
VOLUME 62 2002

Choice of the most appropriate models and estimation procedures of lactation curves for grazing dairy cattle

J.R. BRYANT[1] and N. LOPEZ-VILLALOBOS[2]

[1]AgResearch Limited, Grasslands Research Centre, Tennent Drive,
Palmerston North, New Zealand

ABSTRACT

The objective of this study was to determine the most appropriate lactation curve representation, a 4th-order Legendre polynomial (LP) function or a mechanistic mammary gland (MG) model, for various lactation curve shapes of grazing cattle in New Zealand. The ability of two optimisation methods, evolutionary algorithm (EA) and the Newton method, used to find parameter values that minimised mean prediction error (MPE) was also determined. The 95% confidence intervals for the parameters values and goodness of fitness for each lactation function and numerical algorithm were obtained with a bootstrapping strategy. Milk yields of three different cows were chosen to describe bi-peak, highly variable, and spring-peak lactation curves. The MG function was able to find the best fit between predicted and actual values for the bi-peak and highly variable lactation curves with MPE of 2.7 to 3.7%, compared with MPE of 6.5 to 11.9% for the LP function applying Newton and EA. For the traditional lactation curve, the LP function applying Newton had the lowest MPE (1.7%) with the MG function applying an EA the highest MPE (4.1%). Overall, minimal differences in MPE were observed when solving using the Newton and EA. However, the Newton method fixed some parameters values of the MG function to achieve approximately the same MPE as the EA which manipulated all parameter values. The results of this study illustrate that a MG function solved with an EA is an accurate and efficient way to model non-standard lactation curves but a LP function in combination with the Newton method can be better option for more standard lactation curves.

Keywords: Lactation curve; optimisation; evolutionary algorithm; dairy cattle; bootstrap.

INTRODUCTION

A mathematical description of the milk yield lactation curve is often required to identify and model the effect of breed, age and environmental factors on the shape of the lactation curve (Wilmink, 1987; Wood, 1980), and to provide a quantitative measure of persistency (Lopez-Villalobos *et al*., 2005). The Wood (Wood, 1967), Wilmink (Wilmink, 1987) or spline functions have been applied to model typical lactation curves of dairy cattle. Recently, the Legendre polynomial has become the preferred method in genetic evaluation test-day models (Macciotta *et al*., 2005). Mechanistic representations of the lactation curve, such as the model of Vetharaniam *et al.* (2003b) are available, but these are complex and often require data such as feed intake information. Nevertheless, provided feed intake information is available, such a mechanistic model should fit the lactation curve better than simpler representations. This is especially the case for lactation curves that do not fit the classical shape (a rise in milk yield to peak and then a constant decline in milk yield thereafter). Uneven lactation curves, often with multiple peaks, are common in New Zealand, due to the strong influence of climate on feed supply to the animal.

To find the best estimates of the function describing the lactation curve for the data, optimisation methods are needed. Optimisation seeks to find the optimal solution which satisfies all stated feasibility constraints and minimizes (or maximizes) the value of an objective function (Pintér, 1996). Two common optimisation methods are Newton-Raphson, hereafter described as the Newton method, for relatively simple numerical problems, and evolutionary algorithms (EA) for more complex non-linear, multivariate problems. The most appropriate optimisation algorithm for fitting lactation curve representations is uncertain. In addition, the ability of an optimisation algorithm to find the best parameter values when data is missing, or perceived outliers are present, is important. The objectives of this study were to compare the Newton and EA methods to find the optimum parameter values of a Legendre polynomial (LP) and mechanistic mammary gland (MG) model for three different lactation curve shapes. Confidence intervals of prediction error and parameter values were estimated to determine the efficiency of the optimisation algorithm.

[2]Institute of Veterinary, Animal and Biomedical Sciences, Massey University, Palmerston North, New Zealand

MATERIALS AND METHODS

Animal Data

Three cows to represent bi-peak (peak in spring and a second peak in late summer), spring-peak and variable lactation curves (Figure 1) were chosen from a farmlet trial using Jersey cattle in the 1999/2000 season (for more details on the trial refer to Bryant *et al.,* 2003). Milk, fat and protein yields were measured at approximately two-weekly intervals with a corresponding estimate of intake of pasture and supplements at each test calculated from the area grazed per cow daily, pre- and post-grazing masses of pasture and supplements offered.

Mathematical representations of the lactation curves

Legendre polynomial

The LP method to fit the measured test-day milk yields as a function of days in milk is:

$$Y_t = \alpha_0 P_0 + \alpha_1 P_1 + \alpha_2 P_2 + \alpha_3 P_3 + \alpha_4 P_4$$

where Y_t is predicted milk yield, and $\alpha_0, \alpha_1, \alpha_2, \alpha_3$ and α_4 are Legendre coefficients for P_0, P_1, P_2, P_3 and P_4**,** respectively which represent days in milk on the Legendre scale (Macciotta *et al*., 2005).

Mammary gland model

The MG method is based on the model described by Vetharaniam *et al*. (2003a; 2003b), with key equations outlined by Bryant *et al*. (2007). In brief, the mammary gland model consists of alveoli (groups of secretory cells) in various states of activation or inactivation. At the start of lactation, each animal has an initial pool of active alveoli, A_0. The number of active alveoli, A_t at time t is dependent on a series of equations with the initial condition (t = 0) where $A_t = A_0$. Active alveoli are derived from progenitor cells with the rate determined by the balance between the parameters, k_1 and k_2. Once lactation commences, alveoli can be in one of three states; active, quiescent (non-secretory but able to be reactivated to a secretory state) or senescent with the parameters, k_3 to k_5, determining the rates of quiescence, reactivation and senescence. At any stage of the lactation, milk energy output is influenced by the relative energy status of the animal, which in this case is the ratio of actual feed intake to a theoretical maximum feed intake (4% of live weight), the number of active alveoli, a theoretical maximum secretion rate per alveoli (3 x 10^{-9} mega-Joules/day based on previous data) and a nutritional response factor (L) which is invoked when relative energy status is less than 1.00. See Bryant *et al*. (2007) for a full description and list of the equations.

Optimisation methods

Newton method

The Newton method is often used to find the solution of nonlinear equations (Galantai, 2000). It is based on the Taylor expansion of the objective function around its root (del Piero *et al*., 2006; Mehl, 2006). The desirable property of the Newton method is its rapid convergence to a solution. However, it requires a good initial estimate of model parameter values, and is less suited to complex optimisation problems (Mehl, 2006).

Evolutionary algorithm

Evolutionary algorithms have proved efficient at finding the global optima in a number of agricultural models (Mayer *et al*., 1996). Evolutionary algorithms are based on Darwinian concepts of evolution where two selected individuals, with different genetic codes, are 'mated' or crossed to produce the next generation. In the context of the present study, the individuals are an array of estimates of LP or MG model parameters. Over generations, or iterations, the process combines successful parameter values, which improve the fitness of the population. The dominant genetic operation is crossover; however, mutation is also introduced at each mating to rediscover any potential beneficial parameter values. Through successive mating of selected individuals (arrays), the population structure tends to find a near-optimal solution (Mayer *et al*., 1999).

Estimating model parameter values and confidence intervals

The lactation curve representations, MG and LP, were solved by the Newton method and the EA method using 200 bootstrap samples within Microsoft Excel® using a macro developed in Visual Basic for Applications. The Newton and EA methods were applied using the Solver and Evolver add-ins for Microsoft Excel®. In each case, minimisation of the mean prediction error (MPE) was the objective function. The MPE, as a proportion of the actual mean, was calculated as described by Rook *et al*. (1990):

$$MPE = \frac{\sqrt{\frac{1}{n}\sum_{n=1}^{n}(\hat{y}_i - y_i)^2}}{\bar{y}^a}$$

where $\hat{y}_i$ and y_i are the predicted and actual values, and $\bar{y}^a$ is the actual mean. Values for MPE

< 0.10, and 0.10 to > 0.20 indicate good, moderate and poor simulation adequacy. In the EA, crossover was set at 0.50 with mutation rate at 0.20. A population size of 50 was used with the EA running for 200 generations. The parameter bounds were those specified in Bryant *et al*. (2007).

Witte *et al*. (1998) noted that estimated regression coefficients or model parameters are sometimes heavily influenced by a few data points only, and suggested bootstrapping as a means to explore the stability of model parameters values when removing influential data points. Bootstrapping assumes the distribution of values found within a random sample of size *n* is the best guide of the actual distribution of a population. Therefore, if the population was resampled, it makes sense to resample the sample (Manly, 1997). Confidence intervals on model parameter values and mean prediction error were obtained using bootstrapping for each optimisation method. For the bootstrap sample, 200 samples of *n* observations were selected randomly from each data set, with replacement, so any observation could be selected more than once, observations could be missing and each observation had an equal chance of being selected (Efron, 1979). The most probable model parameter values were calculated from the mean of all 200 sample solutions. The 95% confidence interval was defined using the 0.025 and 0.975 quantiles.

RESULTS

Comparison of the lactation curve representations

For the bi-peak cow (Figure 1A), which may occur for cows reliant on pasture as their sole feed supply, the MG representation was superior to the LP representation with the MPE significantly lower for the MG representation than the LP representation (Figure 2). For the variable lactation curve (Figure 1B), which can occur due to an erratic feed supply or due to illness, the MG representation resulted in a lower MPE, on average, than the LP representation (Figure 2).

Figure 1: Comparison of predicted (solid line) versus actual (open circles) values for daily milk yield (**A**: Bi-peak cow, **B**: Variable cow, **C**: Spring-peak cow) for the Legendre Polynomial function applying the Newton (LP_Newton) and the mammary gland model applying an evolutionary algorithm (MG_EA).

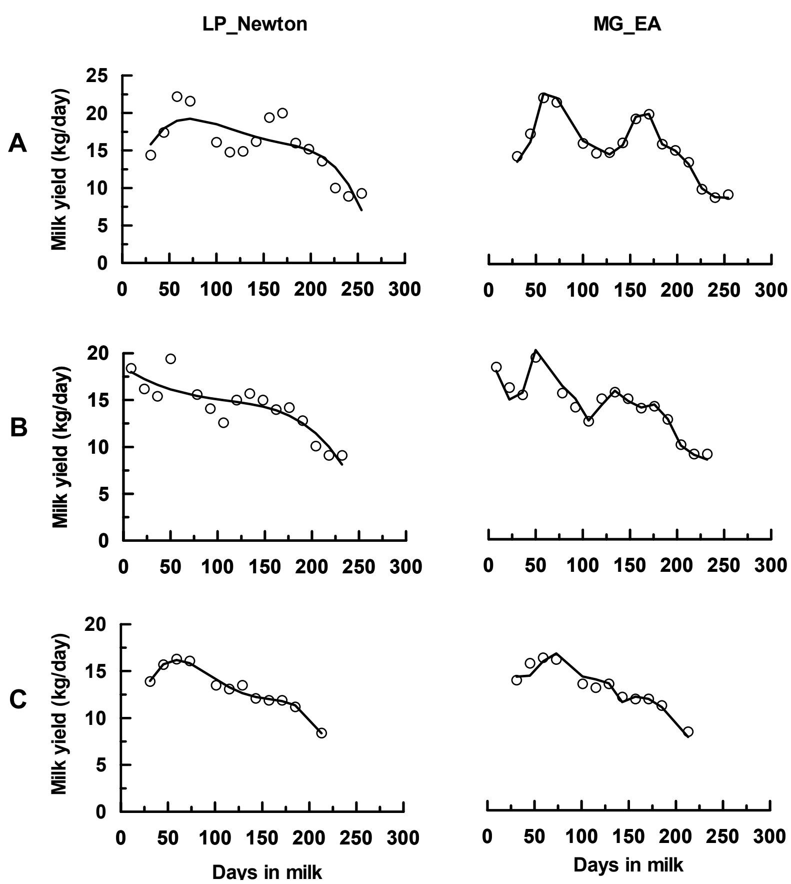
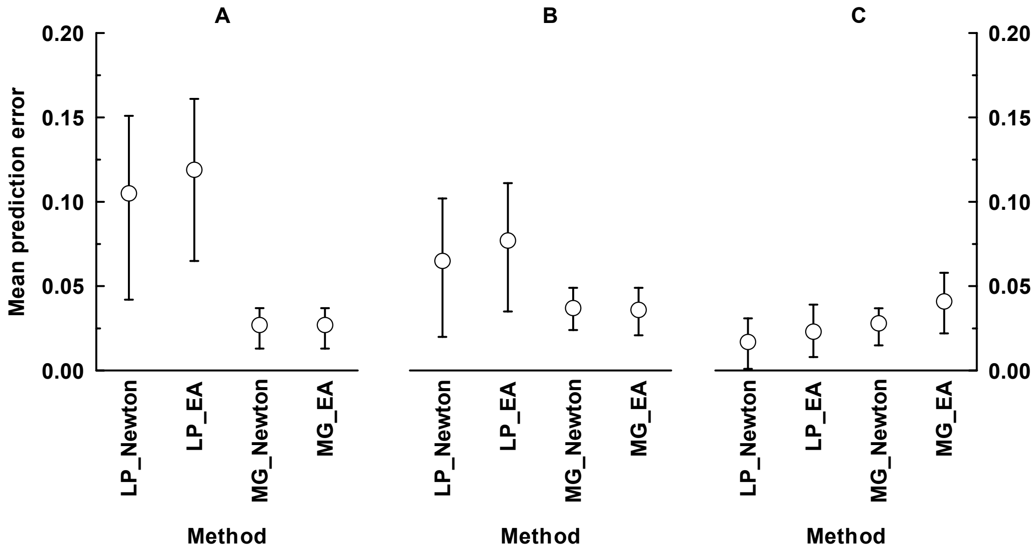


Figure 2: Average (open circles) and confidence intervals (bars) of mean prediction error using the Legendre polynomial function applying Newton (LP_Newton), Legendre polynomial function applying an evolutionary algorithm (LP_EA), mammary gland model applying Newton (MG_Newton) and mammary gland model applying an evolutionary algorithm (MG_EA) for **A:** Bi-peak cow, **B**: Variable cow, and **C**: Spring-peak cow.



However, for some bootstrap samples the LP method resulted in lower MPE than obtained using the MG representation applying both the Newton and EA (Figure 2). The LP representation resulted in a lower MPE, on average, than the MG representation applying either, Newton or EA, for the spring-peak cow (Figure 1C) lactation curve.

Comparison of optimisation methods

For the LP representation of the lactation curve, the Newton method was generally better than the EA method at minimising MPE but narrower confidence intervals of parameter values were obtained using the EA method (Table 1). By contrast for the MG model, the confidence intervals of parameter values applying Newton method were generally narrower than the EA method (Table 2). The MPE of the MG model applying the Newton or EA methods were similar for the bi-peak and variable cow, but the MPE was lower for the spring-peak cow applying the Newton than EA method. In many instances, applying the Newton method to the MG model resulted in the fixation of some parameter values, while still obtaining a very similar level of MPE to the solutions derived from the EA method.

DISCUSSION

The MG model was able to fit the bi-peak and variable lactation curves better than the LP function, but the LP function fitted better the spring-find peak lactation curve. Presumably, this ability to fit variable curves is due to the MG model consisting of a greater number of parameters than the LP function (seven compared to five parameters), as well as considering feed intake and milk composition information. However, the complexity of the MG model in its current form would prevent its use in genetic evaluation methods to identify unusual lactation curve shapes. Nevertheless, the LP function has been shown by Macciotta *et al*. (2005) to be more flexible than the Wood or Wilmink curves at detecting a wider range of lactation curve shapes. The MG model appears to be a valuable research tool to quantify the effect of various environmental factors on mammary gland dynamics. Bryant *et al*. (2007) was able to detect significant effects of age, feeding level, genetic merit and body condition score on specific mammary gland parameters.

As expected due to the relatively simple nature of the LP function, the Newton method was more efficient than the EA at finding the best solution. Solution time was also shorter for Newton than EA, as was observed by Mayer & Huang (1999) and del Piero *et al*. (2006). Very little difference in efficiency was observed when using the Newton and EA methods to solve the MG model. However, applying the Newton method to the MG model resulted in the fixation of some parameter values with other parameter values manipulated to minimise MPE. This is undesirable if attempting to determine the effect of an environmental variable on a mammary gland parameter. Conversely, the finding of no difference in solving efficiency

Table 1: Means and bootstrap confidence intervals of estimates of Legendre polynomial model parameters after applying a Newton and evolutionary algorithm to bootstrapped samples of the actual values.

		Newton		Evolutionary algorithm	
Cow descriptor	**Parameter**	**Mean**	**CI**[1]	**Mean**	**CI**
Bi-peak	α_0	14.77	[7.87, 22.98]	14.99	[13.37, 16.33]
	α_1	-5.76	[-21.89, -1.38]	-3.88	[-7.14, -1.33]
	α_2	-4.88	[-22.58, -13.39]	-4.35	[-8.34, -0.98]
	α_3	-4.09	[-25.27, 4.93]	-1.00	[-7.94, 3.78]
	α_4	-4.82	[-24.50, 15.53]	-4.15	[-10.54, 3.87]
Variable	α_0	12.59	[6.92, 16.99]	12.88	[11.64, 14.03]
	α_1	-6.65	[-17.98, -3.23]	-5.35	[-8.11, -2.68]
	α_2	-2.97	[-15.07, 6.45]	-2.36	[-4.69, 0.03]
	α_3	-4.68	[-18.82, 0.31]	-2.96	[-6.44, 0.58]
	α_4	-0.83	[-10.31, 9.63]	-0.62	[-4.48, 3.03]
Spring-peak	α_0	8.98	[-0.03, 16.03]	9.39	[8.95, 11.05]
	α_1	-9.42	[-25.93, 3.80]	-8.03	[-8.96, -5.34]
	α_2	-7.51	[-23.41, 5.44]	-7.15	[-8.07, -3.66]
	α_3	-5.53	[-23.88, 6.14]	-4.25	[-5.45, -0.83]
	α_4	-6.65	[-15.95, 3.89]	-7.07	[-8.25, -3.74]

[1]95% confidence interval [Lower confidence interval, Upper confidence interval]

Table 2: Means and bootstrap confidence intervals of estimates of mechanistic mammary gland model parameters after applying a Newton and evolutionary algorithm to bootstrapped samples of the actual values.

		Newton		Evolutionary algorithm	
Cow descriptor	**Parameter**	**Mean**	**CI**[1]	**Mean**	**CI**
Bi-peak	A_0 (x 10^{10})	2.00	[2.00, 2.00]	2.43	[2.23, 2.66]
	k_1 (x 10^9 d^{-1})	1.29	[1.29, 1.29]	1.71	[0.51, 2.97]
	k_2 (x 10^{-1} d^{-1})	0.26	[0.20, 0.31]	2.42	[0.79, 4.00]
	k_3 (x 10^{-1} d^{-1})	0.20	[0.20, 0.20]	3.80	[2.06, 5.94]
	k_4 (x d^{-1})	3.00	[3.00, 3.00]	5.27	[3.18, 6.99]
	k_5 (x 10^{-3} d^{-1})	3.44	[0.01, 7.87]	3.55	[0.10, 10.25]
	L (x 10^{-1})	3.00	[3.00, 3.00]	3.00	[3.00, 3.00]
Variable	A_0 (x 10^{10})	2.00	[2.00, 2.00]	2.42	[2.24, 2.67]
	k_1 (x 10^9 d^{-1})	1.29	[1.29, 1.29]	1.86	[0.50, 2.97]
	k_2 (x 10^{-1} d^{-1})	0.31	[0.23, 0.40]	2.33	[0.61, 3.99]
	k_3 (x 10^{-1} d^{-1})	2.04	[2.00, 2.69]	3.69	[2.09, 5.91]
	k_4 (x d^{-1})	6.96	[6.30, 7.00]	5.29	[3.09, 6.93]
	k_5 (x 10^{-3} d^{-1})	12.15	[0.10, 24.59]	5.31	[1.03, 12.78]
	L (x 10^{-1})	3.00	[3.00, 3.00]	3.00	[3.00, 3.00]
Spring-peak	A_0 (x 10^{10})	2.00	[2.00, 2.00]	2.57	[2.27, 2.94]
	k_1 (x 10^9 d^{-1})	1.29	[1.29, 1.29]	1.67	[0.51, 2.97]
	k_2 (x 10^{-1} d^{-1})	0.20	[0.13, 0.31]	2.01	[0.56, 3.98]
	k_3 (x 10^{-1} d^{-1})	5.11	[2.00, 6.00]	4.00	[2.07, 5.96]
	k_4 (x d^{-1})	6.80	[3.00, 7.00]	4.98	[3.04, 6.88]
	k_5 (x 10^{-3} d^{-1})	10.42	[0.10, 37.26]	9.49	[1.76, 2.58]
	L (x 10^{-1})	3.00	[3.00, 3.00]	3.00	[3.00, 3.00]

[1]95% confidence interval [Lower confidence interval, Upper confidence interval]

between the Newton and EA methods, despite the fact that some parameters could be unchanged, suggests a degree of over-parameterisation in the MG model.

The EA applied in this study is not the most efficient evolutionary-based method available. For instance, Elbeltagi *et al.* (2005) found a memetic algorithm that allows offspring to gain experience via a local search before being involved in an evolutionary process, and a particle swarm optimisation that mimics birds communicating together to find the best location (solution) to be more efficient than the Evolver® method for 20 to 100 variable optimisation problems. Differential evolution is also another optimisation technique with considerable promise (Lopez-Villalobos *et al.*, 2004).

CONCLUSION

This study of three cows has found the MG model was significantly more efficient at minimising MPE than a LP function for bi-peak and variable lactation curves, but not for a spring-peak lactation curve. Minimal differences in MPE were obtained when solving with the Newton or EA methods. Bootstrap-derived confidence intervals identified that the MG model was significantly more efficient at minimising MPE for the bi-peak lactation curve than the LP function.

ACKNOWLEDGEMENTS

The authors greatly appreciate the data provided by Graeme Pitman from the Stratford Demonstration farm and suggestions provided by Dr Val Snow of AgResearch Ltd.

REFERENCES

Bryant, J.R.; Lopez-Villalobos, N.; Holmes, C.W.; Pitman, G.D.; Brookes, I.M. 2003: Effect of genetic merit on the estimated partitioning of energy towards milk production or liveweight gain by Jersey cows grazing on pasture. *Proceedings of the New Zealand Society of Animal Production* **63**: 69-72.

Bryant, J.R.; Lopez-Villalobos, N.; Holmes, C.W.; Pryce, J.E.; Pitmam, G.D.; Davis, S.R. 2007: The effect of level of feeding, genetic merit, body condition score and age on biological parameters of a mammary gland model. *Animal* **1**: 175-183.

del Piero, S.; Melchior, A.; Polese, P.; Portanova, R.; Tolazzi, M. 2006: A novel multipurpose Excel tool for equilibrium speciation based on Newton-Raphson method and on a hybrid genetic algorithm. *Annali di Chimica* **96**: 29-49.

Efron, B. 1979: Bootstrap methods: Another look at the jackknife. *The Annals of Statistics* **7**: 1-26.

Elbeltagi, E.; Hegazy, T.; Grierson, D. 2005: Comparison among five evolutionary-based optimization algorithms. *Advanced Engineering Informatics* **19**: 43-53.

Galantai, A. 2000: The theory of Newton's method. *Journal of Computational and Applied Mathematics* **124**: 25-44.

Lopez-Villalobos, N.; Harris, B.L.; Pryce, J.E. 2004: Mate selection to control inbreeding and maximise farm profit of commercial dairy cattle. *Proceedings of the New Zealand Society of Animal Production* **64**: 122-126.

Lopez-Villalobos, N.; Spelman, R.J.; Harris, B.L.; Stanley, G.; Harcourt, S. 2005: Estimation of lactation persistency in crossbred grazing dairy cattle. *Proceedings of the New Zealand Society of Animal Production* **65**: 300-305.

Macciotta, N.P.P.; Vicario, D.; Cappio-Borlino, A. 2005: Detection of different shapes of lactation curve for milk yield in dairy cattle by empirical mathematical models. *Journal of Dairy Science* **88**: 1178-1191.

Manly, B.F.J. 1997. Randomization, bootstrap and Monte Carlo methods in biology. Chapman and Hall, Bodmin, Corwall, UK.

Mayer, A.S.; Huang, C. 1999: Development and application of a coupled-process parameter inversion model based on the maximum likelihood estimation method. *Advances in Water Resources* **22**: 841-853.

Mayer, D.G.; Belward, J.A.; Burrage, K. 1996: Use of advanced techniques to optimize a multi-dimensional dairy model. *Agricultural Systems* **50**: 239-253.

Mayer, D.G.; Belward, J.A.; Widell, H.; Burrage, K. 1999: Survival of the fittest-genetic algorithms versus evolution strategies in the optimization of systems models. *Agricultural Systems* **60**: 113-122.

Mehl, S. 2006: Use of Picard and Newton iteration for solving nonlinear ground water flow equations. *Ground Water* **44**: 583-594.

Pintér, J.D. 1996: Continuous global optimization software: A brief review. *Optima* **52** 1-8.

Rook, A.J.; Dhanoa, M.S.; Gill, M. 1990: Prediction of voluntary intake of grass silages by beef cattle 3. Principal component and ridge regression analyses. *Animal Production* **50**: 439-454.

Vetharaniam, I.; Davis, S.R.; Soboleva, T.K.; Shorten, P.R.; Wake, G.C. 2003a: Modelling the interaction of milking frequency and nutrition on mammary gland growth and lactation. *Journal of Dairy Science* **86**: 1987-1996.

Vetharaniam, I.; Davis, S.R.; Upsdell, M.; Kolver, E.S.; Pleasants, A.B. 2003b: Modelling the effect of energy status on mammary gland growth and lactation. *Journal of Dairy Science* **86**: 3148-3156.

Wilmink, J.B.M. 1987: Adjustment of test-day milk, fat and protein yield for age, season and stage of lactation. *Livestock Production Science* **16**: 335-348.

Witte, H.J.L.; Coope, G.R.; Lemdahl, G.; Lowe, J.J. 1998: Regression coefficients of thermal gradients in north-western Europe during the last glacial-Holocene transition using beetle MCR data. *Journal of Quaternary Science* **13**: 435-445.

Wood, P.D.P. 1967: Algebraic model of the lactation curve in cattle. *Nature* **216**: 164-165.

Wood, P.D.P. 1980: Breed variations in the shape of the lactation curve of cattle and their implications for efficiency. *Animal Production* **31**: 133-141.