Comparison of Random Regression Test-Day Models for Production Traits of South African Jersey Cattle

M.G. Kinghorn, E.D. Cason, V. Ducrocq, F.W.C. Neser

Comparison of Random Regression Test-Day Models for Production Traits of South African Jersey Cattle

M.G. Kinghorn[1*], E.D Cason[1], V. Ducrocq[2], F.W.C Neser[1]

[1]University of the Free State, Bloemfontein, South Africa

[2][1]Université Paris Saclay, INRAE, AgroParisTech, GABI, 78350 Jouy-en-Josas, France

[*]matthew.kinghorn566@gmail.com

Abstract

This study aims to identify the most appropriate random regression test-day model (RRTDM) for the South African Jersey dairy breed, using Legendre Polynomials (LP) to model the lactation curve. The data comprised of production test-day records for milk yield (MY), fat yield (FY) and protein yield (PY) from 33,852 multiparous cows. Genetic parameters were estimated using single trait RRTDM using AIREML90. The fixed effects fitted were Herd x Test-day x Birth Year of the Cow x Calving year x Parity x Number of Milkings (HTDBCPNM), production system (PROD) (TMR, pasture or both), age at first calving (AFC) and inter-calving period (ICP), and fixed regressions on age class-season-district. Random effects were direct additive, permanent environmental and residual effects. The estimated heritabilities for MY, FY and PY were 0.254, 0.182 and 0.170, respectively. The genetic correlations between total milk production and persistency were low (0.06 to 0.24). Overall, the heritability estimates of all the traits analysed were low to moderate.

Introduction

The accurate estimation of genetic parameters and (co)variance components associated with RRTDM is important for the estimation of accurate breeding values (Danell, 1982). The genetic parameter estimates obtained are largely dependent on the regression functions used to model the lactation curve, as well as the covariance structure for additive, permanent environmental, and residual effects (Bignardi *et al.*, 2011). The most common functions used to model the lactation curve include the ones proposed by Ali & Schaeffer (1987), Wilmink (1987), splines (White *et al.*, 1999), and orthogonal polynomials (Kirkpatrick *et al.*, 1990). Parametric functions such as those of Ali & Schaeffer (1987) and Wilmink (1987) have mostly been abandoned because the estimated covariance matrices show high correlations between random regression coefficients which can hinder the estimation process (Schaeffer, 2004). Therefore, orthogonal polynomials such as the Legendre Polynomials (LP) introduced by Kirkpatrick are more commonly used since their parameters tend to be less correlated than the aforementioned functions (Guo *et al.*, 2002), and they provide estimates of covariance matrices that are more robust over different datasets (Schaeffer, 2004). This study makes use of LP to estimate lactation curves, based on additional benefits outlined by Pool (2000).

In South Africa, dairy cattle breeding values are estimated for production traits based on a fixed regression test-day model (Mostert, 2007). Dzomba *et al.* (2010), outlines how the current South African model is criticised for assuming a standard lactation curve for cows in similar contemporary groups, homogenous additive genetic variances across lactations, and for its inability to account for persistency. The adoption of a RRTDM would overcome these limitations and has other advantages such as improved accuracy of breeding value estimates, extra possibility to decrease generation interval, and consequently improved genetic response to selection. This study aims to identify the most appropriate RRTDM for the South African Jersey dairy breed, using Legendre Polynomials to model the lactation curve.

Materials and Methods

Data ranging from 1988 to 2017 were obtained from the South African Jersey Association. A total of 429 940; 420 691 and 420 691 test-day records were analysed for MY, FY, and PY respectively. The data analysed was from 33 852 multiparous cows, of 1 558 sires, in 51 herds located across South Africa. Age at calving ranged between 15-36, 26-53, and 38-67 months for first, second, and third parity respectively (Interbull, 2015). Test-day yields from 5 to 305 days of lactation were considered. Lactations included at least five test-day records per cow. Contemporary groups were defined as a concatenation of Herd x Test-day x Birth Year of the Cow x Calving year x Parity x Number of Milkings (HTDBCPNM), with a total number of 12,033 groups containing at least 10 individuals and 2 sires per group. Only animals containing information on production system (i.e., pasture, total-mixed ration, or a combination of these systems) were included. Test-day yields for milk, fat, and protein were adjusted according to the ICAR guidelines (2017).

Analyses were performed using single trait RRTDM, including the direct additive, permanent environmental, and residual effects. The fixed effects included in the model were contemporary group (as defined above), PROD, AFC, and ICP, and age average lactation curve modelled with fifth-order LP within age class x season x district classes where age classes were assigned per lactation according to Mostert (2007). The chosen statistical model is as follows:

$$y_{tijklmn} = \text{HTDBCPNM}_i + \text{PROD}_j + \text{AFC}_k + \text{ICP}_l + \sum_{r=1}^{5} ASD_{mn0}\, x_n + \sum_{r=1}^{5} a_{mn} x_n + \sum_{r=1}^{5} pe_{mn} x_n + e_{tijklmn}$$

$$(1)$$

where $y_{tijklmn}$ is the tth test day record (milk yield, fat and protein contents) of the mth cow in i^{th} contemporary group (HTDBCPNM) effect, j^{th} production system and k^{th} age at first calving, l^{th} inter-calving period; ASD_{mn0} is the nth fixed regression coefficient of the oth class of cows age class-season-district; a_{mn} and pe_{mn} are regression coefficients n^{th} for additive genetic and permanent environmental effects on m^{th} cow respectively; x_n is the nth Legendre Polynomial for the tth day; $e_{tijklmn}$ random residual effect associated with $y_{tijklmn}$.

For each trait, a single trait RRTDM was run using Average Information Restricted Maximum Likelihood methodology with the AIREMLF90 program (Misztal *et al.*, 2002). Numerous models with varying fixed effects were tested, and AIC values and variance components were compared for each model.

Results

The first coefficient of LP can be related to the total yield in each lactation and the second regression coefficient can be viewed as a persistency measure. The first two coefficients tend to account for most of the genetic variation (Strabel & Jamrozik, 2006). The first two coefficients accounted for approximately 94% of the genetic variation for MY, whereas for FY and PY the first four coefficients accounted for approximately 90% of the genetic variation, only results of the first two coefficients shown in Table 1.

Table 1. Genetic variances of the random coefficients (diagonal, bold face), correlations (upper diagonal), covariances (lower diagonal), phenotypic variances (σP), residual error variances (σr) and direct ratios.

	Direct Additive Effect				
	Random Regression Coefficients				
Trait	0	1	σP	σr	h^2
Milk Yield	5.4297 0.4757	0.1669 1.4957	21.4090	6.1314	0.2536 0.0699
Fat Yield	0.0167 0.0008	0.0678 0.0073	0.0920	0.0300	0.1816 0.0792
Protein Yield	0.0163 0.0034	0.2410 0.0123	0.0960	0.0100	0.1700 0.1286
	Permanent Environmental Effect				
	Random Regression Coefficients				
Trait	0	1	σP	σr	Ratio
Milk Yield	7.1516 0.0825	0.0444 0.4829	21.4090	6.1314	0.3341 0.0226
Fat Yield	0.0117 0.0005	0.0660 0.0047	0.0920	0.0300	0.1272 0.0511
Protein Yield	0.0102 0.0021	0.2398 0.0075	0.0960	0.0100	0.1072 0.0785

Discussion

Assessment of AIC values showed that Model 1 produced the lowest values, and thus best fit the data for all traits (results not shown). The first random regression coefficient (intercept) of orthogonal polynomials describes the total yield in lactation, thus the heritability of this coefficient is equal to the heritability of 305-day lactation yield. The second random regression coefficient (linear) can be viewed as a persistency measure (Strabel & Jamrozik, 2006). The largest genetic variances were associated with the intercept and linear coefficients, while all other variances (not shown) were close to zero for all traits. The genetic correlation for MY and persistency was low and positive (0.1669), however not as low as the genetic correlation between total FY and persistency of fat production (0.0678). The genetic correlation between PY and persistency of protein production, although still lowly correlated (0.2410), appeared to have the strongest genetic relationship between the first two random regression coefficients.

The variances obtained in this study were different to those obtained by Interbull (2015) for the South African Jersey population, likely do due to differences in the model, as well as the data used. Despite this, the heritabilities for the direct additive effects appeared to be similar to those obtained in this study; 0.23, 0.14 and 0.20 for MY, FY and PY respectively. The Interbull variance ratios for permanent environmental effects also appeared to be similar, ranging between 0.27-0.31, 0.13-0.18 and 0.24-0.26 for MY, FY and PY respectively. Heritabilities of 0.18 and 0.11, and variance ratios of 0.44 and 0.88 for the first two random regression coefficients for total milk production were obtained by Strabel & Jamrozik (2006). The same study also found a genetic correlation between the intercept and linear coefficients of first lactation milk production to be 0.06. Other studies have obtained heritabilities of 0.23-0.25 for milk yield (Li *et al.*, 2020). Variance ratios obtained by a fixed regression model used on South African Jersey data were 0.39, 0.21 and 0.34 for additive genetic effects of milk, fat and

protein respectively, and 0.23,0.16 and 0.20 for permanent environmental effects of milk, fat and protein yields respectively. The current South African test-day model is a fixed regression model, and therefore does not have a comparative genetic measure for persistency of production traits (Dzomba *et al.*, 2010).

References

Ali, T. E., & Schaeffer, L. R. 1987. Accounting For Covariances Among Test Day Milk Yields In Dairy Cows. Can. J. Anim. Sci. 67, 637–644.

Bignardi, A. B., El Faro, L., Torres Junior, R. A. A., Cardoso, V. L., Machado, P. F., *et al.,* 2011. Random regression models using different functions to model test-day milk yield of Brazilian Holstein cows. Genetics and Molecular Research 10, 3565–3575 http://dx.doi.org/10.4238/2011.October.31.4.

Danell, B. 1982. Studies on lactation yield and individual test-day yields of Swedish dairy cows. Acta. Agric. Scand. 32, 82–92 https://doi.org/10.1080/00015128209435733.

Dzomba, E. F., Nephawe, K. A., Maiwashe, A. N., Cloete, S. W. P., Chimony, M., *et al*., 2010. Random regression test-day model for the analysis of dairy cattle production data in South Africa: Creating the framework. S.Afr.J.Anim.Sci. 40, 273–284.

Guo, Z., Lund, M. S., Madsen, P., Korsgaard, I., & Jensen, J. 2002. Genetic Parameter Estimation for Milk Yield over Multiple Parities and Various Lengths of Lactation in Danish Jerseys by Random Regression Models. J Dairy Sci 85, 1596–1606.

ICAR. 2017. Section 2 - Guidelines for Dairy Cattle Milk Recording. ICAR The Global Standard For Livestock Data. Available at: https://www.icar.org/Guidelines/02-Overview-Cattle-Milk-Recording.

Interbull. 2015. Description of National Genetic Evaluation Systems. Interbull.Available at: https://interbull.org/ib/geforms.

Kirkpatrick, M., Lofsvold, D., & Bulmer, M. 1990. Analysis of the Inheritance, Selection and Evolution of Growth Trajectories. Genetics 124, 979–993.

Li, J., Gao, H., Madsen, P., Li, R., Liu, W., *et al*., 2020. Impact of the Order of Legendre Polynomials in Random Regression Model on Genetic Evaluation for Milk Yield in Dairy Cattle Population. Frontiers in Genetics 11, 1–8. https://doi.org/10.3389/fgene.2020.586155.

Misztal, I., Tsuruta, S., Strabel, T., Auvray, B., Druet, T., *et al*. 2002. BLUPF90 and Related Programs. Proc. 7th WCGALPP, Montpellier, France.

Mostert, B. E. 2007. The suitability of test-day models for genetic evaluation of dairy cattle in South Africa. PhD Thesis, University of Pretoria, South Africa http://hdl.handle.net/2263/25127.

Pool, M. H. 2000. Test-day Models: Breeding value estimation based on individual test-day records.Universal Press, Veenendaal, The Netherlands.

Schaeffer, L. R. 2004. Application of random regression models in animal breeding. Livest. Prod. Sci. 86, 35–45. https://doi.org/DOI: 10.1016/S0301-6226(03)00151-9.

Strabel, T., & Jamrozik, J. 2006. Genetic Analysis of Milk Production Traits of Polish Black and White Cattle Using Large-Scale Random Regression Test-Day Models. J. Dairy Sci. 89, 3152–3163.https://doi.org/ 10.3168/jds.S0022-0302(06)72589-9

White, I. M. S., Thompson, R., & Brotherstone, S. 1999. Genetic and Environmental Smoothing of Lactation Curves with Cubic Splines. J Dairy Sci 82, 632–638. https://doi.org/10.3168/jds.S0022-0302(99)75277-X.

Wilmink, J. B. M. 1987. Adjustment of Lactation Yield for Age at Calving in Relation to Level of Production. Livest. Prod. Sci. 16, 321–334. https://doi.org/10.1016/0301-6226(87)90002-9.