

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

**Modelling Community Productivity, Species
Abundance and Richness in a Naturalised
Pasture Ecosystem**

Baisen Zhang

2005

**Modelling Community Productivity, Species
Abundance and Richness in a Naturalised Pasture
Ecosystem**



A thesis presented in partial fulfilment of the
requirements for the degree of

Doctor of Philosophy

in Plant Science

at Massey University, Palmerston North
New Zealand

Baisen Zhang

2005

Abstract

This study focuses on modelling community productivity, species abundance and richness, and the impact of climate change and alternative phosphorous fertiliser application strategies on pasture productivity by integration of decision tree and regression modelling approaches with a geographical information system (GIS) in a naturalised hill-pasture ecosystem in the North Island, New Zealand, using data derived from research conducted on hill-pastures over the last several decades.

The results indicated that the decision tree models had a high predictive capability and clearly revealed the relative importance of environmental and management factors in influencing community productivity, species abundance and richness. Spring rainfall was the most significant factor influencing annual pasture productivity in the North Island, while hill slope was the most significant factor influencing spring and winter pasture productivity. Annual P fertiliser input and autumn rainfall were the most significant factors influencing summer and autumn pasture productivity, respectively. For species functional group abundance, soil Olsen P was the most significant factor influencing the relative abundance of low fertility tolerance grasses (LFTG) and moss, while soil bulk density, slope and annual P fertiliser input were the most significant factors influencing the relative abundance of legume, high fertility response grasses (HFRG) and flatweeds, respectively. Legume abundance was the most significant factor influencing species richness in the hill-pasture. Species richness increased with an increase in legume abundance and showed a tendency for a hump-shaped response. Grazing animal species also had a significant effect on species richness; pasture grazed by sheep had more species than pasture grazed by cattle. Climate change scenarios of temperature increases of 1-2 °C and rainfall changes of -20% to +20% would have a great impact (-46.2% to +51.9%) on pasture production in the North Island. Pasture in areas with relatively low rainfall had a higher response to increased P fertiliser input than pastures in areas with a relatively high rainfall.

In conclusion, the integration of a GIS with decision tree and regression models in this study provided an approach for effective predictive modelling of community productivity, species abundance and richness in the hill-pasture. This modelling

approach can also be used as a tool in pasture management such as in assessing the impact of climate change and alternative fertiliser management on pasture production.

Key words: climate change, community productivity, data mining, decision tree, functional group, hill-pasture, geographical information system, GIS-based modelling, multivariate analysis, pasture production, predictive modelling, regression, relative abundance, species richness.

Acknowledgements

I would like to express my sincere thanks to my supervisors Dr. Ian Valentine and Assoc. Prof. Peter Kemp for their guidance, encouragement, friendship and endless reading of my manuscripts throughout my doctoral study.

My thanks to those who generously provided their unpublished and/or raw data to me for developing my models; without this data, I would definitely have not finished my thesis before the scheduled time. They are Dr. Jamie Blennerhassett, Dr John Brock, Dr. Mike Dodd, Dr. Allan Gillingham, Mr Dave Grant, Mr. Maurice Gray, Dr. Greg Lambert, Dr. Ignacio Lopez, Dr. Jim Moir, and Dr. Phillipa Nicholas. Thanks to National Institute of Water & Atmospheric Research (NIWA) New Zealand, with the assistance of Ms. Elaine Fouhy, for providing climate data. I am very appreciative of their contributions toward my study and thesis.

I would like to acknowledge the support of people who helped me in various ways during my study. Thanks to Prof. Richard Aspinall, Assoc. Prof. Alex Chu, Dr. Allan Gillingham, Dr. David Gray, Assoc. Prof. Mike Hedley, Dr. Greg Lambert, Dr. Kerry Harrington, Prof. John Hodgson, Assoc. Prof. Cory Matthew, Prof. Russ Tilman and Dr. Todd White for discussion and comments on my work and thesis manuscripts. Thanks to Mr. Mike Tuohy for providing the DEM of the North Island, Mr. Matthew Irwin and Mr. Des Costall for their help with GIS analyses, Dr. Siva Garnish and Mr. Judi Scheffer for advice on statistics, and Dr. Guisan Antonia and Dr. Louis Iverson for advice on issues of developing ecological models. Thanks to the administrative staff and computer lab staff in Institute of Natural Resources for their assistance during my study.

My thanks to my fellow students Tehseen Aslam, Bhoj Bahadur, Zulfiqar Butt, Zaker Hussain, Jian He, Xiongzhaoh He, Mofakkarul Islam, Congqi Liu, Qianhe Liu, Ernest Okorley, Tara Pande, Endang Savitri, Haoran Sun, Zulkefly Sulaiman, Edmundo Viegas, Minhua Xu and others for their friendships and/or helpful discussions on my thesis. Special thanks to Entin Daningsih for frequent discussion on statistics, and Tri Priantoro on GIS.

I would like to express my thanks to the Massey Doctoral Scholarship and the Helen E Akers Ph.D. Scholarship for financial support during this study.

I wish to thank my parents, my sister and brothers for their love, encouragement and support. Very special thanks to my wife Airong and my son Liyu (Oliver) for their love, patience and understanding during my three years' study.

Table of Contents

Abstract	i
Acknowledgements	iii
Table of Contents	v
List of Tables	ix
List of Figures	xi
Chapter 1: General Introduction	1
Chapter 2. Literature Review	7
2.1. Introduction	8
2.2. Hill-pasture in New Zealand	8
2.2.1. Introduction	8
2.2.2. Species composition and diversity in hill-pasture.....	10
2.2.3. Pasture productivity in hill-pasture	12
2.3. Plant ecology in relation to species diversity, distribution and abundance.....	13
2.3.1. Biodiversity	13
2.3.2. Diversity and ecosystem function	14
2.3.3. Patterns of species diversity	17
2.3.4. Species distribution and abundance	19
2.4. GIS and GIS-based ecological and environmental modelling	22
2.4.1. GIS in ecological and environmental study	22
2.4.2. GIS in ecological and environmental modelling.....	23
2.4.3. GIS-based predictive modelling of species distribution and abundance	25
2.5. Data mining and decision tree.....	26
2.5.1. Data mining.....	26
2.5.2. Technologies used in data mining:.....	27
2.5.3. Decision tree	27
2.6. Conclusion.....	31
2.7. References	32

Chapter 3. Modelling the Productivity of Naturalised Hill-pasture in the

North Island, New Zealand: a Decision Tree Approach	42
3.1. Introduction.....	43
3.2. Materials and methods	45
3.2.1. Study area.....	45
3.2.2. Database setup.....	47
3.2.3. Models development	48
3.2.3.1. Decision tree	48
3.2.3.2. Regression.....	49
3.2.4. Model assessment and empirical validation.....	51
3.3. Results.....	54
3.3.1. Decision trees and regression models	54
3.3.2. Model assessment and empirical validation.....	56
3.4. Discussion	63
3.4.1. Performance of model.....	63
3.4.2. Insights from the decision trees.....	66
3.4.2.1. General pattern.....	66
3.4.2.2. Effects of environmental and management variables	67
3.4.2.3. Limitations of decision tree.....	70
3.5. Conclusion.....	70
3.6. References.....	71

Chapter 4. Predictive Modelling of Hill-pasture Productivity: Integration

of a Decision Tree and a Geographic Information System	76
4.1. Introduction.....	77
4.2. Methods.....	80
4.2.1. Study area.....	80
4.2.2. Model development.....	81
4.2.2.1. Dataset generation.....	81
4.2.2.2. Decision tree development and assessment	82
4.2.3. GIS-based prediction and model validation.....	84
4.3. Results.....	86
4.3.1. Decision trees.....	86
4.3.2. GIS-based prediction and model validation.....	93

4.4. Discussion	93
4.5. Conclusion.....	96
4.6. References.....	96

Chapter 5. Modelling the Impact of Climate Changes and Alternative Phosphorus Fertiliser Management on Pasture Production in the North Island.....

5.1. Introduction.....	102
5.2. Methods.....	104
5.2.1. Study area.....	104
5.2.2. Dataset.....	105
5.2.3. DEM and GIS surfaces.....	106
5.2.4. Model development and assessment	107
5.2.5. Climate change scenarios and GIS-based model prediction	108
5.3. Results.....	110
5.3.1. The regression model	110
5.3.2. GIS-based model prediction.....	110
5.4. Discussion	118
5.5. References.....	121

Chapter 6. A Decision Tree Approach to Modelling Functional Group Abundance of Plant Species in a Pasture Ecosystem.....

6.1. Introduction.....	126
6.2. Methods.....	129
6.2.1. Study area.....	129
6.2.2. Model development.....	130
6.2.2.1. Dataset.....	130
6.2.2.2. Decision tree development and assessment	132
6.2.2.3. GIS-based prediction and model validation	133
6.3. Results.....	135
6.3.1. Decision trees	135
6.3.2. GIS-based prediction and model validation	139
6.4. Discussion	139
6.6. References.....	145

Chapter 7. Modelling and Investigating Species Richness Patterns and Underlying Factors in a Pasture Ecosystem	150
7.1. Introduction	151
7.2. Methods	153
7.2.1. Study area	153
7.2.2. Model development	154
7.2.2.1. Dataset	154
7.2.2.2. Correlation analysis	154
7.2.2.3. Decision tree development and assessment.	155
7.3. Results	158
7.3.1. Correlation	158
7.3.2. Decision trees	159
7.4. Discussion	163
7.6. References	167
Chapter 8. General Discussion and Conclusions	172
8.1. Introduction	173
8.2. The decision tree approach	173
8.3. Advantages of applying geographic information system (GIS)	175
8.4. Model insights and implications in ecological study and pasture management	176
8.4.1. Pasture productivity	176
8.4.2. Impact of climate change and alternative P fertiliser application strategies on pasture production	177
8.4.3. Functional group abundance	178
8.4.4. Species richness	178
8.5. Conclusions	179
8.6. References	181
Appendix 1. List of species in the hill-pasture studied in Chapter 7 [after Lambert et al. (1986) and Nicholas et al. (1998)]	185

List of Tables

Chapter 3

Table 1. Variables used in the decision tree and regression analyses.....	50
Table 2. Eigenvectors (CORR) for the first nine principal components.....	52
Table 3. Regression model for annual pasture productivity	56
Table 4. Summary of the regression models for annual and seasonal pasture productivity. See Table 1 for variable symbols.	57

Chapter 4

Table 1. Variables used in the decision tree and GIS analyses.....	83
--	----

Chapter 5

Table 1. Climate zones in the North Island.....	106
Table 2. Scenarios of climate changes and P fertilizer input.	108
Table 3. Summary of the regression model for pasture production. The variable symbols are: Cos (slope) (cosine slope), P_fert (P fertilizer input, kg/ha), Temp_y (annual daily mean temperature, °C), Rain_sp (spring total rainfall, mm), Rain_y (annual total rainfall, mm).....	112
Table 4. Predicted pasture production (mean ± standard deviation) (kg/ha) for the whole study area and the five climatic zones under different scenarios. See Table 1 for the climate zones and Table 2 for scenarios.	114
Table 5. Changes in pasture production (mean ± standard deviation, kg/ha) in amount (mean ± standard deviation kg/ha) and in percentage (%) between the climate change scenarios and the normal scenarios and between the two P fertilizer input scenarios and the normal scenario for the whole study area and the five climatic zones. See Table 1 for the climate zones and Table 2 for scenarios.	115

Chapter 6

Table 1. Input variables used in the decision tree and the GIS analyses	131
Table 2. Functional groups of plant species at Ballantrae (after Lambert et al. (1986) and Nicholas et al. (1998))	132

Chapter 7

Table 1. Environmental, pasture management and vegetation variables used in model analyses	155
Table 2. Eigenvectors (Corr) for the first ten principal components. Numbers in bold font indicate the main contributions of the variables to the principal components.	156
Table 3. Pearson correlation coefficients and significance among input variables that had a significant ($P < 0.05$) correlation with species richness, and the partial correlation coefficient and significance between these variables and species richness when each of the soil Olsen P, aboveground biomass, legume abundance, total N or a combination of them were used as partial variables. See Table 1 for variable descriptions.....	158

List of Figures

Chapter 2

Fig 1. General structure of a decision tree.	28
---	----

Chapter 3

Fig. 1. The study area and the main sampling locations	46
Fig. 2. Model assessment in developing the decision tree model for annual pasture productivity using average squared error (ASE) of the validation data. The vertical line indicates the smallest ASE from the validation data and the number of leaves selected for the final tree.....	53
Fig. 3. The decision tree model for annual pasture productivity. Predicted productivities are in the unshadowed rectangles, splitting variables and split-points are in the shadowed rectangles. Prediction goes to the left-side branch when the splitting variable is less than the split-point, and goes to the right-side branch when the splitting variable is equal to, or more than, the split-point (in the case of aspect, pasture with aspect set before the comma go to left branch, others go to right branch). See Table 1 for variable descriptions and units.	58
Fig. 4. The decision tree model for spring pasture productivity. See caption of Figure 3 for the interpretation of decision tree.....	59
Fig. 5. The decision tree model for summer pasture productivity. See caption of Figure 3 for the interpretation of decision tree.....	60
Fig. 6. The decision tree model for autumn pasture productivity. See caption of Figure 3 for the interpretation of decision tree.....	61
Fig. 7. The decision tree model for winter pasture productivity. See caption of Figure 3 for the interpretation of decision tree.....	62
Fig. 8. Comparison of average square error (ASE) between the decision trees and the regression models for annual and seasonal pasture productivity	63
Fig. 9. Empirical validation of the decision tree (A) and the regression model (B) for annual pasture productivity. Deviation is the difference between	

predicted and observed pasture productivity. The two spreading lines show the $\pm 20\%$ of the observations, and 91% of predictions in the decision tree and 80.1% of predictions in the regression are within the $\pm 20\%$ of the observations of the validation data. Both observation and deviation are transformed data for pasture productivity. 64

Fig. 10. Percentage of adequately predicted cases of the decision trees and the regression models for annual and seasonal pasture productivity in the model empirical validation..... 66

Chapter 4

Fig. 1. Study area, digital elevation model (DEM) and the farmlets' used in validating the GIS prediction. 81

Fig. 2. Diagram of the procedures for generating the model predictions in GIS. The rectangles in the GIS environment area represent the GIS map layers. The arrows indicate the orders and the resources in generating predictions on pasture productivity. Names of input variables and their description are in Table 1..... 85

Fig. 3. The decision tree model for annual pasture productivity. Predicted productivity are in the unshaded rectangles, splitting variables and split-points are in the shadowed rectangles. Prediction goes to the left-side branch when the splitting variable less is than the split-point, and goes to the right-side branch when the splitting variable is equal to, or more than, the split-point (in the case of aspect, pasture with aspect set before the comma go to left branch, others go to right branch). See Table 1 for variable symbols and unit descriptions..... 87

Fig. 4. The decision tree model for spring pasture productivity. See caption of Figure 3 for the interpretation of decision tree..... 88

Fig. 5. The decision tree model for summer pasture productivity. See caption of Figure 3 for the interpretation of decision tree..... 89

Fig. 6. The decision tree model for autumn pasture productivity. See caption of Figure 3 for the interpretation of decision tree..... 90

Fig. 7. The decision tree model for winter pasture productivity. See caption of Figure 3 for the interpretation of decision tree..... 91



Fig. 8. Map layers of the predicted annual (A), spring (B), summer (C), autumn (D) and winter (E) pasture productivity (kg/ha). Note the different scales for productivity in the map keys.	92
Fig. 9. Comparison of the predicted and the observed annual (A), spring (B), summer (C), autumn (D) and winter (E) pasture productivity of the farmlets LL, LN, HN and HH. Bars indicate the 95% confidence interval of the observed productivity means.	94

Chapter 5

Fig 1. Study area, climate zone and land cover. See table 1 for the climate zone explanations.	105
Fig. 2. The procedure for generating GIS-based predictions for pasture production under different climate scenarios. The rectangles represent data files and results, the ellipses represent the GIS functions and the parallelograms represent the GIS map layers.....	109
Fig. 3. Predicted pasture production (kg/ha) under normal scenario for the North Island. The white areas on the map are forest.....	113
Fig. 4. Changes in pasture production (mean \pm standard deviation, kg/ha) under the climate change scenarios (difference between the climate change and the normal scenario) (a-h) and P fertiliser input scenarios (difference between the P fertiliser input scenarios and the normal scenario) (i-j), for the North Island. The white areas on the map are forest. Note the different scales in the map key.....	116

Chapter 6

Fig. 1. Study area, digital elevation model (DEM) and the four farmlets used in model validation.....	130
Fig. 2. Diagram of the procedures for generating the model predictions in GIS. The rectangles in the GIS environment area represent the GIS map layers. The arrows indicate the orders and the resources in developing the predictions of relative abundance of the functional groups in GIS. Names of input variables and their description are in Table 2.....	134

Fig. 3. The decision tree model for relative abundance of high fertility response grasses (HFRG). Predicted relative abundance (in percentage) is in the unshaded rectangles, splitting variables and split-points are in the shadowed rectangles. Prediction goes to the left-side branch if the splitting variable is less than the split-point, and goes to the right-side branch if the splitting variable is equal to, or more than, the split-point. See Table 2 for variable symbols and unit descriptions.....	136
Fig. 4. The decision tree model for relative abundance of low fertility tolerance grasses (LFTG). See the caption of Figure 3 for the description of decision tree interpretation.....	137
Fig. 5. The decision tree model for relative abundance of flatweeds. See the caption of Figure 3 for the description of decision tree interpretation.....	138
Fig. 6. The decision tree model for relative abundance of legume. See the caption of Figure 3 for the description of decision tree interpretation.....	139
Fig. 7. The decision tree model for relative abundance of moss. See the caption of Figure 3 for the description of decision tree interpretation.	140
Fig. 8. The responses of functional groups HFRG, LFTG, legume, moss and flatweeds to Olsen P and slope gradients.....	141
Fig. 9. Map layers of predicted relative abundance (%) for high fertility response grasses (HFRG), low fertility tolerance grasses (LFTG), legume, moss and flatweeds. Note the different scales on map legends.....	142
Fig. 10. The predicted  and the observed  relative abundance of the functional groups in the farmlets LL, LN, HN and HH. Bars indicate the 95% confidence interval of the observed relative abundance means. Note the different scale on the y-axis for LFTG.....	143

Chapter 7

Fig. 1. The decision tree model for species richness using the environmental, management and vegetation variables. Predicted species richness is in the unshaded rectangles, splitting variables are in the shadowed rectangles. Prediction goes to the left-side branch if the splitting variable is less than the split-point, and goes to the right-side branch if the

splitting variable is equal to, or more than, the split-point. See Table 1 for variable and unit descriptions	159
Fig. 2. The decision tree model for species richness using principal components as input variables. Princ 1, Princ 4 and Princ 5 represent principal components 1, 4 and 5, respectively. The splitting points are values of principal components. Princ 1 was mainly composed of annual, spring and summer temperature (all negative), and spring and summer rainfall. Princ 4 was mainly composed of total N (negative), slope and legume abundance. Princ 5 was mainly composed of annual P fertiliser input and legume abundance. See the caption of Figure 1 for the description of decision tree interpretation.....	160
Fig. 3. The decision tree models for species richness using total nitrogen (a), Olsen P (b), grazing animal species (c), legume abundance (d), aboveground biomass (e) and five-year cumulative P fertiliser input (f) as input variables, respectively. See the caption of Figure 1 for the description of decision tree interpretation.....	161
Fig. 4. Responses of species richness to soil total N, soil Olsen P, aboveground biomass, legume abundance, and principal components 1, 4, and 5.....	162