Empirical Algebraic Modelling of Liveweight of Irish Dairy Cows over Lactation

N. Quinn[1,2], L. Killen[1], and F. Buckley[2]

[1] School of Computing, Dublin City University, Dublin 9, Ireland

[2] Dairy Production Research Centre, Teagasc, Moorepark, Fermoy, Co.Cork, Ireland

Corresponding author:

Noreen Quinn, School of Computing, Dublin City University, Dublin 9, Ireland.

E-mail Address: nquinn@computing.dcu.ie

Tel: +353 (0)1 7008449 Fax: +353 (0)1 7005442

ABSTRACT

The aim of this study is to derive an equation that has the ability to model liveweight of Irish dairy cows over lactation. The dataset consisted of 6899 cows from 63 herds of which 428 were from experimental herds and 6471 cows were from commercial herds. An initial examination focussed on time series techniques, as the data is of a time series nature. Splines were also examined to determine the dimensions of a model required to represent the data. As an incomplete gamma function, which was previously used to model milk yield, has been used in other studies to model liveweight, various milk yield models were investigated. Finally, liveweight changes between two calvings were modelled as a function of age, lactation and pregnancy. As multicollinearity was evident in this function, the variance inflation factor was examined and principal component analysis was carried out on the variables responsible for multicollinearity. The proposed liveweight model has a better fit than previous models, weak multicollinearity and the residuals are homoskedastic, independent and normally distributed. This liveweight model therefore provides an acceptable level of accuracy in representing the shape of the liveweight curve for Irish dairy cows, and can be easily modified for different environmental scenarios.

(Keywords: Liveweight, Modelling, Analysis of Residuals.)

INTRODUCTION

Accurate estimates of liveweight of individual animals can be beneficial when making management and nutritional decisions both at herd level and for individual cows (Forbes, 1983; Walter et al*.,* 1984). The Moorepark Dairy Systems Model (MDSM) (Shalloo et al., 2004) is a farm simulation model that requires precise representations of the liveweight profiles of cows under Irish production conditions. The loss and gain of liveweight has a net cost in energy within the production system, and it is necessary to include the change in liveweight in an economic model to accurately reflect the production system. A realistic model to estimate and predict liveweight change of an animal throughout the year is therefore worthy of investigation. Liveweight has been modelled using three approaches; modelling liveweight from birth to maturity (Brown et al., 1976; Bakker and Koops, 1978; Taylor, 1980; Moore, 1985; Perotto et al*.,* 1992; Berry et al*.,* 2004), using body measurements (Gravir, 1967; Heinrichs et al., 1992; Wicks, 2001; Madalena et al., 2003) and modelling liveweight over a lactation period (Wood et al., 1980; Korver et al., 1985; Berglund and Danell 1987; Lopez-Villalobos et al., 2001). The focus of interest in this study is the evolution of the liveweight of a dairy cow throughout a lactation.

In contrast to the modelling of lactation curves for both milk production and composition, relatively few researchers have contributed to the progression of work in modelling patterns in liveweight change of individual animals; Wood et al. (1980) examined the liveweight changes of several breeds of British dairy cows using an incomplete gamma function. However, this analysis was restricted to the first 20 weeks after calving. Korver et al. (1985) constructed a function from the incomplete gamma function incorporating liveweight level (scale) together with variables representing pregnancy status, the maximum decrease of liveweight during lactation, and the time during lactation at which minimum liveweight occurred. Berglund and Danell (1987) and Lopez-Villalobos et al. (2001) also used the model of Wood et al. (1980) to predict liveweight change in their respective studies.

The principal model to describe liveweight over a lactation is the model of Wood et al. (1980). As this model form was previously used to describe milk yield, other models that were used by Quinn et al. (2005a) to describe milk yield were also investigated. The suitability of the models was primarily judged on the basis of goodness-of-fit and a residual analysis was carried out to test the validity of the assumptions of regression analysis, namely autocorrelation, homoskedasticity, multicollinearity and normality of distribution of error terms. The effect of environmental and seasonal factors, independent of stage of lactation, was also examined.

MATERIALS AND METHODS

Data

The study consisted of 63 dairy herds comprising of both spring and autumn calving cows, with a potential 6,899 cows available for inclusion in the dataset. Trained Teagasc personnel visited the commercial farms up to nine times annually. Visits were carried out at two-and-a-half to four weekly intervals, with visits being more frequent in early lactation. During the visits all cows in the herd were recorded electronically, using a portable weighing scales and Winweigh software. The scales were calibrated with known weights on arrival at each farm. In all cases, recordings were taken after milking so as to minimize variations due to changing weights of gut fill. All herds were included into the Dairy Management Information System (DairyMIS) run by Moorepark (Crosse, 1986). Test day records for each individual cow were obtained from the Irish Dairy Records Cooperative. Milk recording was carried out in each herd at approximately four-week intervals and calving date records were captured through the DairyMIS system. The calving date and drying date (also obtained from the IDRC files) were used to validate lactation number and test-day records for a given lactation. Beginning 40 or 50 days after the start of the breeding season, all herds were visited on three or four occasions, at approximately 40-day intervals, to perform pregnancy diagnosis by transrectal ultrasound imaging (Aloka 210D*II, 7.5 MHz) as described by Buckley et al. (2003). Total milk yield was estimated for the commercial herds using the Ali-B model as proposed by Quinn et al. (2005a) and then quartiled for use in this analysis.

Data Editing

For the purposes of this study, lactation number was categorized as lactation 1, lactation 2 and lactation 3 or greater. Records that contained fewer than five weighings during lactation and those with no weighing post-confirmed pregnancy were removed. Thus, the edited data set consisted of 5,331 cow records of which 428 were from experimental herds and 4,903 were from commercial herds.

Models and Statistical Analysis

There are very few models to describe the pattern of liveweight change over a lactation in the literature. In recent years Korver et al. (1985), Berglund and Danell (1987) and Lopez-Villalobos et al. (2001) examined liveweight curves over a lactation, but their analyses were all based on the model of Wood et al. (1980). Initially, to investigate the type of model required to fit the data well time series techniques were examined. The dimensions of the model required to fit the liveweight data was approximated using splines. As cubic splines are the most widely used splines, they were invoked initially; a cubic spline is a third-order curve applied to a set of *m* control points. If there are one or more splines, the abscissa values of the join points are called knots. The general form of a third-order or cubic polynomial is given by the functional form:

$$f(x) = ax^3 + bx^2 + cx + d$$

where x is the variable and *a, b, c* and *d* are constant coefficients. A condition of a cubic spline is that its derivative and its second derivative are continuous at the knots and the second derivative is commonly set to zero at the endpoints to provide the boundary conditions. By subtracting the number of continuous derivatives from the total number of degrees in the spline, the dimensions of the dataset are calculated.

Once the dimensions of the data were determined, the task of deriving an equation to represent the data could be explored. The Ali-B model (Quinn et al*.*, 2005a) and the model of Ali and Schaeffer (1987) were eliminated as possible models to represent the liveweight curve as they are polynomial expressions and thus keep their concave shape. The models under consideration; Wood et al. (1980), Wilmink (1987) and Guo and Swalve (1995), (See Table 1) were tested on the basis of their goodness-of-fit and their ability to adhere to the assumptions of regression analysis. These models were fitted to pooled data using nonlinear regression and the effects of lactation number, calving month, herd and total milk yield were removed from the parameter estimates. Cows of high milk production potential have been shown to lose more liveweight in early lactation and gain less liveweight from nadir to end of lactation (Horan et al., 2005). It was therefore considered plausible to investigate the impact of lactation number and the level of production on the parameter estimates.

(Insert Table 1 near here)

The Mean Square Prediction Error (MSPE) value was used as a measure of goodness-of-fit (Kvanli, Guymes and Pavur, 1986). The Durbin-Watson statistic, *d,* was calculated for each model to test for the existence of autocorrelation between the residuals; the decision rules for autocorrelation used in this study are those outlined by Mendoza (1999). Initially, first order autocorrelation was examined and if this proved to be inconclusive, higher order autocorrelations were tested. A condition index was calculated to test for the presence of multicollinearity, values ranging from 30 to 100 indicating that moderate to strong multicollinearity was present (Belsley, Kuh and Welsch, 1980). Multicollinearity gives rise to two problems: the computation of the parameter estimates may be slow and nonconvergent, and the parameter estimates may have inflated variances (Belsley et al., 1980). To reduce multicollinearity at least one of the variables should be removed, but if this is insufficient principal component analysis (PCA) can be utilised. White's test (Sen and Srivastava, 1990) was used to test for heteroskedasticity and a mean value was computed after accounting for calving month, lactation number, herd and total milk yield. Additional tests included the Kolmogorov-Smirnov statistic (*D*), a test for normality of the distribution of the residuals as well as tests for kurtosis and skewness.

Principal component analysis (PCA) is a technique that involves the formation of new variables that are linear components of the original variables. The maximum number of new variables that can be formed is equal to the number of original variables. The first principal component (or new variable) accounts for as much of the variability in the data as possible and each succeeding component accounts for as much of the remaining variability as possible. The normal convention is to standardise the data before carrying out principal component analysis so that each recording makes an equal contribution to the total variance. Finding the principal components for two variables involves an orthogonal rotation of the axes. The first principal component will be in the direction of greatest variance and this is found by minimising the sum of the squared perpendicular distances from the observations to the first component. Once the first component is positioned, the second component is fixed since it must be orthogonal to the first. The principal components are, as a result, uncorrelated among themselves.

Once a model was identified, the parameter estimates for each lactation category were tested to see if they were significantly different. A test of homogeneity of variances was initially performed before a one-way analysis of variances could be carried out. Also the deviations found by comparing each data point with the corresponding value as estimated by the model were cumulated for each month of the year. This enabled an average effect of calendar month on liveweight, regardless of stage of lactation, to be calculated. These effects were averaged over several seasons (1995 to 2002).

RESULTS AND DISCUSSION

The preliminary examination dismissed the use of time series techniques because they require that data points occur at equal time intervals (Bowerman and O'Connell, 1987). The dimensions of the model required to fit the data were investigated and Table 2 shows that a cubic spline without a knot fitted 73 per cent of the records available in this study with an average R^2 of 0.68. A cubic spline with one knot fitted 71 per cent of the records with an average R^2 of 0.75. Although a model with four cubic splines with three knots had an average R^2 of 0.81 this could only be applied to 56 per cent of the data due to the problem of overfitting (the ratio of observations to variables being too low). Therefore, the most robust model which accurately depicted the liveweight curve over a lactation was deemed to be a four dimensional equation. Two cubic splines, (one before and one after the knot) has a total degree of six and involves two derivatives, which reduce the dimensions of the equation to four.

(Insert Table 2 near here)

The analysis of residuals showed that there was strong multicollinearity present (Table 3) when fitting the models of Wood et al. (1980), Wilmink (1987) and Guo and Swalve (1995) to pooled data, having adjusted for lactation number, calving month, total milk yield and herd effects. It can also be seen from Table 3 that there was no first order autocorrelation present in any of the models and that the residuals were homoskedastic and normally distributed. In addition the MSPE values show that there was no significant difference in the fit of the three models. The effect of multicollinearity was a severe problem with these models but removing a variable was not an option. If a variable was removed it would significantly reduce the fit of these models to the data because there were only two variables in them. As a result, other factors affecting liveweight and other techniques to reduce multicollinearity (i.e. PCA) were investigated.

(Insert Table 3 near here)

The effect of factors such as age and pregnancy, were examined firstly and it was concluded that liveweight changes of a dairy cow could be modelled as a function of age, lactation and pregnancy in the following way:

$$LW_n = f_1(age) + f_2(lactation) + f_3(pregnancy)$$

where LW_n= the liveweight in lactation week *n*. As dairy farmers in Ireland operate a strict calving pattern, the age at calving within lactation does not vary to any great extent and thus, a constant multiplied by lactation number was considered to be appropriate as the measure of age. As all of the models described in this analysis were functions of lactation, any of the three models could be used, but it was decided to use the model of Guo and Swalve (1995). This model has a slightly better MSPE value than the others (Table 3) and was overall more consistent in explaining other measures such as milk yield (Quinn et. al., 2005a), protein content and fat content (Quinn et. al., 2005c). The function described by Huggett and Widdas (1951) to represent the effect of pregnancy on liveweight was incorporated into our model. Thus the total function describing the combined effects of age, lactation and pregnancy on liveweight is as follows:-

$$LW_n = \underset{\substack{\big\downarrow \\ a(lactation\ \ number)}}{f_1(age)} + \underset{\substack{\Big\downarrow \\ b + c\sqrt{n} + d\ \ln(\ n\ )}}{f_2(lactation)} + \underset{\substack{\Bigg\downarrow \\ g(days\ \ pregnant\ \ - h)^3}}{f_3(pregnancy)}$$

where LW_n= the liveweight in lactation week *n* and *a, b, c, d, g* and *h* are parameters. As the lactation number is constant for each cow, the function of age was combined with the constant term to give the following model:-

$$LW_n = a + c\sqrt{n} + d\ln(n) + g(days\ \ pregnant - h)^3$$

When regression was performed on this equation a strong presence of multicollinearity was evident and therefore the variance inflation factor was examined to determine which variables were correlated. Not surprisingly, the terms $\sqrt{n}$ and $\ln(n)$ were found to be highly correlated, with variance inflation factor values of

25.88 and 22.58, respectively. Thus, PCA was carried out and the two correlated terms were replaced with two independent linear components:-

$$\sqrt{n} = \alpha_{11}PC1 + \alpha_{12}PC2$$

$$\ln(n) = \alpha_{21}PC1 + \alpha_{22}PC2$$

where *PC1* and *PC2* are principal component scores one and two, respectively and α_{ij} are the eigenvectors associated with the i^{th} variable and the j^{th} principal component. These two independent linear components describe all the variation in the two original variables leading to the following functional form:-

$$LW_n = a + c[\alpha_{11}PC1 + \alpha_{12}PC2] + d[\alpha_{21}PC1 + \alpha_{22}PC2] + g(m-h)^3$$

$$= a + (c\alpha_{11} + d\alpha_{21})PC1 + (c\alpha_{21} + d\alpha_{22})PC2 + g(m-h)^3$$

$$= \beta_0 + \beta_1 PC1 + \beta_2 PC2 + \beta_3 (m-h)^3$$

where $\beta_0 = a$, $\beta_1 = (c\alpha_{11} + d\alpha_{21})$, $\beta_2 = (c\alpha_{21} + d\alpha_{22})$, $\beta_3 = g$, *PC1* and *PC2* = principal component scores 1 and 2, respectively, *m*=days pregnant, and *a, c, d, g* and *h* are the original parameters.

When regression analysis was performed on this function, it was found that the parameter *h* varied considerably and it was therefore decided to keep this figure constant. Various values of the parameter *h* were tested (See Table 4) and it was found that the most satisfactory value was *h =65*. The model incorporating *h=65* had the lowest condition index, kurtosis and skewness values for the best MSPE value. Thus, the function to describe liveweight became:-

$$LW_n = \beta_0 + \beta_1 PC1 + \beta_2 PC2 + \beta_3 (m-65)^3$$

where *PC1* and *PC2* = principal component scores 1 and 2, respectively, m=days pregnant, and β_1, β_2 and β_3 are regression parameters. It was found that the effect of multicollinearity is weak when fitting this model and that the residuals were homoskedastic, independent and normally distributed (Table 4).

(Insert Table 4 near here)

Finally the values for *a, c, d* and *g* were calculated using the values of β_1, β_2 and β_3 and the eigenvectors, α_{ij}, associated with the i^{th} variable and the j^{th} principal component and were then tested to find out if the parameter estimates are significantly different for each lactation category. It was found that the assumption of homogeneity of variances was violated for parameters *a* and *g* (See Table 5). However, one-way analysis of variance is still reasonably robust when this assumption is violated if the sample sizes for the groups are equal (LeBlanc, 2004). As the sample sizes of the groups are equal in this study, ANOVA was applied and it was found that there was not a significant difference between lactations for parameters *a, c* and *d* (See Table 6), but a significant difference, at five per cent significance level, was found for parameter *g*. In order to find where exactly the differences occur, a post-hoc test assuming that the variances are not equal was performed. The Dunnett's T3 post-hoc test was carried out (Table 7) and it was found that a difference occurred between lactations 2 and 3+ for parameter *a* and between lactation 1 and 2 for parameter *g*. Although ANOVA showed that there was no significant difference between the parameter estimates for each lactation category for *a,* Dunnett's T3 post-hoc test is more reliable as it accounts for the possibility of unequal variances. Thus, a significant difference occurs between the lactation categories for at least one of the parameter estimates and it is considered necessary to have a separate model of liveweight for each lactation as follows:

$$\text{For Lactation } 1: LW_n = 538.24 - 12.73\sqrt{n} - 0.92\ln(n) + 0.000023\,(m-65)^3$$
$$\text{For Lactation } 2: LW_n = 545.26 + 18.12\sqrt{n} - 2.00\ln(n) - 0.000040\,(m-65)^3$$
$$\text{For Lactation } 3+: LW_n = 580.86 - 8.52\sqrt{n} + 0.61\ln(n) + 0.000094\,(m-65)^3$$

(Insert Tables 5, 6 and 7 near here)

The impact of the level of milk production potential on the parameter estimates was also examined. The Levene statistic (Table 8) showed that there was no evidence that the variances were not equal, at five per cent significance level, for the parameter estimates. It was found using ANOVA (Table 9) that there was no significant difference between the parameter estimates for the different milk production potential groups which is as expected as the data did not consist of many high yielding or low yielding cows.

(Insert Tables 8 and 9 near here)

As environment factors are known to have a significant effect on liveweight throughout the year (Wood et al, 1980), Table 10 shows the incremental adjustment for environmental and seasonal effects on the liveweight model. The implication of these seasonal effects is that although the function can predict liveweight at any stage in lactation, actual liveweight at any time is also influenced by a seasonal component. Table 10 shows that from January to September the liveweight function overestimates the liveweight by between 0.3 and 3.5 per cent. In the months of October, November and December the model underestimates the actual liveweight by approximately two, four and five per cent, respectively. The seasonal effects in Table 10 were added to the liveweight functions for each lactation category to account for the seasonal variations attributable to production month. Figure 1 shows the comparisons between the predicted liveweight curves for each lactation number category. It shows that the shapes of the curves are different and that three equations for each lactation category are required.

(Insert Table 10 and Figure 1 near here)

CONCLUSIONS

The aim of this study was to arrive at a well-fitting and robust form of model to represent the shape of the liveweight curve for Irish dairy cows. An examination of the liveweight data using splines indicated that a four dimensional model was required. A number of models cited in the literature were tested, and their suitability was judged on the basis of adherence to the regression assumptions and goodness-of-fit. The only assumption that was not satisfied by the original models examined was that of the explanatory variables being independent (multicollinearity). Other factors were then examined such as age and pregnancy and liveweight was deduced as being a function of age, lactation and pregnancy. It was evident by examining the variance inflation factor values that there was a strong correlation between two of the variables and as a result these were replaced by two linear independent components. Before using this model to predict the liveweight of a specific cow, adjustments are made to account for seasonal effects on liveweight.

In conclusion, the liveweight function, which accounts for the effect of age, lactation and pregnancy, is the best fitting model to explain the liveweight curve of Irish dairy cows; the effect of multicollinearity is weak and the residuals are normal, homoskedastic and independent.

REFERENCES

Bakker, H., Koops W.J. 1978. An approach to the comparison of growth curves of Dutch Friesian, British Friesian and Holstein Friesian cows. Current Topics in Veterinary Medicine. Pattern of Growth and Development in Cattle 2, 705-715

Belsley, D., Kuh, E. and Welsch E., 1980. “Regression Diagnostics: Identifying Influential Data and Sources of Collinearity”. John Wiley and Sons.

Berglund, B., Danell, B. 1987. Live weight changes, feed consumption, milk yield and energy balance in dairy cattle during the first period of lactation. Acta Agric. Scand. 37, 495-509.

Berry, D.P., Horan, B., Dillon, P. 2004. Comparison of growth curves of three strains of female dairy cattle. Anim. Sci. (In Press)

Bowerman, B., and O’Connell, R. 1987. Time Series Forcasting – Unified Concepts and Computer Implications. Duxbury Press, Boston.

Brown, J.E., Fitzhugh, H.A., Jr and Cartwright, T.C. 1976. A comparison of nonlinear models for describing weight-age relationships in cattle. J. Anim. Sci. 42, 810-818

Forbes, J.M. 1983. Models for the prediction of food intake and energy balance in dairy cows. Livest. Prod. Sci. 10, 149-157

Gravir, K. 1967. Studies on Different Body Measurements as Estimators of live and carcass weight in young NRF bulls. Acta Agric. Scand. 17, 217-227

Guo, Z. and Swalve, H.H. 1995. Modelling of lactation curve as a sub-model in the evaluation of test day records. Paper presented at the INTERBULL open meeting, 7-8 September, 1995, Prague, Czech Republic.

Heinrichs, A.J., Rogers, G.W., Cooper, J.B. Predicting Body weight and wither height in Holstein heifers using body measurements. J. Dairy Sci. 75, 3576-3581

Horan, B., Dillon, P., Faverdin, P., Delaby, L., Buckley, F. and Rath, M. 2005. Interaction of strain of Holstein-Friesian cows and pasture-based feeding systems on milk yield, body weight and body condition score. J. Dairy Sci. 88, 1231-1243

Huggett, A. and Widdas W. 1951. The relationship between mammalian foetal weight and conception age. J. Physio. 114:306-317

Kvanli AH., Guynes C.S. and Pavur R.J. 1986. Introduction to Business Statistics. Fourth Edition. West Publishing Company

Korver, S., Van Arendonk, J.A.M., Koops, W.J. 1985 A function for live-weight change between two calvings in dairy cattle. Anim. Prod. 40, 233-241.

Lopez-Villalobos, N., Lemus-Ramirez, V., Holmes, C.W., Garrick, D.J. 2001. Lactation curves for milk traits, live weight and body condition score for heavy and light Holstein-Friesian cows. Proc. New Zealand Soc. Anim. Prod. 61, 217-220

LeBlanc, D. 2004. Statistics – Concepts and Applications for Science. Jones and Bartlett Publishers.

Madalena, F.E., Teodoro, R.L., Madureira, A.P. 2003. Relationships of weight and height with age in hybrid Holstein Friesian /Guzera females. Genet. Mol. Res. 2, 271-278

Mendoza, R.U. 1999. Serial Autocorrelation: Why it's there, what it does, and how to get rid of it. Fordham University. New York.

Moore, A.J. 1985. A mathematical equation for animal growth from embryo to adult. Anim. Prod. 40, 441-453

Perotto, D., Cue, R.I., Lee, A.J. 1992. Comparison of nonlinear functions for describing the growth curve of three genotypes of dairy cattle. Can. J. Anim. Sci. 72, 773-782

Quinn N., Killen L., Buckley F. 2005a. Empirical Algebraic Modelling of Lactation Curves using Irish Data. J. Ir. Agric. & Food Res. 44 (In Press)

Quinn N., Killen L., Buckley F. 2005b. Modelling the Fat and Protein Content Curves for Irish Dairy Cows. J. Ir. Agric. & Food Res. (Submitted)

Sen, A. and Srivasta, M. 1990. "Regression analysis-Theory, Methods and Applications." Springer-Verlag. New York.

Shalloo, L. Dillon, P., Rath, M. and Wallace, M. 2004. Description and validation of the Moorepark Dairy Systems Model (MDSM). *J. Dairy Sci.* 87, 1945-1959.

Taylor, St. C.S. 1980. Genetically Standardized Growth Equations. Anim. Prod. 30, 167-175

Walter, J.P., Mao, I.L., Emery, R. S. 1984. Simultaneous fitting of energy requirements and intake, lactation response and body weight changes in dairy cattle, with different forage bases rations. J. Dairy Sci. 70, Supplement, Abstract P221.

Wicks, H.C.F. 2001. The influence of genetic merit and environment on dairy cattle performance on farm. PhD Thesis, Imperial College of Wye, University of London

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

Wood, P.D.P., King, J.O.L., Youdan, P.G. 1980. Relationships between size, liveweight change and milk production characters in early lactation in dairy cattle. Anim. Prod. 31, 143-151

Table 1: Selection of models investigated

Year	**Author**	**Model***
1967	Wood	$LW_n = an^{-b}e^{cn}$
1987	Wilmink	$LW_n = a + be^{-kn} + cn$
1995	Guo & Swalve	$LW_n = a + b\sqrt{n} + c\ln(n)$

* LW_n is the liveweight in lactation week n.

Table 2: Fit of Cubic Splines with different numbers of knots

No. of Knots	**R^2**	**% of data lost due to overfitting**
0	0.68	27
1	0.75	29
2	0.78	35
3	0.81	44
4	0.83	48

Table 3: Comparison of Models

Test	**Wood**	**Wilmink**	**Guo & Swalve**
MSPE	924.513	921.87	918.69
R^2	0.40	0.40	0.40
Autocorrelation	None	None	None
Heteroskedasticity	None	None	None
Normality	Normal	Normal	Normal
Kurtosis	0.66	0.64	0.67
Skewness	-0.01	0.04	0.01
Multicollinearity (Condition Index)	Strong 227341	Strong 232083	Strong 227525

[*] LW_n is the liveweight of a cow in lactation week *n*

Table 4: Goodness-of-fit and Analysis of Residuals of Liveweight Function for various values of *h*

Test	**h=25**	**h=65**	**h=100**	**h=150**
MSPE	685.30	682.93	693.06	685.40
R^2	0.54	0.54	0.55	0.54
Autocorrelation	None	None	None	None
Heteroskedasticity	None	None	None	None
Normality	Normal	Normal	Normal	Normal
Kurtosis	0.53	0.44	0.45	0.58
Skewness	-0.14	-0.14	-0.18	-0.13
Multicollinearity (Condition Index)	Weak 12.25	Weak 15.66	Weak 24.01	Weak 31.27

Table 5: Test of Homogeneity of variances for the parameter estimates of the proposed model for different lactation groups

Parameter Estimate	**Levene Statistic**	**df1**	**df2**	**p-value**
A	7.56	2	19	0.004
C	2.19	2	19	0.139
D	1.66	2	19	0.217
G	8.93	2	19	0.002

Table 6: One-way Analysis of Variance to compare the parameter estimates of the proposed model for each lactation group

Parameter Estimate	**Source**	**df**	**Sum of Squares**	**Mean Square**	**F-value**	***p-value***
A	Between Groups	2	7755.29	3877.65	2.66	0.096
	Within Groups	19	27752.66	1460.65		
	Total	21	35507.66			
C	Between Groups	2	4180.94	2090.47	2.86	0.082
	Within Groups	19	13882.54	1460.65		
	Total	21	35507.66			
D	Between Groups	2	27.52	13.76	0.83	0.452
	Within Groups	19	315.59	16.61		
	Total	21	343.11			
G	Between Groups	2	0.000000073	0.000000036	4.20	0.031
	Within Groups	19	0.000000016	0.0000000086		
	Total	21	0.00000024			

Table 7: Using Dunnett's T3 post-hoc test to find where the differences occur between lactation for each parameter estimate in the proposed model

Parameter Estimate	**Lactation Number (i)**	**Lactation Number (j)**	***p-value***
A	1	2	0.990
	1	3+	0.386
	2	3+	0.034
C	1	2	0.201
	1	3+	0.993
	2	3+	0.138
D	1	2	0.846
	1	3+	0.898
	2	3+	0.565
G	1	2	0.024
	1	3+	0.471
	2	3+	0.094

Table 8: Test of Homogeneity of variances for the parameter estimates of the proposed model for different milk production potential groups

Parameter Estimate	**Levene Statistic**	**df1**	**df2**	**p-value**
A	0.72	3	21	0.550
C	1.73	3	21	0.192
D	1.60	3	21	0.220
G	1.00	3	21	0.411

Table 9: One-way Analysis of Variance to compare the parameter estimates of the proposed model for each milk production potential group

Parameter Estimate	**Source**	**df**	**Sum of Squares**	**Mean Square**	**F-value**	***p-value***
A	Between Groups	3	5058.36	1686.12	1.29	0.303
	Within Groups	21	27421.82	1305.80		
	Total	24	32480.18			
C	Between Groups	3	3763.59	1254.53	1.45	0.258
	Within Groups	21	18231.48	868.17		
	Total	24	21995.07			
D	Between Groups	3	2703.56	901.19	1.73	0.191
	Within Groups	21	10928.24	520.39		
	Total	24	13631.80			
G	Between Groups	3	0.000	0.000	0.39	0.760
	Within Groups	21	0.000	0.000		
	Total	24	0.000			

Table 10: Seasonal Deviations on the proposed model, independent of stage of lactation

Month	**Liveweight (%)**
January	-3.03
February	-3.01
March	-3.53
April	-2.95
May	-2.95
June	-1.57
July	-1.50
August	-3.27
September	-0.33
October	2.06
November	4.27
December	5.14

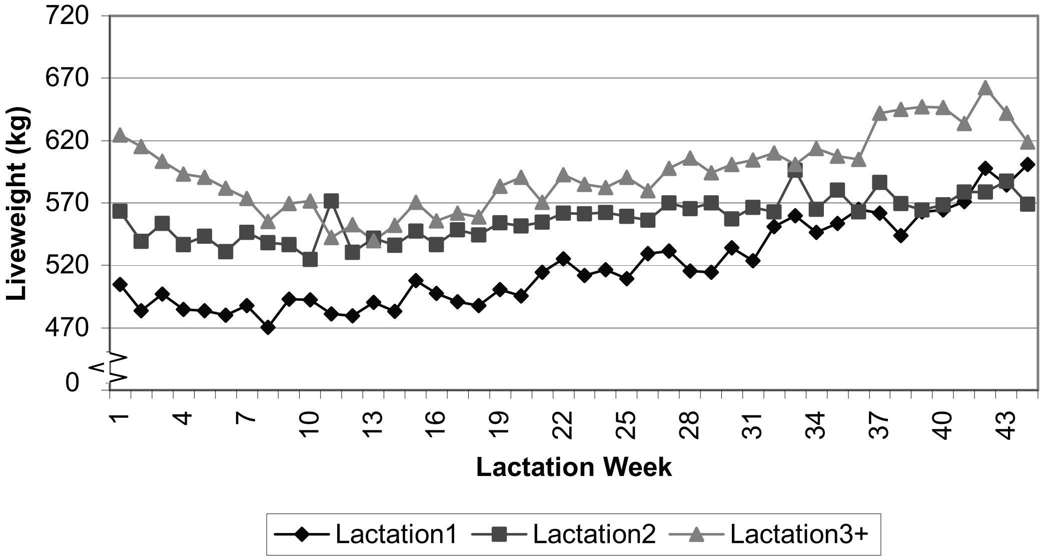

Figure 1: Comparison of the predicted liveweight curves for each lactation category