Best Management Strategies for the Australian Dairy Industry Using Multiple Linear Regression Techniques

Daniel Zamykal[1], Michael Steele[1,*], Don Kerr[2] and Janet Chaseling[2]

[1]James Cook University, [2]Griffith University

[*]Corresponding author: Mike.Steele@jcu.edu.au

Abstract

Dairy farmers in Australia are under increasing pressure to demonstrate that their production systems are environmentally friendly (Armstrong et al. 2000). These farmers need assistance in evaluating the impact on their profitability of a range of environmental and social management practices. In 2000, the Australian Dairy Research and Development Corporation conducted a large-scale telephone survey of 1826 Australian dairy farms to examine current on-farm management practices in relation to milk production and farm and farmer demographics (DRDC 2001). The questionnaire results from the 196 dairy farms in the sub-tropical region of South East Queensland and Northern New South Wales were analysed in three steps. Firstly, multiple linear regression models were produced for all dairy farms within the region. Secondly, farms that produced less than 750kL per annum were treated as a separate population and a specific model was developed for these cases. Finally, farms that produced in excess of 750 kilolitres (kL) were analyzed and a linear model was developed for members of this population. An important outcome of the analysis was the construction of management profiles of farmers in different milk production classes, in particular, identification of the specific management practices which define a farmer's ability to produce optimal yield.

KEY WORDS: Multiple linear regression models, dairy farms.

Introduction

Within the Australian rural sector, the Dairy Industry is a major contributor to the national economy by employing over 60,000 people and generating $2 billion dollars annually. However, the industry has undergone significant rationalization in the last 30 years resulting in the decline of dairy farms from 30,000 in 1975 to less than 15,000 today (DRDC, 2001).

To satisfy the national demand for dairy products while improving environmentally sustainable practices a partnership between the Dairy Research and Development Corporation, National Land and Water Resources Audit, Australian Dairy Farmers Federation and the Australian Dairy Products Federation was formed to initiate the program 'Sustaining our Natural Resources-Dairying for Tomorrow'. The specific purpose of this cooperative was to survey dairy farmers in order to asses and quantify management practices on Australian dairy farms.

In order to investigate the management practices on South East Queensland and Northern New South Wales dairy farms, three multiple linear regression models were developed. Firstly, analyses of industry survey results provided a method for estimating the response rates of milk production from all surveyed farms based on demographics and management practices. Secondly, models where developed for operations producing less than 750 kilolitres (kL) and greater than 750 kL total yields. Comparison of farms producing greater than 750 kL and farms producing less than 750 kL was used to isolate specific contributors to milk production.

Materials and methods

Multiple Linear Regression

The primary goal of developing a linear model was to estimate milk production from all 196 farms surveyed and the two production classes within the greater survey population. To this end two dependent variables were considered: (i) Total milk production on farm and (ii) total milk produced per cow. Previous studies have used milk production per cow (King *et al.*, 1980), however, these measures are difficult to interpret in a meaningful way due to the variety of land and soil types across the surveyed region (Kerr *et al.*, 1995; Kerr *et al.*, 1998). Therefore, the dependent variable selected for this study was total milk produced per farm. A stepwise selection technique was applied to all continuous demographic and management related variables to select the most important factors affecting milk production. Variables were considered to have a significant effect on the total milk production if the regression coefficient was significant at the $\alpha = 0.05$ level. For the developed model to be considered appropriate, the residuals should be approximately normally distributed and exhibit constant variance.

Results

All Producers from the Surveyed Region

The initial model was developed for the total 196 farms in the survey region. The variables that significantly contributed to total milk production were: (i) total number of cows per farm; (ii) the amount of money spent on fertilizer in an average year; (iii) the amount of nutritional supplements given; (iv) the size of the farm in hectares and (v) the years of experience a dairy farmer possess (Table 1). Investigation of the residuals revealed approximate normality and constant variance.

Table 1: Unstandardized regression coefficients, standard errors and model R^2 values for the total survey population, *n =196*

Variables added	Unstandardized coefficients	Standard error	R^2 value
Number of cows per farm	3057.155	234.413	0.616
Money spent on fertilizer	6.662	0.890	0.702
Nutritional supplements	13.196	3.890	0.715
Size of farm (ha)	313.534	117.410	0.723
Years of dairy experience	-1836.749	729.070	0.732

*constant = -31068.8

The order of importance each variable contributes to the model is reflected in the cumulative R^2 values. The most influential variable to affect the model is the total number of cows per farm which accounts for 61.6% of the variation in the total milk production. The remaining variables only contribute a further 11.6% of the variation in the total volume of milk produced. Unexpectedly, the years of experience a farmer has negatively affects the milk production.

The multiple regression model for all producers is as follows:

$$\begin{aligned}\text{Milk production (kL)} &= -31068.8 + 3057.155(\text{number of cows}) + 6.662(\text{amount spent on fertiliser (p.a.)}) \\ &+13.196(\text{nutrional supplements}) + 313.534(\text{property area (ha)}) - 1836.749(\text{years of experience})\end{aligned} \quad [1]$$

Milk Production Dichotomy

The South East Queensland and Northern New South Wales dairy region is assumed to consist of two separate milk-producing populations. The first is dairy farmers producing less than 750 kL of milk per year and the second consist of farms producing in excess of 750 kL of milk per year. This dichotomy is based on analyses of survey results conducted in 1986-87, 1990-91 and overseas (Kerr *et al.*, 1998). The results of these studies concluded that high producing farms appear to be more efficient due to improved economies of scale and should be considered as a separate population (Kerr *et al.*, 1995; Kerr *et al.*, 1998)

Producers with Yield Less Than 750 kL Per Annum

Multiple regression analysis was performed for all continuous demographic and management related variables for farms producing less than 750 kL per annum. The variables selected by the stepwise selection technique were: (i) total number of cows per farm; (ii) the amount of money spent on fertilizer in an average year; (iii) the amount of nutritional supplements given; (iv) and the percentage of debt in relation to farm worth (Table 2). Further investigation revealed that the residuals were approximately normally distributed and exhibit constant variance.

Table 2: Unstandardized regression coefficients, standard errors and model R^2 values for milk producers less than 750kL, $n = 165$

Variables added	Unstandardized coefficients	Standard error	R^2 value
Number of cows per farm	2516.727	235.093	0.431
Money spent on fertilizer	4.429	1.094	0.480
Nutritional supplements	9.120	3.485	0.500
Debt percentage	1050.171	501.509	0.513

*constant = 35822.331

The order of importance each variable contributes to the model is reflected in the cumulative R^2 values. The most influential variable to affect the model is again total number of cows per farm which accounts for 43.1% of the variation in the total milk production. The remaining variables only contribute a further 8.2% of the variation in the total volume of milk produced.

The multiple regression model for farms that produce less than 750kL is as follows:

$$\text{Milk production (kL)} = 35822.331 + 2516.727(\text{number of cows}) + 4.429(\text{amount spent on fertiliser (p.a.)}) + 9.120(\text{nutrional supplements}) + 1050.171(\text{debt percentage}) \quad [2]$$

Producers with Yield Greater Than 750 kL Per Annum

Multiple regression analysis was performed for all continuous demographic and management related variables for farms producing greater than 750 kL of total milk production. The variables selected by the stepwise selection technique are: (i) total number of cows per farm; (ii) and the use of phosphorus (Table 3). Subsequent investigation revealed that the residuals where approximately normally distributed and exhibited constant variance.

Table 3: Unstandardized regression coefficients, standard errors and model R^2 values for milk producers greater than 750kL, $n = 34$

Variables	Unstandardized coefficients	Standard error	R^2 value
Number of cows per farm	3250.596	606.223	0.369
Phosphorus	-571853.000	216228.900	0.485

*constant = 371614.500

The most influential variable to affect the model is the total number of cows per farm which accounts for 36.9% of the variation in the total milk production. The remaining variable only contributes to a further 11.6% of the variation in the total milk produced. Intriguingly, the use of phosphorus negatively affects the amount of milk produced by these farmers.

The multiple regression model for farms that produce greater than 750 kL is as follows:

$$\text{Milk production (kL)} = 371614.500 + 3230.596(\text{number of cows}) - 571853(\text{phosphorus usage}) \quad [3]$$

Model Comparison

The initial model for all producers' attempts to predict milk production based on shared demographic and management characteristics. This model can account for a significant amount of milk production by the dairy farmers of South East Queensland and Northern New South Wales. Clearly, any subdivision of producers into sub-populations will alter the overall model and will produce a specific model for that sub-class. Any reduction in the R^2 values for each sub-population model is expected due to the reduction in sample size.

For the majority of milk producers who yield less than 750 kL of milk per annum the first three variables are similar to the initial model with the addition of dept levels as a percentage of total farm worth. The minority of farms that produce in excess of 750 kL of milk per annum is significantly affected by the number of cows per farm and the use of phosphorus which negatively affects the potential yield. However, the later model may be unduly affected by the small sample size of $n = 34$ which may represent anomalies present with the sample sub-population.

Conclusion

For the individual dairy farmer within the South East Queensland and Northern New South Wales dairy region it is possible to model their expected milk production based on the model parameters. The simplistic nature of the three linear models allows a farmer to 'plug in' the required information and produce an expected total yield. He or she can either use a specific model tailored for a specific milk producing population or use an overall model for all dairy farms within the region. The usefulness of such an approach removes the guess work for each producer and highlights the important demographic and management practices that influence the total milk production within their region.

References

Armstrong, D.P., J.E. Knee, P.T. Doyle, K.E. Pritchard and Gyles, O.A. 2000. Water-use efficiency on irrigated dairy farms in northern Victoria and southern New South Wales Australian Journal of Experimental Agriculture 40: 643-653.

DRDC. 2001. Sustaining our Natural Resources—Dairying for Tomorrow. Dairy Research and Development Corporation, Melbourne, Australia

Kerr, D.V., Chaseling, J., Chopping, G.D., Davison, T.M., and Busby, G. 1998. A study of the effect of inputs on level of production of dairy farms in Queensland - a comparative analysis of survey data. Australian Journal of Experimental Agriculture 38: 419-425

Kerr, D.V., Davison, T.M., Cowan, R.T., and Chaseling, J. 1995. Factors affecting productivity on dairy farms in tropical and sub-tropical environments Asian Australasian Journal of Animal Science 8: 505-513.

King, K.R., Stockdale, C.R., and Patterson, I.F. 1980. The effect of restriction of pasture intake in late lactation on the milk production and body condition of dairy cows. Australian Journal of Experimental Agriculture and Animal Husbandry 20: 389-393.