Validating a Model that Predicts Daily Growth and Feed Quality of New Zealand Dairy Pastures

Simon J. R. Woodward, Systems Modelling Group, AgResearch, Private Bag 3123, Hamilton, New Zealand.
Tel +64 7 838 5914. Fax +64 7 838 5117. Email woodwards@agresearch.cri.nz.

Abstract The Pasture Quality (PQ) model is a simple mechanistic dynamical system model that was designed to capture the essential biological processes in grazed grass-clover pasture, and to be optimised to derive improved grazing strategies for New Zealand dairy farms. While the individual processes represented in the model (photosynthesis, tissue growth, flowering, leaf death, decomposition, worms) were based on experimental data, this did not guarantee that the assembled model would accurately predict the behaviour of the system as a whole (i.e., pasture growth and quality). Validation of the whole model was thus a priority, since any strategy derived from the model could impact a farm business in the order of thousands of dollars per annum if adopted. This paper describes the process of defining performance criteria for the model, obtaining suitable data to test the model, and carrying out the validation analysis. The validation process highlighted a number of weaknesses in the model which will lead to the model being improved. As a result, the model's utility will be enhanced. Furthermore, validation was found to have an unexpected additional benefit, in that despite the model's poor initial performance, support was generated for the model among field scientists involved in the wider project.

1. INTRODUCTION

On New Zealand dairy and drystock farms, profitability depends on the farmer's ability to match the seasonal energy demands of his or her animals to the seasonal pattern of pasture growth and quality. If grain or silage has to be purchased in order to supplement animals' pasture intakes, profits are affected. While the seasonal pattern of animal requirements is well understood [Holmes and Wilson 1984], pasture growth and quality are difficult to predict, being climate driven, and affected by grazing. In particular, the timing and intensity of grazing in spring affects the extent to which pasture grasses go to seed. Unless grass plants are close-grazed early on in their reproduction development, seed heads will grow tall before dying, leaving large quantities of dead leaf and stem material in the pasture through summer and autumn. As well as being unpalatable to animals, this material is of low quality, and animal production can be poor as a result.

This paper outlines the context and design of a pasture model built to aid in pasture management decisions on farm. Particular attention is given to testing the model against empirical data. Almost all published grassland models have been subject to some degree of validation [e.g., Hacker et al. 1991; Bacsi and Zemankovics 1995]. However, because appropriate pasture growth data is usually hard to come by, especially if the model requires accompanying soil and climate information, comparison of total biomass production is often all that is attempted. For some models, this may be sufficient, since the purpose of validation is to demonstrate "that a model is acceptable for its intended use because it meets specified performance requirements" [Rykiel 1996]. In our case, the model must be shown to produce acceptable predictions of pasture quantity and quality for use in feed budgeting, so more stringent validation is required.

2. PASTURE QUALITY MODEL

The pasture quality (PQ) model [Woodward et al. unpublished] is being developed to support grazing management decisions on New Zealand farms. An earlier model developed by our group modelled only green and dead fractions in hill country pastures [McCall 1984]. In the PQ model, as well as predicting the quantity of pasture production through the year, we wished to predict the

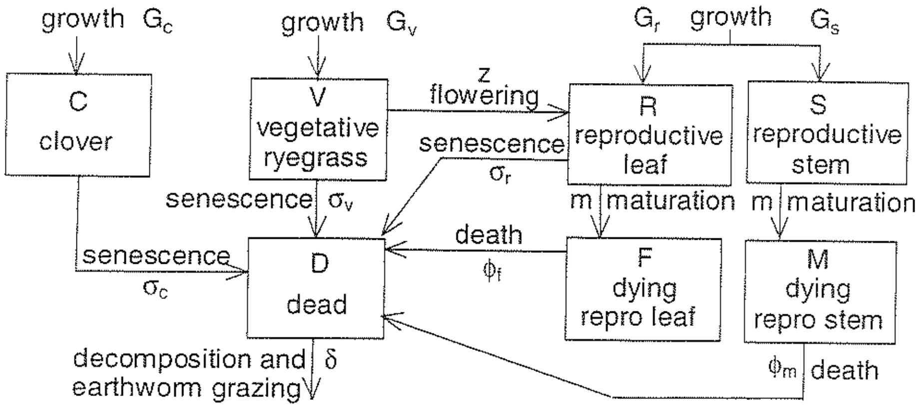

Figure 1 Pasture fractions and processes represented in the PQ model.

accumulation of reproductive stem and dead fractions in summer pastures. This is influenced by grazing patterns in spring, because grazing of developing reproductive stem is desirable for the reasons outlined above. In the PQ model, therefore, *pasture quality* refers to the relative quantities of grass leaf and stem, clover and dead matter in pasture at a particular time, rather than the overall nutritional value of the pasture, which is commonly expressed as megajoules metabolisable energy per kilogram dry matter (MJME $kgDM^{-1}$).

To achieve this purpose, the model is designed to predict daily growth of vegetative ryegrass, reproductive ryegrass, clover and dead fractions (Figure 1) in New Zealand pastures, based on daily temperature, rainfall and radiation data. We initially focused on high quality, high fertility, perennial ryegrass–white clover pastures such as those found on most dairy farms.

The model is defined by a set of differential equations:

$$
\begin{aligned}
C' &= G_c - \sigma_c C \qquad (1)\\
V' &= G_v - zV - \sigma_v V\\
R' &= G_r + zV - \sigma_r R - mR\\
S' &= G_s - mS\\
F' &= mR - \phi_f F\\
M' &= mS - \phi_m M\\
D' &= \sigma_c C + \sigma_v V + \sigma_r R + \phi_f F + \phi_m M - \delta D
\end{aligned}
$$

which are used to calculate net daily growth of each of the seven pasture fractions. The growth functions G_c, G_v, G_r and G_s are complex, since the growth of one component depends on shading within the whole pasture canopy. The parameters for the various biological processes (z, m, σ_i, ϕ_i, δ) are dependant on temperature, soil water and time of year. Soil water is calculated using a modification of the model of Scotter et al. [1979] which gives the soil water deficit in a surface soil layer and in the whole soil profile separately.

3. VALIDATION

A certain degree of validation naturally takes place as the structure and functions of the model are designed and constructed. For example, the state variables were chosen to ensure that the model would perform its intended use, and the functions describing the particular physiological processes (growth, senescence, decomposition, etc.) were based on empirically observed relationships, as well as being calibrated to experimental data.

However, validation of model inputs does not guarantee that the *outputs* will perform to specification. For this reason, further validation was required to show that the model gives good quantitative predictions of pasture growth and quality *at the scale at which it is to be used*, that is, in this case, to establish grazing management recommendations.

Rykiel [1996] points out that model validation can take a large number of forms. There are a number of statistics available for comparison of model predictions to data [Bacsi and Zemankovics 1995], and validation may also include subjective assessment by "knowledgeable individuals" [Rykiel 1996]. In this paper we are concerned with validation against historical data other than that used to build the model.

4. VALIDATION DATA

If possible, data for model evaluation should be chosen to isolate model components and validate these independently. In terms of the PQ model, ideal validation data would have been from an irrigated, well fertilised, pure ryegrass dairy pasture. This would have removed the fertility limiting, soil water, and clover components of the model from consideration, and thus focused attention on the performance of the ryegrass and dead components.

However, the data actually used in the first round

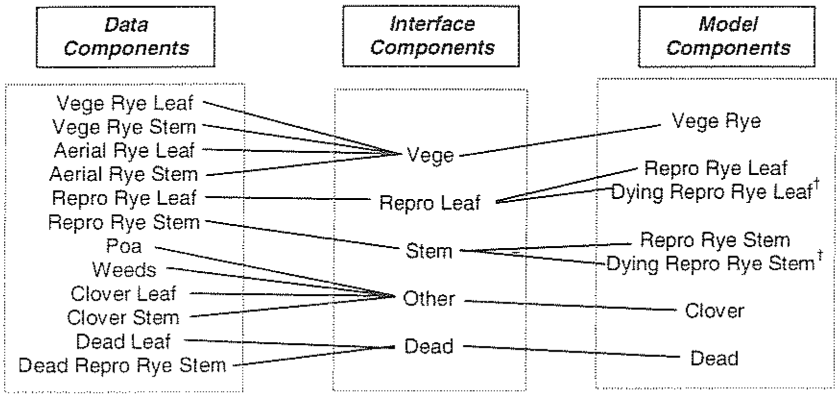

Figure 2 Matching model structure to data structure.
†"Dying" refers to reproductive ryegrass tillers which have been decapitated by grazing or whose seed head has matured.

of validation did not match these criteria. These data came from a plot experiment described in Korte et al. [1984]. Sixteen ryegrass-white clover plots were intensively sampled from September 1976 to May 1977, while subject to four grazing treatments (4 replicates each) with contrasting timing and intensity of grazing. Twelve herbage components were measured at approximately weekly intervals. The best weather and soil information available was from Palmerston North, 14 km to the west of the trial site. Because the data covered the reproductive and post-reproductive seasons, and because reproductive leaf and stem fractions for measured separately, this data set was particularly suitable for testing the reproductive predictions made by the model.

5. BASIS OF COMPARISON

The pasture components measured in the data set differed from those simulated by the model. This made it necessary to construct a set of "interface components" which matched both data sets (Figure 2). The model predictions were converted into interface components, and these were then compared with the interface components calculated from the data.

In testing a model, it is desirable to be explicit about how accurate we require its predictions to be [Rykiel 1996]. Rykiel [1996] suggests that model outputs should fall within the 95% confidence interval 75% of the time. However, the use of statistical comparison may be restricted by the accuracy of the data we are using. Pasture data is notoriously variable, even from plots receiving identical treatments, and Korte's study consisted of only four replicates. The coefficients of variation for the data points averaged 53%. Thus the data themselves were not very reliable. Any statistic based on confidence intervals [Law and Kelton 1991; Rykiel 1996] would therefore be of little use, since a model may easily predict within the confidence interval of the data but still not yield useful predictions [Klepper and Rouse 1991].

A more suitable method of evaluating the model in this case was to compare it to another model, by comparing the residual sums of squares. This generates an R^2 value [Kvålseth 1985],

$$R^2 = 1 - \frac{\sum \frac{(\hat{y}_i - y_i)^2}{\sigma_i}}{\sum \frac{(y_i^* - y_i)^2}{\sigma_i}} \tag{2}$$

where y_i and σ_i are the observation means and standard errors, $\hat{y}_i$ are the model predictions, and y_i^* are the predictions from a comparison model. Typically, the comparison is against the mean of the data, i.e., $y_i^* = \overline{y}$. The R^2 value indicates how much of the residual sum of squares from the first model (y_i^*) is explained by the second model

Interface Component	*Mean Accumulation Rate (kg DM ha^{-1} day^{-1})*	*Mean Accumulation Rate Sum of Squares*	*Model Sum of Squares*	*n*	R^2
Vege	22.9	17.6	18.3	98	−0.03
Repro Leaf	3.0	50.6	12.3	71	0.76
Stem	4.8	128.6	4.9	100	0.96
Other	12.8	11.6	11.9	98	−0.02
Dead	2.7	9.0	17.0	98	−0.88

Table 1 Results of comparing the mean accumulation rates of each interface component and the model predictions to the data. *n* is the number of observations compared, and R^2 indicates the improved predictions from using the model compared to using the mean accumulation rate.

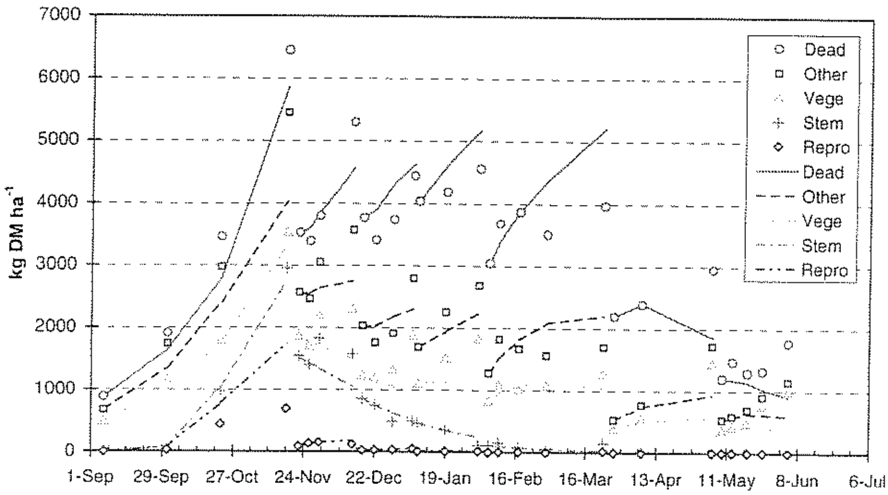

Figure 3 Initial comparison of model outputs (lines) against Korte et al. [1984] treatment T2L data (symbols).

($\hat{y}_i$).

In our case, we would consider the model to be doing quite well if could predict the seasonal mean rate of accumulation of each herbage fraction. Calculating the mean accumulation rate of each interface component in Figure 2 from the data and using these as our y_i^* therefore provides a useful test of the model (Table 1). An R^2 greater than zero in this case implies that the model not only predicts the mean accumulation rate over the season but also some of the seasonal pattern of herbage accumulation as well. An R^2 of 0.50 or greater would almost certainly be sufficient for the model's intended purpose here.

6. RESULTS

Without calibrating the model by adjusting its parameters, initial comparisons were made with the data (all four treatments simultaneously). Figure 3 shows one of the data treatments and the model predictions (both converted to interface components). The performance of each component was analysed separately (Table 1). The model performed better than the mean for the Repro Leaf and Stem components, similarly to the mean for the Vege and Other components, and worse than the mean for the Dead component.

One method of isolating errors in the model is to simulate only one model component at a time, obtaining values for the other state variables by interpolating the data. This approach has previously been used by McCall [1984].

In Figure 4, only the dead fraction was determined by the model and the other fractions were linearly interpolated from the data. This allowed us to focus on the equations that predict flow into and out of the dead herbage category only, i.e. senescence, death and decomposition (Figure 1). The model overestimated accumulation of dead material from November to March. After some testing, it was realised that the likely cause of this was that the senescence functions used by the model were constructed using only winter

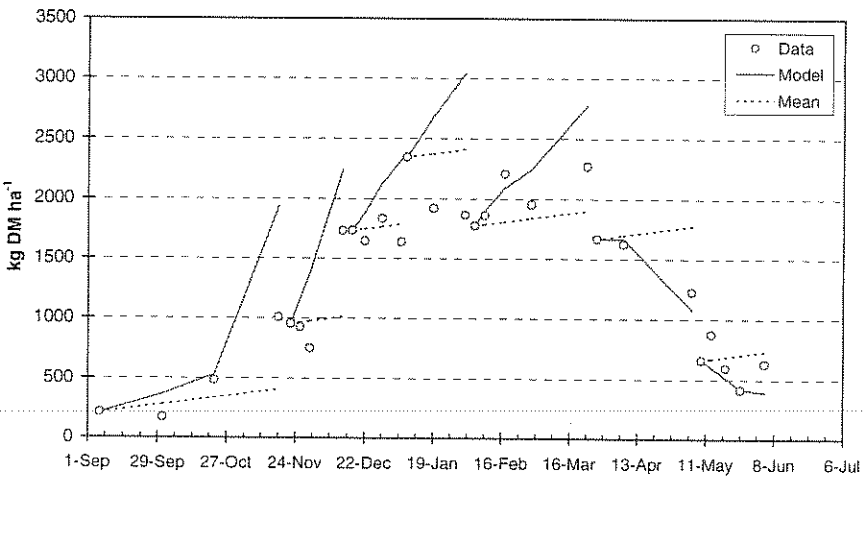

Figure 4 Model predictions of dead accumulation only (other components estimated from data), compared to the mean observed rate of dead matter accumulation, and to the data themselves.

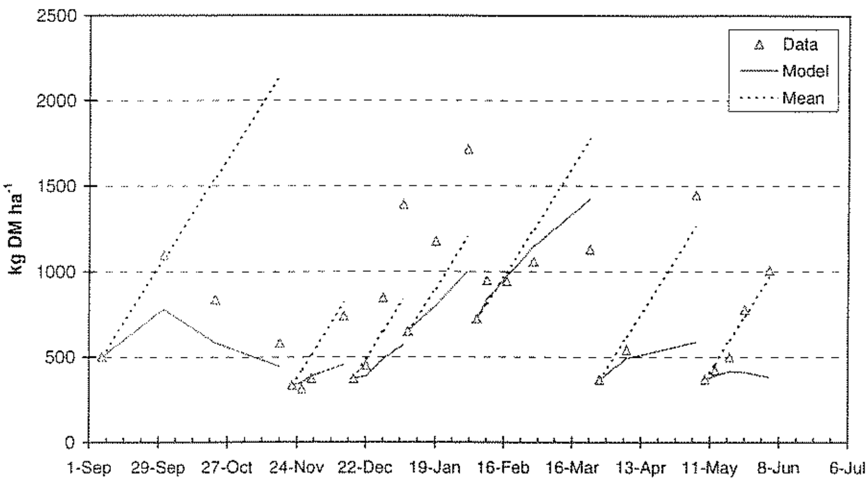

Figure 4 Model predictions of vegetative ryegrass accumulation (other components estimated from data), compared to the mean observed rate of accumulation, and to the data themselves.

senescence data, and greatly over-predicted senescence through the summer period [Woodward 1998]. As a result, the senescence functions are currently being revised to give better predictions of leaf death in summer.

Similarly, in Figure 5, only the vegetative ryegrass fraction was simulated, and the other fractions were interpolated from the data. In general the model underestimated accumulation of vegetative material over this period. This may partly be due to the overestimate of leaf senescence rates as mentioned above. It is also possible that the leaf photosynthesis parameters used in predicting vegetative growth are incorrect, as during model construction it was difficult to find appropriate data for estimating these under field conditions.

7. CONCLUSIONS

Although simulation modelling has become a popular tool in agricultural science, model validation remains an underrated and unpopular discipline. Perhaps this is understandable given the difficulty in obtaining appropriate validation data and the pessimism of some researchers regarding the value, *and even the possibility*, of validating models of natural systems [Oreskes et al. 1994]. Furthermore, lack of, or minimal, validation of a model does not necessarily mean that the model cannot be used—this would depend on the particular use to which the model is to be put [Hodges 1991; Rykiel 1996].

Nevertheless, the process of defining performance specifications for a model and comparing the model to real data yields deep insights into the performance of the model relative to the real world. Seeing comparisons between model predictions and field data also often has a persuasive effect on less mathematically inclined researchers, who are then better able to appreciate both the weaknesses in the field data and the real achievements, no matter how small, made by models in predicting biological dynamics.

The model presented in this paper is receiving an uncomfortably thorough evaluation as a result of the validation process described here. While earlier simulations had suggested that the model predictions were qualitatively correct, comparison to data has revealed several significant flaws, as we have seen. Rather than being destructive, this has presented an opportunity to improve the model, and thus to greatly increase its utility in the long run.

Furthermore, disseminating the results-in-progress of the validation process has captured the imagination of a number of field agronomists, who as a result have come out surprisingly supportive of the model. This in itself is sufficient payback for undertaking model validation.

8. REFERENCES

Bacsi, Z. and Zemankovics, F., Validation: An objective or a tool? Results on a winter wheat simulation model application. *Ecological Modelling*, 81, 251-263, 1995.

Hacker, R.B., Wang, K.M., Richmond, G.S. and Lindner, R.K., IMAGES: An integrated model of an arid grazing ecological system. *Agricultural Systems*, 37, 119-163, 1991.

Hodges, J.S., Six (or so) things you can do with a bad model. *Operations Research*, 39, 355-365, 1991.

Holmes, C.W. and Wilson, G.W., *Milk Production from Pasture*. Butterworths, New Zealand, 1984.

Klepper, O. and Rouse, D.I., A procedure to reduce parameter uncertainty for complex models by comparison with real system output illustrated on a potato growth model. *Agricultural Systems*, 41, 375-395, 1991.

Korte, C.J., Watkin, B.R. and Harris, W., Effects of the timing and intensity of spring grazings on reproductive development, tillering, and herbage production of perennial ryegrass dominant pasture. *New Zealand Journal of Agricultural Research*, 27, 135-149, 1984.

Kvålseth, T.O., Cautionary note about R^2. *American Statistician*, 39, 279-285, 1985.

Law, A.M. and Kelton, W.D., *Simulation Modeling and Analysis*. McGraw-Hill, New York, 1982.

McCall, D.G., *A systems approach to research planning for North Island hill country*. Ph.D. thesis, Massey University, Palmerston North, New Zealand, 1984.

Oreskes, N., Shrader-Frechette, K. and Belitz, K., Verification, validation, and confirmation of numerical models in the earth sciences. *Science*, 263, 641-646, 1994.

Press, W.H., Flannery, B.P., Teukolsky, S.A. and Vetterling, W.T., *Numerical recipes in Pascal: The art of scientific computing*. Cambridge University Press, New York, 1989.

Rykiel, E. J., Testing ecological models: The meaning of validation. *Ecological Modelling*, 90, 229-244, 1996.

Scotter, D.R., Clothier, B.E. and Turner, M.A., The soil water balance in a fragiaqualf and its effect on pasture growth in central New Zealand. *Australian Journal of Soil Research*, 17, 455-465, 1979.

Woodward, S.J.R., Quantifying different causes of leaf and tiller death in grazed perennial ryegrass swards. *New Zealand Journal of Agricultural Research*, 41, 149-159, 1998.