THE ANALYSIS OF BINARY DATA IN QUANTITATIVE PLANT ECOLOGY

by

THOMAS WILLIAM YEE

A thesis submitted for the
degree of Doctor of Philosophy
at the University of Auckland.

Department of Mathematics & Statistics
Department of Botany
June 1993

Abstract

The analysis of presence/absence data of plant species by regression analysis is the subject of this thesis. A nonparametric approach is emphasized, and methods which take into account correlations between species are also considered. In particular, generalized additive models (GAMs) are used, and these are applied to model species' responses to greenhouse scenarios and to examine multispecies interactions. Parametric models are used to estimate optimal conditions for the presence of species and to test several niche theory hypotheses.

An extension of GAMs called vector GAMs is proposed, and they provide a means for proposing nonparametric versions of the following models: multivariate regression, the proportional and nonproportional odds model, the multiple logistic regression model, and bivariate binary regression models such as the bivariate probit model and the bivariate logistic model. Some theoretical properties of vector GAMs are deduced from those pertaining to ordinary GAMs, and its relationship with the generalized estimating equations (GEE) approach elucidated.

Acknowledgements

I wish to express my gratitude to my supervisors, Drs Neil Mitchell and Chris Wild, for their guidance and encouragement throughout the course of this study. Special thanks to Chris for doing the bulk of the proofreading. Thanks to a number of people with whom I had discussions—Drs Alan Lee and Brian McArdle, John Leathwick, and members of the Statistics Unit. I wish to thank Professors Gauld and Scott for granting me a part-time lecturing position over the last five years, and for the environment conducive to study provided by the Department of Mathematics & Statistics to work in.

I wish to thank Professor R. J. Tibshirani and Dr T. J. Hastie for sending me copies of GAIM during the early stages of this work, and Dr J. A. Fessler for VSPLINE; all three kindly answered all my e-mail enquiries. I also wish to thank Dr M. F. Hutchinson for referring me to vector splines and for CUBGCV.

My family, especially my mother, have been very supportive over many years and I wish to thank them too.

Finally, I thank God for giving me the grace and strength to complete this work.

I gratefully acknowledge the financial assistance provided from the following scholarships: U.G.C. Postgraduate Scholarship, William Georgetti Scholarship, C. Alma Baker Postgraduate Scholarship.

Contents

NOTE: Errata sheet enclosed at back

List of Tables

List of Figures