The Use of GIS for Biodiversity Mapping in New Zealand

Lars Brabyn

Department of Geography
University of Waikato, Hamilton, New Zealand.
Phone: +64 7 8384466, Fax: +64 7 8384633
Email: larsb@waikato.ac.nz

ABSTRACT

This paper discusses potential applications and current limitations of GIS for mapping and modelling biodiversity in New Zealand. Three functions of GIS that are important for biodiversity modelling are demonstrated. These are terrain analysis, data integration and data visualisation. Terrain analysis can be used to identify micro, meso and macro terrain indices. Data integration can be used to determine the environmental characteristics of known habitats of species. Data visualisation uses maps, graphs and statistics to make the enormous amount of data that can be derived on a species' habitat easy to understanding. This research has attempted to use "Public Good" New Zealand species location data sets but had problems with access and format. It is argued that the ability of computer hardware and the functionality of GIS software is being under utilised in New Zealand for biodiversity mapping because of poor sharing and storing of data, and an overall lack of co-ordination and leadership in environmental information management.

Keywords and phrases: GIS, Biodiversity, Mapping, Environment, New Zealand, Species, Information

1.0 INTRODUCTION

The New Zealand Biodiversity Strategy (Department of Conservation and The Ministry for the Environment, 2000) makes it clear that biodiversity information is a critical component of effective conservation management. A Geographical Information System (GIS) and its associated data sets are very powerful tools for mapping biodiversity and the intention of this paper is to demonstrate this using three important GIS functions: terrain analysis, data integration, and data visualisation. In the course of applying these GIS functions to New Zealand data sets, many barriers and issues regarding data access have been encountered. These issues, in themselves, are an interesting result of this research and reinforce the following point raised in the New Zealand Biodiversity strategy (Department of Conservation and The Ministry for the Environment, 2000, p.109):

> "Barriers to effective sharing of information mean that biodiversity information (including new techniques to manage biodiversity) is not necessarily informing people who are facing similar issues elsewhere. Often, we simply do not know what biodiversity information is available elsewhere. Systems to coordinate existing data and information about biodiversity, in a way that is accessible to resource managers and the wider community, are lacking".

Biodiversity mapping using GIS may seem like a technical challenge but it is just as much a political challenge because of data access issues and differences in the interpretation of "public data".

During recent years there have been some innovative research projects in biodiversity mapping in New Zealand using GIS. These projects have been mainly limited to Crown Research Institutes (CRIs) with access to species location databases, and include the mapping of Environmental Domains (www.landcare.cri.nz/science/biodiversity) and assessing fern density (Lehmann, et al, 2000). Typically these approaches apply a range of statistical techniques to a range of environmental data sets. These data sets include the National Vegetation Survey (NVS) data bank, Climate Surfaces, and the Land Resource Inventory (LRI). The NVS data bank provides information on the locations of where species are known to exist. Environmental character databases such as the LRI and the Climatic Surfaces are used to derive environmental descriptive parameters of these locations which are then used to develop predictive classification models at a generalised ecosystem level. It is important to note that species location databases are essential for calibrating these biodiversity models.

The implementation of GIS requires coordination of software, hardware, data, personnel, and an overall information management strategy. Software and hardware are generally not the issue for many research organisations. The use of GIS for biodiversity mapping no longer requires expensive high-end computer hardware. The largest GIS environmental data sets are typically less than 100 Mbytes and the modern desktop computer can manipulate and analyse these within an acceptable time period. The modern desktop computer now has hard drive space exceeding 20 Gbytes and can easily store a wide range of data sets for GIS analysis. The price of GIS software may still be a financial barrier for many researchers, but never the less, many organisation use sophisticated GIS software. There is now a wide selection of GIS software to cater for a range of user abilities and the easy to use graphical user interface (GUI) GIS has a selection of powerful analysis and visualisation functions. It is the biodiversity data, and skilled personnel that are limiting progress with biodiversity mapping. These points will be discussed later in this paper, but first the functionality of GIS for biodiversity mapping will be discussed and illustrated.

2.1 TERRAIN ANALYSIS

Terrain analysis commonly refers to the identification of elevation, slope and aspect of a site. However it can also be used to calculate the curvature of the surface, distinguish upslope or downslope areas, define catchment boundaries, determine runoff, and identify morphological landform types, such as terraces and gullies (Band, 1986; Dikau, 1989; Dymond et al., 1995; and Brabyn, 1997).

Terrain, along with climate, soil and landcover, is an important determinant of species' habitat (Tivy, 1993). At a species habitat scale, terrain has an important impact on the micro climate of an area, such as whether a site is a sheltered hollow prone to frosts or a exposed ridge exposed to heavy rain and wind (Geiger, 1971). Terrain is also an important abiotic determinant of the soil properties and moisture content of a site and is therefore a variable used in soil landscape modelling (Francis, 1985).

In New Zealand we have data sets that describe climate, soil, geology, and landcover. These are at a generalised (macro) scale that can not be used for modelling accurately habitats of species that are effected by micro changes in climate, soil properties and vegetation cover. However, we now have widely available 20m contour interval terrain data for New Zealand, thanks to the Government relinquishing its copyright. From this terrain data it is now possible to automatically identify terrain features, such as small gullies, spurs, and terraces, that are at a scale relevant to the identification of the habitats of many species.

Terrain features at a multiple of scales affect soil properties, drainage, climate and ultimately species habitat. For example, macro scale terrain features, such as the large open valleys of the eastern Southern Alps, create wind tunnels and are subject to large mass movements and erosion deposits. Meso and micro scale valleys may provide a more sheltered environment from winds but be subjected to more frequent frosts. It is important therefore to identify a range of terrain features that vary in scale if terrain is used for species habitat modelling.

Rather than identify a large range of possible terrain features and work with a nominal terrain classification, it is more practical for biodiversity mapping to work with terrain indices that take into consideration the important habitat properties of different terrain features. Important habitat properties of a terrain feature are whether it is subjected to erosion attrition or erosion deposits, has good or poor water drainage, and is sheltered or exposed to climate. If terrain indices are able to differentiate these important micro climate and soil properties, then terrain indices could be used as surrogates of these properties.

One such index can be obtained by subtracting the mean elevation from the actual elevation. The mean elevation is calculated by using a focal neighbourhood function. This index indicates whether a location is above the surrounding terrain (like a ridge) or below the surrounding terrain (like a valley floor) or somewhere in between. By varying the extent of the neighbourhood when calculating the mean it is possible develop a series of indices that identify terrain properties at different scales. Figures 1 -3 shows three terrain indices that distinguish terrain properties at three different scales. These figures show 3D perspective views with the vertical dimension being the actual elevation and the relative values of the terrain indices draped over the elevation using a continuous grey scale. It is therefore possible to compare the terrain indices with the actual terrain surfaces. Figure 1 shows a terrain index based on a 10 km neighbourhood radius for a section of the Wilberforce Valley (viewed from near Lake Coleridge). It can be seen that it distinguishes the large valley floors from the mountain ranges. Figure 2 shows a terrain index based on a 1km-neighbourhood radius for Mt Karioi (viewed from the South West). As you can see, it distinguishes the gullies from the main ridges and spurs. Figure 3 shows the same view of Mt Karioi, but displays a terrain index based on a 100m neighbourhood radius. This distinguishes small indentations on the side of spurs and ridges. These terrain indices could be used as surrogates for micro climate and soil properties.

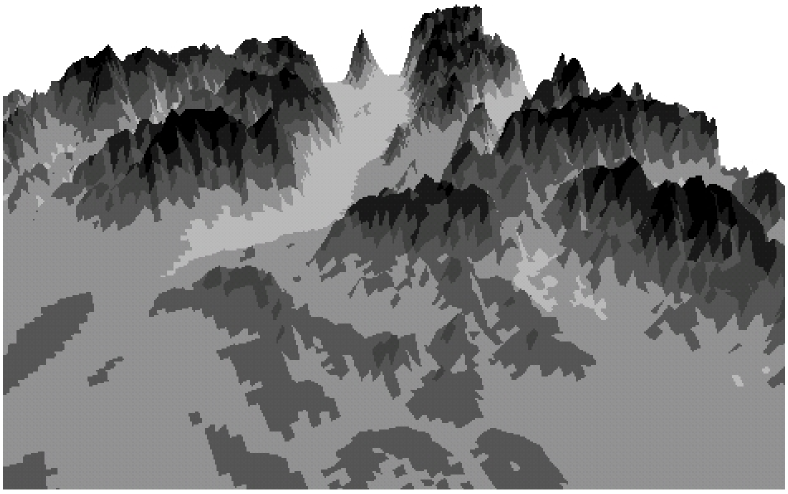


Figure 1 Section of the Wilberforce valley showing a Terrain Index based on a 10km Neighbourhood (Cellsize 500, Continuous grey scale from -829 to 1124. Light is low and dark is high. The mean value is -13)

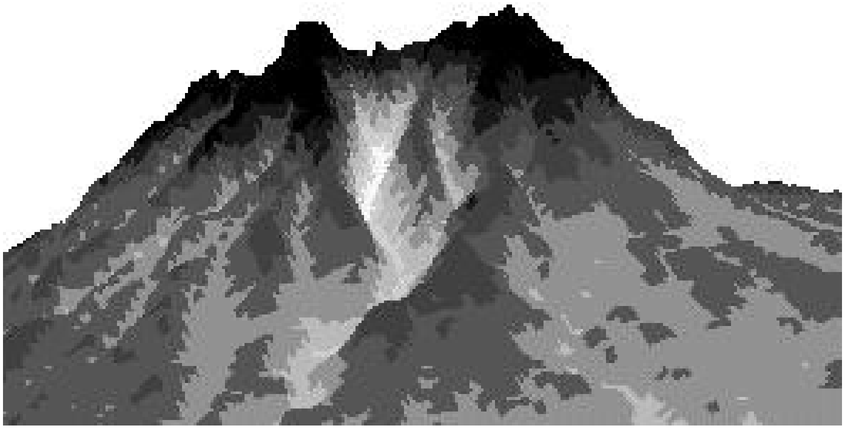


Figure 2 Mt Karioi showing a Terrain Index based on a1km Neighbourhood (Cellsize 50m.Contiunous grey scale from -115 to 190. Light is low and dark is high. The mean value is -3)

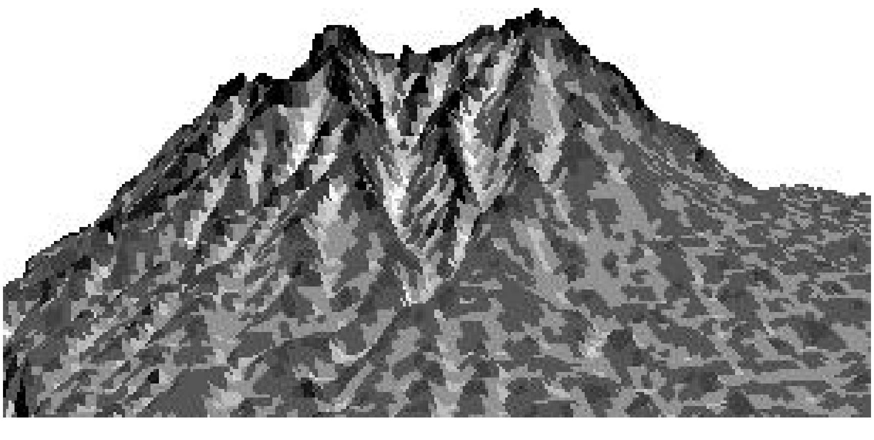


Figure 3 Mt Karioi showing a Terrain Index based on a 100m Neighbourhood (Cellsize 50m. Continuous grey scale from -34 to 61. Light is low and dark is high. The mean value is -3)

2.2 DATA INTEGRATION

In order to develop predictive habitat and ecosystem classifications it is necessary to know the environmental characteristics of a habitat or ecosystem. GIS can play a key role in this process because it has the ability to integrate environmental character information with known locations of species or ecosystems.

In New Zealand, one of the most important sources of information on vascular plant species locations is contained in the National Vegetation Survey (NVS) databank. This databank includes information from descriptions made at approximately 14 000 permanently marked locations (may include repeat measures) and approximately 52 000 unmarked locations throughout New Zealand. Each sample records limited information on the site (e.g. elevation, aspect, slope, and drainage) and all vascular plant species present at the sample location. This, in effect, also provides information on what species are absent from the sample location. Absence information is essential in developing predictive distribution models.

There is now a range of GIS data sets that can be used to derive environmental characteristics of each plot. These include the terrain indices previously described, a range of climate parameters (such as minimum and maximum temperatures, rainfall, and solar radiation), geology, distance from coast, and the main landcover type. It is a relatively simple and routine operation to integrate geographically referenced databases, based on a common location. The NVS data set contains the NZMG co-ordinates of each plot, so it is therefore possible to identify a range of environmental characteristics of each plot location and combine this information to the NVS data set. Table 1 illustrates this concept. It shows a sample of information obtained from NVS, with 19 records of known locations of *Fushica execelsar* expressed as NZMG co-ordinates (Eastings and Northings). GIS, in combination with a range of environmental data sets, was used to derive environmental characteristics of each known location. This data integration can be implemented for large data sets containing 20,000 records within a few minutes.

Once species location information has been integrated with environmental character data, such as in Table 1, it is then possible to export the table into statistical software and do complex multivariate analysis (this research is still in progress). It is also possible to use GIS to graphically represent the data as described in the next section.

2.3 DATA VISUALISATION

It is difficult to interpret information from a table containing a large number of records. By applying the functionality of GIS, information on the location of species can be ascertained using a variety of visualisation tools. These tools can be used to make information on the relationships between the location of the species and the environmental characteristics more apparent. The use of GIS to map data is well known, as a map is a basic spatial analysis tool (Wadsworth and Treweek 1999). However, the development of GIS has lead to functions such as statistical analysis and charting also being available as visualisation tools (Gahegan, 1999). Recent work by Lehmann *et al.* (2000) employed visualisation tools to display fern distributions derived from using generalised regression analysis and spatial prediction (GRASP). The current project differs from GRASP as it uses one software package (ARC/INFO) and the visualisation of biodiversity data is completed automatically, using a macro programming language, so that it can be produced for any species recorded in the data set.

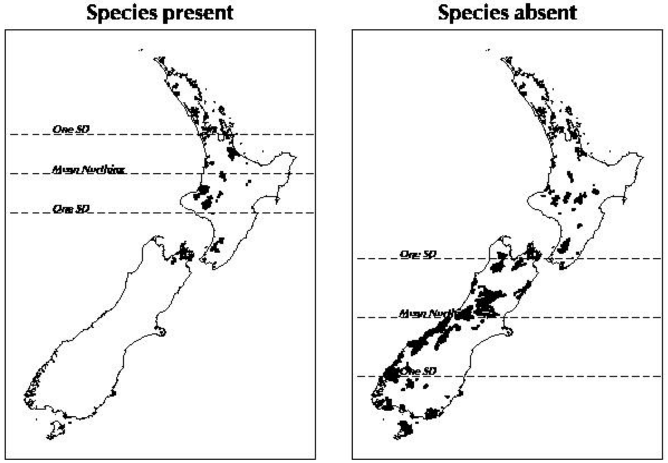
Figure 4 shows an example of the visualisation functions of a GIS using the species Tawa. This information was extracted from a 20,000 record data set that was derived from integrating a subset of the NVS data bank with environmental character data. The resulting graphic contains both presence and absence information and uses a combination of maps, graphs and univariate statistics. This graphic takes approximately 10 minutes to produce and could be completed for all the species recorded in the data set. Currently there are only 33 species in this relational data set.

The graphical display of environmental characteristics of the habitat of a species helps people understand the complex nature of species distribution patterns. It perhaps raises more questions than are answered but the technique is very useful for encouraging further research and analysis. The graphic displayed in Figure 4 could be improved with the inclusion of multivariate statistics, a predicted distribution map, and statistical error information. It is conceivable that if species distribution data were made available in a relational format, it would be possible to produce graphics for every species. These graphics could be compiled into a publication on the Internet that was updated as new data distribution information was received. Such a publication could be used to ascertain priority areas for future data collection and analysis.

PLOT_NO	EASTING	NORTHING	FUCEXC	Elevation	Slope	Aspect	Mean Temperature	Mean Rainfall (monthly mm)	Mean Solar Radiation	Distance from Coast (km)	Terrain Index 10000	Terrain Index 1000	Terrain Index 500
174	2748676	6182027	1	1045	6	304	8	107	14	80	291	10	4
188	2747737	6181139	1	1008	9	21	8	105	14	79	53	15	9
189	2748650	6181113	1	1000	3	108	8	106	14	79	279	5	9
193	2743609	6180707	1	948	18	345	9	101	14	75	232	8	-5
194	2743603	6180525	1	984	7	235	9	101	14	75	240	40	25
195	2743598	6180342	1	979	11	293	9	101	14	75	240	35	13
200	2743865	6180060	1	989	6	169	9	101	14	75	240	30	9
201	2743959	6180149	1	979	7	161	9	101	14	75	240	12	-4
203	2744467	6181963	1	892	13	341	9	102	14	77	151	-67	-42
207	2744817	6181405	1	1060	8	339	8	103	14	77	303	47	12
208	2744904	6181219	1	1060	6	246	8	103	14	77	303	32	-4
210	2745946	6179269	1	878	28	126	8	106	14	76	176	-43	-21
216	2745055	6180118	1	1020	6	184	8	108	14	76	265	8	0
221	2747068	6180152	1	903	9	308	8	104	14	77	202	-50	-42
222	2747243	6179873	1	1005	18	322	8	103	14	78	235	46	33
225	2746930	6181711	1	964	9	151	8	104	14	79	228	-22	-16
228	2746768	6182447	1	1040	7	215	8	102	14	79	212	45	18
332	2761396	6411685	1	186	10	171	13	200	15	9	67	-72	-49
594	2767912	6383421	1	384	12	114	12	194	15	10	-126	-30	-26

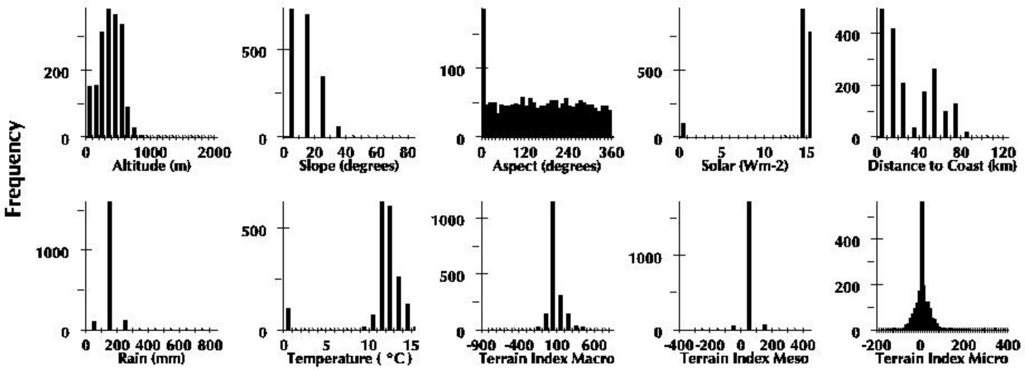
Table 1: Data Integration

Species - Tawa



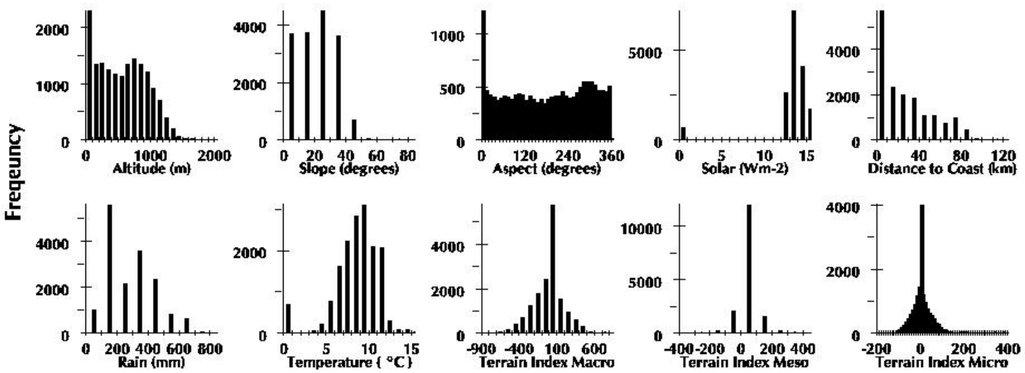
Environmental Characteristics - Species Present

These graphs show the frequency of different values of environmental characteristics where species bietaw are present.



Environmental Characteristics - Species Absent

These graphs show the frequency of different values of environmental characteristics where species bietaw are absent.



Environmental Characteristic Means

Easting	2691049
Northing	6315348
Altitude (m)	365
Slope (degrees)	12
Aspect (degrees)	164
Temperature (°C)	11
Rainfall (mm)	158
Solar Radiation (Wm-2)	13
Distance to Coast (km)	29
Macroscale Terrain Index	52
Mesoscale Terrain Index	3
Microscale Terrain Index	2

Vegetation Type	Frequency
Mixed Indigenous Forest	1588
Broadleaved Forest	92
Beech Forest	5
Indigenous Scrub	65
Cropland and Pasture	21
Kauri Forest	44
Exotic Forest	31
Podocarp Forest	1

Environmental Characteristic SD

Easting	61591
Northing	149433
Altitude (m)	174
Slope (degrees)	8
Aspect (degrees)	109
Temperature (°C)	2
Rainfall (mm)	43
Solar Radiation (Wm-2)	3
Distance to Coast (km)	23
Macroscale Terrain Index	125
Mesoscale Terrain Index	50
Microscale Terrain Index	33

Geology Type	Frequency
Tertiary sediments	855
Quaternary sediments	24
Jurassic sediments	53
Tertiary volcanics	444
Jurassic-Lwr Cretaceous seds	22
Cretaceous sediments	3
Quaternary volcanics	308
Tertiary intrusives	5
Triassic-Jurassic sediments	57
Permian sediments	52
Carboniferous-Permian sediments	21
Permian plutonics	3

Number of records where species bietaw present: 1847
Number of records where species bietaw absent: 16360
Total number of records: 18207

Figure 4: Data Visualisation

3.0 DATA ISSUES

The previously described applications of GIS demonstrate the potential of GIS for biodiversity mapping. If the appropriate data was available, it is conceivable that analysis could be taken further and the potential habitats of species could be predicted and mapped. To model biodiversity, it is important to have large data sets on the known location of species. These data sets are necessary to calibrate predictive models. The more information researchers have on the known location of species, the more accurate they can predict species distribution and map potential biodiversity. It is not the hardware or the software that is limiting the development of biodiversity models; it is access to environmental and species distribution data. The development of large species location data sets requires large-scale collaboration and coordination between environmental managers and research agents. Issues regarding data quality, metadata and standards are important (Burley, 1998). Data collectors perhaps need to be certified and there needs to be consensus on the type and format of the information collected. Above all, there needs to be a central organisation to coordinate such a project and a Biodiversity Strategy needs to address this.

If New Zealand is serious about conservation, then it should ideally be making biodiversity data available to anyone who cares to use it. With the current electronic and networked communication structures in place, the transfer of data should be of little cost and subsidised to encourage its use. An FTP site where data can be downloaded would cost very little in relation to the millions of dollars spent collecting and collating environmental data. It is important to recognise that there are a range of GIS users with different analysis abilities and requirements. Data sets provided using internet GIS may be suitable for projects at high school but are of limited use for environmental research analyst who need access to the entire data set as relational tables.

In 1999, Landcare Research priced the cost of transfer of the NVS data bank at $3110 and the format would be as individual plot files. The Herbarium Specimen database was priced at $11,109. These two data sets are very important biodiversity data sets. Even though the collection of the data in these data sets have been a substantial investment of state funding and the conversion of this data to a digital format has used additional state funding, they are not easily accessible to researchers. Not only are these data sets expensive to obtain they are also in an unfriendly individual plot file format, rather than a relational table format that can be easily used by standard database software. Landcare Research charged $1300 for a copy of the Land Resource Inventory and $700 for the climate data. Not only were there charges for the transfer of data, but signed contracts were required implying that Landcare Research has copyright on this "Public Good" data. This contract understandably includes a disclaimer but also prevents the sharing of data and is restrictive on use. On a more positive note, Landcare Research provided a subset of the NVS data bank for free which was in a user friendly relational format, and the Institute of Geological and Nuclear Sciences provided a Geology GIS layer for free. It is this more collaborative, "Public Good" approach that will lead to the development of biodiversity analysis tools and a more informed and educated public. Biodiversity is a "Public Good", therefore any agent, whether publicly or commercially funded, that improves biodiversity information and its dissemination is providing a "Public Good". These agents, whether public or private, should be encouraged and assisted to use the best biodiversity data available.

The maintenance and distribution of biodiversity data sets need to be managed by an agent that has a clear and willing agenda to do so. Under the Crown Research Institutes Act 1992, CRIs have conflicting agendas. Under section 5 of this Act, subsection 1(d) requires a CRI to promote and facilitate the application research results, while subsection 3(a) requires CRIs to generate profit to provide "an adequate rate of return on shareholders' funds". These conflicting "principles of operation" become manifested in "guiding philosophies" such as "We have important roles as information providers and knowledge brokers" (Landcare Research's Annual Report, 1999 p. 1). A CRI's requirement to generate profit conflicts with an agenda to share information with other researchers and environmental managers. Under the current legislation, CRIs are not the most appropriate agent to disseminate biodiversity data. Their expertise may make them appropriate to collect and document data but then this data and metadata should be deposited with a publicly funded agent that has a clear and willing agenda to promote and distribute the data.

4.0 CONCLUSION

This research has demonstrated some applications of GIS for biodiversity mapping, which will hopefully encourage ecologist to use GIS as a research tool. The availability of relevant data sets is a major catalyst for encouraging people to use GIS, so it is important that environmental character and species location data sets are easily accessible and usable. This is currently not the case.

Given that GIS requires coordination and calibration at a strategic management level, it is important that high level managers also have a general understanding of GIS functionality and application. Although it can not be expected that strategic managers be capable of using GIS, they need to be aware of its functionality and the issues relating to its implementation. There is a need for coordination of environmental information collection and dissemination.

With the appropriate databases made accessible in New Zealand and with the functionality of GIS analysis, the development of useful biodiversity and species habitat distribution models are possible. Such initiatives need to be instigated by independent research teams that can validate and critique each other's work to ensure such models are of scientific quality. The independent validation of research requires data to be shared and results and findings to be made known. Without such a process, there is no science.

ACKNOWLEDGEMENTS

I would like to thank the assistance of Duane Wilkins and Hugh Blackstock for the development of species data visualisation graphics. I would also like to thank Landcare Research, LINZ, and the Institute of Geological & Nuclear Sciences for providing data.

REFERENCES

Band, L. E. 1986. Topographic partitioning of watersheds with digital elevation models. Water Resource Research., 22: pp 15-24.

Brabyn, L.K. 1997 Classification of macro landforms using GIS. ITC Journal, 1997-1, pp 26-40.

Burley, J. 1998 Joining the revolution. A strategy for the standardization, integration and dissemination of biodiversity information- as a prospective model for the management of other kinds of environmental information. An internal Landcare Research publication.

Department of Conservation and The Ministry for the Environment 2000 The New Zealand Biodiversity Strategy. Our chance to turn the tide. ISBN 0-478-21919-9.

Dikau, R. 1989 The application of a digital relief model to landform analysis. In: Raper, J.F. (ed.) 1989 Three dimensional applications in Geographical Information Systems. Taylor and Francis, London, pp 51-77.

Dymond, J.R. DeRose, R.C. and Harmsworth, G.R. 1995 Automated mapping of the land components from digital elevation data. Earth Surface Processes and Landform, vol. 20, pp 131-137.

Francis, H. D. 1985 Soil landscape analysis, London : Routledge & Kegan Paul.

Gahegan, M., 1999 Four barriers to the development of effective exploratory visualisation tools for the geosciences. Int. Jrnl. Geographical Information Science 13(4): 289-309.

Geiger, R. 1971 The climate near the ground. Harvard University Press, Cambridge.

Lehmann, A., Leathwick, J. and Overton, J. 2000 Assessing hotspots of New Zealand fern diversity using Generalized Regression Analysis and Spatial Prediction (GRASP) 4th International Conference on Integrating GIS and Environmental Modeling (GIS/EM4. Banff, Alberta, Canada, September 2 - 8, 2000.

Stafford, S.G., J.W. Brunt and W.K. Michener, 1994: Integration of scientific information management and environmental research. In: Michener, W.K., J.W. Brunt and S.G. Stafford (editors), Environmental Information Management and Analysis. Taylor and Francis, Bristol.

Tivy, J. 1993 Biogeography - A study of plants in the ecosphere. Longman, Harlow.

Wadsworth, R. and Treweek, J. 1999 Geographic Information Systems for ecology. Longman, London.