent environment than for additive genetic effects gave biologically reasonable genetic correlations estimates. Genetic correlations between test-days from RR3 and RR4 in general ranged from 0.37 to 0.96 and 0.38 to 0.96, respectively. Cobuci et al. (2005), Santos et al. (2013) and Torshizi et al. (2013) reported genetic correlations between test-days that were higher than 0.06, 0.47 and 0.51, respectively. In our study, genetic correlation estimates from RR3 and RR4 models were in general positive between different test-days and were higher for adjacent test-days and decreased with increasing
distance between test-days. In all cases, towards the extreme part of the lactation, a slight increase in the genetic correlations was observed. This might be due to lower number of observations at the extreme parts of the lactation and the assumption of homogenous residual variances across different stages of lactation. Overall, the RR4 model gave better estimates of genetic correlations especially at the extreme periods of lactation and the unexpected slight increment towards the end of lactation was not pronounced. Santos et al. (2013) has also reported slight increment of genetic correlations at the ends of lactation.

Conclusions

Determining the order of Legendre polynomial fits in test-day models is crucial for the accurate estimation of genetic parameters. The results from this study clearly showed that a constant or lower order Legendre polynomial fits for permanent environment than for the additive genetic effects resulted in highly erratic variance component estimates, particularly at both extremes of the lactation curve. As a result, RR1 and RR2 models, predicted basically different shape of variance components implying that they are less robust in describing genetic and environmental variances across the different stages of lactation. This could be due to the inability of such models to differentiate between the variance estimates correctly across different stages of lactation resulting in unrealistic heritability estimates. Therefore, the indications from this study are that at least a second or higher order Legendre polynomial fits are needed to model the permanent environmental variance structure for the accurate estimation of genetic parameters for milk yield in first-lactation Ethiopian Holstein Friesian.

Conflict of interests

The author(s) did not declare any conflict of interest.

ACKNOWLEDGMENTS

The authors would like to acknowledge the Ethiopian Ministry of Agriculture, National Artificial Insemination Center for providing the data. The authors are also grateful to the Natural Resources Institute Finland, Department of Biometrical Genetics for their support in the data analyses. Addis Ababa University and Wolaita Sodo University are acknowledged for their financial support.

REFERENCES

Abdullahpour R, Shahrbabak MM, Nejati-Javaremi A, Torshizi RV (2010). Genetic Analysis of Dairy milk, fat and protein percentage of Iranian First lactation Holstein Cattle. World Appl. Sci. J. 10(9):1042-1046.

Bignardi AB, El Faro L, Torres Júnior RAA, Cardoso VL, Machado PF, Albuquerque LG (2011). Random regression models using different functions to model test-day milk yield of Brazilian Holstein cows. Genet. Mol. Biol. 10(4): 3565-3575.

Cobuci JA, Euclydes RF, Lopes PS, Costa CN, Torres RdeA, Pereira CS (2005). Estimation of genetic parameters for test-day milk yield in Holstein cows using a random regression model. Genet. Mol. Bio. 28(1):75-83.

Cobuci JA, Costa CN, Braccini NJ, de Freitas AF (2011). Genetic parameters for milk production by using random regression models with different alternatives of fixed regression modeling. R. Bras. Zootec. 40(3): 557-567.

Druet T, JaffrezicF, Boichard D, Ducrocq V (2003). Modeling lactation curves and estimation of genetic parameters for first lactation test-day records of French Holstein cows. J. Dairy Sci. 86(7):2480-2490.

El Faro L, Cardoso VL, de Albuquerque LG (2008). Variance component estimates for test-day milk yield applying random regression models. Genet. Mol. Biol. 31(3):665-673.

Interbull (2012). Description of National Genetic Evaluation Systems for dairy cattle traits as applied in different Interbull member countries.

Jamrozik J, Schaeffer LR (1997). Estimates of genetic parameters for a test-day model with random regressions for yield traits of first lactation Holsteins. J. Dairy Sci. 80(4):762-770.

Jamrozik J, Schaeffer LR (2002). Bayesian comparison of random regression models for test-day yields in dairy cattle. 7th World Congr. Genet. Appl. Livest. Prod. Montpellier, France.

Jensen J (2001). Genetic evaluation of Dairy Cattle Using Test-Day models. J. Dairy Sci., 84(12):2803-2812.

Kirkpatrick M, Hill WG, Thompson R (1994). Estimating the covariance structure of traits during growth and aging illustrated with lactations in dairy cattle. Genetic Research, 64(1):57-69.

Lidauer M, Mäntysaari E (1999). Multiple trait reduced rank random regression test day model for production traits. Zurich, Switzerland, August 26-27, 1999. Interbull Bulletin, No. 22, 74.

Lidauer M, Mäntysaari EA, Strandén I (2003). Comparison of test-day models for genetic evaluation of production traits in dairy cattle. Livest. Prod. Sci. 79(1):73-86.

Liu YX, Zhang J, Schaeffer LR, Yang RQ, Zhang WL (2006). Optimal Random Regression Models for Milk Production in Dairy Cattle. J. Dairy Sci. 89(6):2233-2235.

Liu Z, Reinhardt F, Reents R (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.
Madsen P, Jensen J (2013). DMU A Package for Analysing Multivariate Mixed Models. Version 6, release 5.2. Center for Quantitative Genetics and Genomics Dept. of Molecular Biology and Genetics, University of Aarhus Research Centre Foulum Box 50, 8830 Tjele Denmark.

Negussie E, Stranden I, Mäntysaari EA (2008). Genetic Association of Clinical Mastitis with Test-Day Somatic Cell Score and Milk Yield during First Lactation of Finnish Ayrshire Cows. J. Dairy Sci. 91(3):1189-1197.

Pool MH, Meuwissen THE (1999). Reduction of the number of parameters needed for a polynomial random regression test day model. Livest. Prod. Sci. 64:133-145.

Pool MH, Janss LLG, Meuwissen THE (2000). Genetic parameters of Legendre polynomials for first-parity lactation curves. J. Dairy Sci. 83(11):2640-2649.

Reents R (1996). The use of test day models in genetic evaluation. INTERBULL Bulletin No. 12.

Rekaya R, Carabaño MJ, Toro MA (1999). Use of test day yields for the genetic evaluation of production traits in Holstein-Friesian cattle. Livest. Prod. Sci. 57(3):203-217.

Santos DJA, Peixoto MG, Borquis RR, Verneque RS, Panetto JC, Tonhati H (2013). Comparison of random regression models to estimate genetic parameters for milk production in Guzerat (Bosindicus) cows. Genet. Mol. Biol. 12(1):143-153.

Schaeffer LR, Jamrozik J (2008). Random regression models: a longitudinal perspective. J. Anim. Breed. Genet. 125(3):145-146.

Strabel T, Szyda J, Ptak E, Jamrozik J (2003). Comparison of Random Regression Test-Day Models for Production Traits of Dairy Cattle in Poland. Rome, Italy, August 28-30, 1999. Interbull Bulletin, No. 31, 197-201.

Swalve HH (2000). Theoretical basis and computational methods for different test-day genetic evaluation models. J. Dairy Sci. 83(5):1115-1124.

Torshizi ME, ASlamenejad AA, Nassiri MR, Farhangfar H, Solkner J, Kovac M, Meszaros G, Malovrh S (2013). Analysis of Test Day Milk Yield by Random Regression Models and Evaluation of Persistency in Iranian Dairy Cows. Iranian J. Appl. Animal Sci. 3(1):67-76.

Wall E, Simm G, Moran D (2010). Developing breeding schemes to assist mitigation of greenhouse gas emissions. Animal 4(3):366-376.

Wilmink JBM (1987). Adjustment of test-day milk, fat and protein yields for age, season and stage of lactation. Livest. Prod. Sci. 16(4):335-348.