THE USE OF REGRESSION TREE ANALYSIS TO IDENTIFY INTERACTIONS BETWEEN ON-FARM FACTORS AFFECTING MILK PROTEIN CONTENT

D.G. BARBER[A,B], N.R. GOBIUS[B], I.J.C. HANNAH[B] and D.P. POPPI[A]

[A] Schools of Animal Studies and Veterinary Science, University of Queensland, St Lucia, Qld 4072
[B] Department of Primary Industries, Mutdapilly Research Station, Peak Crossing, Qld 4306

On-farm milk protein content (%, m/v) is affected by a number of factors, including breed, genetics, environment, management practices, disease, stage of lactation and nutritional management (De Peters and Cant 1992). Large-scale on-farm monitoring programs provide the opportunity to include a large number of factors in an analysis, however, there are limited statistical methods available to evaluate the interactions between these on-farm factors. Regression tree analysis is a new method to analyse large data sets and provides a map, in the form of a tree, of the interactions between factors and the quantitative branch points. The tree can then be utilised on subsequent datasets to predict milk protein content from a known set of variables.

Twelve commercial dairy farms in south-east Queensland were monitored for nutritional, environmental, animal and management factors associated with milk protein content over a 13-month period from September 2000 to September 2001. The dietary composition for each farm was analysed and evaluated from nutritional data using NRC (2001). Average monthly data for each farm were analysed using the BRegression procedure in Genstat.

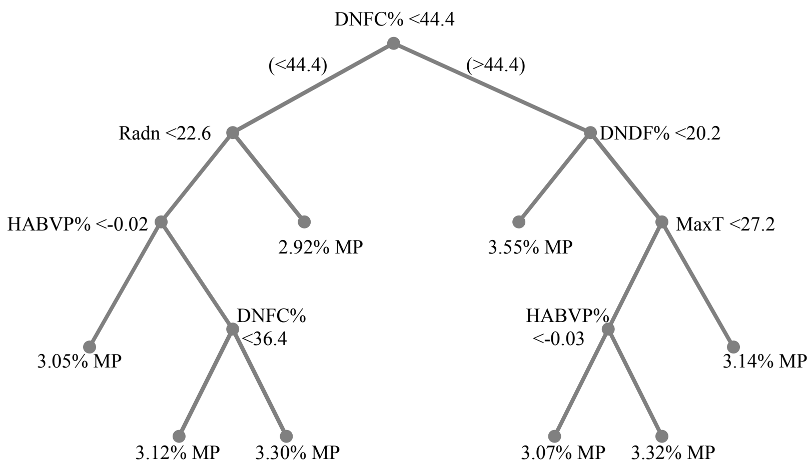

Figure 1. Regression tree of the 5 main factors affecting on-farm milk protein content (%, m/v). (Key: DNFC% = diet non-fibre carbohydrate content (% of DM), Radn = radiation (watts/m^2), DNDF% = diet NDF content (% of DM), HABVP% = herd ABV for milk protein content, MaxT = maximum temperature (^{o}C) and % MP = milk protein content (%, m/v)).

The primary factor affecting milk protein content was the dietary non-fibre carbohydrate (DNFC) content of the diet. Regression tree analysis also identified an interaction between nutrition (DNFC and NDF content), environment (radiation and maximum temperature) and genetics (average herd ABV for milk protein content) as the main factors associated with milk protein content (Figure 1). Figure 1 shows that low milk protein content can arise from quite different combinations of factors, and that the tree analysis provides a means of identifying and developing different strategies for alleviating low milk protein. In conclusion, regression tree analysis highlighted an interaction between nutrition, genetics and climate as having the greatest effect on milk protein content, and the method identified quantitative values and major components by which to assess on-farm data and suggest the quantitative changes needed in various variables to increase milk protein content.

The study is part of a Dairy Australia funded national project.

DE PETERS, E.J. and CANT, J.P. (1992). *J. Dairy Sci.* **75**, 2043-2070.

Email: David.Barber@dpi.qld.gov.au