Proc. Assoc. Advmt. Anim. Breed. Genet. 20:74-77

MODIFICATION OF LACTATION YIELD ESTIMATES FOR IMPROVED SELECTION OUTCOMES IN DEVELOPING DAIRY SECTORS

D. M. McGill[1], P. C. Thomson[2], H. A. Mulder[3], and J. J. Lievaart[1]

[1] School of Animal and Veterinary Sciences, Charles Sturt University, NSW 2650, Australia
[2] Faculty of Veterinary Science, The University of Sydney, Camden NSW 2570, Australia
[3] Animal Breeding and Genomics Centre, Wageningen University, The Netherlands

SUMMARY

Animals are extremely important to the livelihood of smallholder dairy farmers in developing countries. However, due to limited resources, investment into long-term genetic improvement is rare and herd recording is minimal. Therefore, obtaining adequate performance records for genetic evaluation and selection is difficult and hence it is important to optimize the selection outcomes from any data that are collected. The aim of this study is to determine a robust and efficient method for estimating lactation yield for low producing dairy cattle and their subsequent genetic evaluation. Using Sahiwal cattle as an example, simulated data sets, based on lactation data from Pakistan, were used to compare different methods of lactation yield estimation (i.e., test-interval method, and three nonlinear models). Furthermore, these estimates were analysed to explore their implications on the subsequent estimated breeding value (EBV) ranking and selection outcomes. Utilising these results, different test-day sampling schedules were compared to investigate possible recording regimes involving few records that can accurately estimate lactation yield without significantly affecting selection. Results indicate that the lactation models proposed by Wood (1967) and Wilmink (1987) yield similar selection outcomes to the recommended test-interval method. These results provide opportunities for further research into test-day scheduling which could reduce the number of records required and have considerable implications on progeny testing systems of low producing dairy cattle and developing dairy sectors.

INTRODUCTION

Breed improvement and selection in developing dairy systems can be challenging as field conditions are generally constrained by a lack of infrastructure for regular test-day recording. For this reason, regular twice daily recording of milk yield for entire lactations is not feasible (Khan *et al.* 2008). The limited resources and data exacerbate the need to utilise each record efficiently to maximize their contribution to the evaluation process (Bajwa *et al.* 2002). Therefore, for any developing dairy sector there is a need to develop a genetic evaluation system which optimizes selection outcomes given the current resources for test-day recording.

There are numerous methods for genetically evaluating milk production based on test-day records. In developed nations complex methods such as test-day models are commonly used. These models require accurate estimates of genetic and phenotypic parameters based on many daily milk yields from large populations of animals which are unlikely to be available in a developing dairy sector (Ilatsia *et al.* 2007). For this reason, simpler methods such as a two-step approach can be used. This is where test-day records are first used to estimate lactation yield and then these values are used as the phenotype for genetic evaluation. Methods of lactation yield estimation from test-day records are well researched. In a developing country scenario, the Test-Interval Method [TIM] (Sargent *et al.* 1968) is recommended by the International Committee for Animal Recording (ICAR 2009). Other approaches involve fitting a mathematical model to lactation data and using the model outputs to estimate yield. Many models have been proposed for describing the lactation curve of dairy animals (Dongre *et a.l* 2011). A handful of studies have investigated the ability of lactation curve models to depict Sahiwal cattle lactation data. Kolte *et* *al.* (1986), found that the inverse polynomial function proposed by Nelder (1966) was the superior fitting model, followed by the gamma function proposed by Wood (1967). Contrary to this, Rao and Sundaresan (1979), reported that Wood's (1967) function was the most appropriate. The Wood (1967) model is one of the most widely accepted lactation models and is commonly used in research (Swalve 2000). Similarly, the Wilmink (1987) model is frequently used within test-day evaluations to model the lactation curve of dairy cattle (Naranchuluum *et al.* 2011).

This current study is concerned with Sahiwal cattle in Paktistan and will focus on how different lactation models behave when fitted to the lactation characteristics of this particular breed. Specifically, this study aims firstly to determine which lactation model is the most robust at modelling the lactation curve of Sahiwal cattle at different test-day recording schedules. The second aim is to discuss what implication this may have on the future of test-day sampling in Pakistan and how it can be used to improve their current progeny testing program.

METHODS AND MATERIALS

Lactation Estimation Models and Methods. The lactation estimation methods used within this study were:

1. The test-interval method (TIM) described in Sargent *et al.* (1968) which is based on an approximation of the area under a curve
2. The inverse polynomial model proposed by Nelder (1966): $\mathbf{yield}_i = \frac{1}{a + b * \mathbf{dim}_i + c * \mathbf{dim}_i^2}$
3. The gamma function proposed by Wood (1967) $\mathbf{yield}_i = a * \mathbf{dim}_i^b * \mathrm{e}^{-c*\mathbf{dim}_i}$
4. The lactation model proposed by Wilmink (1987): $\mathbf{yield}_i = a + b^{\mathbf{dim}_i} + c * \mathrm{e}^{-0.05*\mathbf{dim}_i}$

where *a, b* and *c* are different parameters to be estimated separately within each model and **dim** are the days in milk (i = 1, …, 280) for a lactation length of 280 days.

Data. Weekly test-day Sahiwal lactation records from 839 lactations from 464 dams, collected during 2005-2010 from the Livestock Production Research Institute (LPRI), Bahadurnagar Okara, were used as the basis for data simulation in this study. Data were simulated using three different lactation models (Wood, Wilmink and Nelder). Variance and covariance matrices of the parameters (*a*, *b* and c) and a residual variance of each of these models was determined based on the raw Pakistani data. Using these variance structures and the pedigree relationship matrix (**A**), phenotypic lactation yields were simulated for entire lactations for all the dams in the population. This was repeated 100 times for each of the simulation models to yield three batches of one hundred data sets for comparison. Data were simulated using three different lactation models because it allows for a more thorough comparison of lactation yield estimation methods as it gives an indication of their robustness across different lactation curve shapes.

Model Comparison. Four lactation yield estimation methods were used to fit and calculate the lactation yield for every dam for each set of simulation data. These included the recommended TIM as well as three lactation models, Wood, Wilmink and Nelder, fitted and estimated using the nonlinear mixed effects (NLME) model function in R Version 2.13.0 (R Development Core Team 2008) following a similar process outlined by Raadsma *et al.* (2009). This was carried out for four different test-day scheduling regimes (weekly, monthly, five test-days; random selection and five test-days; stratified selection). For each method, the percentage of models which successfully converged was recorded as well as the lactation yield estimates. The lactation yield estimates were compared with the true simulated lactation yield and summed to calculate a mean square error (MSE) of estimation for each simulated data set. The MSE was then used to directly compare between the lactation yield estimation methods. Lastly, the lactation yield estimates for each simulated data set were used to calculate estimated breeding values (EBVs) for each of the animals in the data set using ASReml-R Discovery Edition 1.0 (Butler *et al* 2009). The outputs of this analysis allowed further comparison between models to determine if the lactation yield estimation method had any effect on the ranking and subsequent selection of animals.

RESULTS AND DISCUSSION

The robustness of each of the lactation models for fitting Sahiwal test-day data can be determined by comparing the percentage of success rates of each model's ability to be fitted to the different simulated data sets (Table 1). These results show that overall the Wood model is superior to the Wilmink and Nelder models as it generally has higher rates of success, most importantly when fitting data from both a random and stratified selection of five test-day records. This has an important practical implication, as in the field conditions of Pakistan, test-day recording is likely to be irregular and infrequent.

Table 1. Percentage of lactation yield estimation models that were successfully fitted to each set of simulated lactation data at each of the four different test-day recording regimes (weekly, monthly, 5 test-days: random sample and 5 test-days: stratified sample).

Data Simulation Model	Fitted Model	Test-Day Recording Regime			
		Weekly	Monthly	Random	Stratified
Wood	Wood	100	100	82	92
	Wilmink	100	100	76	83
	Nelder	100	100	70	72
Wilmink	Wood	100	100	88	86
	Wilmink	100	100	74	83
	Nelder	100	100	78	75
Nelder	Wood	97	98	60	67
	Wilmink	94	100	69	82
	Nelder	75	71	82	83

Using the MSE values from each of the lactation yield estimation methods we can directly compare between models for the same simulated lactation. The average MSE values across lactation yield estimates can be seen in Table 2. These are presented for only two of the data simulation methods (Wood and Wilmink). The results from the Nelder simulated data are not reported here as the number of failed models caused unreliable values. From Table 2 the MSE values show that the Wilmink and Wood models were superior to the TIM and Nelder methods. Furthermore, the Wilmink model has a lower average MSE than the Wood model in both sets of simulated data (5,124,550 vs 5,327,934 for the Wilmink simulated data and 5,234,436 vs 5,235,715 for the Wood simulated data). This suggests that the Wilmink model is superior to the Wood model in its ability to accurately estimate lactation yield on different types of lactation data.

Despite the differences in the MSE seen in Table 2, the important outcome of this analysis relates to the animals in the top proportion of the population that would be selected for breeding and how they compare with the true (simulated) superior animals. For the different methods of lactation yield estimation, using the Wood simulated data, the average number of corresponding animals with the true top fifty superior animals were; TIM 39.2±2.22, Wood 39.7±2.27 and Wilmink 39.6 ±2.28. For the Wilmink simulated data sets, the results were very similar; TIM 36.8±2.33, Wood 37.1±2.38 and Wilmink 37.1 ±2.41. The results show that the average number of corresponding animals with the true top fifty were all within one animal of the other estimation methods. This suggests that these methods of estimating lactation yield, for a given test-day scheduling regime, are each capable of selecting the superior animals from a given population.

Table 2: Average Means Squared Error values (± st.dev) for four different methods of lactation estimation (TIM, Nelder, Wilmink and Wood) when calculated using monthly records from two methods of data simulation (the Wilmink and Wood models)

Model used for lactation yield estimate	Model used for data simulation			
	Wilmink		Wood	
	Average MSE (±sd)		Average MSE (±sd)	
TIM	6,102,273	(±385,457.5)	6,143,607	(±377,174.1)
Nelder	5,962,774	(±413,556.1)	5,696,461	(±387,866.5)
Wilmink	5,124,550	(±311,253.8)	5,234,436	(±347,630.8)
Wood	5,327,934	(±342,708.3)	5,235,715	(±368,825.4)

CONCLUSIONS

The benefit of modelling test-day yields is the ability to subsequently estimate lactation yield on fewer records. This then provides an opportunity to record more animals fewer times which will help to improve the accuracy of evaluations as well as increase the population of animals from which selection can take place. The outcomes of this study show that although the Nelder model is capable of fitting and modelling low producing dairy cattle like the Sahiwal, it is unreliable with different lactation curves and test-day sampling regimes. The results from the other lactation models tested, the Wood or Wilmink, show that they are both robust in different scenarios with the Wood model better fitting irregular and infrequent test-day recording regimes. Despite this, both the Wood and Wilmink models provide an opportunity to further investigate their use in estimating lactation yield in Sahiwal cattle and the possibility of reducing the number of required test-day records whilst maintaining the accuracy of selection outcomes.

REFERENCES

Bajwa I. R., Khan M.S., Ahmad Z. and Gondal K.Z.. 2002. Pages 1-56 in Proc. 7th World Congr. Genet. Appl. Livest. Prod. , Montpellier, France.

Butler D. G., Cullis B.R., Gilmour A.R. and B. J. Gogel . 2009. Department of Primary Industries and Fisheries, Queensland.

Dongre V. B., Gandhi R.S., Avtar S., and Atul G. 2011. *Intl. J. Agric. Res. and Rev.* **1:** 6

ICAR. 2009. International Committee for Animal Recording, Niagra Falls, USA.

Ilatsia E. D., Muasya T.K., Muhuyi W.B., and . Kahi A. K. 2007. *Animal* **1**:185

Khan, M. S., Bilal G., Bajwa I.R., Rehman Z. and Ahmad S. 2008. *Pakistan Veterinary J.* **28**:131

Kolte D. V., Gore A.K., and Deshmukh S.N. 1986. *PKV Research Journal* **10**:145.

Naranchuluum, G., Ohmiya H., Masuda Y., Hagiya K., and Suzuki M. 2011. *Anim. Sci.***82:** 383

Nelder J. A. 1966. *Biometrics* **22**:128.

R Development Core Team 2008. R. F. f. S. Computing, ed, Vienna, Austria.

Raadsma H. W., Jonas E., McGill D., Hobbs M., Lam M.K., and Thomson P.C. 2009. *Genet. Sel. Evol.* **41**

Rao M. K. and Sundaresan D. 1979. *J. of Agric. Sci.* **92**:393

Sargent F. D., Lytton V.H., and Wall O.G. Jr. 1968. *J. Dairy Sci.* **51**:170

Swalve H. H. 2000. *J. Dairy Sci.* **83**:1115

Wilmink J. B. M. 1987. *Livest. Prod. Sci.* **17**:1

Wood P. D. P. 1967. *Nature* **216**:164