GDM Terrestrial Ecosystem Classifications

Jake Overton and Robbie Price, Landcare Research, Hamilton New Zealand
Simon Ferrier, CSIRO, Canberra Australia
Glenn Manion, Office of Environment and Heritage, New South Wales Australia

Summary

We have produced a new generation of quantitative ecosystem classification, providing improvements and increased functionality to the existing and widely used Land Environments of New Zealand (LENZ). As with LENZ, these new classifications are designed to capture potential (or natural) ecosystem characteristics, and can define the natural extent of these ecosystems, even in areas that have been converted to agricultural or other anthropogenic landscapes. This new classification uses a relatively new analytical approach to solve several shortcomings of the LENZ classification. It can be used in the same manner as LENZ for a wide range of conservation and land management issue, and also provides additional uses not available from LENZ.

Introduction

This project, funded by TFBIS, investigates and develops the use of Generalised Dissimilarity Modelling (GDM, Ferrier et al. 2007) to produce a next generation of ecosystem classification that solves several limitations of the widely used Land Environments of New Zealand (LENZ, Leathwick et al. 2003). Both the existing LENZ classifications and the new GDM classification are designed to predict potential or natural ecosystem potential and allow estimation of natural ecosystem patterns across New Zealand, including agricultural or other highly modified landscapes.

The LENZ classification of New Zealand's potential ecosystems provided a dramatic increase in the ability to analyse quantitatively New Zealand ecosystem patterns and loss, and has been widely used for a broad range of conservation and other management purposes.

A number of limitations of the methods used in the LENZ classification, along with advances in quantitative methods for ecosystem depiction, suggest that a next generation of LENZ is appropriate. Assumptions or limitations of LENZ include that it used no method for assessing the relative importance of the different underlying environmental variables, but assumed that all were equally important (although some soil variables were given reduced importance). In addition, it assumed that differences along any environmental gradient were equally important along any portion of the gradient, for example, the difference between 500 mm and 1000 mm of rainfall was assumed as important as the difference between 4500 mm and 5000 mm. In addition, because LENZ did not consider geography in the classifications, differences between ecosystems due to biogeographic influences could not be accounted for by the LENZ classification.

A relatively new approach, generalised dissimilarity modelling (GDM), solves many of these issues. GDM uses a set of biotic data to weight and transform environmental variables so that the resulting model best explains differences in biotic composition. This solves the problem of weighting and differences along gradients. Geographic distance can also be considered in the GDM, so that biogeographic patterns can be incorporated. The transformed environmental variables can be produced as grids and have a number of applications. The GDM technique solves the problem of how to weight and transform environmental and geographic variables in making ecosystem classifications, but opens the question of which data and which taxa to use to define the ecosystem classification.

GDM has been trialled and used in New Zealand for snails, beetles and freshwater communities, as well as trial runs on vegetation. Here we draft a next generation of terrestrial ecosystem classification, using vascular plant data from forests. Two subsets, a) trees and shrubs, and b) ferns, are used to provide classifications tailored for certain physiognomic or taxonomic groups. These are useful in their own right, as well as for comparisons with each other and the overall classification, and provide examples of the types of tailored classifications that are possible for other physiognomic or taxonomic groups.

In the sections that follow, we detail 1) the classifications and their potential uses, 2) where and how to access the data, 3) examples of specific uses, 4) the data and methods used to model and derive the classifications, and 5) detailed results from models and classifications and a discussion of the underlying and supporting information that is available.

The classifications and their use

General Ecosystem Classification

We have produced a new classification of natural ecosystem character for New Zealand. This can be considered to be a new generation of the LENZ type of quantitative ecosystem classification (Figure 1). This classification is based on the community composition of all native vascular forest plant species observed on the plots. This classification can be expressed at any number of classes between 2 and 400. We have produced and made available grids at 400, 100, 20, and 5 groups (e.g., Figure 2). Each of these grids has an extra high elevation group. In addition to the classification, the transformed environmental variables are available as grids and can be used in various ways (see example below). Since this is a new product, it does not have the same level of documentation that is available for LENZ, which was achieved through sizeable funding for a full implementation. This level of documentation is certainly possible for the new classification given appropriate funding.

This classification can be used in any manner that the LENZ levels I, II, III or IV classifications have been used (see examples in Leathwick et al. 2003), often in combination with other information such as land cover or the protected natural areas (e.g., DOC reserves). For instance, the classes can be used to calculate the natural extent of each class and compare this to the current extent that remains in natural vegetation as estimated from the Land Cover database (LCDB). Together, these can be used to calculate the representation of New Zealand ecosystems in remaining natural vegetation by calculating the proportion of each GDM class that remains in native vegetation as has been done by Rutledge et al. (2004), Walker et al. (2006), and Overton et al. (2010). Further examples of the use of these classifications can be seen in Walker et al. (2008) and Overton et al. (2009).

Specific Classifications

In addition to the general ecosystem classification, the GDM approach to classifications allows producing classifications tailored to specific taxonomic or physiognomic groups. We have produced two tailored classifications:

Trees and shrubs This classification uses only the (non-fern) tree and shrub species to define the classification. As such, the classification is designed to best capture the community composition of trees and shrubs across New Zealand.

Ferns This classification uses only the fern species to define the classification. As such, the classification is designed to best capture the community composition of ferns across New Zealand.

These specific classifications can be used in the same manner as the general ecosystem classification, but where the interest is in these specific groups more than overall ecosystem character. Comparison of the GDM models and classifications resulting from different groups may provide insight into ecological differences between groups, or the stability of classifications to the consideration of different groups.

Underlying information

The information underlying the classifications also has a range of uses which are not available for LENZ. For instance, the GDM models of compositional turnover (see detailed results below) provide ecological insights into the important drivers of ecosystem and plant community characteristics. The GDM model and program also produces transformed environmental layers that can be used to predict species compositional differences. For each transformed environmental layer, the GDM program uses the biotic data to scale the units of each environmental layer (e.g., mean annual temperature, degrees C) into units of species compositional difference. This scaling is non-linear to account for the differing importance along different parts of the environmental gradient. When done for each environmental variable chosen by the model, this results in a set of transformed variables that can be used to predict compositional difference. The classifications use this set of variables to classify all New Zealand into areas of similar potential biotic composition; but these layers can also be used directly without classification. For instance, Overton et al. (2009) used the GDM model of snail composition to produce maps of estimated similarity of snail composition to a focal location. In each map, the value for each location is the estimated similarity of snail composition to the focal location. In the example of Overton et al. (2009), two maps were produced using different focal locations – one near Hamilton and one in Fiordland. Such maps could be useful for a variety of applications in which assessing biotic or ecosystem similarity is important such as designing a biodiversity offset, species translocation, or assessing areas that would be suitable for ecosourcing for restoration.

While the use of these transformed environmental layers provides a range of applications not available from LENZ, their use requires more technical expertise and is expected to be of most value for users with moderate to high levels of technical and ecological expertise. Assessments of similarity (or dissimilarity) can be done using the GDM software or programs developed by the user.

Figure 1. 100 Groups GDM classification of New Zealand based on community composition of all terrestrial vascular plants. As with all figures in this document, each ecosystem group is given a distinct colour, and similarity in colour between groups does not indicate similarity in ecosystem character.

Figure 2. 20 Groups GDM classification of New Zealand based on community composition of all terrestrial vascular plants.

Accessing the data

The classification layers are currently available for viewing or download as Arcview GIS grids at:
http://lris.scinfo.org.nz
The layers can be found by searching for the term GDM. Links for particular classifications are given below.

Grids may be downloaded as ESRI grids in ASCII format, or as geo-tiff files. A range of projections are available, including NZMG, NZTM, and WGS84. We recommend the use of these grids in NZMG, since this is the projection in which they were created. Other projections should be used with caution, in case anomalies have been created in the conversion process. If users want to do analyses using the GDM tool, as described in the previous section, they will need the underlying transformed environmental grids in DIVA format. These can be obtained by converting the ESRI grids or by contacting Landcare Research to request the grids in DIVA format.

Each has an additional high elevation group, adding one to the number of groups.

The general ecosystem classification based on all plant taxa is available in several versions, with different numbers of groups:
400 groups
100 groups
20 groups
5 groups

The tree and shrub classification is also available in several versions, with different numbers of groups.
400 groups
100 groups
20 groups
5 groups

The fern classification is also available in several versions, with different numbers of groups.
400 groups
100 groups
20 groups
5 groups

The transformed environmental grids underlying each classification can be found by searching the LRIS portal for the name of the classification and the term 'transformed'. For example, the transformed environmental layers for the fern classification can be found by searching for 'fern transformed'.

Methods

Environmental data

The environmental data layers used for these analyses were the same data used for LENZ, with a few exceptions. These layers include mean annual temperature, minimum temperature, slope, rainfall to potential evapotranspiration, mean annual solar radiation, June solar radiation, soil calcium, soil age. and soil acid phosphorus. Root zone water deficit was excluded from the analyses because the current analyses suggested the spatial prediction of the variable might have anomalies that should be investigated further. In contrast, two variables were added -- mean annual rainfall and distance to coast. The latter variable has shown itself important in a range of previous work, so was also included here.

All analyses were done with grids of 100-m resolution.

Vegetation data

The vegetation data used for these analyses (Figure 1) was plot data of three types:
1) NVS recce data. We used NVS recce data consisting of over 19 000 plots from both public access data and a few large Level 2 access surveys.
2) Pollen data. We used recreated community compositions from pollen data. Pollen data and boosted regression tree models were used to estimate community composition for all plant taxa at pollen sites. These estimated communities were then sampled to create pseudoplots with richness similar to recce plots. Five pseudoplots were created for each of the approximately 80 pollen sites.

The pollen pseudoplots were included to provide increased information on natural community composition in areas that have very high levels of human disturbance, and very little natural vegetation remaining.

For recce data, we used environmental naturalness layers from the Vital Sites and Actions model to estimate the naturalness of the plots. This naturalness is based on LCDB2 and predicted potential vegetation. Plots with low naturalness were excluded from the analyses. Because of the data limitations of the GDM program, the data were subsampled. Recce data were subsampled inversely proportional to local plot density to reduce the geographic bias in the data. Only species native to New Zealand were used in the analyses.

The final data in the analyses below includes 4684 recce plots and 400 pollen pseudoplots.

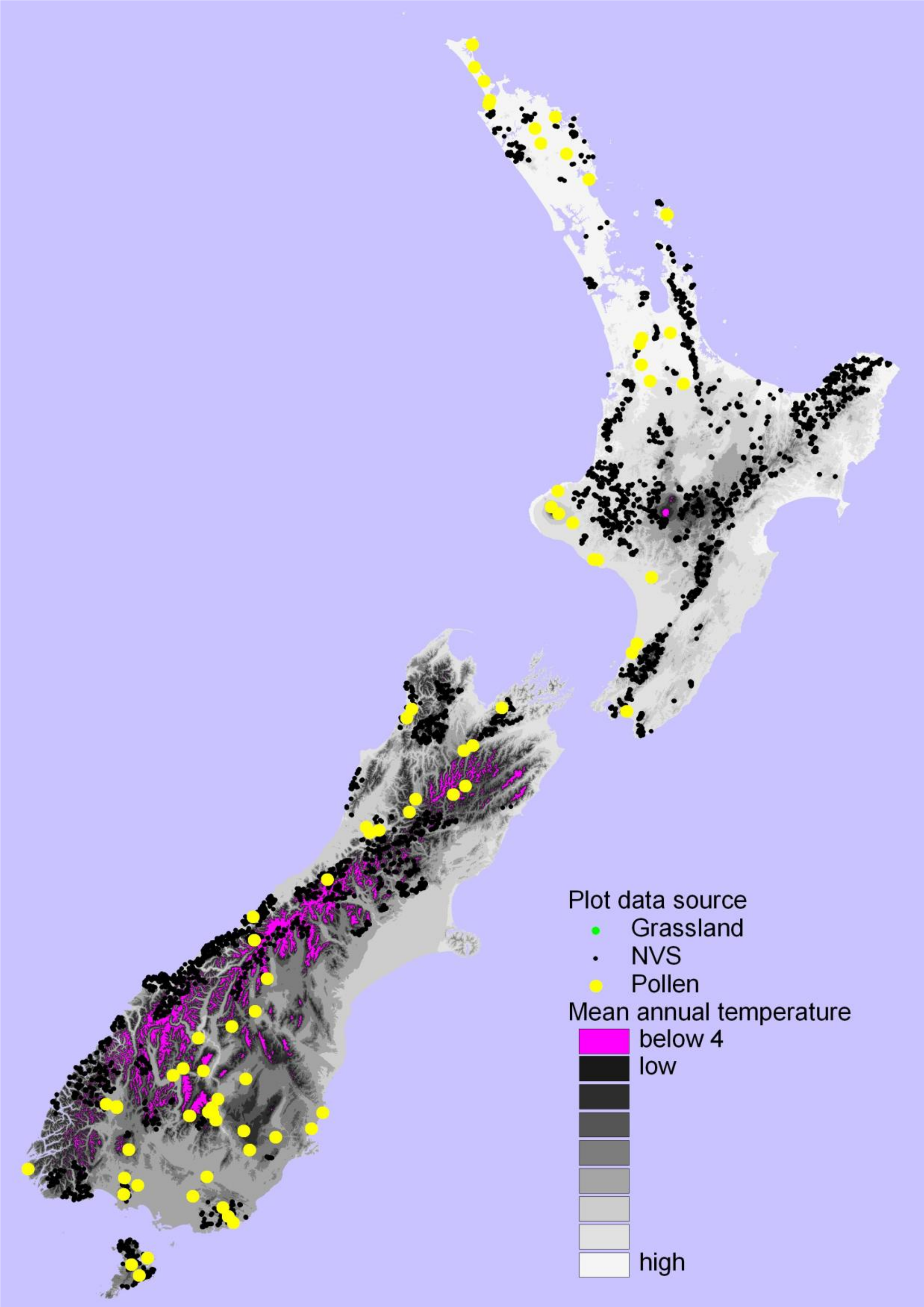

Figure 3. Map of data and sources. Grassland data were not used in the analyses. The area of mean annual temperature less than 4C is shown.

Analysis

Generalised dissimilarity modelling (GDM, Ferrier et al. 2007) was used to model species compositional turnover in relation to environmental variables. A new version of GDM was used which runs as a standalone application. This version of GDM provides classifications, but we performed classifications of GDM results in R and grids were produced from the classifications using purpose-built software.

Because of the strong effect of temperature on the model, we did not extrapolate the classification procedure to low temperatures that were not represented in the data. We used only areas with mean annual temperature greater than or equal to 4C for the classifications. An additional low temperature (high elevation) class was defined by all areas with less than 4C mean annual temperature for all classifications. This avoided extrapolation and classification into very cold areas with little vegetation. Instead, we assigned an additional class that incorporated all of these cold, high elevation areas.

Like LENZ, the classification is first done with a non-hierarchical algorithm to a large number of groups (in this case 400) and then hierarchically classified into smaller numbers of groups. This means the overall classification scheme has a nested, hierarchical structure.

Results Overview

Initial models showed that when grassland data were included, the models showed very strong effects of temperature, resulting from the strong effects in the model of the alpine-forest transition on community composition. In the classifications, this resulted in many small classes in low richness alpine areas, and relatively few in forest areas. For the purposes of this work, it was judged better to focus on non-alpine communities. A future classification should be considered on alpine communities to supplement this work.

Models and classifications for three taxonomic/physiognomic groups are shown in detail below. For each group, the classification is shown, together with the model and the relative importance of each environmental variable, and the shape of the function between the environmental variable and community composition in that group.

The GDM models for all groups show a strong influence of temperature and moisture variables, together with some influence of geography. In comparison with LENZ, the new classifications show greater geographic coherence and greater influence of temperature and moisture variables relative to solar radiation and soils.

We summarize the results below:

1) All vascular plant taxa. The GDM model explained 44% of turnover in plant community composition. Temperature and moisture variables dominate the model, with mean annual temperature and mean monthly minimum temperature the two most important variables and vapour pressure deficit and the ratio of rainfall to potential evapotranspiration the next two most important variables. Geographic distance is the fifth most important variable.

The all plant taxa classification is the most likely candidate as a next generation of LENZ.

2) Trees and shrubs (excluding tree ferns). The GDM model explained 41% of turnover in tree and shrub composition. As for all plant taxa, temperature and moisture variables dominate the model, and the importance of the different variables is similar.

3) Ferns. The GDM model explained only 22% of turnover in ferns, suggesting that ferns are less poorly predicted by environment than are plants overall. As for all plant taxa, temperature and moisture variables dominate the GDM model, but minimum temperature (barely) surpasses mean annual temperature in importance, vapour pressure deficit has become relatively more important, whereas the importance of geographic distance is less.

Together, these results provide the strong basis for a next generation of terrestrial ecosystem classification that considers the relative importance of environmental variables and the variable influence of the variables across different portions of their gradients. They also highlight that geographic distance is important and should be included in ecosystem classifications.

Detailed Results for each model

For each of the three plant groups, detailed results are shown below. Classifications of 100-group and 20-group are mapped for each. A bar graph is presented showing the relative importance of each environmental variable in determining plant compositional turnover in that group. The estimated effect of each variable on compositional turnover in that plant group is also shown. For each variable, the relative overall importance can be judged by the total height of the curve. The amount of turnover along any portion of the environmental gradient can be seen by the slope of the curve. Portions of the environmental gradients with steep curve slopes are more important than portions with flat slopes. These functions are used to transform each environmental layer into a new variable that is scaled and warped into units of community dissimilarity and output into a new GIS grid. In this way, environmental variables with very different units can be transformed into one common unit of species compositional turnover. This process accounts for the different importance of the different variables, or the different importance of different portions of any one particular variable.

All Plant Taxa

Figures 1 and 2 showed the 100-group and 20-group classification of plant communities from the GDM model based on all terrestrial vascular plants. Figure 4 shows the environmental variables chosen by the GDM model, and their relative importance in determining plant species composition. Figures 5 to 10 show the relationship between plant community compositional turnover and the chosen environmental variables.

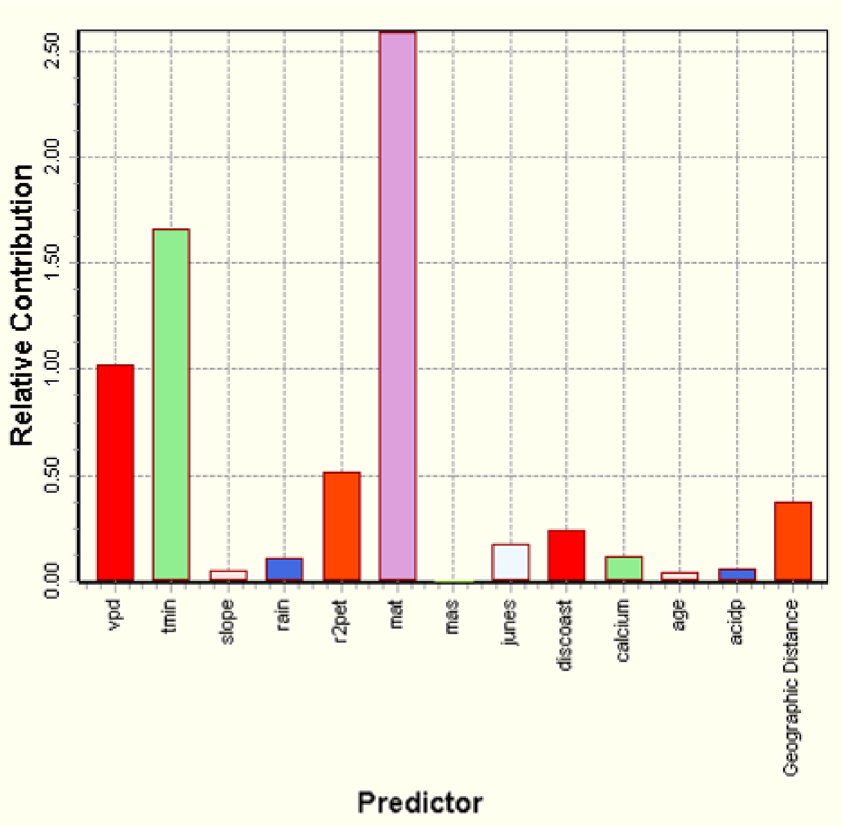

Figure 4. The relative importance of each predictor variable in the GDM model for determining turnover in plant community composition. Temperature and moisture variables dominate the model. Variable abbreviations are: vapour pressure deficit (vpd), mean minimum monthly temperature (tmin), topographic slope (slope), mean annual rainfall (rain), rainfall to potential evapotranspiration (r2pet), mean annual temperature (mat), mean annual solar radiation (mas), June solar radiation (junes), distance to coast (discoast), soil calcium (calcium), soil age (age), soil acid phosphorus (acidp).

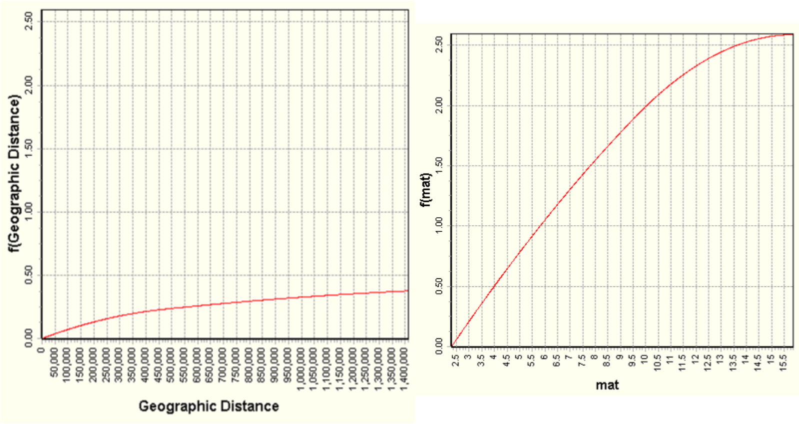

Figure 5. The effect of geographic distance and mean annual temperature (mat) in determining plant compositional differences. The importance of mat decreased at higher values.

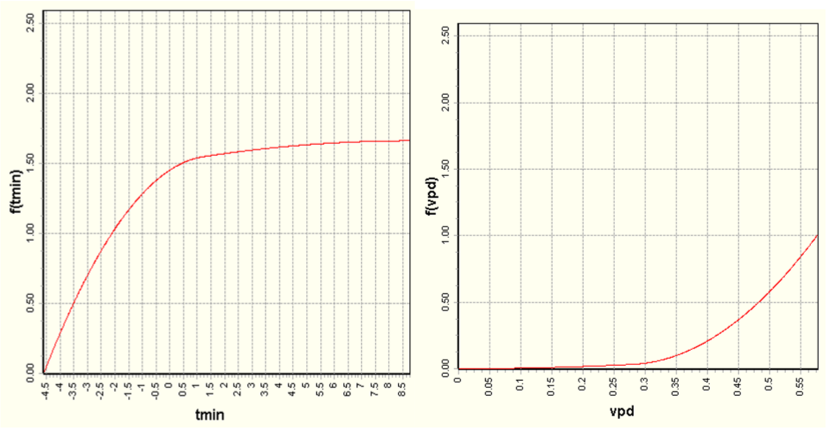

Figure 6. The effect of minimum temperature (tmin) and vapour pressure deficit (vpd) in determining plant compositional differences. Tmin is important mostly at low values, and vpd important mostly high values.

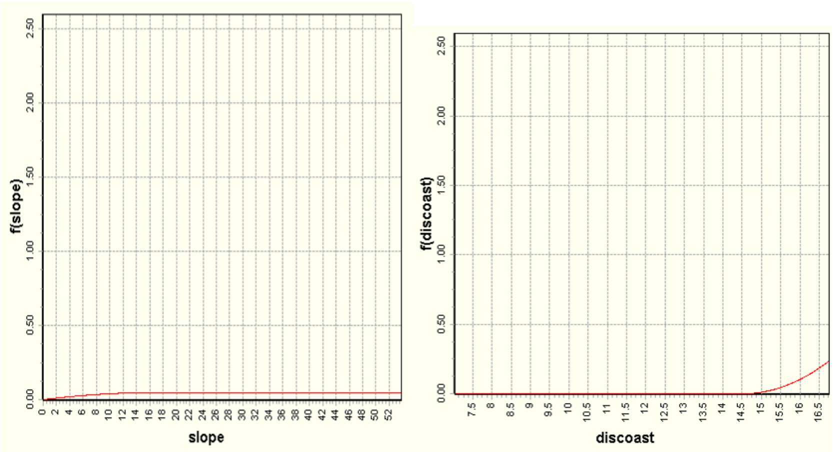

Figure 7. The effect of slope and distance to coast in determining plant compositional differences. Overall, these variables are of little effect. The importance of distance to coast at high values may represent an effect of more continental climates.

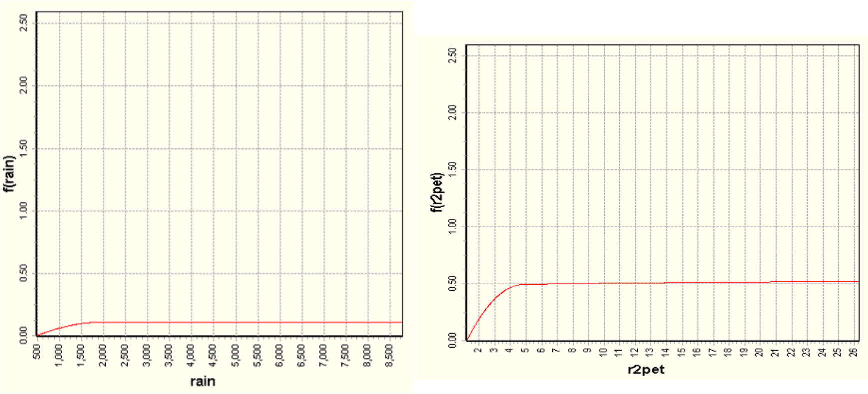

Figure 8. The effect of rain and the ratio of rainfall to evapotranspiration in determining plant compositional differences. Both show the expected effect of being important at low values, but not important at high values.

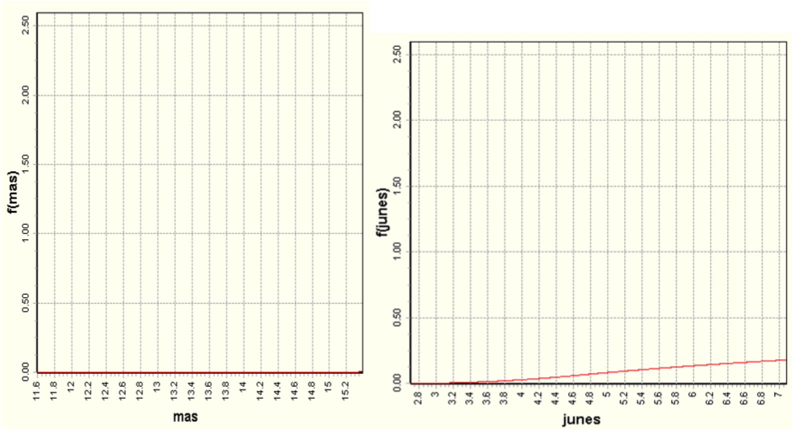

Figure 9. The effect of mean annual solar radiation and june solar radiation (junes) in determining plant compositional differences.

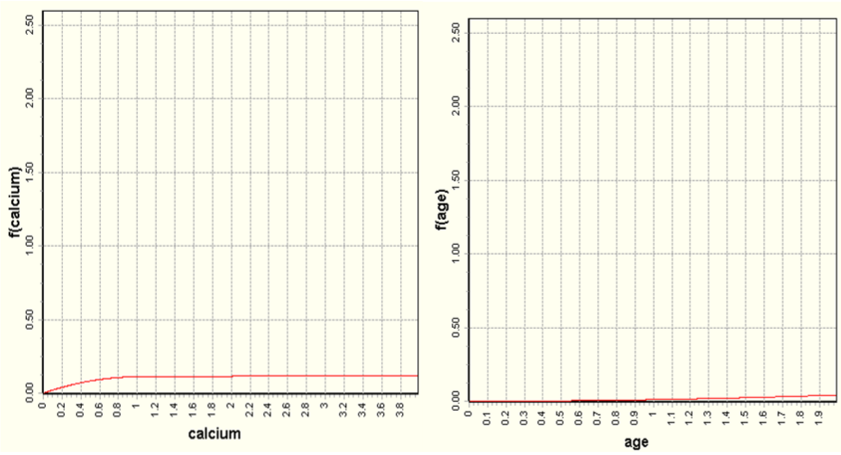

Figure 10. The effect of soil calcium and age in determining plant compositional differences.

Trees and Shrubs

Figures 11 and 12 show the 100-group and 20-group classification of tree and shrub communities from the GDM model. Figure 13 shows the environmental variables chosen by the GDM model, and their relative importance in determining plant species composition. Figures 14 to 18 show the relationship between plant community compositional turnover and the chosen environmental variables.

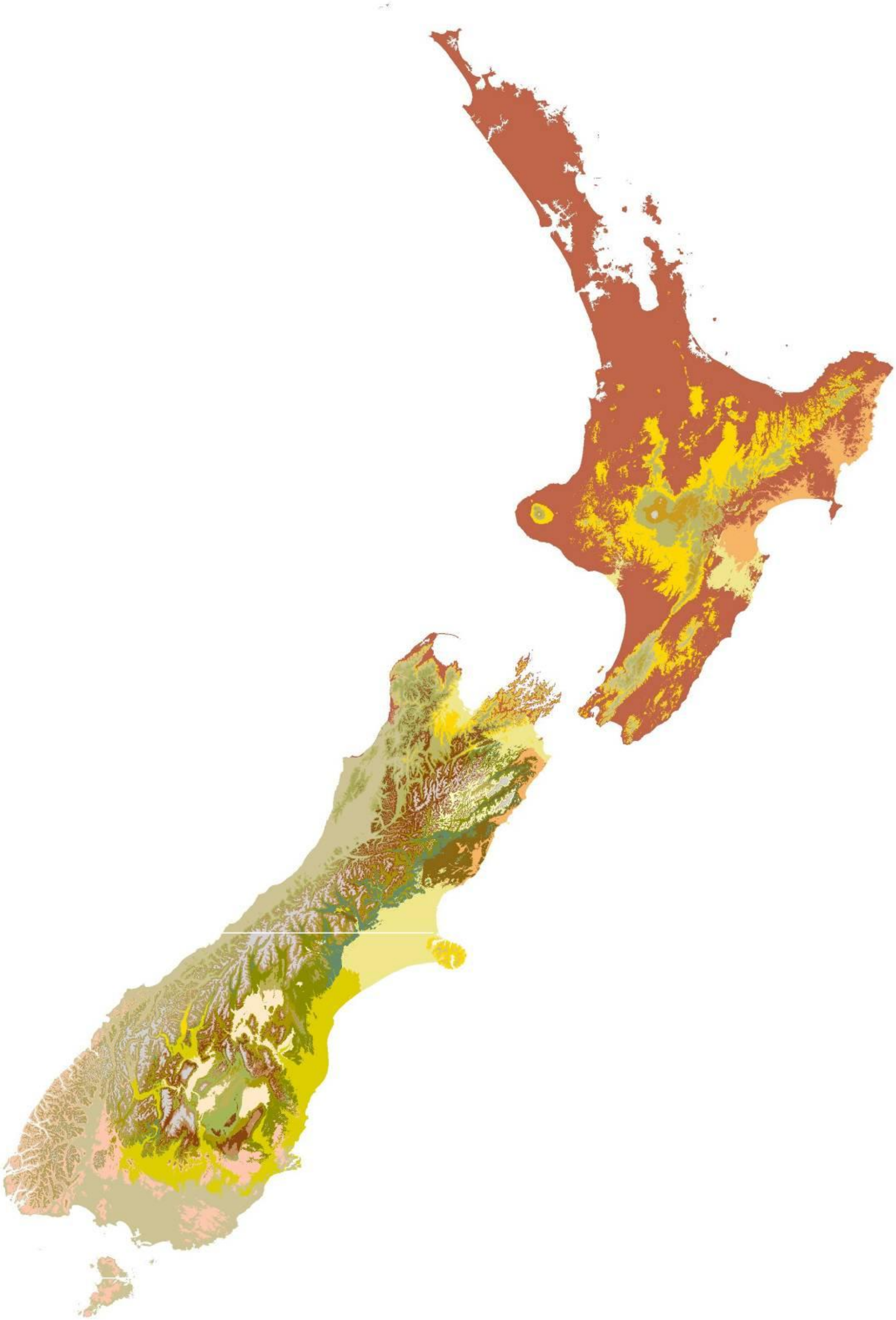

Figure 11. 100-Groups GDM classification of New Zealand based on community composition of trees and shrubs.

Figure 12. 20-Group GDM classification of New Zealand based on community composition of trees and shrubs.

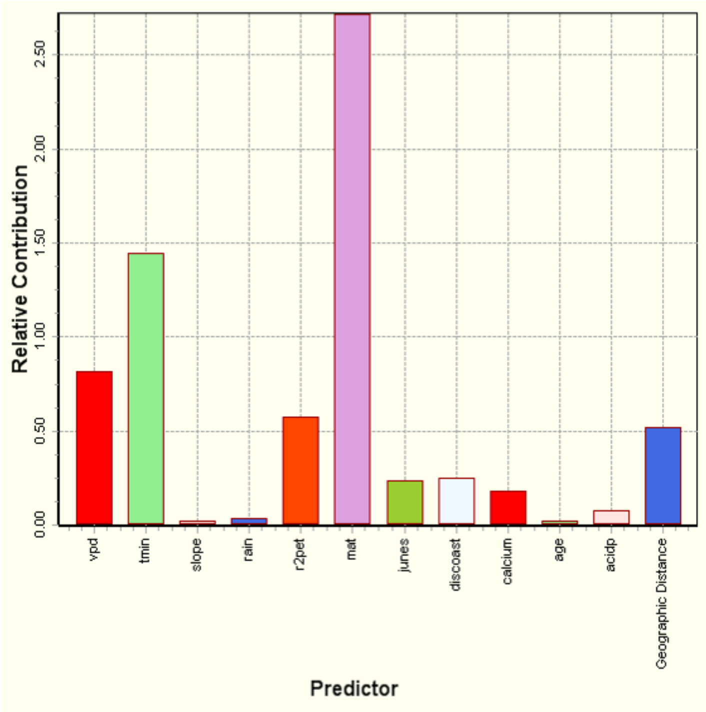

Figure 13. The relative importance of each predictor variable in the GDM model for determining turnover in tree and shrub community composition. As for all plant taxa, temperature and moisture variables dominate the model. Variable abbreviations are: vapour pressure deficit (vpd), mean minimum monthly temperature (tmin), topographic slope (slope), mean annual rainfall (rain), rainfall to potential evapotranspiration (r2pet), mean annual temperature (mat), Junne solar radiation (junes), distance to coast (discoast), soil calcium (calcium), soil age (age), soil acid phosphorus (acidp).

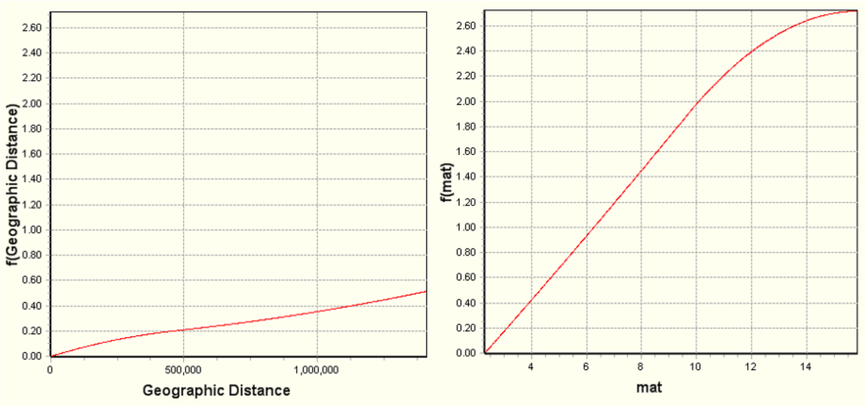

Figure 14. The effect of geographic distance and mean annual temperature (mat) in determining tree and shrub compositional differences. The importance of mat decreased at higher values.

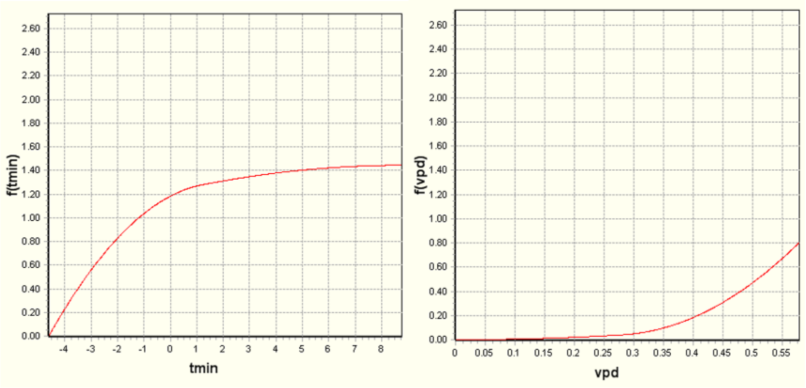

Figure 15. The effect of minimum temperature (tmin) and vapour pressure deficit (vpd) in determining tree and shrub compositional differences. Tmin is important mostly at low values, and vpd important mostly high values.

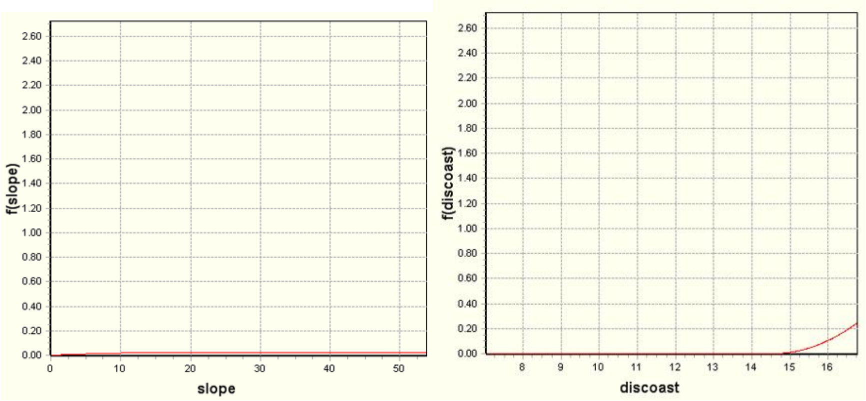

Figure 16. The effect of slope and distance to coast in determining tree and shrub compositional differences. Overall, these variables are of little effect. The importance of distance to coast at high values may represent an effect of more continental climates.

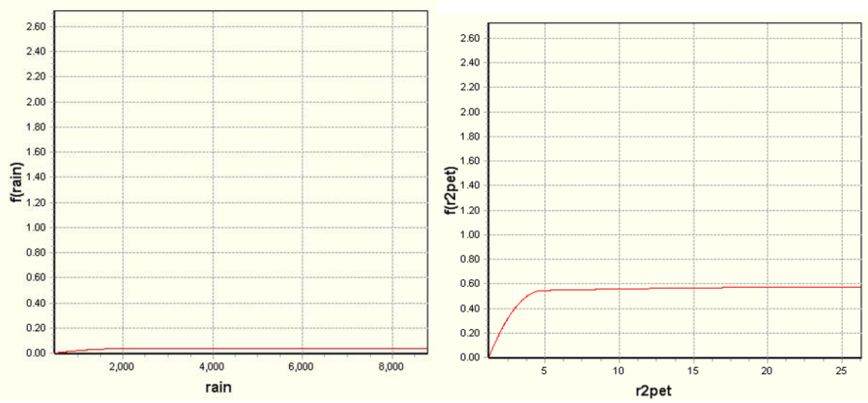

Figure 17. The effect of rain and the ratio of rainfall to evapotranspiration in determining tree and shrub compositional differences. Both show expected effect of being important at low values, but not important at high values.

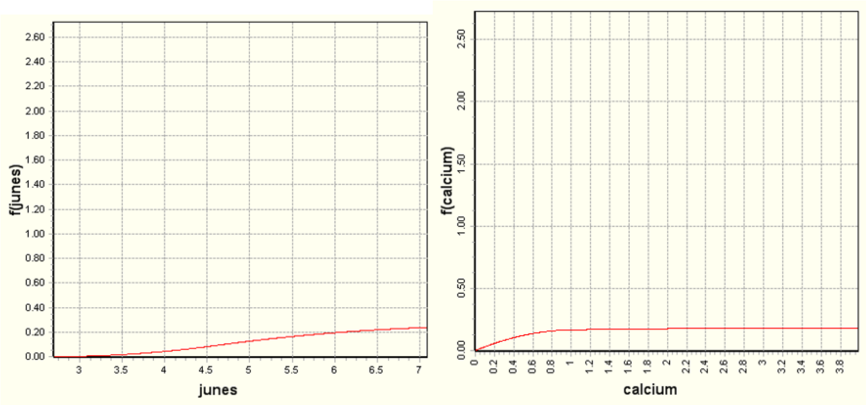

Figure 18. The effect of june solar (junes) radiation and soil calcium (calcium) in determining tree and shrub compositional differences. Mean annual solar radiation was not chosen for this GDM model.

Ferns

Figures 19 and 20 show the 100-group and 20-group classification of tree and shrub communities from the GDM model. Figure 21 shows the environmental variables chosen by the GDM model, and their relative importance in determining plant species composition. Figures 22 to 26 show the relationship between plant community compositional turnover and the chosen environmental variables.

Figure 19. 100-Group GDM classification of New Zealand based on community composition of ferns.

Figure 20. 20-Groups GDM classification of New Zealand based on community composition of ferns.

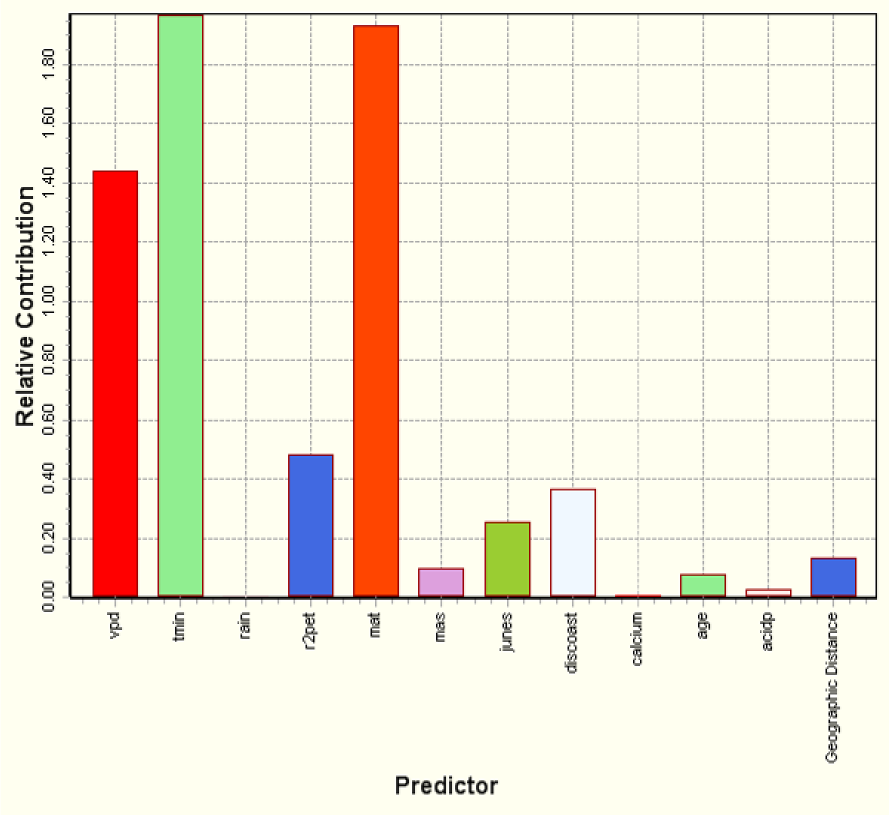

Figure 21. The relative importance of each predictor variable in the GDM model for determining turnover in fern community composition. As for all plant taxa, temperature and moisture variables dominate the model, but minimum temperature surpasses mean annual temperature in importance and vapour pressure deficit has become relatively more important, whereas the importance of geographic distance is less. Variable abbreviations are: vapour pressure deficit (vpd), mean minimum monthly temperature (tmin), mean annual rainfall (rain), rainfall to potential evapotranspiration (r2pet), mean annual temperature (mat), mean annual solar radiation (mas), June solar radiation (junes), distance to coast (discoast), soil calcium (calcium), soil age (age), soil acid phosphorus (acidp).

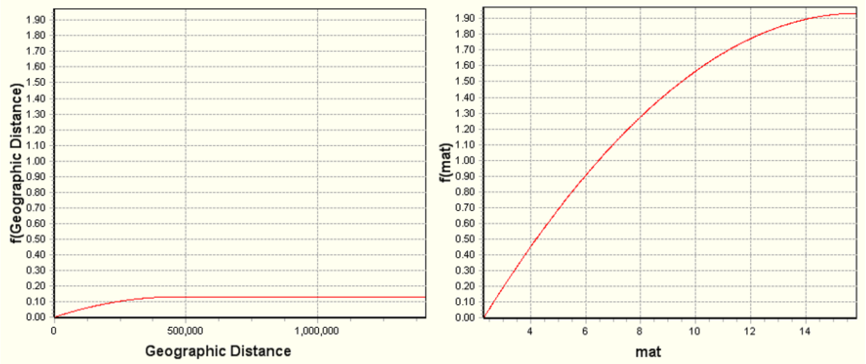

Figure 22. The effect of geographic distance and mean annual temperature (mat) in determining fern compositional differences. The importance of mat decreased at higher values.

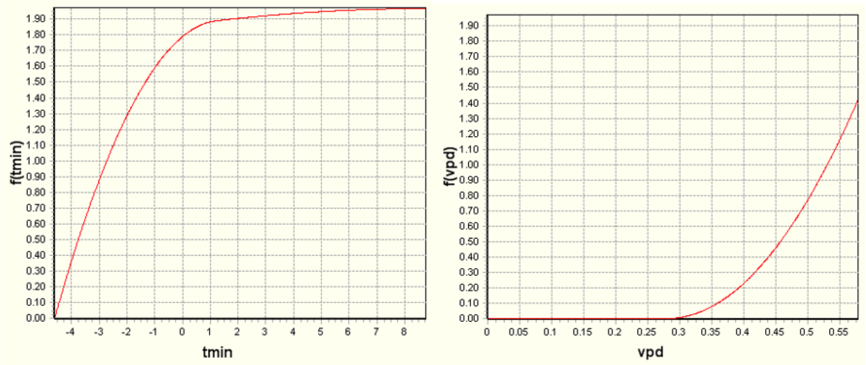

Figure 23. The effect of minimum temperature (tmin) and vapour pressure deficit (vpd) in determining fern compositional differences. Tmin is important mostly at low values, and vpd important mostly high values.

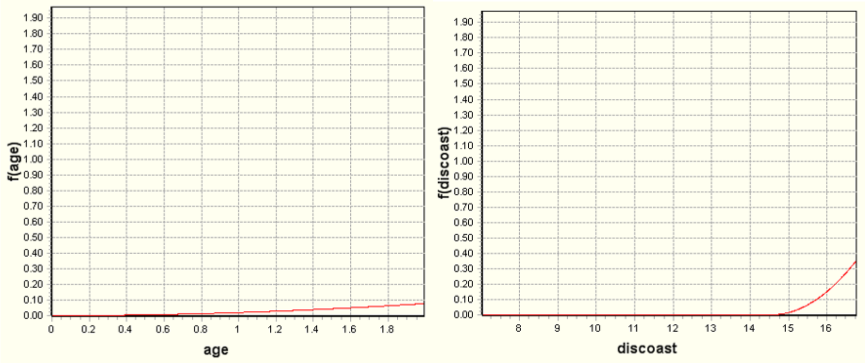

Figure 24. The effect of soil age (age) and distance to coast (discoast) in determining fern compositional differences. Overall, these variables are of little effect. The importance of distance to coast at high values may represent an effect of more continental climates.

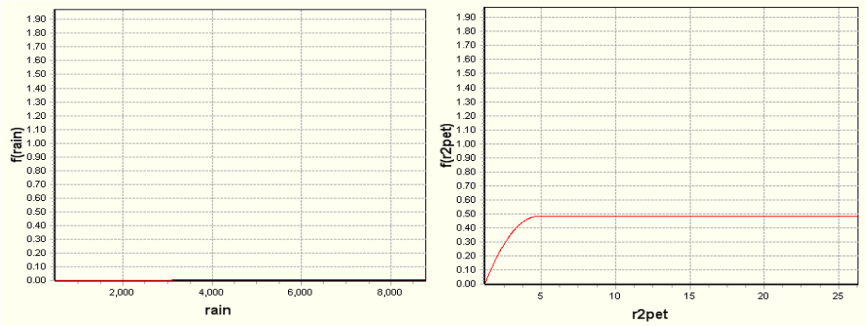

Figure 25. The effect of rain and the ratio of rainfall to evapotranspiration (r2pet) in determining fern compositional differences. Both show expected effect of being important at low values, but not important at high values.

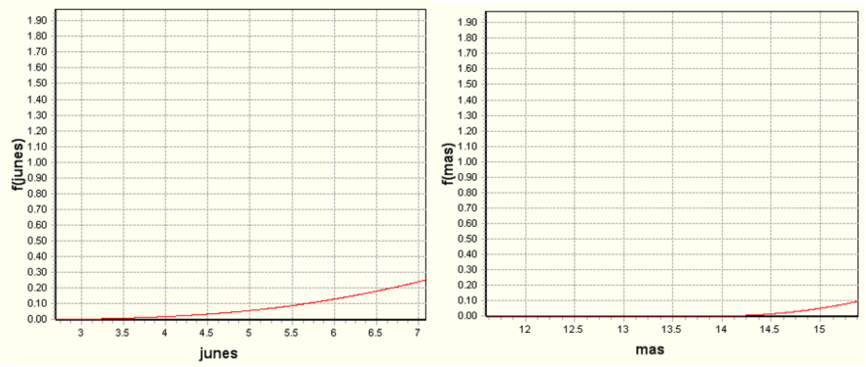

Figure 26. The effect of june solar (junes) radiation and mean annual solar radiation (mas) in determining fern compositional differences.

Acknowledgements

NVS recce data were sourced from NVS. We acknowledge all the contributors to the public access data. Larry Burrows, Glenn Stewart, and Alan Rose granted permission for use of Level 2 data. Janet Wilmhurst and Matt McGlone provided pollen data and advised on its use. Norm Mason provided vegetation distributions predicted from the pollen data. Theo Stephens and Fleur Maseyk commented on the manuscript.

References

Ferrier S, Manion G, Elith J, Richardson K 2007. Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. Diversity and Distributions 13: 252–264.

Leathwick JR, Wilson GH, Rutledge D, Wardle P, Morgan FJ, Johnston K, McLeod M, Kirkpatrick R 2003. Land environments of New Zealand: Nga Taiao o Aotearoa. Auckland, New Zealand, Batemans. 184 p.

Leathwick JR, Moilanen A, Ferrier S, Julian K 2010. Complementarity-based conservation prioritization using a community classification, and its application to riverine ecosystems. Biological Conservation 143: 984–991.

Overton JM, Barker GM, Price R 2009. Estimating and conserving patterns of invertebrate diversity: a test case of New Zealand land snails. Diversity and Distributions 15(5): 731–741.

Overton JMcC, Price R, Stephens RTT, Cook S, Earl R, Wright E, Walker S 2010. Conservation planning and reporting using the Vital Sites model. Landcare Research Report No. LC0910/064, prepared for the Department of Conservation.

Rutledge D, Price R, Heke H, Ausseil A-G 2004. National analysis of biodiversity protection status: methods and summary results. Landcare Research contract report number 0405/042 to the Ministry for the Environment.

Walker S, Price R, Rutledge D, Stephens RTT, Lee WG 2006. Recent loss of indigenous cover in New Zealand. New Zealand Journal of Ecology 30: 169–177.

Walker S, Price R, Stephens RTT 2008. An index of risk as a measure of biodiversity conservation achieved through land reform. Conservation Biology 22: 48–59.