The development of predictive models to enhance biological assessment of riverine systems in New Zealand

A thesis presented in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Ecology
At Massey University, Palmerston North, New Zealand

Michael Kevin Joy

2003

Massey University
COLLEGE OF SCIENCES

Institute of Natural Resources
- Ecology
Massey University
Private Bag 11 222,
Palmerston North,
New Zealand
Telephone: 64 6 356 9099
Facsimile: 64 6 350 5623

Ph.D. CANDIDATE DOCTORAL EXAMINATATION APPLICATION

Candidate's Name: ***Joy, Michael Kevin***

Academic Unit: ***Institute of Natural Resources (Ecology)***

Provisional registration date 14.12.1999 – F; Thesis submission deadline 14.12.2003

Thesis title: "The development of predictive models to enhance biological assessment of riverine systems in New Zealand"

Statement regarding the nature and extent of any assistance received during the doctoral research:

For all chapters, my input was the greatest, I planned the research, undertook all fieldwork, analysed all data, and wrote all manuscripts. My main supervisor Russell Death gave assistance in the following fields: the development of the original project concept, editing manuscripts, overseeing project administration and funding and discussing ongoing developments. My other two supervisors Prof Brian Springett and Dr R. M. McDowall gave no direct assistance apart from general discussions and administrative duties.

None of the material in this thesis has been used for any other degree or diploma.

The first 4 chapters of the thesis have been published or are in press, the other two chapters are under review with journals. Details on publications:

Chapter 1 Joy, M. K., and R. G. Death. 2002. Predictive modelling of freshwater fish as a biomonitoring tool in New Zealand. Freshwater Biology **47**:2261-2275.

Chapter 2 Joy, M. K., and R. G. Death. 2003. Biological assessment of rivers in the Manawatu-Wanganui region of New Zealand using a predictive macroinvertebrate model. New Zealand Journal of Marine and Freshwater Research **33**:367-379.

Chapter 3 Joy, M. K., and R. G. Death. In Press. Neural network modelling of freshwater fish and macro-crustacean assemblages for biological assessment in New Zealand. *in* S. Lek, M. Scardi, and S. E. Jorgensen, editors. Modelling community structure in freshwater ecosystems. Springer Verlag.

Chapter 4 Joy, M. K., and R. G. Death. In Press. Assessing biological integrity using freshwater fish and decapod habitat selection functions. Environmental Management.

Candidate Michael K. Joy

Massey University
COLLEGE OF SCIENCES

Institute of Natural Resources
- Ecology
Massey University
Private Bag 11 222,
Palmerston North,
New Zealand
Telephone: 64 6 356 9099
Facsimile: 64 6 350 5623

Ph.D. CANDIDATE DOCTORAL EXAMINATATION APPLICATION

Candidate's Name: ***Joy, Michael Kevin***

Academic Unit: ***Institute of Natural Resources (Ecology)***

Provisional registration date 14.12.1999 – F; Thesis submission deadline 14.12.2003

Thesis title: "The development of predictive models to enhance biological assessment of riverine systems in New Zealand"

Statement regarding doctoral thesis:

This statement confirms that the candidate has pursued the Doctoral Course in accordance with the University's Doctoral regulations.

Supervisor Russell G. Death

Massey University
COLLEGE OF SCIENCES

Institute of Natural Resources - Ecology
Massey University
Private Bag 11 222,
Palmerston North,
New Zealand
Telephone: 64 6 356 9099
Facsimile: 64 6 350 5623

Ph.D. CANDIDATE DOCTORAL EXAMINATATION APPLICATION

Candidate's Name: ***Joy, Michael Kevin***

Academic Unit: ***Institute of Natural Resources (Ecology)***

Provisional registration date 14.12.1999 – F; Thesis submission deadline 14.12.2003

Thesis title: "The development of predictive models to enhance biological assessment of riverine systems in New Zealand"

Statement regarding thesis:

1. Reference to work other than that of the candidate, has been appropriately acknowledged,
2. Research practice, ethical and genetic technology policies have been complied with as appropriate, and
3. The thesis does not exceed 100,000 words (excluding appendices)

Supervisor Russell G. Death

Candidate Michael K. Joy

Abstract

A suite of new regional and national lotic freshwater bioassessment tools were developed for New Zealand. This work permits the inclusion of freshwater fish in bioassessment, a component of the fauna previously largely ignored. The multivariate predictive models developed gave a number of advantages over the existing albeit overextended single-index approach (the macroinvertebrate community index) used by regional authorities. To acquire the data for constructing the models more than 500 sites were sampled over three North Island regions. The sites were selected to represent least impacted conditions known as reference sites so that the biotic communities sampled would representing the best attainable or the goal for resource managers. Models were constructed to predict the biota representing best available conditions based on the non human influenced physicochemical variables defining the sites. The predicted and observed assemblages were then compared using an observed over expected ratio *(O/E)* so that scores less than 1 represent less species observed than expected. This *(O/E)* ratio is more than simply the assessment of species richness, as only those species predicted are included in the ratio. Reference site multivariate predictive models using fish and macroinvertebrate assemblage groups were developed for bioassessment in the Manawatu-Wanganui Region. Two reference site multivariate predictive models using individual fish and decapod species were developed for the Auckland region. The first used traditional linear discriminant function analysis and the second used artificial neural networks (ANNs). A model to predict the spatial occurrence of fish and decapods was developed for fish in the Wellington Region using Geographic Information Systems (GIS) and ANNs. The remotely sensed data was available for all rivers in the region so the predictions could be extended over the entire stream network to produce a fish map. Finally an index of biotic integrity (IBI) using fish was developed for the entire country and evaluated using remotely assessed environmental data. Exhaustive evaluations of predictions from all the models confirmed their credibility as a biomonitoring.

Acknowledgments

Each Chapter has its own acknowledgments section. Here I thank those people whose assistance was not specific to any one part of the project.

First, I thank my first supervisor, Russell Death for giving me the confidence to tackle this project right from its beginnings in my Masters research. His encouragement and support from inception to final write-up has been comprehensive and inspirational. I also owe his wife Fiona a debt of gratitude for her support and missing out on family time with Russell so he could spend time on this project.

Thanks to my second supervisor Bob McDowall for his support throughout the entire project and in the preparation. His association with the project undoubtedly helped the progression through the initial funding applications. I thank him for his patience and help over the years sending reprints and innumerable telephone conversations.

The project would not have been possible without the support of three regional councils: Manawatu-Wanganui (now horizons.mw), Auckland (ARC) and Wellington (WRC) and the New Zealand Ministry for the Environment (MfE). In particular through the support of staff who supported the initial application for funding through the Sustainable Management Fund (SMF). In particular I thank John Phillips (horizons), Chris Hatton, John Maxted and Brent Evans (ARC) and Murray McLea (WRC).

Thanks must go to all who helped with fieldwork, especially Allison Hewitt who was fishing partner for the Wellington and Auckland surveys and Grant Blackwell who was my electro-fishing partner and best mate for the 40,000 km Manawatu-Wanganui survey. Thanks to Rosemary Miller, Dr Ian Henderson, Scott Bowie, Sjaan Charteris and Murray McLea for helping with fish surveys when the regulars couldn't make it.

The Ecology Group staff provided support throughout the project, thanks especially to Barbara Just, Jens Jorgensen, Cathy Lake, Erica Reid, Paul Barrett, and Carol Nicholson

Thanks to friends, family and colleagues who have lent support and or discussed ideas over the last 4 years. In particular, thanks to the Saturday morning coffee group Suzanne Bassett, Grant Blackwell, Halema Flannagan, Peter Ritchie, Alistair Robertson, Matt Lowe, Åsa Berggren, Becky Lewis, Cath Morrison, Viv Nichols, Tony McGlynn, John Holland and Brian Springett.

Thanks to the student team of the Massey University Centre for Freshwater Ecosystem Modeling and Management for many discussions and editing, especially Alex, Zoe, Kiryn, Mark, Murray, Pepe and Erna.

The project benefited immeasurably from the support and assistance of my partner and soul mate Allison Hewitt. Who assisted with the fishing of more than half of the >500 sites sampled and put up with my mad obsession for more than 4 years. For this and much more, I am eternally grateful.

Table of contents

Note on text and authorship

Each Chapter is set out largely in the style of the journal it was published in or has been submitted to. Consequently there is some inevitable repetition, especially in the methods sections, and there are minor stylistic differences between the chapters. For each paper the only co-author is my principal supervisor Russell Death this recognises his contribution in developing the original project concept, editing manuscripts, overseeing project administration and funding and discussing ongoing developments. For all chapters, my input was the greatest. I planned the research, undertook all fieldwork, analysed all data, and wrote all manuscripts.

Introduction

The use of biological communities to assess the state of riverine ecosystems has a long history and has in general replaced chemical assessment worldwide (Karr & Chu, 2000; Karr, 1999). In developed countries, the general trend has been for the development of a single response, invertebrate index to be replaced or expanded over time to multimetric or multivariate approaches (Reynoldson et al. 1997). There has long been a parallel trend for inclusion of other riverine biological elements in assessment including fish (e.g. Karr 1981) and periphyton (Chessman et al. 1999). This progression reflects and is a response to the growing requirement for a more holistic, ecosystem level assessment of freshwater systems (e.g. Karr 1995, Allan and Johnson 1997, Johnson and Gage 1997, Postel and Carpenter 1997, Rapport et al. 1998, Boulton 1999, Bunn et al. 1999).

In New Zealand this move towards ecosystem level assessment has been enshrined in resource management legislation with the Resource Management Act (RMA) (1991) (Winterbourn 1999). Specifically the RMA refers to: "safeguarding the life supporting capacity of...water... and ecosystems" and "having regard to intrinsic value of ecosystems". There is thus, not just a will, but a requirement in New Zealand for assessment tools to encompass multiple impacts, multiple variables, and multiple biotic groups at an ecosystem level.

Unfortunately however, the assessment tools used by New Zealand regulatory authorities (Regional Councils) in no way match this legislative requirement for multi response, multivariate, ecosystem level assessment. Although the majority of the 16 Regional Councils include bioassessment in their freshwater resource management, this bioassessment has almost invariably been limited to the Macroinvertebrate Community Index (MCI) and its derivatives (Stark 1993, Winterbourn 1999). This index was developed for one region on streams flowing from Mt Taranaki, to assess the response of invertebrate communities to organic enrichment from intensive agriculture with no initial intention of extending its use beyond the Taranaki Ring Plain. The MCI and its more recent quantitative derivatives is not suitable for national bioassessment for a number of reasons: (i) it was developed on only one stream type (Mt Taranaki streams with cobble substrates)

(ii) it was calibrated on a single impact (organic enrichment) and (iii) it uses only a single component of stream biology (invertebrates).

There is therefore, a manifest requirement for new tools to upgrade freshwater bioassessment in New Zealand. My overriding goal was to provide a suite of new tools to address the shortcomings of the present freshwater bioassessment processes used in New Zealand and to this end emphasis was placed on providing user friendly computer software to implement the models so that they could be readily employed by non-specialists. Specifically the responses to the present shortcomings to be addressed were: (i) the use of more than a single biotic group by including fish as well as invertebrates, (ii) the use of reference sites to give a level of biotic condition ("biocriteria") to be aimed for, (iii) to give assessment a statistical objective basis to bioassessment as opposed to the subjective assignment of scores to taxa and (iv) the inclusion of prediction to models, to give a more "rigorously scientific, more informative, and more useful ecology" (Peters 1991; p. 274).

To achieve the aims outlined above, I described in Chapter 1 the development of a predictive reference site bioassessment model using fish. This was the first published application of a long established bioassessment approach previously only used with invertebrates in the United Kingdom, Europe, and Australia (RIVPACs and AUSRIVAS, Simpson and Norris 2000, Wright 2000) and diatoms (Chessman, 1999). Chapter 2, described the first published, peer-reviewed application of a predictive reference site bioassessment model using invertebrates in New Zealand. Chapters 3 and 4, described the development of predictive reference site bioassessment models using individual fish species models rather than the clustering approach used in chapters 1 and 2 and the existing RIVPACs and AUSRIVAS models. The models described in Chapters 3 and 4 were developed for the same region using the same data but differ in the modelling process used. Chapter 3 was the first application of artificial neural networks to predictive bioassessment in New Zealand and combined individual species models while Chapter 4 described the use of an individual species model using discriminant function analysis. Chapter 5 describes the development of a predictive fish model combining artificial neural networks and geographic information systems (GIS) to produce a spatial prediction map to predict the entire assemblage with one model. In Chapter 6 I described the first application of an index of biotic integrity (IBI) approach to bioassessment using freshwater fish in New Zealand with a national model.

The overall aim was to make these models available for non-specialist biologists and this was accomplished by supplying software to access the models produced in all six chapters to Auckland, Wellington, and the Manawatu-Wanganui Regional Councils.

The modeling process

The individual models were described in detail in each chapter but the approaches used in the first 5 chapters all have a similar underlying basis and this is best described using the diagram in Fig 1. The process follows the steps numbered in the boxes and circles in the figure. The first step in the process is the conceptual model, developed from field knowledge, ecological literature, and existing models. The conceptual models for Chapters 1 – 5 were based on existing predictive models namely the River Invertebrate Prediction and Classification System (RIVPACS) approach, originally developed in the U.K. by Wright and colleagues (Wright, 1995) and later advanced by Simpson & Norris (2000) in Australia with the Australian River Assessment Scheme (AUSRIVAS). Two statistical models were used in box 2; a discriminant function model (DFA) for chapters 1, 2 and 4 whereas artificial neural networks (ANN) models were used for Chapters 3 and 5. The left side of the diagram describes the internal calibration and evaluation and the right side describes the external evaluation.

Step 3 is the model calibration step; this is the mechanism to get data for adjustment of model parameters to improve agreement between model outputs and the data set. The relationship between predictions and observed data was evaluated using a number of tests. The calibration of the DFA models involved finding the set of environmental variables that best discriminate between the groups being modeled, either into biotic groups or individual presence/absence groups. To achieve this DFA optimisation, stepwise variable reduction techniques and iterative trials were used. The calibration of the ANN models however, was carried out to optimize the network architecture with the aim of maximising prediction success.

Step 5 was the generation of predictions from the evaluation dataset, this was either a repeated random selection of 20% of the data or jackknifing of all data. A number of measures of the match between observed and predicted values were then used. This process was used as a measure of prediction success or model credibility.

Not shown on the diagram but explained in the individual chapters is the final step to validate the bioassessment models used in chapters 1 – 4 and 6.

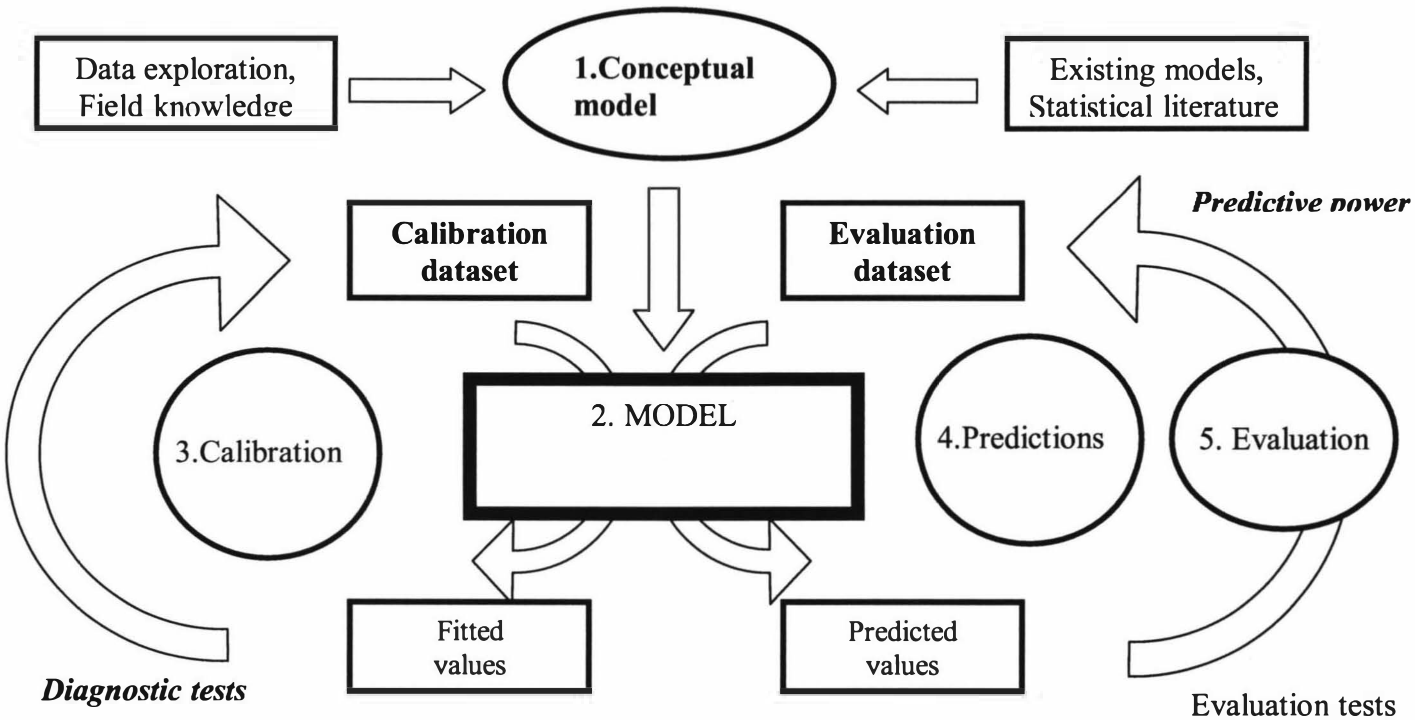

Fig. 1 Overview of the modelling process (steps 1-5), when using two data sets or jackknife techniques. See text for details. After Guisan and Zimmerman (2000).

References

Allan, J. D., and L. B. Johnson. 1997. Catchment-scale analysis of aquatic ecosystems. Freshwater Biology **37**:107-111.

Boulton, A. J. 1999. An overview of river health assessment: philosophies, practice, problems and prognosis. Freshwater Biology **41**:469-479.

Bunn, S. E., P. M. Davies, and T. D. Mosisch. 1999. Ecosystem measures of river health and their response to riparian and catchment degradation. Freshwater Biology **41**:333-345.

Chessman, B. C., I. O. Growns, and N. Plunket-Cole. 1999. Predicting diatom communities at genus level for the biological management of rivers. Freshwater Biology **41**:317-331.

Johnson, L. B., and S. H. Gage. 1997. Lanscape approaches to the analysis of aquatic ecosystems. Freshwater Biology **37**:113-132.

Karr, J. R. 1981. Assessments of biotic integrity using fish communities. Fisheries **6**:21-27.

Karr, J. R. 1995. Protecting freshwater ecosystems: clean water is not enough. *in* W. S. Davis & T. P. Simon, editors. Biological Assessmant and Criteria: tools for water resource planning and decision making. Lewis publishers, Boca Raton.

Karr, J. R. 1999. Defining and measuring river health. Freshwater Biology **41**:221-234.

Karr, J. R., and E. W. Chu. 2000. Sustaining living rivers. Hydrobiologia **422/423**:1-14.

Postel, S. L., and S. Carpenter. 1997. Freshwater ecosystem services. Pages 195-214 *in* G. Daily, editor. Natures services: societal dependence on natural ecosystems. Island Press, Washington D.C.

Rapport D. J., Costanza R. & McMichael A. J. 1998 Assessing ecosystem health. Trends in Ecology & Evolution, **13**, 397-402.

Reynoldson, T. B., R. H. Norris, V. H. Resh, K. E. Day, and D. M. Rosenberg. 1997. The reference condition: a comparison of multimetric and multivariate approaches to assess water-quality impairment using benthic macroinvertebrates. Journal of the North American Benthological Society **16**:833-852.

Simpson, J., and R. H. Norris. 2000. Biological assessment of water quality: development of AUSRIVAS models and outputs. Pages 125-142 *in* J. F.

Wright, D. W. Sutcliffe, and M. T. Furse, editors. Assessing the biological quality of freshwaters. RIVPACS and other techniques. Freshwater Biological Association, Ambleside, UK.

Stark, J. D. 1993. A macroinvertebrate community index of water quality for stony streams. New Zealand Journal of Marine and Freshwater Research **27**:463-478.

Winterbourn, M. J. 1999. Recommendations of the New Zealand Macroinvertebrate Working Group: Monitoring for sustainable river ecosystem management and the role of macroinvertebrates. *in* M. J. Winterbourn, editor. The use of macroinvertebrates in Water Management. Ministry for the Environment, Wellington.

Wright, J. F. 2000. An introduction to RIVPACS. *in* J. F. Wright, D. W. Sutcliffe, and M. T. Furse, editors. Assessing the biological quality of freshwaters. RIVPACS and other techniques., Ambelside, UK.

CHAPTER ONE

Predictive modeling of freshwater fish as a biomonitoring tool in New Zealand

Joy, M. K., and R. G. Death. 2002. Predictive modeling of freshwater fish as a biomonitoring tool in New Zealand. Freshwater Biology **47**:2261-2275

Abstract

A challenge has been issued to ecologists to find quantitative ecological relationships that have predictive power. A predictive approach has been successful when applied to biomonitoring using stream invertebrates with the River Invertebrate Prediction and Classification System (RIVPACS). This approach, to our knowledge, has not been applied to freshwater fish assemblages. This paper describes the initial results of the application of a regional predictive model of freshwater fish occurrence using 200 reference sites sampled in the Manawatu-Wanganui region of New Zealand over late summer autumn 2000. In brief, (i) sites were classified into biotic groups, (ii) the physical and chemical characteristics that best describe variation among these groups were determined and (iii) the relationship between these environmental variables and fish communities was used to predict the fauna expected at a site. Reference sites clustered into six groups based on fish density and community composition. Using 14 physical variables least influenced by human activities, a discriminant model allocated 70% of sites to the correct biological classification group. The variables that best separated the site groups were mainly large-scale variables including altitude, distance from the coast, lotic ecoregion and map co-ordinates. The model was further validated by randomly removing 20% of the sites, rebuilding the model and then determining the number of removed sites correctly allocated to their original biotic groups using environmental variables. Using this process 67% of the removed sites were correctly reassigned to the six predetermined groups. A further 30 sites were used to determine the ability of the model to detect anthropogenic impact. The observed over expected taxa (*O/E*) ratios were significantly lower than the reference site *O/E* ratios, indicating a response of the fish assemblages to the known stressors.

Keywords bioassessment, biotic integrity, diadromy, freshwater fishes, multivariate analysis, New Zealand, RIVPACS-type predictive models

Introduction

Biological communities are sensitive indicators of the relative health of their aquatic ecosystems and the surrounding catchment (Fausch, Lyons, Karr and Angermeier, 1990; McCormick, Peck and Larsen, 2000). The relationship between the biological and physical and chemical components of ecosystems is the basis of biological monitoring. Fish are potentially effective indicators of the condition of aquatic ecosystems because different species exhibit diverse ecological, morphological and behavioral adaptations to their natural habitat (Karr, Fausch, Angermeier, Yant and Schlosser, 1986; Fausch *et al.,* 1990; McCormick *et al.,* 2000). Fish communities integrate the ecological processes of streams across both temporal and spatial scales (Hynes, 1970; Harris, 1995), they can be useful indicators of aquatic degradation (Karr, 1981; Berkman, Rabeni and Boyle, 1986; Steedman, 1988; Fausch *et al.,* 1990; Karr, 1991). Furthermore, because fish are a visible part of stream biological integrity they represent a measure of stream quality easily and intuitively understood by the public (McCormick *et al.,* 2000). Despite this potential however, fish have not been used in the past in biological assessment in New Zealand due to the overwhelming influence of altitude and distance from the sea on fish distribution (McDowall, 1998b, 1999; Joy, Henderson and Death, 2000; McDowall and Taylor, 2000; Joy and Death, 2001). The importance of these variables is related to the predominance of diadromy in the fauna resulting in distributions primarily influenced by altitude and distance from the sea with species richness decreasing as altitude increases (McDowall, 1990; McDowall, 1998a; Joy, Henderson and Death, 2000). This overwhelming influence of diadromy has negated the application of between site comparisons in the past.

Globally, freshwater systems are rapidly deteriorating and hence these systems are receiving increasing attention (Allan and Flecker, 1993; Matson, et al., 1997; Postel, 2000). In New Zealand human impacts on rivers have been considerable since European settlement. Nutrient balances have been altered both with agricultural run-off and urban sewage discharges. Sediment inputs have increased through a combination of deforestation and land clearance resulting in a 10-fold increase in grassland, the removal of 75% of the forest cover and 85% of all wetlands (Anonymous, 1997). These changes, particularly the removal of riparian vegetation and wetland drainage have had profound effects on rivers, wetlands and their ecologies (Anonymous, 1997). However, few land-users acknowledge the presence of a problem. Thus, there is a need for the development of practical tools that provide accurate ecological assessment of river health or condition without requiring a high level of expertise, effort, and time from end users.

The basis of the analyses used in this paper is the River Invertebrate Prediction and Classification System (RIVPACS) approach, originally developed in the U.K. by Wright and colleagues (Wright, 1995) and later advanced by Simpson and Norris (2000) in Australia with the Australian River Assessment Scheme (AUSRIVAS). The two models are similar and will be referred to in this paper as RIVPACS models. RIVPACS models assess biological status by comparing the biotic condition at sites being evaluated with the biota expected to occur in the absence of stress (Wright, 1995). A predictive model is built using biological, physical and chemical data collected at a number of unimpacted or minimally impacted sites, generally referred to as reference sites. The reference sites are classified into groups based on the homogeneity of their fauna, and the physical and chemical characteristics that best describe variation among the groups are determined. Finally, discriminant analysis is used to predict the biotic communities expected to occur in the absence of environmental stress. A detailed account of the background to this type of predictive modeling has been covered in numerous publications (e.g. Wright *et al.*, 1984; Moss *et al.*, 1987; Wright, Furse and Armitage, 1993; Wright, 1995), and hence will not be explained in detail here.

The primary goal of this paper is to describe the application of a RIVPACS-type approach to freshwater fish communities in a large region of New Zealand. The process was to assess whether fish assemblages can be characterised by a limited set of environmental variables and then to convert those associations to predictions of assemblage composition. Here, the aim is to give an objective site-specific assessment of biological integrity based on the best available conditions in the region. Successful predictions will provide confidence in the potential for application of this approach in other regions of New Zealand and elsewhere.

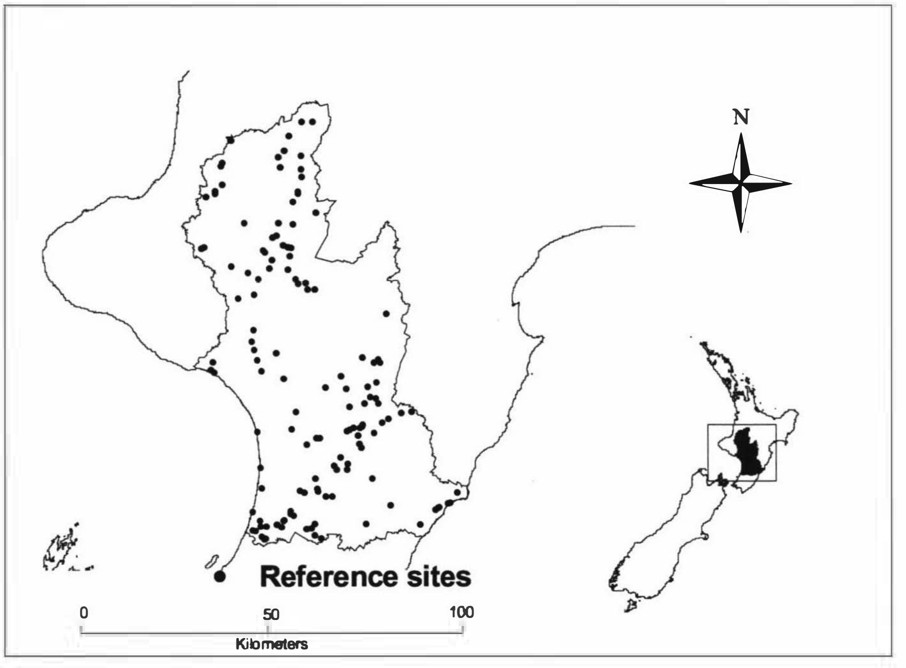

Fig . 1 Location of reference sites in the Manawatu-Wanganui region, North Island New Zealand.

Methods

Study area and sampling sites.

The Manawatu-Wanganui Region covers a large portion (22,179 km^2) of the south west of the North Island in New Zealand (Fig. 1). It is a political region delineated mainly by the catchment boundaries of the three major rivers in the region, the Wanganui, Rangitikei and Manawatu rivers. A volcanic plateau dominates the northern part of the region and to the south the axial Tararua and Ruahine ranges divide the region running approximately north-south. The Manawatu River is unusual in that it bisects this axial mountain range and drains the axial ranges on both their eastern and western flanks. The region includes considerable areas of uplifted ranges, steep mudstone country and rich alluvial plains resulting in a diversity of stream types, ranging from braided cobble to silt dominated lowland rivers, and from an acidic volcano fed river to small spring fed mountain streams. Study sites spanned 40° 45' to 39° 30' south and from sea level to 820 m elevation. The predominant land use in the region is pastoral farming with some cropping while most of the region above 500 m a.s.l. has relatively unmodified native vegetation.

Reference site selection.

Two hundred candidate reference sites were initially selected using 1:50 000 topographic maps and expert knowledge of local regulatory authority staff. Emphasis was placed on sites having the most natural catchment vegetation, channel morphology, and minimal human impacts (e.g. Hughes, 1995). The reference sites were concentrated in areas and altitudes of greatest management concern, where potentially the model will be used. Accordingly sites were stratified over elevational gradients, with fewer sites at greater altitude. In order to have sites over the full range of altitudes and stream types some moderately disturbed but best available catchments were included at lower altitudes (Hughes, Larsen and Omernik, 1986). New Zealand has been classified into lotic ecoregions (Harding, 1994) and seven of these ecoregions in the Manawatu-Wanganui region were used to further stratify natural variability.

Data collection

Fish. The survey was carried out from late summer until early autumn (24 January–30 May 2000). This is the period when maximum fish species diversity occurs in New Zealand streams, because all diadromous fish species are present (McDowall, 1990). Each site was sampled by single-pass electro fishing involving two people using a battery-powered electric fishing machine (EFM300; NIWA Instrument Systems, Christchurch, New Zealand), operated at 150–300 V depending on the water conductivity. Single pass electrofishing is a suitable survey method for assessing distribution of freshwater fish in New Zealand, because all species are collected with equal probability on the first pass (Jowett and Richardson, 1996). Fish were collected in dip nets or stop nets held downstream and identified to species, counted and returned to the water. The length of stream sampled was set to ensure that at least two examples of each habitat type were covered (e.g. riffles, runs, pools, rapids). Fish seen and positively identified to species, but not captured, were also recorded to give a minimum estimate of density.

Fish assemblages. Relative fish abundance at all sites was standardised by dividing the number of fish by the area fished (mean width × length × 100) to give the density of fish per 100 m2. The percentage of river fished was estimated for large rivers where current or depth prevented full access and the area fished was reduced proportionally. Larger pools not sampled at sites where access was limited were considered unlikely to influence species richness estimates as comparatively few of the native species are present in large pools because the combination of fine substrate and deep water provides unsuitable habitat (Jowett and Richardson, 1995, 1996). Juvenile eels (<300 mm length) were treated as separate operational taxonomic units (OTUs) in the analysis, because they are known to shift their habitat niche during ontogeny (Werner and Gilliam, 1984; Hayes, Leathwick and Hanchet, 1989; Glova, 1998). This process was applied only to the above species, because there is no information available on within-species habitat requirement differences for other species. The two non-migratory bully species, Cran's and upland (*Gobiomorphus basalis* Gray and *G. breviceps* Stokell), were combined into a single OTU because their habitat requirements are similar (McDowall, 1990) and differences in occurrence would be indicative of biogeography rather than habitat. Two trout species were encountered, predominantly brown trout (*Salmo trutta* Linnaeus) and occasionally rainbow trout (*Oncorhychus mykiss* Richardson). Again these two trout species were combined into a single OTU because differences in distribution relate to introductions by sport fisheries managers rather than habitat requirement differences.

Macroinvertebrates. At each of the 200 sites macroinvertebrates were collected prior to electro fishing using a 1-min kick-net sample from a riffle with a D net (50 cm wide, 20 cm high, 250 μm mesh). Invertebrate samples were preserved in 10% formalin. In the laboratory, samples were rinsed using a 250-μm mesh sieve to remove fine sediment and preservative and placed in a Marchant subsampling box (Marchant, 1989). The first 100 invertebrates were then extracted from randomly selected cells, sorted and identified. Identifications were to the lowest reliable taxonomic level (usually species) using existing keys (e.g. Winterbourn and Gregson, 1989).

Environmental measures. Forty-five physical and chemical variables were measured at each site (see below) to complement the fish and invertebrate species lists using a number of methods (Table 1 lists the environmental and chemical measures collected together with notes on data collection procedures). At each site, the percentage of riparian vegetation in five categories (native forest, exotic forest, grass, tussock and scrub) and the percentage of overhead shade were visually assessed. The relative proportions of catchment land use and vegetation (e.g. native forest, exotic forest, tussock, predominant farming type), altitude and distance from the sea were calculated using geographic information systems (GIS) (Arc-View, 1999) and 1: 50 000 topographic maps (New Zealand Survey and Land Information, 1987). Conductivity (automatically adjusted to 25 °C), temperature, dissolved oxygen and salinity were measured on-site with a YSI model 85 meter (YSI Inc., Yellow Springs, OH, USA) and pH was measured with an Orion Quickcheck model 106 pocket meter (ThermoOrion, Beverley, MA, USA). A water sample was collected from each site and frozen within 12 h for laboratory analysis of nutrients and alkalinity.

Table 1. Physical and chemical variables measured at 200 potential reference sites in the Manawatu-Wanganui region. Variables in bold were those initially considered for use in model construction as they are unlikely to be affected by anthropogenic activity. Variables with asterisk were used in the final discriminant model.

Environmental variable	Min	Max	Mean	Foot-note
Geographic variables				
Easting* (Co-ords.)	26635	28113	27266.11	1
Northing* (Co-ords.)	60033	62882	61445.46	1
Altitude* (m) a.s.l.	1	820	270.68	1
Distance inland* (km)	0.1	330	141.28	2
River reach sampled				
Length fished (m)	20	120	78.13	3
Mean depth* (cm)	62	1480	39.85	4
Mean width* (m)	1	115	8.12	4
Temperature* (° C)	4	24	14.12	5
PH	7	10	8.13	5
Conductivity mS/cm^{-1}	< detection level	864	189.67	5
Chlorophyll *a* mg/cm^2	< detection level	276	17.63	6
Total Alkalinity (mgCaCO$_3$ l^{-1})	< detection level	302	58.06	7
NO_3 mg/l^{-1}	< detection level	18	0.58	8
PO_4 mg/l^{-1}	< detection level	2	0.06	8
Velocity m/sec^{-1}	0	1.4	0.35	9
Flow type (%)				
Still	0	100	1.60	10
Backwater*	0	15	3.63	10
Pool*	0	100	18.05	10
Run*	0	100	42.20	10
Riffle*	0	100	27.40	10
Rapid	0	80	7.08	10

Habitat characteristics (%)				
Over stream cover	0	95	19.14	10
Undercut banks	0	75	11.94	10
Debris jam	0	40	4.50	10
Exposed bed	0	80	23.49	10
Macrophyte cover	0	90	2.79	10
Catchment vegetation (%)				
Native forest	0	100	56.20	2,10
Exotic forest	0	100	5.50	2,10
Pasture	0	100	35.95	2,10
Tussock	0	15	0.13	2,10
Swamp	0	5	0.03	2,10
Catchment landuse				
Forestry	0	100	6.35	2,10
Native forest	0	100	56.10	2,10
Dairy farming	0	100	7.75	2,10
Sheep/Beef farming	0	100	29.45	2,10
Urban	0	40	0.35	2,10
Riparian Vegetation (%)				
Native	0	100	44.98	10
Exotic forest	0	90	13.53	10
Pasture	0	100	27.95	10
Willow	0	75	9.95	10
Raupo	0	90	3.10	10
Exposed bed	0	80	23.49	10
Substrate and ecoregion				
Pfankuch stability score	32	128	80.22	11
Embededness*	1	4	1.45	13
Median substrate size (cm)*	0	40	12.43	12
Lotic ecoregion*	1	7		14

1. Obtained from 1:50,000 NZMS maps
2. Geographic information systems (ARCVIEW) using 1:50 000 vector data
3. Of water measured over length of reach fished.
4. Of water mean of 5 measures over length of reach fished.
5. Measured at time of fishing with YSI model 85 meter.
6. 6.15 cm^2 periphyton scrubbed from 5 randomly selected stones at each site Chlorophyll *a* extracted with 90% acetone total pigment concentration were calculated using the method of (Steinman and Lamberti, 1996).
7. Potentiometric titration (Horizons.mw laboratory, Palmerston North)
8. Massey University chromatograph
9. Calculated by time taken for a slug of dye to travel length fished
10. Visually estimated at site
11. Pfankuch (Pfankuch, 1975b) stability index which involves scoring 15 variables (weighted in relation to their perceived importance) according to the observers evaluation of predetermined criteria. Three totals relate to three regions of the stream channel upper banks, lower banks, and stream bottom.
12. Median substrate size index from 50 – 100 stones collected at random over the reach fished and measured in 11 phi classes (Wolman, 1954)
13. Subjectively assessed at site after moving substrate (1 = loosely packed; 4 = tightly packed)
14. Lotic ecoregions (Harding, 1994)

Maximum water depth and stream width was measured at five equidistant points longitudinally over the reach fished. Cover for fish was visually assessed as the percentage area consisting of undercut banks, macrophyte cover and debris jams. The median substrate size was calculated by summing the median values of the size classes weighted by their proportional cover from measurements of 70–100 individual particles collected in 11 phi (log2) size classes over the fished reach (Wolman, 1954; Quinn and Hickey, 1990). The percentage of backwater, pool, run, riffle or rapid, was visually estimated over each reach surveyed. Riffles were classified as areas of fast, shallow water with a broken-surface appearance; pools were areas of slow, deep water with a smooth surface appearance, whereas runs were intermediate in character. Rapids were classified as areas of fast cascading deep water. Mean water velocity was calculated by timing the movement of the modal concentration of a slug of dye over the length of reach fished. Channel stability at each site was assessed using the method of Pfankuch (1975), which involves scoring 15 variables (weighted in relation to their perceived importance) in three sections of the stream channel (substrate, lower and upper banks) according to the observer's evaluation of predetermined criteria. The ratings were combined to give an overall stream stability score that can range from 40 (most stable) to 160 (least stable). Periphyton samples were taken at each site using a 6.15-cm^2 circular sample from each of five arbitrarily selected stones and these were frozen for later pigment extraction (Davies and Gee, 1993). Pigments were extracted in 90% acetone and absorbencies were read on a Jenway 6105 UV. Visible spectrophotometer (Jenway Limited, Essex, UK) and converted to pigment concentration (chlorophyll a and phaeophytins) following Steinman and Lamberti (1996).

Model construction

The statistical procedures used in constructing RIVPACS and AUSRIVAS type predictive models using macroinvertebrates have been described elsewhere (Wright *et al.*, 1984; Moss *et al.*, 1987; Wright *et al.*, 1993; Wright, 1995; Clarke, Furse, Wright and Moss, 1996; Simpson and Norris, 2000) and the implementation described below followed similar steps using elements from both procedures.

Reference sites were divided into groups with similar fish communities and densities using two-way indicator species analysis (TWINSPAN) (Hill, 1979; Gauch and

Whittaker, 1981) using PC-ORD (McCune and Mefford, 1997). This process classifies both samples and species simultaneously based on hierarchical divisions of reciprocal averaging ordination space. A number of studies have shown that this technique is useful as a preliminary analysis tool in freshwater biology both with stream invertebrates (Townsend, Hildrew and Francis, 1983; Ormerod and Edwards, 1987) and fish (Hayes *et al.*, 1989; Joy and Death, 2001). Five pseudospecies cut levels, 0, 2, 5, 10 and 20 fish per 100 m^2, were used in the TWINSPAN analysis, and the analysis was taken to three levels. Two groups were not divided further than the second level, because they contained less than five sites when taken to level three, this resulted in six groups.

Discriminant analysis (SAS, 1996) was used to determine how well environmental variables account for the structure of biological groupings. Variables were excluded from consideration if they were likely to be influenced by human activity (e.g. nutrient enrichment, catchment or riparian vegetation). This reduced the number of environmental variables for use in model construction from 45 to 19 (shown in bold, Table 1). The 19 variables were further reduced by selecting only those that contributed significantly ($P < 0.05$) to the separation of groups using stepwise discriminant analysis (SAS, 1996). Final selection of the optimal number of environmental variables was carried out by iteration to minimise error rates (variables with asterisk in Table 1). The optimal selection was the combination of environmental variables that produced the lowest overall error rate using crossvalidation (Parsons and Norris, 1996). Linear discriminant functions are calculated on an assumption of equal within-group variances. To help satisfy this assumption and improve the discrimination the environmental variables were log transformed [$\ln(x + 1)$].

After optimisation of the discriminant functions, which allocate sites to the predetermined biological classification using their environmental characteristics, the next step was to predict the fish communities expected at a test site. To predict the assemblage expected at a site, the frequency with which individual taxa occur in each TWINSPAN group (i.e. the relative group frequency) was calculated as the number of sites where that taxon occurs divided by the total number of sites in the group. This is referred to as the probability of finding that taxon in that group. The overall probability of finding a taxon at a site is the relative group frequency weighted by the probability of membership in each of the six groups (see Wright, 1995).

The final step in site assessment is to compare observed and expected faunas. The predicted fauna is compared with the observed taxa list following the procedure originally described by Wright *et al.* (1984). The probabilities of the predicted taxa are summed to give the 'expected number of taxa' (E). The number of species actually captured at a site, providing they were predicted to occur is the 'observed number of taxa' (O). The ratio of the observed to the expected number of taxa (*O/E*) and taxonomic composition is the output from the model (see Moss *et al.*, 1987).

The number of taxa observed at reference sites were compared with model predictions generating a distribution of reference site *O/E* ratios. The distribution of *O/E* ratios was assessed for each group separately to evaluate any group specific bias. Low *O/E* ratios are used to indicate sites under stress, while high ratios indicate sites with more species than expected, which may indicate sites of high conservation value (Wright, 1995). Determination of whether a site is impacted is judged based on the site's *O/E* ratio compared with the distribution of *O/E* ratios for the reference sites.

Model validation

While crossvalidation provides a robust measure of the predictive ability of the discriminant analysis within the data set, two further tests were applied to validate the model using independent data. First, 30 sites were randomly removed from the data set and the remaining 112 sites used to construct a new discriminant model. The 30 sites were then run through the discriminant model to assess the ability of the model to assign the sites to their predetermined groups, as judged by the posterior crossvalidation error rate (Reynoldson and Rosenberg, 1996; Fielding and Bell, 1997; Hawkins *et al.*, 2000).

Potentially impacted test sites were sampled in December 2000 employing the same sampling and data collection methods used for the reference sites. The test sites were selected to cover a wide range of potential impacts including nutrient enrichment, sedimentation and urbanisation; one site had a natural stressor: volcanic acidification.

The second validation involved running these 30 potentially impacted test sites through the model to calculate the expected assemblage at each site. This test was carried out to determine the discriminating power of the model to detect a range of stresses known to occur in the region. Analysis of variance was used to determine if test site *O/E* values differed significantly from those of the reference sites.

Validation of reference sites using macroinvertebrates

As an independent validation of reference status, the macroinvertebrate community index (MCI; Stark, 1993) was calculated from the invertebrate sample taken at each site. The MCI scores are based on tolerance values to organic enrichment allocated to macroinvertebrate taxa similar to the U.K. BMWP Score System (Armitage *et al.*, 1983). MCI values respond to many environmental factors, including, but not confined to, water quality (Boothroyd and Stark, 2000).

MCI values are calculated as:

$$MCI = 20 \times \sum_{i=1}^{i=S} a_i / S$$

where S is the number of taxa in the sample and *ai* is the score of the *i*th taxon. Scores greater than 120 are considered 'pristine' and scores less than 80 are considered 'severely polluted' (Stark, 1993). A minimum MCI score criteria was used to determine reference site retention. However, the minimum criterion was relaxed for sites at lower altitudes, because the index has limitations when used in streams with fine grain substrates often found at low altitudes (Death, McWilliam and Rodway, 1999). Thus, at altitudes below 100 m a.s.l sites with a score less than 100 were rejected, at 100–200 m a.s.l. sites with a score less than 110 were rejected, and above 200 m a.s.l. a score of greater than 120 was the criterion for inclusion.

RESULTS

Species composition

Nineteen species of fish, 16 native and three introduced, were collected. Four of these species were found at only three sites (1.5% of sites) or less, and therefore they were excluded from further analyses. The longfin eel (*Anguilla dieffenbachii* Gray) was most widely distributed, being found at 80% of the sites, or almost twice as many sites as the next most widely distributed species. Trout and the two non-migratory bully species, Cran's and upland bully, occurred at 43 and 30% of the sites, respectively. The migratory redfin bully (*Gobiomorphus. huttoni* Ogilby) was found at 19% of the sites. Elvers, shortfin eels (*Anguilla. australis* Richardson), inanga (*Galaxias maculatus* Jenyns), torrentfish (*Cheimarrichthys fosteri* Haast) and common bully (*G. cotidianus* McDowall) occurred at 10–15% of the sites, while shortjaw kokopu (*G. postvectis* Clarke), koaro (*G. brevipinnis* Günther), dwarf galaxiids (*G. divergens* Stokell) and common smelt (*Retropinna retropinna* Richardson) occurred at 5.5, 5, 3 and 2% of the sites, respectively. At around 2000 individuals, the non-migratory bullies and the longfin eel were at least twice as abundant as the other species. However, Cran's bully and torrentfish were the most abundant when only sites at which the taxon occurred are considered.

Classification and discriminant analysis

Validation of reference site status using invertebrate community MCI scores and updated information obtained in the field resulted in 142 sites being retained as reference sites. Removal of rare species resulted in 13 taxonomic units used for model construction. From this data set, six main groups were identified using TWINSPAN analysis. The six groups contained different assemblages, densities and the number of sites (Table 2).

Of the 26 environmental variables considered least likely to be influenced by human impacts, 14 were subsequently selected for use in the discriminant model after iteration (variables with asterisks Table 1). Of the 142 reference sites, 100 (70%) were assigned correctly to their predetermined biological groups by the discriminant analysis using crossvalidation. However, of the 42 sites misclassified by the model using crossvalidation, 28 (66%) were assigned to the correct group with the next highest probability.

Relationships between fish assemblages and physical/chemical data

Examination of the standardised canonical coefficients revealed the relative contributions of the environmental variables to the separation of the site groups (Table 3). The first two canonical variates (functions) accounted for 73 and 11% of the variation, and the following three variates combined explained the rest of the variation. Canonical variate 1 revealed the strong negative influence of longitudinal co-ordinates and ecoregion and positive influence of altitude and distance from the sea. The second variate revealed the opposing influences of altitude and distance from the sea, and the positive influence of stream depth. To visualise the separation of groups in two dimensions the centroids (mean co-ordinates) for the six groups are shown in discriminant space (Fig. 2). In general, altitude and distance inland increase from left to right along axis 1, while longitudinal co-ordinates and the identification number arbitrarily allocated to the ecoregions decrease. On canonical variate axis 2, distance from the sea and mean depth were positively correlated, while altitude and ecoregion number were negatively correlated with this axis. The coefficients indicate that all of the variables had some influence in separating site groupings (Table 3).

Table 3. Pooled Within-Class Standardized Canonical Coefficients for the 14 environmental variables used in the discriminant model.

Variable	Can 1 (73.6%)	Can 2 (11.7%)
Altitude	0.797	-0.906
Easting	-0.834	-0.515
Ecoregion	-0.826	-0.814
Northing	0.046	-0.179
Distance inland	0.662	1.101
Mean width	0.168	-0.347
Mean depth	-0.161	0.613
Median substrate size	0.242	0.430
%Pool	-0.179	0.453
%Run	-0.145	0.493
%Rapid	-0.079	-0.130
%Temperature	-0.142	0.293
%Backwater	-0.022	-0.224
Embededness	-0.002	-0.127

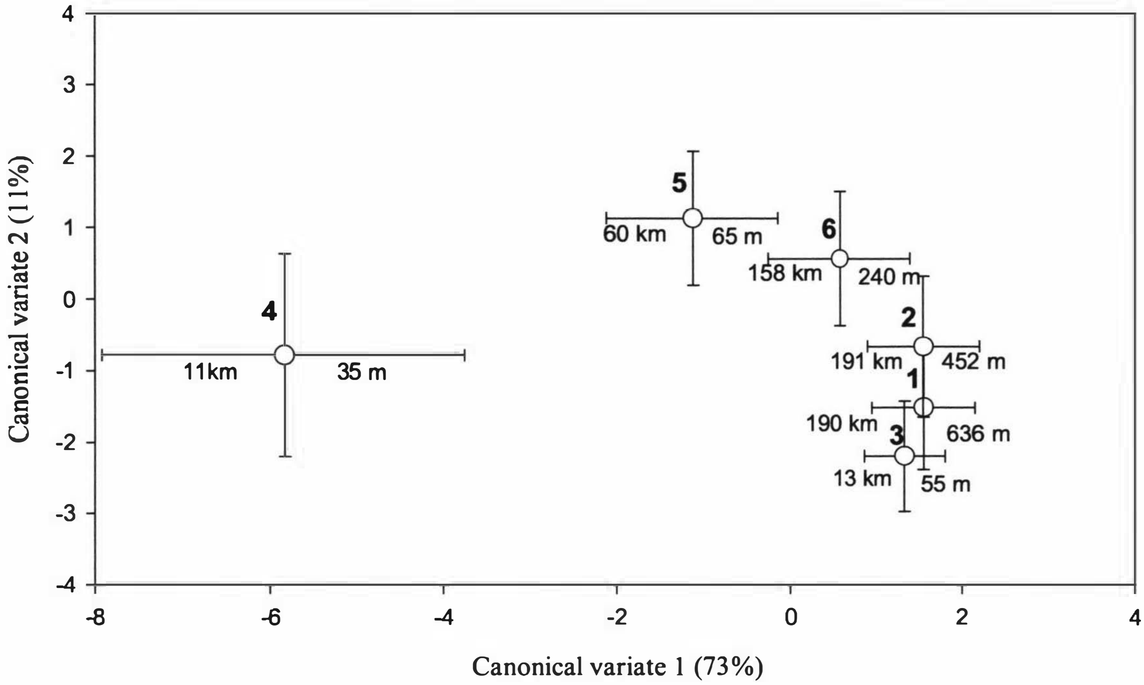

Fig. 2. Position of the mean co-ordinates of the six TWINSPAN groups in discriminant space obtained by linear discriminant analysis on 14 environmental variables unlikely to be influenced by human impacts (marked with asterisk in Table 1). (Number above centroid is group number, number on the left of group centroid is the mean distance from the sea and number on right is the mean elevation, error bars denote standard deviation of site co-ordinates).

Table 4. Example of the method used for calculating observed over expected ratios for site no. 46 in the Mangahao River. Species with an asterisk were captured at the site (modified from Moss *et al.*, 1987; Wright, 1995).

Taxa	Probability of capture percentage
*Anguilla dieffenbachii**	90.66
Gobiomorphus spp. (non-migratory)*	58.82
Salmo or Oncorhynchus spp.*	43.35
Gobiomorphus huttoni	35.23
Cheimarrichthys fosteri	26.88
Anguilla spp.	22.48
Anguilla australis	11.44
Gobiomorphus cotidianus	9.63
Galaxias postvectis	7.09
Galaxias maculatus	6.62
Galaxias brevipinnis	4.37
Retropinna retropinna	3.76
Galaxias divergens	1.46
Number of taxa captured	3.0
Number of taxa expected (sum of probabilities)	3.21
Observed over expected ratio	0.93

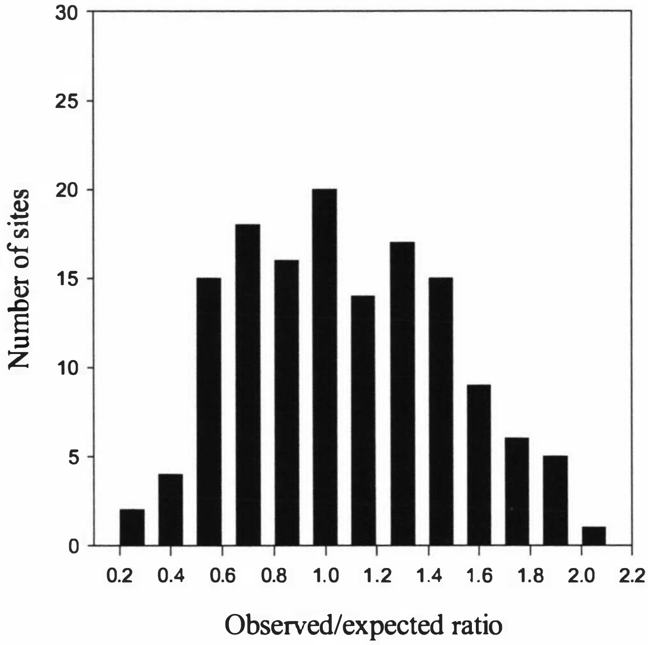

Fig. 3 Spread of observed over expected ratios (mean of two cut-off thresholds) for the 142 reference sites.

The *O/E* ratios of the predicted and observed fish faunas were calculated using all probabilities >0 (an example is given in Table 4). The expected number of taxa was obtained by summing the probability values for each of the taxa from the ranked list of weighted probabilities. This process was repeated for all reference sites, to give the distribution of *O/E* ratios (Fig. 3). The mean *O/E*-value for all reference sites was close to unity (0.99), which suggested that the model produced unbiased estimates of the number of taxa expected to occur at a site. The mean *O/E* ratios were also close to unity for each of the TWINSPAN groups, indicating that all groups produced relatively unbiased estimates of taxon richness (Fig. 4).

Model validation

The first validation procedure involved randomly removing 30 reference sites. When discriminant analysis was applied to the remaining 112 sites, 77 (69%) were correctly assigned to their original groups using crossvalidation. Of the 30 removed validation sites, 20 (67%) were correctly assigned to their group. Three of the remaining sites were placed in the next most probable group. These results suggest reasonable confidence in sites being accurately assigned to the correct group using this discriminant model.

Differences among reference and test site O/E-values

A range of potential stresses were evident at the 30 test sites, most were related to pastoral farming with removal of riparian vegetation, stock access to streams and diffuse nutrient run-off (Table 5). One site was below farming and industrial land use, two had town effluent discharges upstream. Two sites were in exotic forestry areas and two sites were upstream of small impoundments. These test sites were run through the model as described above and the *O/E* ratios calculated. The mean *O/E*-value for the test sites was 0.68 (SE = 0.06) and was significantly less than the mean *O/E*-value (mean = 0.99, SE = 0.038) for the reference sites ($F_{1,170} = 12.5$; P = 0.000).

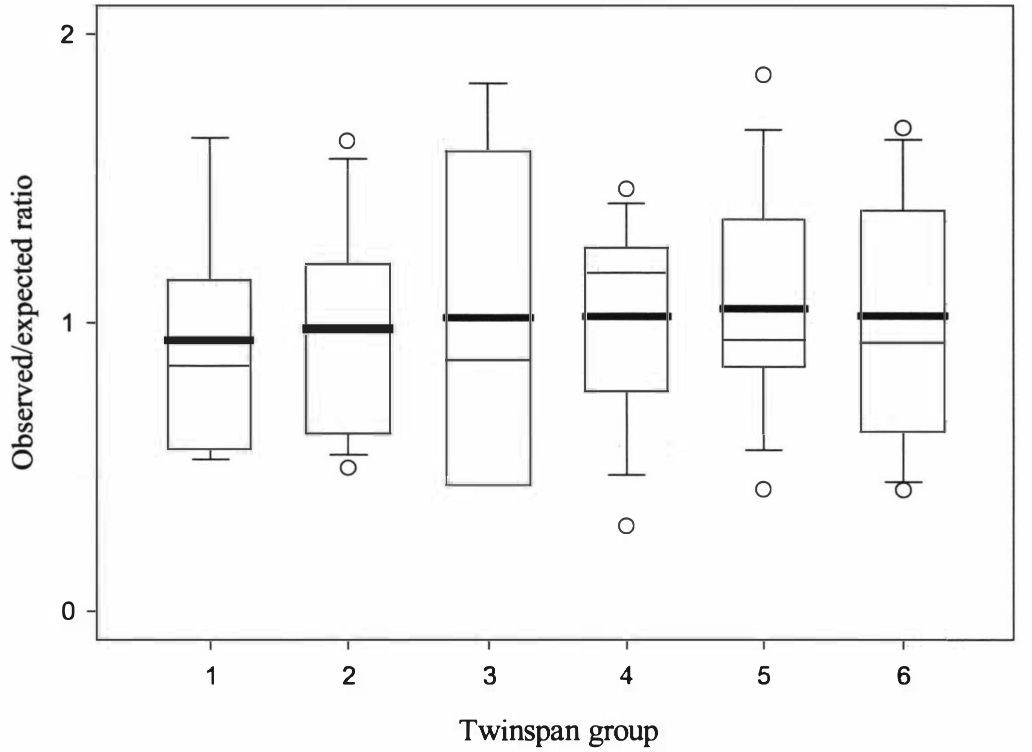

Fig. 4 Box plots showing the distribution of *O/E* values for each of the TWINSPAN groups (box contains interquartile range, circles denote 5^{th} and 95^{th} percentiles, heavy centerline denotes mean, fine centerline is median, and whiskers extend to 1.5 * interquartile range).

Table 5. The relationship between the observed and predicted fauna for potentially stressed sites in the Manawatu–Wanganui region surveyed during December 2000. See text for calculations of observed and expected taxa.

Site no.	Waterway name	Landuse/potential stressors	No. taxa observed	No. taxa expected	*O/E* ratio
T30	Koputaroa Stream	Farming/runoff	7	3.93	1.783
T17	Mangapapa Stream	Farming/runoff	4	3.03	1.319
T25	Returuke River	Farming/runoff	3	2.95	1.018
T16	Mangatera Stream	Farming/runoff	3	3.00	1.001
T8	Matarua Creek	Farming/runoff	3	3.12	0.961
T11	Ponui Stream	Farming/runoff	3	3.52	0.853
T4	Tutanui Stream	Farming/runoff/Town effluent	3	3.76	0.799
T26	Mangartorua Stream	Forestry	2	2.57	0.780
T20	Matarawa Stream	Farming/runoff	3	3.87	0.774
T1	Mangaone Stream	Farming/runoff/Urban	3	3.90	0.769
T15	Akitio River	Farming/runoff	3	4.14	0.725
T7	Lake Horowhenua inflow	Market gardens/orchard	3	4.16	0.722
T28	Hautapu River	Farming/runoff/Town effluent	2	2.80	0.714
T23	Mangaroa Stream	Farming/runoff	2	2.88	0.695
T24	Hikumutu Stream	Farming/runoff	2	2.91	0.687
T29	Porewa Stream	Farming/runoff	2	2.92	0.684
T13	Waihi Stream	Dam/farming	3	4.40	0.681
T14	Table Stream	Farming/runoff	3	4.46	0.672
T5	Manaone Stream	Farming/runoff	2	3.41	0.586
T18	Makirikiri Stream	Farming/runoff	2	3.62	0.552
T3	Makowhai Stream	Farming/runoff	2	3.92	0.510
T2	Oroua River Tributary	Farming/runoff	2	3.97	0.504
T12	Makukupara Stream	Farming/runoff	2	4.45	0.450
T6	Waiwiri Stream	Farming/runoff	2	4.45	0.449
T19	Brunswick Stream	Industrial/farming	2	4.47	0.448
T27	Turakina River	Dam/farming	1	2.66	0.377
T22	Ongarue River	Forestry/sheep	1	2.88	0.347
T9	Makakahi River	Farming/runoff	1	3.48	0.287
T10	Mangaramarama Stream	Farming/runoff	1	3.83	0.261
T21	Whangehu River	Industrial waste/volcanic	0	2.32	0.0

Discussion

The objective of this study was to describe the potential utility of a RIVPACS type approach to biomonitoring using freshwater fish communities in New Zealand. In general this predictive model approach allows for improvement in the precision of bioassessment by explaining as much variation in the reference communities as possible prior to the assessment of test sites. In countries such as New Zealand with a mainly diadromous fish fauna it offers the ability to use fish in biological assessment where this was not previously possible due to the confounding influence of migration on between site comparisons. Furthermore, potentially subjective decisions made when applying scores to species attributes using indices are avoided. When a predictive model approach is taken using invertebrate communities an assessment is made of stream conditions in the vicinity of the site and upstream. In contrast the approach described here takes into account the condition of the entire waterway because the predominantly diadromous New Zealand fish fauna is reliant on access both up and downstream. Furthermore, because these models predict the actual taxonomic composition of a site, they also provide information about the presence or absence of specific taxa. This opens up the potential for the association of the presence or absence of particular taxa with specific stresses/stressors where these can be isolated and tested.

Few of the potentially impacted test sites had more species than expected and there was no obvious relationship between ratios and potential impacts. However, the stressors were generally found in combination rather than singularly thus any association between the stress and result is confounded. To further assess the model the response of the *O/E* ratios and the responses of individual taxa to individual stressors will need to be considered particularly where the stressors can be isolated. This further analysis will potentially allow for diagnostic applications of the model.

The relatively low classification error of the model suggests a strong relationship between the fish communities and the environmental variables used in the model. Most of the physical variables used were obtained from maps and GIS, and thus tend to change little over long time scales, making them ideal for predicting the biotic assemblages expected in the absence of human impacts. The canonical coefficients revealed that altitude and distance from the sea were among the most important of the 14 variables used to discriminate between groups. This is similar to results obtained in a number of other New Zealand studies (Hayes *et al.*, 1989; McDowall, 1990; Jowett and Richardson, 1996; Joy and Death, 2000). Thus, the most important factors associated with fish community structure are the catchment-scale variables: altitude, distance from the coast, ecoregion and map co-ordinates, while local-scale variables such as substrate and flow type appear to be of less importance, at least at the scale used in this study.

The influence of sampling strategy is reduced when the results are employed in RIVPACS type models. The critical factor is that the same sampling method employed at the reference sites used to build the model is used when assessing a new site. Any sampling bias related to a particular survey method would be the same for both reference and test sites. Furthermore, the predictions from this model are intended to represent the natural long-term characteristics of the communities and are not intended to predict seasonal variations in the fauna. Therefore, because references site sampling occurred during summer, sampling sites for assessment should also be in summer.

Many studies have demonstrated the ability to accurately predict the structure of invertebrate communities from a set of environmental variables (Furse *et al.*, 1984; Wright *et al.*, 1984; Corkum and Currie, 1987; Moss *et al.*, 1987; Ormerod and Edwards, 1987; Wright *et al.*, 1993; Reynoldson *et al.*, 1995; Hawkins *et al.*, 2000). In this study, 70% of the sites were correctly placed into the appropriate reference classification groups using discriminant analysis. This results in a mean classification error of c. 30%, which is the same error rate associated with AUSRIVAS models (Simpson and Norris, 2000), but is lower than the mean misclassification error rate of 44% produced by RIVPACS II models (Wright, 1991). The comparison of model error rates should, however, be carried out with caution because error rates are relative to the number of groups. However, in RIVPACS approaches, model error is somewhat misleading, because the calculation of the assemblage for a new site incorporates the probability of the site belonging to all of the reference site groups (Simpson and Norris, 2000). This effectively removes the artificially imposed group boundaries and converts the predictions back to a more ecologically realistic continuum of assemblages. Therefore, a site may appear to be misclassified, yet still have adequate community predictions (Reynoldson *et al.*, 1997).

A potential problem identified with the application of the RIVPACS approach to fish assemblages is the low number of fish species compared with invertebrates (J. F. Wright personal communication). Despite a low number of species, meaningful results were obtained, although the range of *O/E* ratios was wider than that for most published invertebrate models. In this paper we chose to include all taxa with probabilities >0, as was used in the RIVPACS models and a similar invertebrate model in North America (Hawkins *et al.*, 2000). In contrast to AUSRIVAS models, where a probability cut-off threshold of 0.75 is used. A number of probability thresholds were examined and shown to have little impact on the ranking of reference and test sites, although *O/E* ratios became unstable at probabilities >0.5 at sites with low species number.

Understanding the response of the entire aquatic community, including all flora and fauna, to human impacts is the best way to assess ecosystem health (Karr, 1991; Metcalfe-Smith, 1996). However, this is generally not economically practical and usually particular biotic elements are chosen for intensive study. In New Zealand, biological monitoring has generally only focused on macroinvertebrates, predominantly using the MCI (Boothroyd and Stark, 2000). Analysis of results from this study revealed a weak relationship between the fish *O/E* ratios and invertebrate MCI scores (Pearson correlation coefficient 0.10; P = 0.14). This lack of association between fish and invertebrate communities can be attributed to a number of factors, including differences in sampling scales, disturbance susceptibility and the influence of migratory access on fish communities. For the most ecologically meaningful results, when considering the biotic integrity of a site as many of the biotic components as possible should be considered. Thus combining two aspects of biotic integrity using both fish and invertebrates at a site will enhance assessment quality.

This project was undertaken to provide a practical tool to enable an assessment of the biological condition of flowing-water sites in New Zealand using freshwater fish, although it has obvious potential for use in other countries. It allows for an objective, statistically robust, site-specific prediction of the fish assemblage to be made without the requirement for extensive analysis. This is because the sophisticated statistical knowledge used to create the model is not required for its implementation. We believe managers and policy makers, even those without a biological or statistical background, should intuitively and easily understand the output from the model (i.e. the comparison with what should be at a site in the absence of stress with that actually found at a test site).

Acknowledgments

We thank Grant Blackwell and Rosemary Miller for assistance with fieldwork, and John Phillips for assistance with site selection and chemical analyses. This research was funded by horizons.mw and the Sustainable Management Fund (Ministry for the Environment) grant No. 5099 ("Nga Ika Waiora": stream health evaluation tool). We thank Kurt Fausch, Richard Johnson and two anonymous referees whose comments greatly improved this manuscript.

References

Allan, J. D., and Flecker, A. S. (1993) Biodiversity conservation in running waters. *Bioscience*, **43**, 32-43.

Anonymous (1997) *The State of New Zealand's Environment*. Ministry for the Environment, GP Publications Wellington.

ArcView. 1999. Getting to know Arc View GIS. Environmental Systems Research Institute, Inc., Redlands California.

Armitage, P. D., D. Moss, J. F. Wright, and M. T. Furse. 1983. The Performance of a New Biological Water-Quality Score System Based on Macroinvertebrates over a Wide-Range of Unpolluted Running-Water Sites. Water Research **17**:333-347.

Berkman H. E., Rabeni C. F. and Boyle T. P. (1986) Biomonitors of stream quality in agricultural areas: fish versus invertebrates. *Environmental Management*, **10,** 413-419.

Boothroyd I. K. G. and Stark J. D. (2000) Use of invertebrates in monitoring. In: *New Zealand stream invertebrates: Ecology and implications for management* (eds. K. J. Collier, and M. J. Winterbourn), pp. 415, Caxton Press, Christchurch.

Clarke R. T., Furse M. T., Wright J. F. and Moss D. (1996) Derivation of a biological quality index for river sites: comparison of the observed with expected fauna. *Journal of Applied Statistics*, **23,** 311-322.

Corkum L. D. and Currie D. C. (1987) Distributional patterns of immature Simuliidae (Diptera) in Northwestern North America. *Freshwater Biology*, **17,** 201-221.

Davies A. L. and Gee J. H. R. (1993) A simple periphyton sampler for algal biomass estimates in streams. *Freshwater Biology*, **30,** 47-51.

Death R. G., McWilliam H. and Rodway M. (1999) Macroinvertebrate monitoring: Lowland streams and other non-shingle river ecotypes. In: *The use of macroinvertebrates in water management*. (ed. M. J. Winterbourn), pp. 33-41, Ministry for the Environment, Wellington.

Fausch K. D., Lyons J., Karr J. R. and Angermeier P. L. (1990) Fish communities as indicators of environmental degradation. In: *Biological Indicators of Stress in Fish*, pp. 123-144, Bethesda, MD, USA.

Fielding A. H. and Bell J. F. (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, **24,** 38-49.

Furse M. T., Moss D., Wright J. F. and Armitage P. D. (1984) The influence of seasonal and taxonomic factors on the ordination and classification of running-water sites in Great Britain and on the prediction of their macro-invertebrate communities. *Freshwater Biology*, **14,** 257-280.

Gauch H. G. and Whittaker R. H. (1981) Hierarchical classification of community data. *Journal of Ecology*, **69,** 537-557.

Glova G. J. (1998) Factors associated with the distribution and habitat of eels (*Anguilla* spp.) in three New Zealand lowland streams. *New Zealand Journal of Marine and Freshwater Research*, **32,** 255-269.

Harding J. S. (1994) Lotic Ecoregions of New Zealand. PhD. Thesis *Zoology dept.* Canterbury University, Christchurch New Zealand.

Harris J. N. (1995) The use of fish in ecological assessments. *Australian Journal of Ecology*, **20,** 65-80.

Hawkins C. P., Norris R. H., Hogue J. N. and Feminella J. W. (2000) Development and evaluation of predictive models for measuring the biological integrity of streams. *Ecological Applications*, **10,** 1456-1477.

Hayes J. W., Leathwick J. R. and Hatchet S. M. (1989) Fish distribution patterns and their association with environmental factors in the Mokau River catchment, New Zealand. *New Zealand Journal of Marine and Freshwater Research*, **23,** 171-180.

Hill M. O. (1979) *DECORANA a FORTRAN program for detrended correspondence analysis and reciprocal averaging.* Ecology and Systematics, Cornell University. Ithaca, New York

Hughes, R. N. 1995. Defining acceptable biological status by comparing with reference conditions. Pages 31-47 *in* W. S. Davis and T. P. Simon., editors. Biological Assessment and Criteria: tools for water resource planning and decision making. Lewis publishers, Boca Raton.

Hughes R. M., Larsen D. P. and Omernik J. M. (1986) Regional reference sites: a method for assessing stream potentials. *Environmental Management*, **10,** 629-635.

Hynes H. B. N. (1970) *The Ecology of Running Waters.* pp. 555 University of Toronto Press, Toronto.

Jowett I. G. and Richardson J. (1995) Habitat preferences of common, riverine New Zealand native fishes and implications for flow management. *New Zealand Journal of Marine and Freshwater Research*, **29,** 13-23.

Jowett I. G. and Richardson J. (1996) Distribution and abundance of freshwater fish in New Zealand rivers. *New Zealand Journal of Marine and Freshwater Research*, **30,** 239-255.

Joy, M. K., and R. G. Death. 2000. Development and application of a predictive model of riverine fish community assemblages in the Taranaki region of the North Island, New Zealand. New Zealand Journal of Marine and Freshwater Research **34**:243-254.

Joy M. K. and Death R. G. (2001) Control of freshwater fish and crayfish community structure in Taranaki, New Zealand: dams, diadromy or habitat quality. *Freshwater Biology*, **46,** 417-429.

Joy M. K., Henderson I. M. and Death R. G. (2000) Diadromy and patterns of upstream penetration of freshwater fish in Taranaki New Zealand. *New Zealand Journal of Marine and Freshwater Research*, **35,** 531-543.

Karr J. R. (1981) Assessments of biotic integrity using fish communities. *Fisheries (Bethesda)*, **6,** 21-27.

Karr J. R. (1991) Biological integrity: a long neglected aspect of water resource management. *Ecological Applications*, **1,** 66-84.

Karr J. R., Fausch K. D., Angermeier P. L., Yant P. R. and Schlosser I. J. (1986) Assessing biological integrity in running waters: A method and its rationale. *Illinois Natural History Survey Special Publication*, **5,** 28 pp.

Marchant R. (1989) A sub-sampler for samples of benthic invertebrates. *Bulletin of the Australian Society of Limnology*, **19,** 49-52.

Matson P. A., Parton W. J., Power A. G. and Swift M. J. (1997) Agricultural intensification and ecosystem properties. *Science*, **277,** 504-508.

McCormick F. H., Peck D. V. and Larsen D. P. (2000) Comparison of geographic classification schemes for mid-Atlantic stream fish assemblages. *Journal of the North American Benthological Society*, **19,** 385-404.

McCune B. and Mefford M. J. (1997) *Multivariate analysis of ecological data*, Version. 3.01 MJM Software. Gleneden Beach Oregon USA

McDowall R. M. (1990) *New Zealand Freshwater Fishes: A Natural History and Guide.* pp. 553. Heinemann Reed, Auckland.

McDowall R. M. (1998a) Fighting the flow: downstream-upstream linkages in the ecology of diadromous fish faunas in West Coast New Zealand rivers. *Freshwater Biology*, **40,** 111-122.

McDowall R. M. (1998b) Driven by diadromy: its role in the historical and ecological biogeography of the New Zealand freshwater fish fauna. *Italian Journal of Zoology*, **65,** 73-85.

McDowall R. M. and Taylor M. J. (2000) Environmental indicators of habitat quality in a migratory freshwater fish fauna. *Environmental Management*, **25,** 357-374.

Metcalfe-Smith J. L. (1996) Biological water-quality assessments of rivers: use of macroinvertebrate communities. In: *River Restoration* (eds. G. E. Petts, and P. Calow), pp. 17-43, Blackwell, Oxford.

Moss D. (2000) Evolution of statistical methods in RIVPACS. In: *Assessing the Biological Quality of Freshwaters. RIVPACS and other techniques.* (eds. J. F. Wright, D. W. Sutcliffe, and M. T. Furse), pp. 25-37, Freshwater Biological Association, Ambleside, UK.

Moss D., Furse M. T., Wright J. F. and Armitage P. D. (1987) The prediction of the macro-invertebrate fauna of unpolluted running-water sites in Great Britain using environmental data. *Freshwater Biology*, **17,** 41-52.

Moss D., Wright J. F., Furse M. T. and Clarke R. T. (1999) A comparison of alternative techniques for prediction of running-water sites in Great Britain. *Freshwater Biology*, **41,** 167-181.

New Zealand Survey and Land Information (NZSLI). (1987) 1:50 000 *Topographical Map Series. NZMS 260* NZSLI,: Wellington.

Ormerod S. J. and Edwards R. W. (1987) The ordination and classification of macroinvertebrate assemblages in the catchment of the River Wye in relation to environmental factors. *Freshwater Biology*, **17,** 533-546.

Parsons M. and Norris R. H. (1996) The effect of habitat-specific sampling on biological assessment of water quality using a predictive model. *Freshwater Biology*, **36,** 419-434.

Pfankuch D. J. (1975a) Stream reach inventory and channel stability evaluation. *United States Department of Agriculture Forest service, Region 1, Missoula Montana, USA.*

Postel S. L. (2000) Entering an era of water scarcity: the challenges ahead. *Ecological Applications*, **10,** 941-948.

Quinn J. M. and Hickey C. W. (1990) Magnitude of effects of substrate particle size, recent flooding, and catchment development on benthic invertebrates in 88 New Zealand rivers. *New Zealand Journal of Marine and Freshwater Research*, **24,** 411-427.

Reynoldson T. B. and Rosenberg D. M. (1996) Sampling strategies and practical considerations in building reference databases for the prediction of invertebrate community structure. In: *Study design and data analysis in benthic macroinvertebrate assessments of freshwater ecosystems using a reference site approach.* (eds. R. C. Bailey, R. H. Norris, and T. B. Reynoldson), pp. 1-31.

Reynoldson T. B., Bailey R. C., Day K. E. and Norris R. H. (1995) Biological guidelines for freshwater sediment based on BEnthic Assessment of SedimenT (the BEAST) using a multivariate approach for predicting biological state. *Australian Journal of Ecology*, **20,** 198-219.

SAS (1996) *SAS User's Guide: Statistics, Version 6.12.*, SAS Institute Inc. Cary, North Carolina

Simpson J. and Norris R. H. (2000) Biological assessment of water quality: development of AUSRIVAS models and outputs. In: *Assessing the biological quality of freshwaters. RIVPACS and other techniques.* (eds. J. F. Wright, D. W. Sutcliffe, and M. T. Furse), pp. 125-142, Freshwater Biological Association, Ambleside, UK.

Stark J. D. (1993) A macroinvertebrate community index of water quality for stony streams. *New Zealand Journal of Marine and Freshwater Research*, **27,** 463-478.

Steedman R. J. (1988) Modification and assessment of an index of biotic integrity to quantify stream quality in southern Ontario. *Canadian Journal of Fisheries and Aquatic Sciences*, **45,** 492-501.

Steinman A. D. and Lamberti G. A. (1996) *Methods in stream ecology*. Academic Press, San Diego.

Townsend C. R., Hildrew A. G. and Francis J. (1983) Community structure in some southern English streams: the influence of physicochemical factors. *Freshwater Biology*, **13,** 521-544.

Werner E. E. and Gilliam J. F. (1984) The ontogenetic niche and species interactions in size-structured populations. *Annual review of ecological systematics*, **15,** 393-425.

Winterbourn M. J. and Gregson K. L. D. (1989) Guide to the Aquatic insects of New Zealand. *Bulletin of the Entomological Society of New Zealand*, **9**.

Wolman M. G. (1954) A method of sampling coarse river-bed material. *Transactions, American Geophysical Union*, **35,** 951-956.

Wright J. F. (1991) Testing and further development of RIVPACS; interim report to the National Rivers Authority, pp. 141, National Rivers Authority, Bristol

Wright J. F. (1995) Development and use of a system for predicting the macroinvertebrate fauna in flowing waters. *Australian Journal of Ecology*, **20,** 181-197.

Wright J. F., Furse M. T. and Armitage P. D. (1993) RIVIPACS- a technique for evaluating the biological quality of rivers in the UK. *European Water Pollution Control*, **3,** 15-25.

Wright J. F., Moss D., Armitage P. D. and Furse M. T. (1984) A preliminary classification of running water-sites in Great Britain based on macro-invertebrate species and the prediction of community type using environmental data. *Freshwater Biology*, **14,** 221-256.

CHAPTER TWO

Biological assessment of rivers in the Manawatu-Wanganui region of New Zealand using a predictive macroinvertebrate model

Joy, M. K., and R. G. Death. 2003. Biological assessment of rivers in the Manawatu-Wanganui region of New Zealand using a predictive macroinvertebrate model. New Zealand Journal of Marine and Freshwater Research **37**:367-379.

Abstract

This study presents a river invertebrate and classification system (RIVPACS) type bioassessment methodology for the Manawatu-Wanganui region of New Zealand. Aquatic macroinvertebrates and related physico-chemical data were collected at 127 sites, with minimal human impacts (reference sites) in 2000. The reference sites were classified into five groups based on their macroinvertebrate data using TWINSPAN. These biotic groupings were then applied to their corresponding physico-chemical data and discriminant functions were obtained to assign sites into the biotic groups using the physico-chemical data. The discriminant functions correctly allocated 72% of the sites to the correct classification group using a jack-knife validation. The probabilities from the discriminant functions were used to predict macroinvertebrate assemblages and these were compared with observed macroinvertebrate assemblages. The model was then used to assess the health of 29 test sites with known impacts. All test sites were assessed as impacted based on the 10th percentile of the reference data. To evaluate the temporal reliability of the model, data available for 11 sites sampled in 1997 and 2000 were run through the model. The results of this comparison showed little variation in *O/E* ratios over time and the two sites classed as impacted in 1997 were also classed as impacted in 2000.

Keywords: biomonitoring; macroinvertebrates; predictive models; RIVPACS

INTRODUCTION

Water management in New Zealand has had its emphasis broadened from the management of water quality to a more holistic view of aquatic ecosystems with the advent of the Resource Management Act 1991 (RMA) (Winterbourn 1999). Consequently, the assessment of aquatic system health must also progress to an ecosystem level rather than a water quality perspective. Macroinvertebrates are currently used in the bioassessment of lotic systems by the majority of regulatory authorities in New Zealand using a single index, the Macroinvertebrate Community Index (MCI) and its derivatives (Stark 1993; Winterbourn 1999). This is despite evidence that the applicability of the MCI outside the region or stream type for which it was developed is unclear (Winterbourn 1999). However, a predictive modeling approach to bioassessment has been proposed to have a number of potential advantages over a single index approach especially where the focus is on overall ecosystem health rather than simply water quality. This is because the predictive modeling approach combines information on environmental variables and macroinvertebrate assemblages in a predictive format (Winterbourn 1999). Therefore, the application of a predictive modeling approach to bioassessment is crucial to the improvement of aquatic ecosystem management in New Zealand.

A multivariate predictive assessment approach to biomonitoring is well established in a number of countries (Wright 2000). Initially developed in the United Kingdom in the early 1980s by Wright and coworkers, the predictive model approach relates lotic macroinvertebrate composition to environmental descriptors (Furse *et al.* 1984; Wright *et al.* 1993; Wright 2000). The approach has been further developed and models constructed for lotic and lentic systems in Canada, Australia, Indonesia, and the United States (e.g., Reynoldson *et al.* 1995; Marchant *et al.* 1997; Hawkins *et al.* 2000; Sudaryanti *et al.* 2001). The output from these models is the ratio of the number of taxa observed at a site to that expected (*O/E* ratio) but only if the specific taxa predicted were observed and is used as a measure of biological impairment. If the observed to expected ratio is low, the implication is that the site is adversely affected by some environmental stress. This approach avoids the allocation of scores to taxa based on their response to a single environmental gradient, as is the case with index approaches such as the MCI. Impairment is simply assessed by the absence of taxa that would normally be expected to be present. However, biotic indices can be included in predictive model outputs by comparing observed and expected index scores.

The approaches described above are multivariate predictive reference site models and are based on the acquisition of an array of reference sites that characterise the biological conditions of the region for which assessments will be made. The criterion used for reference site status is that the sites should be minimally affected by human activities. In reality however, these sites are seldom pristine but they represent the least impaired conditions within the area of interest. The macroinvertebrate assemblages at these reference sites then provide an empirical foundation against which other sites can be compared (Bailey *et al.* 1998; Reynoldson and Wright 2000). The biotic predictions are made from a suite of environmental features unlikely to be influenced by human activity (e.g., latitude, elevation, and distance from the coast). Despite the globally widespread application of this bioassessment methodology over the last 20 years the predictive modeling approach has not been applied in New Zealand except with fish and macro-crustaceans (Joy and Death 2000; Joy and Death 2002).

The primary aim of this study was to assess the feasibility and applicability of a multivariate predictive approach for regional lotic bioassessment in New Zealand using macroinvertebrates. To achieve this we took a predictive reference site approach to bioassessment in the Manawatu-Wanganui region of the North Island of New Zealand. Following approximately the RIVPACS/AUSRIVAS process (Wright 1995; Simpson and Norris 2000) we constructed multivariate predictive models based on data collected at sites minimally disturbed by human activities. To assess the ability of the model to detect biological impairment we collected data from 29 sites with potential land-use related impacts. These test sites were chosen to demonstrate bioassessment at sites experiencing a range of impacts from high nutrient and sediment inputs to exotic forestry. Furthermore, as a measure of the temporal validity of the model we used the approach to assess the impactedness of sites sampled over a 3-year interval.

METHODS

Study area

The Manawatu-Wanganui Region covers a large portion (22 179 km^2) of the south-west of the North Island in New Zealand (Fig. 1). It is a political region delineated mainly by the catchment boundaries of the three major rivers in the region - the Whanganui, Rangitikei, and Manawatu rivers. A volcanic plateau dominates the northern part of the region and to the south the axial Tararua and Ruahine ranges divide the region running approximately north-south. The region includes considerable areas of uplifted ranges, steep mudstone country, and rich alluvial plains resulting in a diversity of stream types, ranging from braided cobble to silt dominated lowland rivers, and from an acidic volcano-fed river to small spring-fed mountain streams. Study sites spanned 40° 45' to 39° 30' south and from sea level to 820 m a.s.l. The predominant land use in the region is pastoral farming with some cropping whereas most of the region above 500 m a.s.l has relatively unmodified native vegetation.

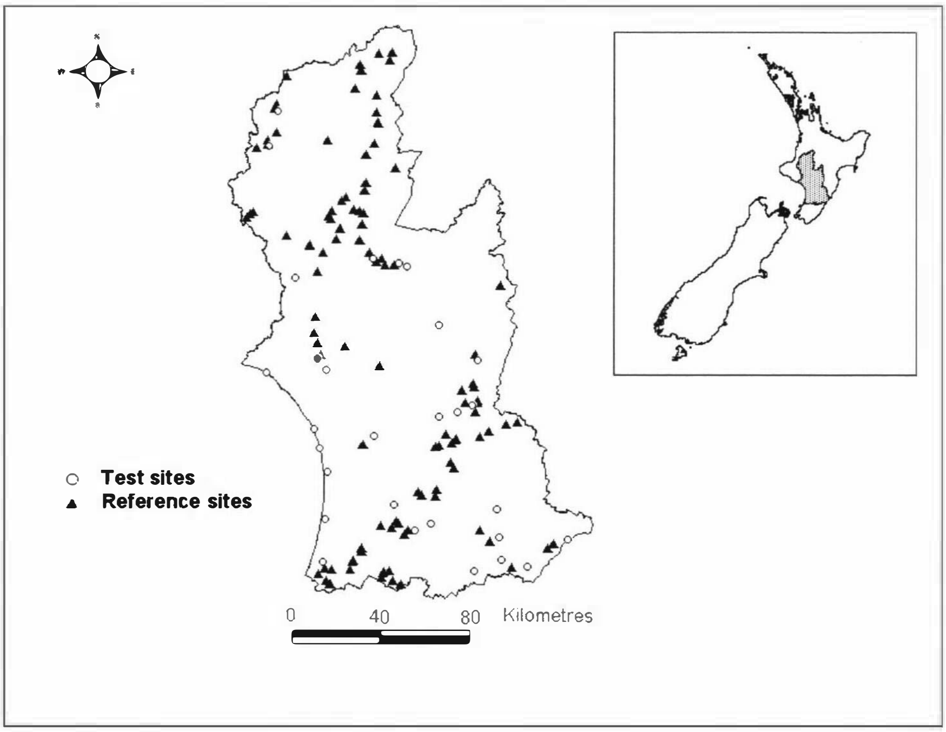

Fig. 1 Location of reference and test sites in the Manawatu-Wanganui region, New Zealand

Site selection

The study was designed to cover all major stream types in the Manawatu-Wanganui region. To achieve this, 200 relatively undisturbed reference sites were selected over the region (Fig. 1). Emphasis was placed on sites having the most natural catchment vegetation, channel morphology, and minimal human impacts (e.g., Hughes 1995). To have sites over the full range of elevations and stream types, some moderately disturbed, but best available catchments were included at lower elevations (Hughes *et al.* 1986). New Zealand has been classified into lotic ecoregions (Harding 1994). The Manawatu-Wanganui political region encompasses portions of seven of these ecoregions and the sites were stratified by each ecoregion. The selection of reference sites occurred in two phases: the first phase, site selection was "desk-based" using expert knowledge of streams combined with geographic information. The second phase occurred after sampling with the inclusion of up to date information from the field. The reference sites were ranked based on the proportion of the catchment in natural vegetation and evidence of best management practice. Seventy-three sites did not reach acceptable standards after field evaluation. These were sites where there was no evidence of best management or had anomalous conditions and these sites were discarded leaving 127 reference sites for further analysis (Hughes *et al.* 1986). Test sites, that were affected by known disturbances were selected and sampled using the same methods as the reference sites.

Collection of macroinvertebrates

Macroinvertebrate samples were collected within a 50-m reach moving in an upstream direction and taken only from riffle habitats (riffles were classified as areas of fast, shallow water with a broken-surface appearance) using a D-shaped net (50 cm wide, 20 cm high, 250 μm mesh). Samples were taken by disturbing the substrate for 1 min with feet and hands and allowing the current to carry the invertebrates into the net. The contents of the net were transferred to sample containers and preserved in 10% formalin. In the laboratory, samples were rinsed using a 500 μm mesh sieve to remove fine sediment and preservative and placed in a Marchant subsampling box (Marchant 1989). The first 100 invertebrates were then randomly extracted, sorted, and identified (Marchant 1989). Identifications were to the lowest reliable taxonomic level (usually genus) using existing keys (e.g., Winterbourn and Gregson 1989).

Environmental measures

Eighty-two physical and chemical variables were measured at each site or gathered from GIS data to complement the invertebrate species lists using a number of methods that are summarised in Table 1.

Table 1 Environmental variables estimated at reference and test sites and details on collection and measurement.

Variable	Units	Footnote
Water		
pH		a
Alkalinity	mg Litre-1 Ca CO3	b
Temperature	$^{\circ}$ C	c
Conductivity	µs cm-1	d
Velocity	m/sec-1	e
Nitrate	mg Litre-1 NO 3-N	f
Reactive phosphorous mg	Litre-1 PO 43-	f
Physical attributes		
Easting	Map co-ords	g
Northing	Map co-ords	g
Distance from coast	km	g
Elevation	m (a.s.l.)	g
Overall slope (site to sea)	m/km	g
Reach slope (over surveyed reach)	m/km	h
Width (mean and SD)	m	i
Depth (mean and SD)	mm	i
Embededness		j
Ecoregion		k
Stream order	Strahler	h
Average catchment elevation	m	h
Total catchment rainfall	annual rainfall × catchment area	h
Total catchment area	km^2	h
Mean and median substrate size	cm	l
Proportion of reach surveyed composed of:		
% Still		n
% Backwater		n
% Pool		n
% Run		n
% Riffle		n
% Rapid		n

Variable		footnote
% Undercut		n

%	Debris jam	n
Pfankuch channel stability index score		
	upper	m
	lower	m
	bottom	m
Catchment geology (entered as baserock and toprock separately)		
%	argillite - crushed	h
%	undifferentiated floodplain alluvium	h
%	argillite	h
%	conglomerate or breccia	h
%	gravels	h
%	greywacke	h
%	lahar deposits	h
%	windblown sands	h
%	loess	h
%	mudstone of fine siltstone - banded	h
%	mudstone of fine siltstone - jointed	h
%	mudstone of fine siltstone - massive	h
%	ashes older than Taupo pumice	h
%	sandstone or coarse siltstone - massive	h
%	Taupo and Kaharoa breccia and volcanic alluvium	h
%	unconsolidated clays, silts, sands, tephra and breccias	h
%	lava, ignimbrite, and other 'hard' volcanic rocks	h
Catchment area landcover		
%	primarily pastoral	h
%	primarily horticultural	h
%	indigenous forest	h
%	planted forest	h
%	scrub	h
%	tussock	h
%	coastal sands	h
%	bare-ground	h

a Determined at site with Orion Quickcheck model 106 pocket meter.
b alkalinity was determined once at each site by standard methods (APHA 1998).
c determined at site with YSI model 85 meter.
d determined at site with YSI model 85 meter (automatically adjusted to 25°C).
e timing the movement of the modal concentration of a slug of dye over approximately 100m of the reach sampled (Joy and Death 2002).
f Hach DR/2010 Portable Data-logging Spectrophotometer. Reactive phosphorus concentrations were obtained using the Ascorbic acid method, and nitrate using the Cadmium Reduction method (Hach Company, P.O. Box 389 Loveland, CO 80539 United States).
g obtained from 1:50,000 maps.
h obtained from River Environment Classification (REC) database (Snelder, *et al.* 2000).
i maximum water depth and width were measured using a staff at five equidistant points longitudinally over the reach sampled.
j subjectively assessed at site after moving substrate (1 = loosely packed; 4 = tightly packed).
k lotic ecoregions (Harding 1994).

l Mean and median substratum particle size was determined using standard granulometry techniques (*sensu* Wolman, 1954; Quinn and Hickey 1990) on 70 - 100 randomly selected rocks at each site using a gravelometer.

m Pfankuch (Pfankuch, 1975) stability index which involves scoring 15 variables (weighted in relation to their perceived importance) according to the observers evaluation of predetermined criteria. Three totals relate to three regions of the stream channel upper banks, lower banks, and stream bottom.

n The percentage of backwater, pool, run, riffle or rapid was visually estimated over 50m above and below the sampled riffle. Riffles were classified as areas of fast, shallow water with a broken-surface appearance; pools were areas of slow deep water with a smooth surface appearance, whereas runs were intermediate in character. Rapids were classified as areas of fast cascading deep water).

Geographic Information Systems data

GIS data on terrain, geology, rainfall, and land cover were obtained from the River Environment Classification (REC) (Snelder *et al.* 1998; Snelder and Biggs 2002). The data underlying the REC describes catchment attributes, in terms of various environmental factors (e.g., geology, elevation, rainfall, and vegetative cover), for individually numbered sections of the river network. For each section of the network (average length = 700 m), each factor is described by the area of catchment occupied by various categories (e.g., geological categories include; greywacke, limestone etc) (Snelder and Guest 2000). All data were associated with each corresponding sample site and converted to catchment proportions. These data were in 33 categories (Table 1).

Model construction

Modeling generally followed the methods used to develop RIVPACS (Wright 1995) and AUSRIVAS models (Smith *et al.* 1999) and proceeded with the application of the following steps:

First, the reference sites were classified into groups containing similar invertebrate communities using two-way indicator species analysis (TWINSPAN) (McCune and Mefford 1999). This process classifies both samples and species simultaneously based on dividing reciprocal averaging ordination space. The TWINSPAN analysis was achieved using five pseudospecies cut levels of 0, 2, 5, 10, and 20 invertebrates per sample. For the site classifications above, but not subsequent predictions of biotic composition all taxa occurring at <5% of the reference sites were omitted

(Hawkins *et al.* 2000). These rare taxa were deleted because they cause noise in the data and have poor predictive capability (Hawkins *et al.* 2000; Marchant 2002).

Second, from the large number of environmental variables measured at sites and the GIS database a subset was chosen using stepwise discriminant function analysis (DFA) that best discriminated between the biological groups from above. Only variables not commonly affected by human activity were included. The number of variables was reduced using the backward STEPDISC procedure in the SAS statistical package (SAS 2000) and by iteration using different combinations of variables to minimise the posterior classification error rate. Using this process a subset of the habitat variables that best discriminate between the site groups obtained from the faunal classification was selected. As one of the assumptions of DFA is that predictor variables have equal within-group variances, the variables were log transformed Log_{10} $(x + 1)$. Variables were entered as both transformed and not transformed and the transformed variables were used if selected by the variable reduction process outlined above (Clarke *et al.* 1996).

Third, the DISCRIM procedure in SAS was used to incorporate the selected variables into a discriminant function to be used to assign sites into the biotic groups based on their environmental characteristics. Crossvalidation was used to check whether sites were allocated to their correct groups. The crossvalidation process (also known as jack-knife or leave-one-out validation) involves leaving out each site in turn, then re-building the model, and testing the held out site to assess whether the site was predicted as belonging to the correct group. This jack-knife procedure has been shown to provide a robust and unbiased assessment when used with other similar models (Manel *et al.* 1999; Manel *et al.* 2001; Olden *et al.* 2002). A site was considered to be correctly classified if the probability of belonging to the correct group is higher than it is for the other groups. However, the actual value of this misclassification rate is not critical because all probabilities of group membership are used for predictions rather than just the group with the highest probability (Simpson and Norris 2000).

Fourth, following the procedure described by Moss *et al.* (1987) and Wright *et al.* (1984), the probability of each taxon occurring at a site was calculated. Initially the probability of a new site belonging to the reference site groups was calculated from the DFA. Then the probability of occurrence of a given taxon at the site was calculated by multiplying the probability from the DFA by the percentage frequency with which the taxon occurred in each site group. The probabilities were then summed for all groups to give a weighted probability of occurrence for that taxon. The sum of the probabilities for all taxa with a probability of occurrence >50% gave the number of predicted taxa expected (*E*). The number of taxa collected at the site from the list of those predicted to occur, represented the number of taxa observed *(O)*. The number of taxa observed was divided by the number of expected taxa (*E*) to give the observed over expected ratio (*O/E*) (Wright *et al.* 1984; Moss *et al.* 1987). The use of a >0.5 cut-off level was first used in AUSRIVAS models (Simpson and Norris 2000) and later by Hawkins *et al.* (2000).

Finally, using a preliminary model, *O/E* ratios were calculated for the reference sites. Sites with an *O/E* ratio <0.75 were then removed from the classification on the assumption they were not true reference sites as 25% of their taxa were absent (following Smith *et al.* 1999; Turak *et al.* 1999). After their removal, steps 2 - 4 above were repeated to give the final model.

Model validation and testing

Two evaluations of the model were made. First, for temporal evaluation data were collected from 11 sites that had been sampled 3 years before the reference site sampling year to determine whether the model would produce consistent results over time. The assumption was that *O/E* ratios should be relatively consistent as land use and thus, potential impacts were considered unlikely to have changed over the 3 years at these sites. For this evaluation *O/E* ratios for the sites sampled over the different years were compared with a Wilcoxon paired rank test (SAS 2000). The second evaluation involved the use of data from 29 test sites with known impacts. The data from the 29 test sites were applied to the model and *O/E* ratios were obtained to determine whether the model would detect a range of human disturbances. The *O/E* ratio values from the model were used as a measure of impact at sites, with lower scores indicating greater impact.

To evaluate the *O/E* scores, the variables likely to be influenced by human impacts and therefore not used in the DFA were used to assess the degree of human impact at the test sites. To achieve this assessment the relationship between physico-chemical variables likely to be influenced by human impacts and *O/E* scores were assessed using Spearman rank correlations. To assist with the interpretation of *O/E* scores produced by the model, the 10th percentile of the distribution of *O/E* ratio values from the reference sites was used as the cut-off point for site assessment, with values less than this classed as failed (following Smith *et al.* 1999; Turak *et al.* 1999).

RESULTS

Faunal characteristics of the groups

Five main groups of sites were identified using TWINSPAN classification of the invertebrate assemblages (Fig. 2). The first division split the group 1 sites from the rest with the snail *Potamopyrgus antipodarum* and Simuliidae as indicator taxa. The mayflies *Deleatidium* spp. and *Coloburiscus humeralis* and the net-spinning caddisfly *Aoteapsyche* sp. were the indicator taxa for the rest of the sites at this first division. These sites were then further divided into a group with *Zephlebia* as the indicator taxa and were distinct from another group containing the stoneflies *Zelandoperla* and *Stenoperla prasina.* The next division formed the other four groups (2 - 5). The group 3 sites indicator taxon was the mayfly *Nesameletus* sp. The indicator species for group 2 sites were the cased caddisflies *Olinga feredayi*, *Helicopsyche* sp., and *Pycnocentrodes* sp. The group 4 indicator taxa were the beetle family Scirtidae and the mayfly *Coloburiscus.* The indicator taxa for group 5 sites were *Pycnocentrodes*, *Aoteapsyche* and the beetle family Elmidae. The group 5 sites contained the most taxa (87), followed by groups 4 (75) and 3 (69) whereas groups 1 and 2 contained 54 and 55 taxa respectively.

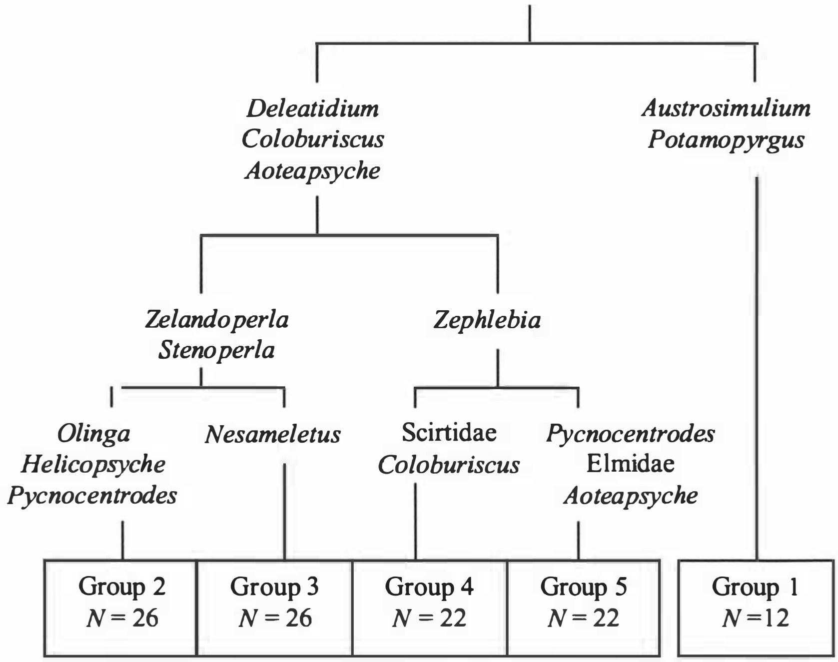

Fig. 2 TWINSPAN classification of reference sites in the Manawatu-Wanganui region, New Zealand, sampled over summer 2000. (*N* the number of sites in each end group).

Fig. 3 Location of the reference sites in the Manawatu-Wanganui region, New Zealand, numbers indicate the five classification groups

Geographic patterns

The group 1 sites were low elevation sites distributed over the whole region, away from the mountain ranges (Fig. 3). The group 2 sites were generally clustered in the Tararua ranges in the south and tributaries of the Whanganui River in the ranges west of Lake Taupo in the north. The group 3 sites were the highest elevation sites, and were generally Whanganui River tributary sites and were clustered mainly around the volcanic plateau and some Manawatu River tributary sites in the south-western Ruahine ranges. The majority of group 4 sites were Whanganui River tributaries and most were clustered around the north-west of the region. The group 5 sites were spread throughout the region but most were in the Tararua ranges.

Environmental variables

Eighty-two variables were initially used in the discriminant analysis (Table 1) and these were reduced to 19 variables after stepwise variable reduction (Table 2). The mean values for the variables in each of the groups revealed the differences between the site groups (Table 2). The environmental variables selected were a mixture of geographic, geological, and site-specific variables. The first four variables (as judged by F value) - conductivity, elevation, longitude and substrate sizes, are generally related to longitudinal gradients from headwaters to the sea. The site-specific variables were the percentages of debris-jams and pool as well as site slope, whereas the rest of the variables were associated with catchment geology and site spatial position.

Model performance

The discriminant function analysis using the 19 selected physico-chemical variables correctly predicted group membership at 70% of sites before removal of sites with an O/E score of <0.75 and 72% afterwards using a jack-knife validation. In its final form, the model was based on 119 reference sites. The 10th percentile O/E score for the reference sites was 0.84 and was used subsequently as the threshold for impact detection.

Effect of between year variability on model performance

Data from 11 sites (six reference sites, four sites not having attained reference standard, and one test site) sampled in 1997 (Polglase 2000) were used to test temporal changes in the model output. These replicate site samples were run through the model and *O/E* ratios calculated. The *O/E* ratios were not significantly different for the two sampling dates (Wilcoxon $Z = 1.19$; $P = 0.23$) (Table 3). Sites that failed the 10th percentile of reference sites (sites 20 and 138) failed for both sample dates, all other sites were equivalent to reference for both sample dates.

Evaluating impacts using the model

The mean *O/E* value for the test sites 0.55 (SE 0.03) was considerably less than the reference site mean *O/E* value of 1.06 (SE 0.02). All of the test sites were assessed as impaired using the 10th percentile of the reference site *O/E* ratios (0.84) as a criterion (Table 4). However, there was no obvious relationship between the type of potential disturbance and *O/E* values (Table 4). Significant correlations between *O/E* values and physico-chemical variables at the test sites are shown in Table 5. The proportion of the catchment in windblown and coastal sand, stream alkalinity and stream width all increased with reducing *O/E* values while the remaining correlated variables decreased. The variables showing positive correlations were those generally associated with longitudinal changes from the coast to headwaters with *O/E* values increasing with increasing distance from the coast, elevation, slope and the proportion of the surveyed reach classed as riffle.

Table 2. Mean density (fish per 100 m^2) for each of the 13 taxonomic units in the 6 groups obtained from the TWINSPAN analysis.

Family	Scientific name	Group number	1	2	3	4	5	6
		Sites in group	(6)	(36)	(5)	(15)	(18)	(62)
Anguillidae	*Anguilla spp.*	Elver		0.03	0.44	5.28	1	0.15
Anguillidae	*Anguilla dieffenbachii*	Longfin eel	0.1	1.51	1.13	2.08	2.02	4.25
Anguillidae	*Anguilla australis*	Shortfin eel				2.73	0.18	0.04
Eleotridae	*Gobiomorphus cotidianus*	Common bully				6.84	0.07	0.12
Eleotridae	*Gobiomorphus huttoni*	Redfin bully				17.08	2.16	0.93
Eleotridae	*Gobiomorphus spp.*	Non-migratory bully		0.75	3.24	1.21	0.98	8.27
Galaxiidae	*Galaxias maculatus*	Inanga				23.41	0.12	0.01
Galaxiidae	*Galaxias brevipinnis*	Koaro	0.8	0.8			0.21	0.01
Galaxiidae	*Galaxias postvectis*	Shortjaw kokopu		0.07				0.04
Galaxiidae	*Galaxias divergens*	Dwarf galaxiid		0.2	4.86			
Mugiloididae	*Cheimarrichthys fosteri*	Torrentfish		0.07			1.6	0.11
Retropinnidae	*Retropinna retropinna*	Common smelt				0.34	0.02	
Salmonidae	*Salmo spp.*	Trout	1.31	2.94	0.54	0.24	0.48	1.15

Table 2 Group means (± standard error) and *F* values for the 19 physico-chemical variables used in DFA model significant at the <0.05 level as measured by partial correlation (for units see Table 1).

Variable	F value	Group 1	Group 2	Group 3	Group 4	Group 5
Conductivity	17.27	314.43 ± 46.71	70.92 ± 3.96	127.05 ± 24.17	147.09 ± 21.10	118.02 ± 13.59
Elevation	13.67	138.33 ± 24.83	301.92 ± 34.00	549.23 ± 36.36	263.64 ± 42.82	274.20 ± 34.45
Longitude	9.06	27354 ± 127	27243 ± 40	27309 ± 42	26960 ± 62	27355 ± 52
Mean substrate size	5.69	5.68 ± 0.97	5.01 ± 0.58	4.39 ± 0.66	6.49 ± 0.72	6.24 ± 0.39
Pfankuch stability (bottom score)	4.84	19.91 ± 1.42	21.23 ± 1.00	22.38 ± 1.28	21.18 ± 0.99	25.72 ± 1.08
Width	4.38	1.33 ± 0.26	2.39 ± 0.21	1.74 ± 0.19	1.22 ± 0.14	0.98 ± 0.12
Toprock lahar deposits	3.47	0.01 ± 0.01				
Site slope (over surveyed reach)	3.93	15.54 ± 6.16	20.73 ± 2.67	42.02 ± 7.20	23.20 ± 5.18	34.72 ± 8.05
% debris jam	3.89	10.17 ± 3.03	1.35 ± 0.65	2.88 ± 0.74	5.23 ± 0.96	3.40 ± 0.85
Baserock mudstone or fine siltstone, jointed	3.23	0.04 ± 0.04	0.02 ± 0.02			
% pool	2.98	23.13 ± 4.62	8.75 ± 2.52	12.14 ± 3.26	29.67 ± 3.92	17.20 ± 3.14
Temperature	2.80	15.77 ± 1.19	12.91 ± 0.51	11.70 ± 0.56	12.57 ± 0.67	14.88 ± 0.70
Sandstone or coarse siltstone, massive	2.61	0.01 ± 0.01				
Toprock conglomerate or breccia	2.42	0.05 ± 0.03				
Taupo and Kaharoa breccia and volcanic alluvium	2.41	0.21 ± 0.07	0.20 ± 0.06	0.07 ± 0.05	0.07 ± 0.03	
Distance to coast	2.35	110.67 ± 22.95	156.42 ± 21.70	200.65 ± 11.78	195.95 ± 17.19	131.68 ± 15.67
Baserock gravels	2.22	0.05 ± 0.03	0.02 ± 0.01	0.01 ± 0.01	0.05 ± 0.03	
Latitude	2.02	61339 ± 219	61240 ± 184	61722 ± 119	62133 ± 113	61308 ± 121
Slope (site to sea)	2	85 ± 0.16	0.51 ± 0.07	0.40 ± 0.04	1.02 ± 0.11	0.55 ± 0.06

Table 3 Number of taxa observed and expected and *O/E* - values for sites sampled in 1997 by Polglase (2000) and again in 2000 (this study). *O/E* values in bold are classed as impaired (*O/E* - value <0.84)

	Sampled 1997 (Polglase 2000)			Sampled 2000 (this study)		
Site No.	No. of taxa expected	Observed no. of taxa	*O/E*	No. of taxa expected	Observed no. of taxa	*O/E*
7	5.82	6	1.03	5.82	5	0.86
18	4.28	6	1.4	4.28	7	1.63
19*	5.76	5	0.86	5.76	6	1.04
20	7.86	6	**0.76**	7.86	5	**0.64**
38*	4.27	4	0.94	4.27	5	1.17
44*	5.76	7	1.21	5.76	7	1.21
75*	5.85	7	1.2	5.85	7	1.2
90	5.78	5	0.87	5.78	6	1.04
138†	7.02	5	**0.71**	7.02	4	**0.57**
139*	5.98	6	1	5.98	6	1
170*	7.89	6	0.86	7.89	7	0.89

* = Reference site
† = Test site

Table 4. Observed over expected (*O/E*) ratio values for the 29 test sites and the potential disturbance upstream of site.

Site	Disturbance	*O/E*
Whitebait Creek, Foxton	Dairy/sheep/beef farming/urban	0.16
Tiraumea River, Wairarapa	Sheep/beef farming	0.23
Bommingrange Stream, Scotts Ferry	Exotic forest/dairy farming	0.23
Koitiata Stream, Scotts ferry	Exotic forest/dairy farming	0.23
Papuka stream, Wairarapa coast	Sheep/beef farming	0.35
Inflow Stream, Lake Papiatonga, Levin	Dairy farming	0.38
Mowhanau Stream, Wanganui coast	Dairy/sheep/beef farming	0.44
Paparata Stream, near Ohura	Sheep/beef farming	0.46
Keebles Stream, near Palmerston North	Dairy farming	0.47
Mangahau River, near Pahiatua	Sheep/beef farming/dam	0.47
Parewanui Drain, near Scotts Ferry	Dairy/sheep/beef farming	0.47
Owahanga River, Wairarapa coast	Sheep/beef farming	0.58
Pongaroa River, Wairarapa	Sheep/beef farming	0.58
London Creek, near Kimbolton	Sheep/beef farming	0.62
Wahiona Stream, near Waiouru	Exotic forest	0.63
Tokiahuru River, near Waiouru	Exotic forest	0.63
Mangakukeke Stream, near Mangaweka	Sheep/beef farming	0.64
Horopito Stream, near Apiti	Sheep/beef farming	0.64
Hautapu River, Taihape	Sheep/beef farming	0.65
Makirikiri Stream, near Fordell	Sheep/beef farming	0.69
Waihoki Stream, near Alfredon	Sheep/beef farming	0.69
Huhatahi Stream, tributary, near Ohura	Sheep/beef farming	0.69
Makiekie Creek, near Utuwai	Sheep/beef farming	0.7
Porewa Stream, near Marton	Sheep/beef farming/sewage	0.7
Mangatainoka River, near Pahiatua	Dairy/sheep/beef farming	0.7
Mongotai Stream, near Wanganui	Exotic forest	0.7
Mangateitei Stream, near Ohakune	Market gardening/vegetable washing	0.72
Mangatoro Stream, near Pahiatua	Sheep/beef farming	0.72
Whanganui River, above Wanganui	Sheep/beef farming/dam	0.76

Table 5 Variables which had significant Spearman rank correlation with test site *O/E* (observed/expected) scores.

Variable	Spearman *r*	*P* value
Toprock windblown sands	-0.57	0.001
Landcover coastal sands	-0.54	0.001
Average catchment elevation	0.52	0.001
Velocity	0.49	0.001
Median substrate size	0.49	0.001
% riffle	0.48	0.001
Distance from the coast	0.47	0.009
Altitude	0.44	0.01
Alkalinity	-0.43	0.01
Mean width	-0.42	0.02
Latitude	0.41	0.02
Reach slope	0.41	0.02
Landcover tussock	0.39	0.03

DISCUSSION

The results presented here are the first published application of the RIVPACS/AUSRIVAS methodology using macro-invertebrates in New Zealand. This application in the Manwatu-Wanganui region shows that the use of a predictive reference site model provided a reliable assessment of biological conditions. The results presented here were remarkably similar to applications of predictive bioassessment models in Western Australia (Smith *et al.* 1999), New South Wales (Turak *et al.* 1999) and Indonesia (Sudaryanti *et al.* 2001). The number of groups, number of reference sites, and the percentage of sites correctly assigned to groups (72%) were almost identical to the Australian and Indonesian AUSRIVAS models.

Predictor variables

The predictor variables selected by the stepwise and iterative variable reduction processes could be placed in six categories - elevation/distance from the sea, location, water chemistry, substratum, river size, and geology. Five of the categories above with the exception of geology are also represented in the majority of RIVPACS models developed in the United Kingdom (Wright *et al.* 1984; Moss *et al.* 1987) and in the Australian and Indonesian applications of AUSRIVAS (Smith *et al.* 1999; Turak *et al.* 1999; Sudaryanti *et al.* 2001). The GIS based geology variables used in this model made up six of the 19 predictor variables in our model but these variables have not been available for use in the other AUSRIVAS applications so are not comparable. However, the consistent selection of variables from the five categories above in a number of different countries and continents worldwide suggests that the distributions of aquatic macroinvertebrates at large geographic scales can be determined by the same set of environmental variables.

Relationship between O/E scores and other measures of impact

When all available variables, not just those used as predictors, were correlated with *O/E* ratios from the test sites there were a number of variables that unexpectedly showed no link with *O/E* ratios (Table 5). These variables included the catchment proportions of both indigenous vegetation and pastoral farming as well as nitrate and phosphorus. However, as the nutrient samples were one-off they may not reflect the true nutrient status of the sites. The reason other variables showed no significant relationship is probably because of a lack of variation in these variables at the test sites or the relative proximity of sites to potential impacts. The percentage of catchment in different land-use categories may also not give a true indication of the intensity and effect of these land-use activities. Dairy farming for example will have a greater impact than low intensity sheep and beef farming, although both may have the same percentage in pasture (Harding *et al.* 1999).

Test site assessment

In this model as with the AUSRIVAS models the probability of assessing a site as impaired when it was not (type I error) was set at 10% by choosing the first decile of the reference *O/E*-values as the cut-off level for impairment. Thus, the classification of the entire set of test sites as impaired (Table 3) suggests that the model can reliably assess impacts. There was no strong relationship between the *O/E*-values and type of impact as the model detects impairment but not the particular type of impact. When the model was applied to data collected over different years and processed by different individuals the model produced consistent outputs (Table 3). The results of this evaluation although tentative suggest we can have confidence that the model will perform accurately on data collected by others and/or at different times.

Probability cut-off levels for inclusion of taxa

The use of the 0.5 probability level followed the AUSRIVAS protocol rather than the more stringent 0.0 (i.e., all probabilities) level used in Britain with the RIVPACS models. Hawkins *et al.* (2000) found that when these two cut-off levels were compared, the 0.5 cut-off level yielded more robust model outputs.

One major advantage of reference site predictive models over index approaches such as the MCI commonly used in bioassessment in New Zealand is that any subjectivity in assigning scores to taxa is removed. Furthermore the scores applied to individual taxa when creating indices are based on their natural or more commonly their perceived, response to environmental gradients, (e.g., MCI and organic enrichment). In contrast, the only prior assumption with the predictive model approach is that impacts alter invertebrate community make-up by leading to the extinction of taxa that would otherwise be present. Thus, predictive models will indicate any form of impact while single indices may not (e.g., MCI and heavy metals Hickey and Clements (1998)). One potential area of subjectivity with predictive models is reference site selection, but this can be mitigated to some extent (as in this application) by the stratification of sites within ecoregions (Omernik 1995) and the use of local expertise in site selection.

CONCLUSIONS

A predictive bioassessment model was successfully developed and applied to a region within New Zealand. This is the first published application in New Zealand of a RIVPACS type predictive reference site model using macroinvertebrates and demonstrated that model development at genus/species level was successful in detecting ecological impacts at the selected test sites. The use of environmental variables to detect human disturbance has been shown in other parts of the world and this example reveals the potential for a similar application throughout New Zealand. This application follows a call for the implementation of predictive models for bioassessment made by the New Zealand Ministry for the Environment through its macroinvertebrate-working group (Winterbourn 1999). Our study also suggests that this approach would be suitable for state of the environment reporting where the degree of deviation from pristine could be presented rather than a single index score that has no intuitive meaning. Furthermore, the legislation related to freshwater management in New Zealand (RMA) emphasises the development of whole ecosystem approaches and these predictive models help meet that requirement by combining information on environmental variables and macroinvertebrate assemblages in a predictive format.

Acknowledgments

We thank Grant Blackwell and Rosemary Miller for assistance with fieldwork; and John Phillips for assistance with site selection and chemical analyses. Thanks to Miranda Oliff for sorting the macroinvertebrate samples and to Kirsty Francis for identifying them. Thanks also to Ian Henderson for many useful discussions on predictive modeling and to Miranda Robinson for making her data available. We thank the two anonymous referees for their constructive comments. This research was funded by horizons.mw and the Sustainable Management Fund (Ministry for the Environment) grant no. 5099.

References

APHA. 1998. Standard methods for the examination of water and wastewater., 20 edition. American Public Health Association, Washington D.C., USA.

Bailey, R. C., M. G. Kennedy, M. Z. Dervish, and R. M. Taylor. 1998. Biological assessment of freshwater ecosystems using a reference condition approach: comparing predicted and actual benthic invertebrate communities in Yukon streams. Freshwater Biology **39**:765-774.

Clarke, R. T., M. T. Furse, J. F. Wright, and D. Moss. 1996. Derivation of a biological quality index for river sites: comparison of the observed with expected fauna. Journal of applied Statistics **23**:311-322.

Furse, M. T., D. Moss, J. F. Wright, and P. D. Armitage. 1984. The influence of seasonal and taxonomic factors on the ordination and classification of running-water sites in Great Britain and on the prediction of their macro-invertebrate communities. Freshwater Biology **14**:257-280.

Harding, J. S. 1994. Lotic Ecoregions of New Zealand. PhD. University of Canterbury, Christchuch.

Harding, J. S. 1999. Changes in agricultural intensity and river health along a river continuum gradient. Freshwater Biology **42**:345-357.

Hawkins, C. P., R. H. Norris, J. N. Hogue, and J. W. Feminella. 2000. Development and evaluation of predictive models for measuring the biological integrity of streams. Ecological Applications **10**:1456-1477.

Hickey, C. W., and W. H. Clements. 1998. Effects of heavy metals on benthic macroinvertebrate communities in New Zealand streams. Environmental Toxicology and Chemistry **17**:2338-2346.

Hughes, R. M., D. P. Larsen, and J. M. Omernik. 1986. Regional reference sites: a method for assessing stream potentials. Environmental Management **10**:629-635.

Hughes, R. N. 1995. Defining acceptable biological status by comparing with reference conditions. Pages 31-47 *in* W. S. Davis and T. P. Simon., editors. Biological Assessmant and Criteria: tools for water resource planning and decision making. Lewis publishers, Boca Raton.

Joy, M. K., and R. G. Death. 2000. Development and application of a predictive model of riverine fish community assemblages in the Taranaki region of the

North Island, New Zealand. New Zealand Journal of Marine and Freshwater Research **34**:243-254.

Joy, M. K., and R. G. Death. 2002. Predictive modeling of freshwater fish as a biomonitoring tool in New Zealand. Freshwater Biology **47**:2261-2275.

Manel, S., J. Dias, and S. J. Ormerod. 1999. Comparing discriminant analysis, neural networks and logistic regression for predicting species distributions: a case study with a Himilayan river bird. Ecological Modelling **120**:337-347.

Manel, S., H. C. Williams, and S. J. Ormerod. 2001. Evaluating presence-absence models in ecology: the need to account for prevalence. Journal of Applied Ecology **38**:921-931.

Marchant, R. 1989. A subsampler for samples of benthic invertebrates. Bull. Aust. Soc. Limnol. **19**:49-52.

Marchant, R. 2002. Do rare species have any place in multivariate analysis for bioassessent. Journal of the North American Benthological Society **21**:311-313.

Marchant, R., A. Hirst, R. H. Norris, R. Butcher, L. Metzeling, and D. Tiller. 1997. Classification and prediction of macroinvertebrate assemblages from running waters in Victoria, Australia. Journal of the North American Benthological Society **16**:664-681.

McCune, B., and M. J. Mefford. 1999. PC-ORD. Multivariate Analysis of Ecological Data., Version 4.14 edition. MjM Software Design, Gleneden Beach, Oregon USA.

Moss, D., M. T. Furse, J. F. Wright, and P. D. Armitage. 1987. The prediction of the macro-invertebrate fauna of unpolluted running-water sites in Great Britain using environmental data. Freshwater Biology **17**:41-52.

Olden, J. D., D. A. Jackson, and P. R. Peres-Neto. 2002. Predictive models of fish species distributions: A note on proper validation and chance predictions. Transaction of the American Fisheries Society **131**:329-336.

Omernik, J. M. 1995. Ecoregions: a spatial framework for environmental management. *in* W. S. Davis and T. P. Simon., editors. Biological Assessmant and Criteria: tools for water resource planning and decision making. Lewis publishers, Boca Raton.

Pfankuch, D. J. 1975. Stream reach inventory and channel stability evaluation. U.S.D.A. Forest Service, Region 1, Missoula, Montana.

Polglase, M. A. 2000. The development of the Waterway self assessment form, a stream management tool for landowners. MSc Thesis. Massey University, Palmerston North.

Quinn, J. M., and C. W. Hickey. 1990a. Characterisation and classification of benthic invertebrate communities in 88 New Zealand rivers in relation to environmental factors. New Zealand Journal of Marine and Freshwater Research **24**:387-409.

Quinn, J. M., and C. W. Hickey. 1990b. Magnitude of effects of substrate particle size, recent flooding, and catchment development on benthic invertebrates in 88 New Zealand rivers. New Zealand Journal of Marine and Freshwater Research **24**:411-427.

Reynoldson, T. B., and J. F. Wright. 2000. The reference condition: problems and solutions. Pages 293-303 *in* J. F. Wright, D. W. Sutcliffe, and M. T. Furse, editors. Assessing the biological quality of freshwaters. RIVPACS and other techniques. Freshwater biological association, Ambleside, United Kingdom.

Reynoldson, T. B., R. C. Bailey, K. E. Day, and R. H. Norris. 1995. Biological guidelines for freshwater sediment based on BEnthic Assessment of SedimenT (the BEAST) using a multivariate approach for predicting biological state. Australian Journal of Ecology **20**:198-219.

SAS. 2000. SAS/STAT Version 8.2. SAS Institute Inc, Cary, North Carolina.

Simpson, J., and R. H. Norris. 2000. Biological assessment of water quality: development of AUSRIVAS models and outputs. Pages 125-142 *in* J. F. Wright, D. W. Sutcliffe, and M. T. Furse, editors. Assessing the biological quality of freshwaters. RIVPACS and other techniques. Freshwater Biological Association, Ambleside, UK.

Smith, M. J., W. R. Kay, D. H. D. Edward, P. J. Papas, K. S. J. Richardson, J. C. Simpson, A. M. Pinder, D. J. Cale, P. H. J. Horwitz, J. A. Davis, F. H. Yung, R. H. Norris, and S. A. Halse. 1999. AusRivAS: using macroinvertebrates to assess ecological condition of rivers in Western Australia. Freshwater Biology **41**:269-282.

Snelder, T., B. J. F. Biggs, U. Shankar, B. McDowall, T. Stephens, and I. K. G. Boothroyd. 1998. Development of a system of physically based habitat classification for water resources management of New Zealand rivers. NIWA, Christchurch.

Snelder, T., and P. Guest. 2000. The 'river ecosystem management framework' and the use of river environment classification as a tool for planning. NIWA Client Report: CHC00/81, Ministry for the Environment, National Institute of Water & Atmospheric Research Ltd.

Stark, J. D. 1993. A macroinvertebrate community index of water quality for stony streams. New Zealand Journal of Marine and Freshwater Research **27**:463-478.

Sudaryanti, S., Y. Trihadiningrum, B. T. Hart, P. E. Davies, C. Humphrey, R. H. Norris, J. Simpson, and L. Thurtell. 2001. Assessment of the biological health of the Brantas River, East Java, Indonesia using the Australian River Assessment System (AUSRIVAS) methodology. Aquatic Ecology **35**:135-146.

Turak, E., L. K. Flack, R. H. Norris, J. Simpson, and N. Waddell. 1999. Assessment of river condition at a large spatial scale using predictive models. Freshwater Biology **41**:283-298.

Winterbourn, M. J. 1999. Reccomendations of the New Zealand Macroinvertebrate Working Group: Monitoring for sustainable river ecosystem management and the role of macroinvertebrates. *in* M. J. Winterbourn, editor. The use of macroinvertebrates in Water Management. Ministry for the Environment, Wellington.

Winterbourn, M. J., and K. L. D. Gregson. 1989. Guide to the Aquatic Insects of New Zealand. Revised edition. Bulletin of the Entomological Society of New Zealand No. 9.:95.

Wolman, M. G. 1954. A method of sampling coarse river-bed material. Transactions, American Geophysical Union **35**:951-956.

Wright, J. F. 1995. Development and use of a system for predicting the macroinvertebrate fauna in flowing waters. Australian Journal of Ecology **20**:181-197.

Wright, J. F., M. T. Furse, and P. D. Armitage. 1993. RIVIPACS- a technique for evaluating the biological quality of rivers in the UK. European Water Pollution Control. **3**:15-25.

Wright, J. F., D. Moss, P. D. Armitage, and M. T. Furse. 1984. A preliminary classification of running water-sites in Great Britian based on macroinvertebrate species and the prediction of community type using environmental data. Freshwater Biology **14**:221-256.

Wright, J. F., D. W. Sutcliffe, and M. T. Furse. 2000. Assessing the biological quality of fresh waters: rivpacs and other techniques. Freshwater biological association, Ambelside, Cumbria, UK.

CHAPTER THREE

Neural network modeling of freshwater fish and macro-crustacean assemblages for biological assessment in New Zealand

Joy, M. K., and R. G. Death. In Press. Neural network modeling of freshwater fish and macro-crustacean assemblages for biological assessment in New Zealand. *in* S. Lek, M. Scardi, and S. E. Jorgensen, editors. Modeling community structure in freshwater ecosystems. Springer Verlag.

Abstract

The comparison between the number of taxa observed and that expected in the absence of human impact is an easily understood and ecologically meaningful measure of biological quality. This comparison has been successfully applied to macroinvertebrates to assess the biological quality of flowing water sites using with the River Invertebrate and Classification System (RIVPACS) and its derivatives. We developed a methodology based on the comparison between observed and expected freshwater fish and macro-crustacean assemblages to assess the biological quality of stream sites in the Auckland region, New Zealand. Freshwater fish and macro-crustacean assemblages were sampled at 118 least impacted (reference) sites and individual artificial neural network (ANN) models developed using environmental measures taken at those sites to predict the presence or absence of the 12 most common species. The ANN models used only the environmental variables least likely to be influenced by human impacts and outputs from models for all species were combined so that the predictions were for the assemblage expected in the absence of human impacts. The output from the model is the ratio of observed to expected species number so that the biological quality of a site can be assessed under the assumption that human impacts reduce species richness.

Keywords. Artificial neural networks; predictive models; New Zealand freshwater fish; fish ecology; reference sites.

Introduction

Biological assessment of flowing water has undergone a conceptual change from the use of biological indicators of water quality to the assessment of 'biological quality' (Wright *et al.* 2000) or 'biological integrity' (Karr 1994). This change from biological indicators to the assessment of biological integrity marks a shift towards ecosystem level evaluation and these approaches are often referred to as measures of 'ecosystem health' (Chessman *et al.* 1999). These new concepts come from a recognition that water quality is the result of many factors including biological interactions, flow regimes and habitat structure and that these are all dependant on modification by human activities (Karr 1991, 1995). It follows from this movement to an ecosystem approach that impacts on the biological condition of rivers exposed to human impacts can be judged by comparing the river biota with that from relatively unimpacted reference ecosystems. The prerequisite is that the sites occur in similar geomorphological and climatic settings referred to as reference conditions (Chessman *et al.* 1999). Thus, we can assess the condition of human disturbed sites by measuring their structural attributes and then comparing them with relevant reference conditions that are pristine or, at worst, relatively undisturbed (Hughes *et al.* 1986). This reference site method of river assessment has been in use for some time although different approaches have been used in different parts of the world. In the USA, a reference site approach has been applied with 'indices of biotic integrity' e.g. (Karr 1981). Whereas in the United Kingdom and more recently Australia, multivariate reference site predictive models have been used (Simpson and Norris, 2000).

The predictive reference condition approach (RIVPACS: River InVertebrate Prediction And Classification System) and its derivatives developed originally by Wright *et al.* (1984) and later advanced by Reynoldson *et al.* (1995), Simpson and Norris (2000) and Hawkins *et al.* (2000) have been successfully applied to streams worldwide. The output from the models is a list of taxa expected to be found at a site in the absence of human impacts predicted from a suite of environmental variables. The predictions come from a database of least impacted sites selected to cover all stream types within a region. In the bioassessment of a site, the final output is a measure of the relationship between the fauna collected at a site and that predicted. That is the observed number of taxa (O) is compared to the number expected (E) in the absence of stress as a measure of departure from expected conditions, and is thus, a measure of biological impairment. Although the predictive models described above have been developed using macroinvertebrates they have the potential for use with other biotic groups (Reynoldson and Wright 2000) and have been developed for use with fish (Joy and Death 2002), diatoms (Chessman *et al.* 1999) and stream habitat features (Davies *et al.* 2000).

New Zealand has a limited native fish and macro-crustacean fauna of c. 39 recognised species, dominated by Galaxiidae and Eleotridae as well as a non-migratory crayfish and a diadromous shrimp. The fauna is characterised by a high proportion of diadromous species such that there are marked longitudinal trajectories of fish distribution with species richness reducing with elevation (Joy *et al.* 2000). These distributional patterns negate the application of bioassessment methods using fish employed elsewhere in the world such as the index of biotic integrity (IBI). These index approaches would be problematical in New Zealand because they rely on relationships between community metrics and habitat quality and would not account for the overriding longitudinal distributional patterns caused by diadromy (McDowall and Taylor 2000). Migratory fish and crustacea can however be used in bioassessment if the method used takes into account their longitudinal distribution patterns. A predictive modeling approach using reference sites allows for the migration driven distributional patterns to be incorporated in the bioassessment of fish and macro-crustaceans (Joy and Death 2002).

We took a reference site predictive modeling approach to the use of fish and macro-crustaceans in the assessment of biological quality in the rivers of Auckland, New Zealand. To achieve this we modeled the presence or absence of 10 fish and 2 macro-crustacean species using individual artificial neural network models, one for each taxon based on environmental variables. To make the predictions from the model independent of human impacts we used only environmental variables unlikely to be influenced by human impacts and used only data from minimally disturbed reference sites. The predictions from these models were combined to predict the fish and macro-crustacean assemblage to be expected at sites.

Methods

Reference sites

The sites designated as reference sites were those that represented the best available natural condition within the region, that is, sites with open access to the sea and the least evidence of human disturbance. These reference sites are used as a standard against which to assess the health of other sites with potential impacts. The reference site selection process involved two phases. The first was essentially "desk based' and was performed in consultation with staff from the local regulatory authority using topographic maps, global information systems (GIS) and local knowledge. From this initial phase, 165 potential reference and 35 impacted sites were selected and these were sampled over the Austral summer of 2000-2001 (Fig. 1). The 35 test sites had impacts included migratory barriers in the form of weirs and dams, eutrophication, and high densities of introduced piscivorous fish.

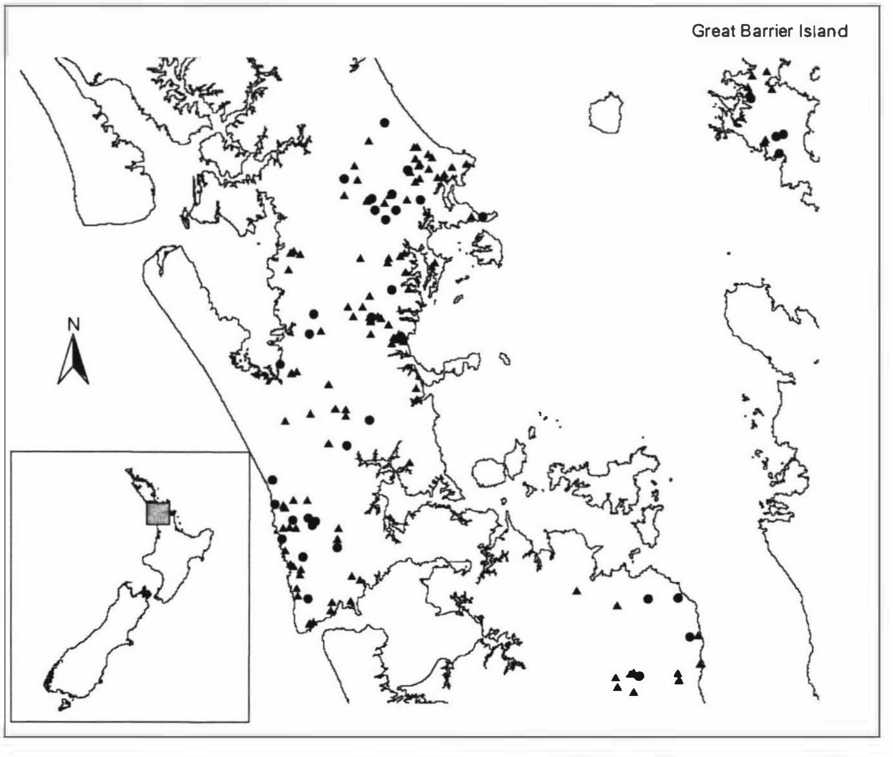

Fig. 1 Site map showing the reference sites (triangles) and test sites (circles) surveyed in the Auckland region between January and May 2001. Inset shows location in New Zealand.

The second phase took place post sampling and included updated information from site visits. To further refine the reference site dataset the sites were ranked based on the following criteria: 1) unimpeded access to the sea, 2) indigenous forested catchment, 3) mature exotic forest catchment (Hughes *et al.* 1986). The sites below the 25th percentile of the rankings were discarded leaving 118 reference sites for further analysis.

The aim of the model was to enable the determination of the biological condition of sites with respect to reference conditions; thus, only predictor variables that are uninfluenced or least influenced by human activity were included. Sixteen predictor variables were selected from the suite of variables available for each site from GIS databases or measures made during sampling (Table 1).

Fish and macro-crustacean data

Fish and macro-crustacean communities were sampled using overnight trapping. This is the most efficient sampling method for New Zealand fish found in the small, low gradient, low water clarity, streams common in this area (McDowall 1990). Although other sampling methods may have been more suitable at a few of the sites, the use of a consistent sampling method for all reference sites is imperative for data used in predictive modeling. At each survey site, two types of fish trap were used. The traps consisted of five pot type 'Gee minnow traps' and three large fyke nets. The pot traps were metal with 5 mm mesh (220 mm diameter) and two of the three nylon fyke nets (660 mm diameter hoops, 3 m long, 3 m gate) at each site had 12 mm mesh and the other had 2 mm mesh (see McDowall 1990 for details). Traps were not baited and were positioned where possible to cover all microhabitats over approximately 50m of the stream reach. Fyke nets were positioned with the entrance facing downstream and the gate angled across the stream. After 24 hrs in situ the traps were retrieved and fish and crustaceans were removed, measured, weighed and returned to the water. Juvenile eels (< 300-mm length) were treated as separate operational taxonomic units (elvers) in the analysis because they are known to shift habitat during ontogeny (Hayes *et al.* 1989, Glova 1998). This process was applied only to eels, because there is no evidence of an ontogenetic shift of within-species habitat requirements for the other species.

Table 1. Physical and chemical variables measured at the 118 reference and 35 test sites in the Auckland region between January 18 and May 2001 and selected for use as predictor variables in ANN models.

		Reference sites			**Test sites**		
Variable	Foot note	Min	Mean	Max	Min	Mean	Max
Geographic Variables							
Altitude (m)	1	1.0	47.9	220.0	2.0	62.1	200
Longitude (grid)	1	26397	26641	27282	26379	26662	27293
Latitude (grid)	1	64419	65074	65611	64522	65082	65560
Distance to coast (river km)	2	0.1	15.3	79.5	0.3	17.5	79
Channel variables							
Mean width (m)	3	0.8	3.3	31.8	0.8	3.8	17.3
Mean depth (m)	4	0.15	.42	1.96	0.02	0.47	1.42
Median substrate	5	0.0	9.1	39.1	0.0	6.2	40
Leaf litter	6	1.0	2.1	4.0	1.0	2.0	4
% Pool	7	0	75.3	100.0	0	59.4	100
% Riffle	7	0	12.7	70.0	0	24.3	100
% Run	7	0	10.8	90.0	0	15.7	100
Cobble packing	8	0	1.8	3.0	0	1.9	3
Water variables							
Temperature (° C)	9	12.9	16.5	24.0	13.5	16.8	20.9
Conductivity ($\mu S/cm^{-1}$)	9	100.00	216.07	974.00	99.4	222.2	897.0
Velocity (m/sec)	10	0	0.2	1.3	0	0.2	1.0
pH	11	7.1	8.1	8.9	6.3	8.0	8.6

1. Obtained from 1:50,000 NZMS topographic maps.
2. Geographic information systems (ARCVIEW) using 1:50,000 vector data
3. Mean of 5 measures over length of reach fished.
4. Mean of the maximum depth measured at the 5 points above.
5. Median substrate size index from 50 – 100 stones collected at random over the reach surveyed and measured in 11 classes (Wolman 1954).
6. Leaf litter visually assessed (0 = absent, 1 = rare, 2 = sparse, 3 = common, 4 = abundant)
7. Visually estimated at site
8. Subjectively assessed at site after moving substrate (1 = loosely packed; 4 = tightly packed)
9. Measured at time of fishing with YSI model 85 meter.
10. Calculated from time taken for a slug of dye to travel 20m over length fished
11. Orion Quickcheck model 106 pocket meter

Predicting fish and macro-crustacean assemblages from habitat data using ANN models

Artificial neural networks are derived from a simple model of the structure and function of the brain, and are characterised by their ability to 'learn'. This is achieved by comparing actual and desired outputs during the model-training phase.

In the training phase an algorithm modifies the internal parameters (weights) until the performance of the network, in this case prediction success, is maximized. For this case the presence or absence of each taxon was predicted using the back-propagation algorithm (Rumelhart *et al.* 1986). The architecture of the layering has been described by other authors (Lek *et al.* 1995, Mastrorillo *et al.* 1997b, Manel *et al.* 1999). The first layer, called the input layer, comprised of 16 cells represented each of the environmental variables. The second or hidden layer, was composed of a further set of neurones the number of which depended on the reliability required and the structure that best optimises bias and variance (Lek *et al.* 2000). In this application we used a network with a single hidden layer of three neurones (more layers and more neurones did not improve performance) trained through 100 iterations (SAS 1999). The third layer or output layer consisted of a single neurone responsible for the prediction of presence or absence of the taxon from the environmental variables.

Model evaluation

For each site output values in the range of 0.5 - 1 were interpreted as presence and values 0 - 0.5 as absence. We used several methods to assess the performance of the individual taxa models. First, all taxa models were assessed based on prediction success, which is the overall percentage of sites at which the presence or absence of each taxon was correctly predicted. This comparison involved using all 118-reference sites as training data and provided a lower boundary for the error probabilities (Fielding and Bell 1997). As an independent test, we employed *k*-fold-partitioning. In this partitioning method we randomly divided the reference sites into a training set of 80% (98 sites) of the sites and an independent validation set of 20% (24) of the sites (Manel *et al.* 1999). This process was repeated five times and the results pooled giving 120 sites for assessment of models. To deconstruct overall prediction success into separate elements, matrices of confusion were derived, after (Fielding and Bell 1997), in which true presence, false presence, true absence, and false absence were identified. From these values we calculated a range of performance measures: 1) sensitivity (percentage of true presence correctly identified), 2) specificity (percentage of true absence correctly identified), 3) false positive (percentage of actual absences wrongly predicted as being present), 4) false negative (percentage of actual presence wrongly predicted as being absent), 5) positive predictive power (percentage of true positives that were real) and 6) negative predictive power (percentage of predicted absences that were real) (Fielding and Bell 1997; Manel *et al.* 1999).

Relationships between taxa and environmental variables

The habitat variables associated with the predictions from the ANN models were assessed by obtaining pair-wise Spearman rank correlations (SAS 1999) between predictions of presence for each of the 12 taxa and the 16 predictor variables. To visualise these relationships between the predictions from the models and the environmental variables the probability of capture for each taxon was plotted against environmental gradients for each of the variables. For this, the environmental variables were arranged in ascending order then split into five groups of 24 sites. The mean probability of occurrence was calculated for each of the groups and this was repeated for all taxa.

Calculation of the expected number of taxa and O/E values

To identify the assemblages expected to occur at sites, the individual models (one for each taxon) were combined using the following protocol of (Wright *et al.* 1984). The probabilities of the predicted taxa are summed to give the expected number of taxa (*E*). The number of species actually captured at a site, providing they were predicted to occur is the observed number of taxa (*O*). The ratio of the observed to the expected number of taxa (*O/E*) and taxonomic composition is the output from the model used for assessing the biological quality of the test sites (see (Moss *et al.* 1987). To calibrate reference site *O/E* variability the reference sites were run through the ANN models and *O/E* ratios were calculated using the process described above.

Test sites

After satisfying the validation criteria the 12 trained ANN models, one for each taxon were applied to assess the status of 35 sites with potential impacts. Observed over expected ratios were calculated for the test sites using the process described above for reference sites. Analysis of variance was used to compare *O/E* ratios for test and reference sites using (SAS 1999). To evaluate how the actual occurrence of individual taxa related to the predicted patterns of occurrence the number of sites each taxon was predicted to occur at was plotted against the number of sites where it was observed.

Results

A total of 11 967 fish and macro-crustaceans from 23 species were caught at 200 sites during this study. Of these, 16 fish and the 2 crustacean species (a diadromous shrimp and a non-migratory crayfish) were native (see Table 2 for names). Eleven fish species, including all non-native species, were found at less than 2% (4) of the sites and were not used in any of the analyses. These rare species were not included because most were introduced and their rarity made association of habitat variables problematic. At the 118 reference sites, ten fish taxa (including elvers) and two crustacean species were used in analyses after the removal of rare species. The most common was the longfin eel occurring at 88% of the sites, and the rarest was the torrentfish at 4% of the sites (Table 2.)

Table 2. The twelve fish and crustacean species found at more than 2% of the 200 potential reference sites and the number of the 118 reference sites at which they were present.

Scientific name	*Common name*	Code	present	absent
Anguilla australis	shortfin eel	ANGAUS	25	93
Anguilla dieffenbachii	longfin eel	ANGDIE	105	13
Anguilla spp.	elvers	ELVERS	7	111
Cheimarrichthys fosteri	torrentfish	CHEFOS	5	113
Galaxias fasciatus	banded kokopu	GALFAC	73	45
Galaxias maculatus	inanga	GALMAC	51	67
Gobiomorphus basalis	Crans bully	GOBBAS	51	67
Gobiomorphus cotidianus	common bully	GOBCOT	23	95
Gobiomorphus huttoni	redfin bully	GOBHUT	42	76
Paranephrops planifrons	koura	PARANE	72	46
Paratya curvirostris	shrimp	PARATA	69	49
Retropinna retropinna	common smelt	RETRET	7	111

Table 3. Prediction assessment results for the ANN models. For each assessment measure the left column contains the results for the 120 pooled validation data set and the right column contains the results for the 118-reference sites used as training data for ANN model. See text for details on assessment measures.

	Prevalence %		% correct classification		Specificity		Sensitivity		False positive		False negative		Positive predictive power		Negative predictive power	
GALFAC	65	62	85	93	82	100	87	90	12	90	18	0	74	82	92	100
GOBCOT	18	19	79	90	87	99	40	67	38	67	13	1	88	88	61	97
GOBBAS	29	43	66	94	88	94	45	94	35	94	12	6	60	96	90	93
ELVERS	5	6	95	100	95	100	0	100	0	100	5	0	100	100	0	100
GALMAC	48	43	78	92	77	93	78	90	18	90	23	7	81	93	78	91
PARANE	58	61	68	91	93	97	64	88	26	88	7	3	25	78	99	99
ANGDIE	89	89	89	92	50	57	90	99	9	99	50	43	8	92	99	92
PARATA	59	58	64	89	62	97	65	85	26	85	38	3	33	76	90	99
GOBHUT	37	36	64	78	67	100	52	62	32	62	33	0	84	66	45	100
ANGAUS	26	21	76	95	78	95	57	95	30	95	22	5	93	99	38	81
RETRET	2	6	98	100	98	100	0	100	0	100	2	0	100	100	0	100
CHEFOS	3	4	98	99	98	99	50	100	33	100	2	1	99	100	50	80
Min	1	4	64	78	50	57	0	62	0	62	2	0	8	66	0	80
Max	89	89	98	100	98	100	90	100	38	100	50	43	100	100	99	100
Mean	36	37	80	93	81	94	52	89	22	89	19	6	70	89	62	94

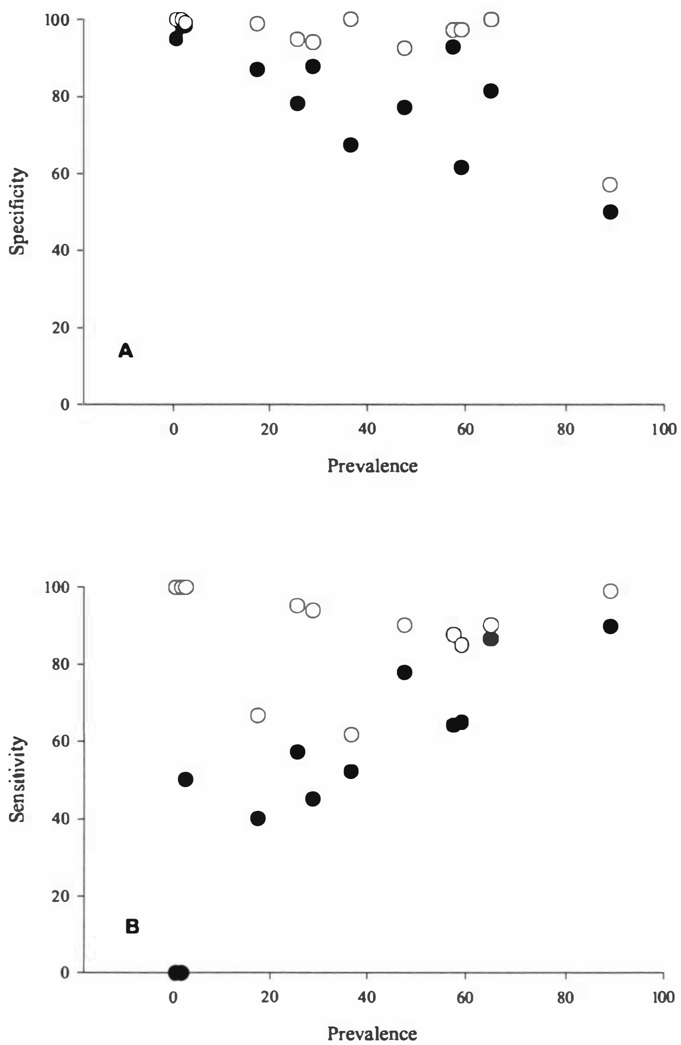

Fig. 2 The relationship between species prevalence and A) and specificity (percentage of true presence's correctly predicted) and B) sensitivity (percentage of true absences correctly predicted) for predictions of the presence of 12 fish and macro-crustacea taxa using ANN models. Open circles are 118 reference sites used in training and filled circles are 120 validation sites pooled from 5 fold partitioning.

Fitting and validating models

Training data-The overall prediction success rate was high for the training dataset with 93% of the predictions correct (ranging from 78% to 100%) (Table 3). The mean values for sensitivity, specificity, and were also high showing that prediction success was not linked to species prevalence (Table 3, Fig. 2). False positive and negative rates were also low but suggested some linkage to prevalence with the most prevalent taxon longfin eel having the highest false negative rate (43%).

Validation data - The average prediction success results for the independent data set pooled from the 5 fold partitioning also revealed a high overall success rate of 80% ranging from 64% to 98% (Table 3). However, inspection of the alternative assessment measures revealed a linkage between prevalence and prediction success. Common smelt and elvers occurred at only one or two sites in this dataset and their presence was not correctly predicted (i.e. sensitivity, false positive and negative predictive power = zero) (Table 3). The percentage of true absence correctly predicted (specificity) was considerably higher (mean 81%) than prediction of true absence (mean 52%) revealing that the model was better at predicting presence than absence (Fig. 2). This bias however, may relate to the imbalance in prevalence, as there are more rare than common species (mean prevalence = 36%, Table 3). The values for false positive, false negative, positive predictive power and negative predictive power revealed similar patterns with prediction success linked to prevalence.

Relationships between taxa and environmental variables

There is a complex mixture of associations between predictor variables and taxa (Table 4). Two variables latitude and pH, however, appeared to have little influence on any predictions. The highest correlations coefficients were between the three bully species and elevation and distance inland; the two migratory bully species were negatively correlated while the non-migratory Cran's bully was positively correlated with distance inland and elevation. Torrentfish show a negative relationship with the percentage of pool and a positive association with the percentage of run and riffle, while in contrast redfin bullies show opposite associations with percentages of pool and run. Banded kokopu show an affinity for small streams evidenced by the negative coefficients for width and depth.

Table 4. Spearman rank correlations > 0.30 between predictor variables and probability of capture for each taxon. Probabilities are from the neural network model using all reference sites (variables ranked by number of correlations > 0.30). (* indicates non-migratory taxa).

Predictor variable	GALFAC	GOBCOT	GOBBAS*	ELVERS	GALMAC	PARANE*	ANGDIE	PARATA	GOBHUT	ANGAUS	RETRET	CHEFOS
Distance inland		-0.75	0.75			0.45	0.35	-0.41	-0.71			
Altitude		-0.57	0.59		-0.48				-0.72	-0.4		
Width	-0.6	0.36			0.41			0.56				
Conductivity		0.4	-0.36				-0.42			0.43		
Temperature					0.44						0.37	
Velocity		-0.39					0.35		-0.36			
Longitude						0.46						-0.4
Depth	-0.3										0.5	
% Pool									0.45			-0.4
% Riffle				0.39								0.35
% Run									-0.42			0.35
Substrate size										-0.37		
Leaf litter												0.36
Cobble packing							0.54					
Latitude												
pH												

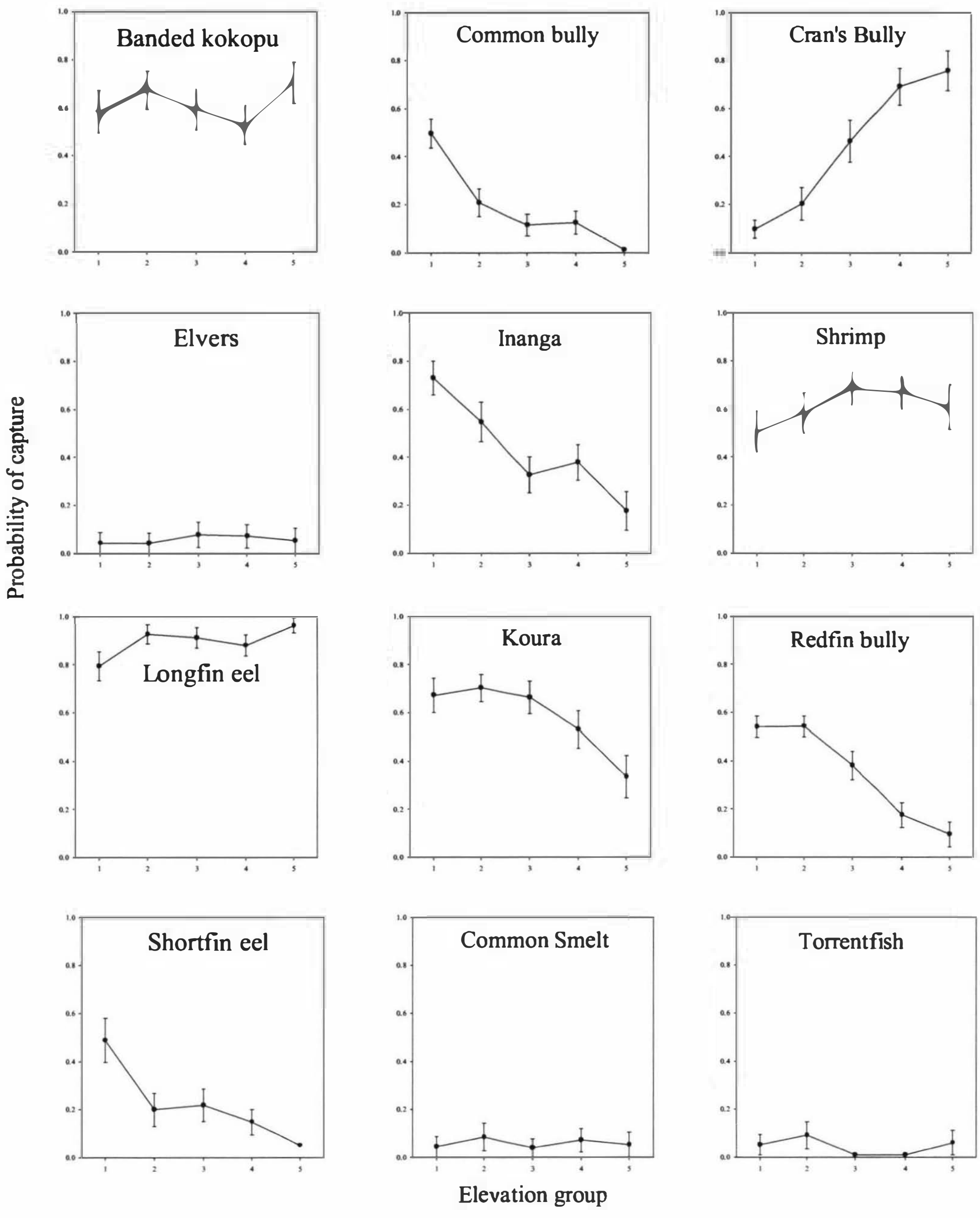

Fig. 3 Mean probability (± SE) of capture of 12 taxa in five groups of 24 sites over an elevational gradient from the sea (left) to 260 m a.s.l. (right). The mean elevation in each of the five groups is 6, 20, 39, 61 and 105 m a.s.l. respectively

To visualise the relationships an example, Fig. 3 shows the probabilities of capture for each of the taxa plotted against the elevational gradient. The three bully species (common, Cran's and redfin), shortfin eel and the crayfish (koura) showed a strong relationship with the elevational gradient. Five taxa showed reducing probability of capture with elevation, they were common and redfin bullies, inanga, koura, and shortfin eel. Longfin eel and the non-migratory Cran's bully showed increasing probability of capture with elevation. There was no discernable relationship with elevation for the rare taxa: smelt, torrentfish and elvers or the more abundant banded kokopu.

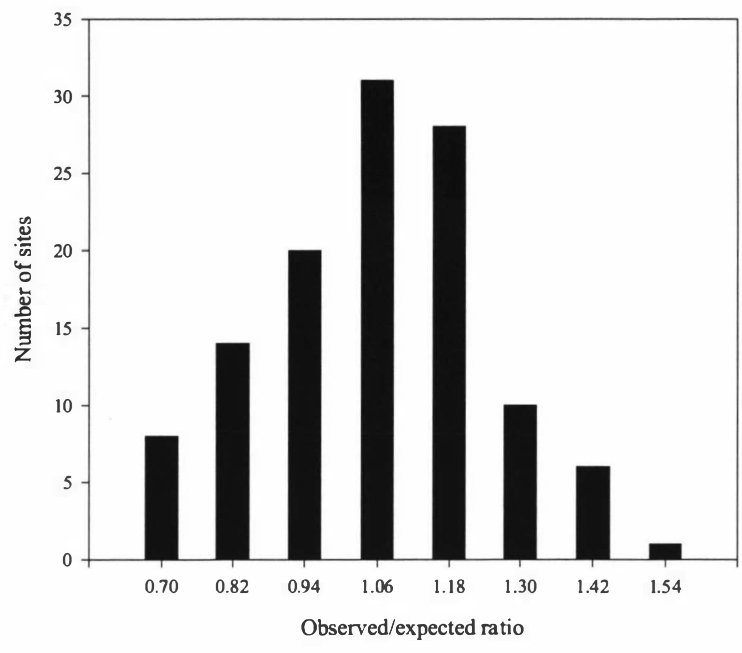

Fig. 4 The distribution of *O/E* ratios from the 118 reference sites using predictions from the ANN models for the twelve taxa.

Calculation of observed over expected ratios

All reference sites were run through the ANN models and predictions were used to calculate *O/E* ratios. The distribution of these *O/E* ratio values for the reference sites approximated a normal distribution (Fig. 4). The *O/E* ratios were centered on unity with a mean of 0.99 (standard error 0.02) and ranged between 0.6 and 1.41. The plot of observed versus expected species number (Fig. 5) shows a strong relationship between the two as revealed by the slope and intercept very close to 1 (observed taxa richness = - 0.11 + 1.02 predicted taxa richness; $r^2 = 0.64$).

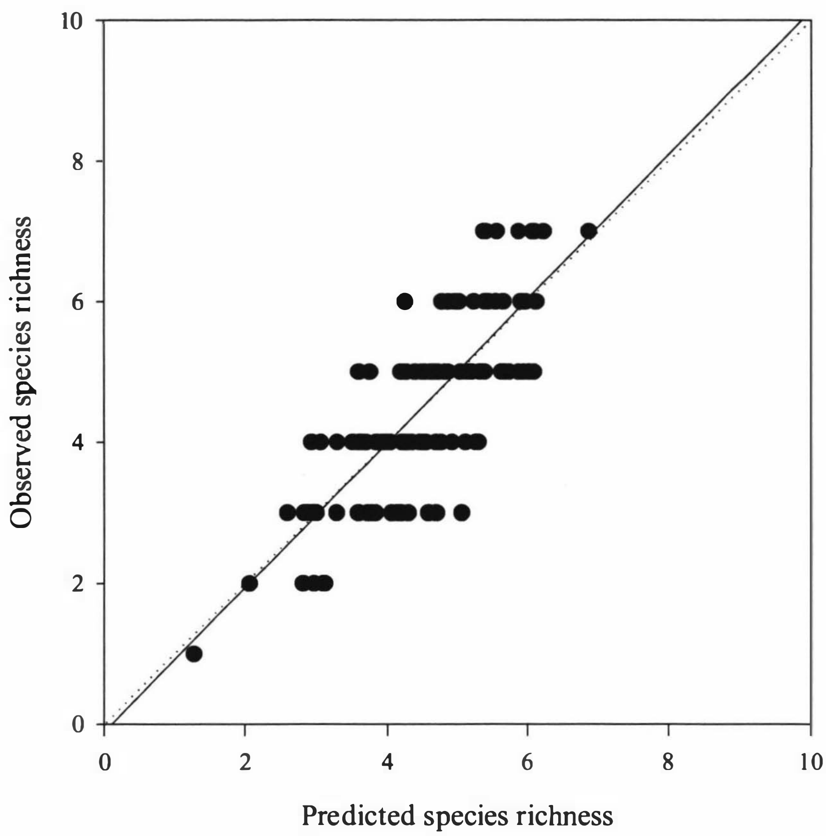

Fig. 5 Observed species richness plotted against expected species richness from the 118 reference sites using predictions from the ANN models for the twelve taxa. The dotted line has a slope of 1.

Assessment of test sites

The 35 test sites with potential impacts were run through the ANN models and predictions were used to calculate *O/E* ratios. The mean *O/E* ratio for the test sites (0.69, SE = 0.04) was significantly lower than the reference site *O/E* mean ($F_{1,152}$ = 62.9, $P < 0.0001$). The plot of observed versus expected number of sites revealed the individual taxa that occurred at fewer sites than expected and thus, contributed to the lower overall *O/E* ratios (Fig. 6). The two galaxiid species (inanga and banded kokopu) as well as the two macro crustaceans (koura and shrimp) and the two migratory bullies (redfin and common) occurred at considerably less sites than expected. Three taxa however, occurred at slightly more sites than expected they were elvers, the non-migratory Cran's bully and shortfin eels.

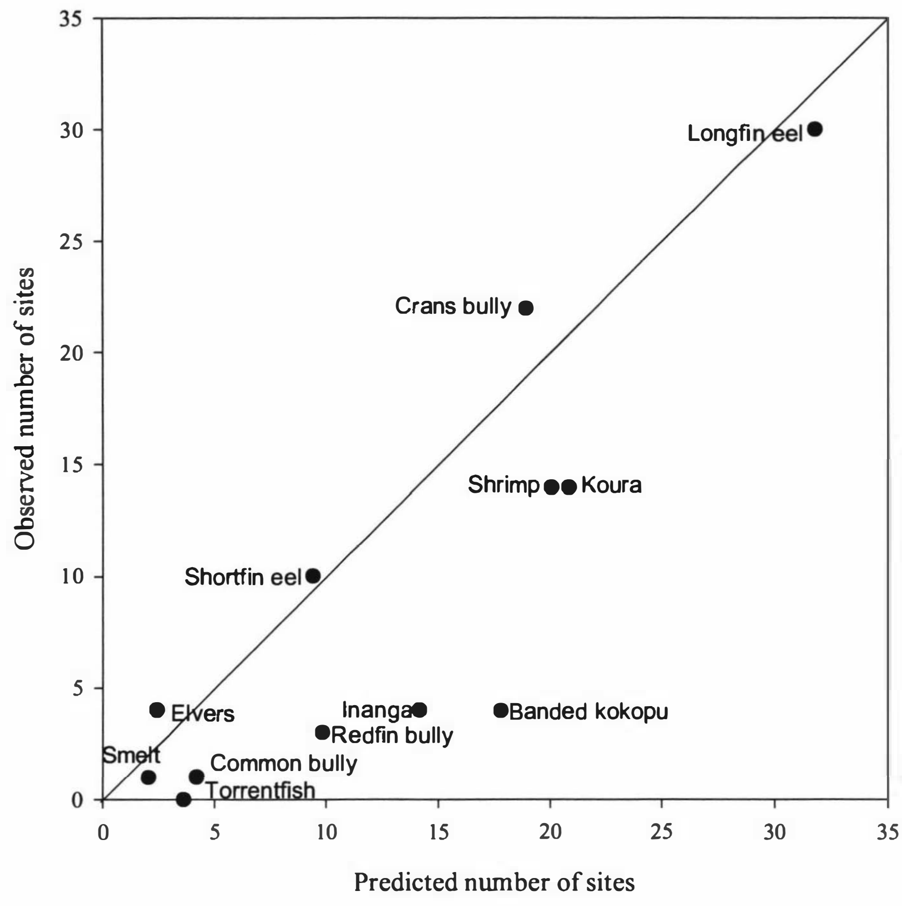

Fig. 6 Predicted number of sites plotted against observed number of sites for the 35 test sites, the diagonal line represents the line of perfect agreement.

Discussion

This project was conceived to make available a practical tool for the assessment of a significant part (fish and macro-crustaceans) of the biological integrity of stream sites in a region of New Zealand. It allows for an objective, statistically robust, site-specific prediction of the fish assemblage to be made without the requirement for extensive analysis. This is because the sophisticated statistical knowledge used to create the model is not required for its implementation. The model can be made available for use by others requiring only the input of environmental variables of a site to give a list of expected taxa. A RIVPACS type predictive model of biological quality using fish has been developed in New Zealand based on biotic site groups (Joy and Death, 2002) and individual species discriminant function analysis (DFA) (Joy and Death, in press). The ANN models reported here were applied to same data used in the DFA models above (Joy and Death, in press) and similar results were obtained. Discriminant function analysis was slightly less accurate than ANN using all training data but DFA was marginally more accurate than ANN when considering holdout validation data (Joy and Death, in press). This similarity of results suggests that the relationships being modeled are linear.

The species-specific models we have developed appear to be methodologically, and ecologically robust for comparing reference and impacted sites. The species that occurred at less test sites than expected (Fig. 6): inanga, banded kokopu koura, shrimp, redfin and common bullies are the same species that have been observed to have reduced densities and/or sensitivity to anthropogenic impacts in other studies (Richardson *et al.* 1994, Dean and Richardson 1997, Richardson 1997, Richardson *et al.* 1998, Rowe *et al.* 1999, Rowe *et al.* 2000, Richardson *et al.* 2001). The species occurring at more sites than predicted (Cran's bullies and shortfin eels) have similarly been shown to be tolerant to factors associated with human impacts (Richardson *et al.* 1994, Rowe *et al.* 1999, Rowe *et al.* 2000). The correlations between the taxa and environmental variables from the reference data also revealed ecologically realistic associations. Examples include the reducing probability of capture with elevation and distance from the sea for the diadromous species, and the association of banded kokopu with small streams (McDowall 1990, Joy and Death 2000).

Model evaluation

The effectiveness of predictive models has in the past generally been judged based on prediction success alone (Mastrorillo *et al.* 1997a, Mastrorillo *et al.* 1997b, Oberdorff *et al.* 2001). Recent work by Fielding and Bell (1997) and Manel *et al.* (1999) have revealed the importance of assessing the different elements of prediction success separately and our results support these recommendations. In this study, the models were much better at predicting absence (mean specificity 81%) than presence (mean sensitivity rate 52%) when using validation data (Table 3; Fig 2). This is despite the apparently high percentage of correct predictions. However, the effect of prevalence on predictions was much more noticeable with the 20% *k* fold validation data than the training data. There are two potential explanations for this pattern. First that the reduction in number of training sites reduces the model precision, especially for the rarer taxa through a lack of experience of the conditions at these sites. The second explanation is that there was some overtraining of the total data set. Overtraining occurs when the network attempts to model noise in the data rather than real patterns (Walley and Fontama, 1998).

The models generally performed well when applied to independent validation data suggesting that the predictions are accurate. The validity of the models was further supported by the ecologically meaningful associations that were found between taxa predictions and environmental variables (Table 4). We are thus confident that the *O/E* ratios produced from the models provide an accurate assessment of the expected biological condition at a site in the absence of human impacts. The predominantly diadromous fauna of New Zealand has negated the application of other bioassessment tools such as the index of biotic integrity (McDowall and Taylor, 2000). This ANN model for site assessment takes into account not only conditions at the site but also from the site to the coast as the diadromous species are dependent on access to and from the sea as well as proximal habitat and catchment quality. Thus, an *O/E* ratio indicates biological quality of the waterway at many scales over the whole catchment from the source to sea.

Conclusions

In this study, we developed an empirical model capable of predicting the stream fish and macro-crustacean assemblages expected to occur in the absence of human impacts using a suite of environmental variables. Expected assemblages based on data from reference sites were compared to those observed at a number of potentially impacted sites, and the deviation between the two measures provided a measure of the magnitude of degradation of biological quality. This process also identifies individual taxa responsible for the deviations between observed and expected assemblages, which then potentially allows for diagnosis of the relationship between individual taxa and impacts. This diagnostic process can be achieved by correlating individual components (taxa) of the biological structure with known impacts.

The model developed here provides a rapid and powerful technique for assessing the biological condition of the fish and macro-crustacean fauna of a stream and can potentially identify targets for management or rehabilitation. Therefore, it has potential for assisting with stream management in New Zealand because the stream is assessed from headwaters to the coast and migration induced trajectories of occurrence are taken into account. Furthermore, the method applied here also has potential for application in other countries with a high proportion of diadromous species.

Acknowledgments

This work was funded by a grant from the sustainable management fund (New Zealand Ministry for the Environment grant No. 5099 "Nga Ika Waiora": stream health evaluation tool) and the Auckland Regional Council. Thanks to Allison Hewitt for assistance with all fieldwork and to Auckland Regional Council staff Brent Evans, Chris Hatton, David Glover, and John Maxted. For assistance with site access, information, and accommodation, we thank Conservation Department staff Chris Roberts, Thelma Wilson, Jo Ritchie, and Mike McGlynn.

References:

Chessman, B. C., I. O. Growns, and N. Plunket-Cole. 1999. Predicting diatom communities at genus level for the biological management of rivers. Freshwater Biology **41**:317-331.

Davies, N. M., R. H. Norris, and M. C. Thoms. 2000. Prediction and assessment of local stream habitat features using large scale catchment characteristics. Freshwater Biology **45**:343-369.

Dean, T., and J. Richardson. 1997. Native fish survival during exposure to low levels of dissolved oxygen. Water and atmosphere **5**:12-14.

Fielding, A. H., and J. F. Bell. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environmental Conservation **24**:38-49.

Glova, G. J. 1998. Factors associated with the distribution and habitat of eels (*Anguilla* spp.) in three New Zealand lowland streams. New Zealand Journal of Marine and Freshwater Research **32**:255-269.

Hayes, J. W., J. R. Leathwick, and S. M. Hanchet. 1989. Fish distribution patterns and their association with environmental factors in the Mokau River catchment, New Zealand. New Zealand Journal of Marine and Freshwater Research **23**:171-180.

Hawkins, C. P., R. H. Norris, J. N. Hogue, and J. W. Feminella. 2000. Development and evaluation of predictive models for measuring the biological integrity of streams. Ecological Applications **10**:1456-1477.

Hughes, R. M., D. P. Larsen, and J. M. Omernik. 1986. Regional reference sites: a method for assessing stream potentials. Environmental Management **10**:629-635.

Joy, M. K., and R. G. Death. 2000. Development and application of a predictive model of riverine fish community assemblages in the Taranaki region of the North Island, New Zealand. New Zealand Journal of Marine and Freshwater Research **34**:243-254.

Joy, M. K., I. M. Henderson, and R. G. Death. 2000. Diadromy and longitudinal patterns of upstream penetration of freshwater fish in Taranaki, New Zealand. New Zealand Journal of Marine and Freshwater Research **34**:531-543.

Joy, M. K., and R. G. Death. 2002. Predictive modeling of freshwater fish as a biomonitoring tool in New Zealand. Freshwater Biology **47**:2261-2275.

Joy, M. K., and R. G. Death. (In Press) Assessing biological integrity using freshwater fish and decapod habitat selection functions. Environmental Management.

Karr, J. R. 1981. Assessments of biotic integrity using fish communities. Fisheries **6**:21-27.

Karr, J. R. 1991. Biological integrity: a long neglected aspect of water resource management. Ecological Applications **1**:66-84.

Karr, J. R. 1994. Using biological criteria to protect ecological health. *in* D. Rapport, C. Gaudet, and P. Calow, editors. Evaluating and Monitoring the Health of Large Scale Ecosystems. Springer-Verlag, New York.

Karr, J. R. 1995. Protecting freshwater ecosystems: clean water is not enough. *in* W. S. D. a. T. P. Simon., editor. Biological Assessment and Criteria: tools for water resource planning and decision making. Lewis publishers, Boca Raton.

Lek, S., A. Belaud, I. Dimopulous, J. Lauga, and J. Moreau. 1995. Improved estimation, using neural networks, of the food consumption of fish populations. Marine and Freshwater Research **46**:1229-1236.

Lek, S., J. L. Giraudel, and J. Guegan. 2000. Neuronal Networks: algorithms and architectures for ecologists and evolutionary ecologists. Pages 262 *in* S. Lek and J. Guegan, editors. Artificial Neuronal Networks. Springer-Verlag, Berlin.

Manel, S., J. Dias, and S. J. Ormerod. 1999. Comparing discriminant analysis, neural networks and logistic regression for predicting species distributions: a case study with a Himalayan river bird. Ecological Modelling **120**:337-347.

Mastrorillo, S., S. Lek, and F. Dauba. 1997a. Predicting the abundance of minnow *Phoxinus phoxinus* (Cyprinidae) in the River Ariege (France) using artificial neural networks. Aquatic Living Resources **10**:169-176.

Mastrorillo, S., S. Lek, F. Dauba, and A. Belaud. 1997b. The use of artificial neural networks to predict the presence of small-bodied fish in a river. Freshwater Biology **38**:237-246.

McDowall, R. M. 1990. New Zealand Freshwater Fishes: A Natural History and Guide. Heinemann Reed, Auckland.

McDowall, R. M., and M. J. Taylor. 2000. Environmental indicators of habitat quality in a migratory freshwater fish fauna. Environmental Management **25**:357-374.

Moss, D., M. T. Furse, J. F. Wright, and P. D. Armitage. 1987. The prediction of the macro-invertebrate fauna of unpolluted running-water sites in Great Britain using environmental data. Freshwater Biology **17**:41-52.

Oberdorff, T., D. Pont, B. Hugueny, and D. Chessel. 2001. A probabilistic model characterizing fish assemblages of French rivers: a framework for environmental assessment. Freshwater Biology **46**:399-415.

Reynoldson, T. B., R. C. Bailey, K. E. Day, and R. H. Norris. 1995. Biological guidelines for freshwater sediment based on BEnthic Assessment of SedimenT (the BEAST) using a multivariate approach for predicting biological state. Australian Journal of Ecology **20**:198-219.

Reynoldson, T. B., and J. F. Wright. 2000. The reference condition: problems and solutions. Pages 293-303 *in* J. F. Wright, D. W. Sutcliffe, and M. T. Furse, editors. Assessing the biological quality of freshwaters. RIVPACS and other techniques. Freshwater biological association, Ambleside, United Kingdom.

Richardson, J. 1997. Acute ammonia toxicity for eight New Zealand indigenous freshwater species. New Zealand Journal of Marine and Freshwater Research **31**:185-190.

Richardson, J., J. Boubee, T. Dean, D. Rowe, and D. West. 1998. Effects of suspended solids on migratory native fish. Water and Atmosphere **6**:22-23.

Richardson, J., J. A. T. Boubee, and D. W. West. 1994. Thermal tolerance and preference of some native New Zealand freshwater fish. New Zealand Journal of Marine and Freshwater Research **28**:399-407.

Richardson, J., D. K. Rowe, and J. P. Smith. 2001. Effects of turbidity on the migration of juvenile banded kokopu *(Galaxias fasciatus)* in a natural stream. New Zealand Journal of Marine and Freshwater Research **35**:191-196.

Rowe, D. K., B. L. Chisnall, T. L. Dean, and J. Richardson. 1999. Effects of land use on native fish communities in east coast streams of the North Island of New Zealand. New Zealand Journal of Marine and Freshwater Research **33**:141-151.

Rowe, D. K., M. Hicks, and J. Richardson. 2000. Reduced abundance of banded kokopu *(Galaxias fasciatus)* and other native fish in turbid rivers of the North Island New Zealand. New Zealand Journal of Marine and Freshwater Research **34**:547-558.

Rumelhart, D. E., G. E. Hinton, and R. J. Williams. 1986. Learning representations by back-propagating errors. Nature **323**:533-536.

SAS. 1999. SAS Enterprise Miner. *in*. SAS Institute Inc., SAS Campus Drive, Cary, North Carolina 27513, USA.

Wolman, M. G. 1954. A method of sampling coarse river-bed material. Transactions, American Geophysical Union **35**:951-956.

Wright, J. F., D. Moss, P. D. Armitage, and M. T. Furse. 1984. A preliminary classification of running water-sites in Great Britain based on macro-invertebrate species and the prediction of community type using environmental data. Freshwater Biology **14**:221-256.

Wright, J. F., D. W. Sutcliffe, and M. T. Furse. 2000. Assessing the biological quality of fresh waters: rivpacs and other techniques. Freshwater biological association, Ambelside, Cumbria, UK.

CHAPTER FOUR

Assessing biological integrity using freshwater fish and decapod habitat selection functions

Joy, M. K., and R. G. Death. In Press. Assessing biological integrity using freshwater fish and decapod habitat selection functions. Environmental Management.

Abstract

Comparison between the number of taxa observed and the number expected in the absence of human impact is an easily understood and ecologically meaningful measure of biological integrity. This approach has been successfully applied to the assessment of the biological quality of flowing water sites using macroinvertebrates with the River Invertebrate and Classification System (RIVPACS) and its derivatives. In this paper, we develop a method similar to the RIVPACS predictive model approach to assess biological integrity at flowing-water sites using freshwater fish and decapod assemblages. We extend the RIVPACS approach by avoiding the biotic classification step and model each of the individual species separately. These assemblages were sampled at 118 least impacted (reference) sites in the Auckland region, New Zealand. Individual discriminant models based on the presence or absence of the 12 most common fish and decapod species were developed. Using the models, predictions were made using environmental measures at new sites to yield the probability of the capture of each of the 12 species and these were combined to predict the assemblage expected at sites. The expected assemblage was compared to that observed using an observed over expected ratio (*O/E*). The models were evaluated using a number of internal tests including jackknifing, data partitioning and the degree to which *O/E* values differed between reference sites and a set of sites perceived to be impaired by human impacts.

Keywords. Diadromy, freshwater fishes, discriminant analysis, biomonitoring, biological quality, biological assessment, fish assemblages.

Introduction

The pressures placed on natural resources and the environment by humans are growing exponentially in line with human population growth. This makes it imperative that we find scientifically robust ways of providing accurate ecological assessment, preferably without the requirement for high levels of end user expertise for their use. Aquatic communities are altered structurally and functionally following human induced disturbance. The impacts take many forms but include altered nutrient, sediment, and flow regimes, biotic interactions, and placement of barriers to migration. To quantify these alterations to communities a benchmark is required against which changes can be compared. Benchmarks in the form of biotic communities occurring in unimpacted areas provide a 'biocriterion' upon which comparisons can be based (Bailey *et al.* 1998). The areas where least impacted conditions occur provide potential for establishment of reference sites (Karr 1981, Fausch *et al.* 1984, Bailey *et al.* 1998). Information available from reference site communities can then be used to quantify expectations at a test site using predictions based on the relationship between the biotic and abiotic components of each site. This process enables an objective assessment to be made of the biological condition expected at other sites.

Biological indicators for freshwater ecosystems can be chosen from a suite of animal or plant assemblages or combinations of them. Fish and macro-crustacea are particularly suitable because they fit many of the criteria required of bioindicators (e.g. Hellawell 1986, Harris 1995, Oberdorff *et al.* 2001). Present in many waterways fish and decapods are: 1) relatively easily identified, 2) have both economic and aesthetic values, 3) are primarily affected by macro-environmental influences, unlike algae and macroinvertebrates that are affected by both micro and macro-environmental influences, 4) relatively long-lived, providing temporal integration in assessments, and 5) often at the apex of stream food webs, they integrate many stream processes (Harris 1995, Oberdorff *et al.* 2001).

The predominantly diadromous fish fauna of New Zealand have rarely been used in biological assessment in the past because of the overwhelming influence of altitude and distance from the sea on their distribution (McDowall 1998a, 1998b, Joy and Death 2000, 2001), except for Joy and Death (2002). These overriding influences result in a breakdown of the relationship between fish occurrence and proximal habitat quality (Joy and Death 2001). Thus, the application of the index of biological integrity (IBI) in New Zealand is problematical because the index relies on the relationship between population metrics and habitat quality (McDowall and Taylor 2000).

A predictive reference condition approach (RIVPACS: *R*iver *I*n*V*ertebrate *P*rediction *A*nd *C*lassification *S*ystem) using macroinvertebrates originally developed by (Wright *et al.* 1984) and later advanced by (Reynoldson *et al.* 1995), Simpson and Norris (2000) and Hawkins *et al.* (2000) has been very successful as a biomonitoring tool in the United Kingdom, and Australia and the United States of America. The predictive approach used in RIVPACS models has rarely been used with stream assemblages other than invertebrates, although proponents have suggested other groups including plants, and fish could be used (Reynoldson and Wright 2000). Moreover, a RIVPACS approach to biomonitoring has been developed in a New Zealand region using freshwater fish (Joy and Death 2002).

The RIVPACS procedure and its derivatives mentioned above have a feature that may limit their performance. This feature is the intermediate step of site classification, which is used to relate biological and physical data by way of a weighting procedure to calculate the probability of site membership to a biotic group using discriminant analysis. In effect, the sites are clustered into groups based on their biota and then joined again by weighted averaging. Chessman (1999) pointed out that this clustering step might be questionable as the biotic communities are a continuum rather than discrete groups, and thus any clustering will be artificial and arbitrary.

The process we outline in this paper uses aspects of the RIVPACS method, but avoids one of the criticisms of the approach by not partitioning the data into biotic groups (Chessman 1999). Instead, we model the presence or absence of the 12 most common fish and crustacean species individually in relation to environmental factors using discriminant analysis. To make our predictions independent of human impacts we use models generated from environmental variables unlikely to be influenced by humans. The discriminant functions are then used to predict, based on the environmental data measured at a test site, the probability of finding each of the species. To assess a site, the expected assemblage is compared with that actually found at the site. The measure of the biological condition of the site is the divergence of the observed assemblage from that expected in the absence of human impacts.

In this paper, we have two aims. First, to investigate if it is possible to accurately predict the occurrence of individual fish and decapod species from environmental data. Second, to develop a method to use those predictions as a tool to assess the biological quality of streams in the Auckland Region of New Zealand. We envisage that this method will have applications in biological surveillance, conservation, pollution control, and other aspects of water management in New Zealand and worldwide.

Methods

Study area

To cover all the stream types, 200 sites were sampled in the Auckland region between January and May 2001 (Fig. 1). This summer/autumn period is the time when all diadromous species are present in New Zealand streams (McDowall, 1990). The region is in the north of the North Island at c. 37° south 176° east. With a land area of approximately 5000 km^2, it contains New Zealand's largest city and 33% of the national human population. It is a long and narrow region approximately 85 km from top to bottom and is less than 13 km from coast to coast at the narrowest point and 55 km at the widest. The region is bisected by numerous tidal inlets and estuaries, and more than 90% of the streams are first or second order.

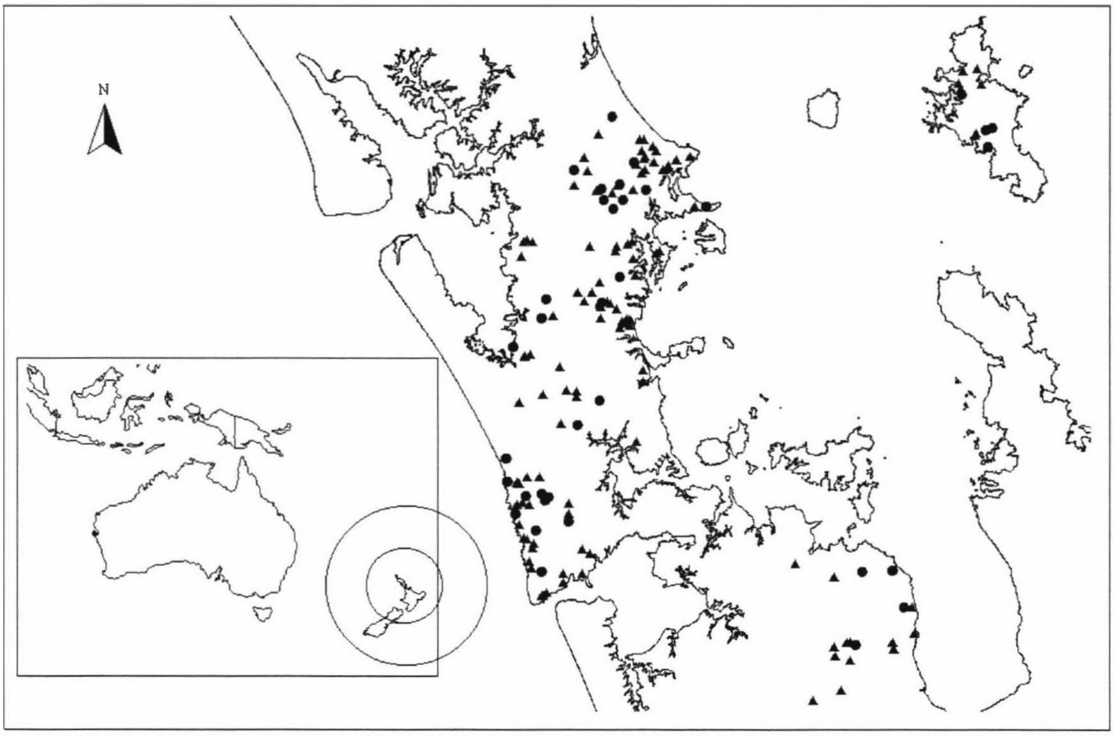

Fig. 1 Site map showing the reference sites (triangles) and test sites (circles) surveyed in the Auckland region between January and May 2001. Inset shows New Zealand in relation to Australia.

Selection of reference sites

Reference sites were initially selected to represent the best available natural condition within the region, that is, those with open access to the sea and least evidence of human disturbance. Emphasis was placed on access to the sea because of the high proportion of diadromy in the New Zealand freshwater fish fauna (McDowall 1990) and one of the two decapods (the diadromous shrimp) and the impact this has on their distribution (McDowall 1998, Joy and Death 2001). The region has been divided into seven broad areas based on stream type by the local territorial authority (Auckland Regional Council, (ARC)) and the potential reference sites were stratified in an attempt to cover these areas over the range of elevations in the region.

Reference site selection occurred in two phases: for the first phase, site selection was 'desk-based' using expert knowledge of streams combined with geographic information. The second phase occurred after sampling with the inclusion of up to date information from the field. The initial phase was conducted in consultation with ARC environmental personnel using 1:50 000 topographic maps and global information systems (GIS) (Arcview 1999). After the field survey, updated site and catchment condition information was used to refine criteria for reference site retention (Hughes *et al.* 1986).

Of the 200 potential reference sites, 35 had on inspection previously unknown impacts and were discarded as potential reference sites but retained as test sites for model calibration. The impacts discovered included migratory barriers in the form of weirs and dams, eutrophication; high densities of introduced piscivorous fish, and a natural impact at one site near a geothermal outflow. Other sites had more subtle impacts such as sedimentation from exotic forest harvesting and/or restricted passage. To refine further the reference site dataset the sites were ranked based on the following criteria in order: 1) unimpeded access to the sea, 2) indigenous forested catchment, 3) mature exotic forest catchment (Hughes *et al.* 1986). The sites in the lower 25% of the ranking were discarded leaving 118 reference sites for further analysis.

Fish and decapod data

Fish and decapod communities were sampled with overnight trapping. This sampling method is the most efficient for the New Zealand fish found in the small, low gradient, low water clarity, streams common in this area of New Zealand (McDowall 1990). Although other sampling methods may have been more suitable at a few of the sites, the use of a consistent method for all reference sites is imperative for predictive modeling. At each survey site, two types of fish trap were used. The traps consisted of five, pot type ‘G-minnow’ traps and three large fyke nets. The pot traps were metal with 5 mm mesh (220 mm diameter) and two of the three nylon fyke nets (660 mm diameter hoops, 3 m long, 3 m gate) at each site had 12 mm mesh and the other had 2 mm mesh (see McDowall 1990 for details). Traps were not baited and were positioned where possible to cover all microhabitats over approximately 50 m reach. Fyke nets were positioned with the entrance facing downstream and the gate angled across the stream. After 24 hrs in situ the traps were retrieved and fish and crustaceans were removed counted measured weighed and returned to the stream. Juvenile eels (< 300-mm length) were treated as separate operational taxonomic units (elvers) in the analysis because they are known to shift their habitat niche during ontogeny (Hayes *et al.* 1989, Glova 1998). This process was applied only to eels, because there is no evidence of ontogenetic shifts within species habitat requirements for the other species.

Environmental variables

Physical and chemical variables were measured at each site to complement species lists. Site elevation, and distance from the sea were calculated using a geographic information system (GIS) (ARCVIEW 3.2 1999) and 1:50 000 NZMS topographic maps (Information 1987). Conductivity (automatically adjusted to 25°C), temperature, and salinity were measured on site with a YSI model 85 meter, and pH was measured with an Orion Quickcheck model 106 pocket meter. Mean water depth was the average of five measures taken at equidistant points along the thalweg surveyed using a staff, and the mean width of the wetted area was averaged from measures taken at each of the five points. The standard deviation of the five measures of stream depth and width were calculated and used as a measure depth heterogeneity and sinuosity respectively. Median substrate size was calculated by summing the mid-point values of the size classes weighted by their proportional cover from measurements of 70-100 individual particles collected in 11 phi (log) size classes over the reach fished (Wolman 1954, Quinn and Hickey 1990). The percentage of backwater, pool, run, riffle or rapid was visually estimated over each reach surveyed. Riffles were classified as areas of fast, shallow water with a broken-surface appearance; pools were areas of slow deep water with a smooth surface appearance, whereas runs were intermediate in character. Mean water velocity was calculated by timing the movement of the modal point of a slug of dye over 20 m of the reach trapped. The amount of leaf litter on the substrate was estimated in 4 abundance classes (absent, rare, common and abundant). Cobble packing was subjectively assessed by moving substrate by hand.

Discriminant analyses

The relationships between environmental variables and the presence or absence of each of the fish and crustacean species were modeled using discriminant analysis (SAS 1996). Discriminant analysis has been extensively used in ecological investigations (Wiegleb 1981, Rice *et al.* 1983, Williams 1983, Manel *et al.* 1999). The object of discriminant analysis is (a) to exhibit optimal separation of groups and (b) to predict the probability of group membership for an observation (Williams 1983). In this study, the analysis was initially used to find the discriminant function, which best separated sites into two groups i.e. sites where a particular taxon was present and sites where they are absent. Later the function was used to calculate the probability of that taxon being captured at a site given a set of environmental conditions. The emphasis in this application was on predictive power as this is the requirement of the biomonitoring tool being developed.

The objective of the final combined predictive model is to predict the fauna expected at sites lacking human impacts, therefore only the variables considered unlikely to be affected by human activities were included in the construction of the models. To evaluate the importance of each environmental variable in discriminating between the presence-absence groups, coefficients of canonical variables were examined to relate the site groupings to the environmental data. The five highest coefficients of the canonical variables were recorded for each taxon and the coefficients were ranked over all taxa to assess the most important variables in the prediction of assemblages. Linear discriminant functions are calculated on an assumption of equal within group variances. To help satisfy this assumption and improve the discrimination the environmental variables were log transformed (ln $(x+1)$) to improve normality and heteroscedasticity.

Validation of models

We used several methods to assess the performance of the individual models. The discriminant model output in each case has a value within the range of 0 - 1 derived from the Mahalanobis distances to the group centroids of presence and absence groups. The outputs from these models were evaluated by considering the posterior probability error rates for each of the species models for all reference sites. The resubstitution comparison involved using all 118-reference sites as training data and provided a lower boundary for the error probabilities (Fielding and Bell 1997). This assessment was included to enable comparison with RIVPACS models that use resubstitution for model validation (Wright *et al.* 1984). As a further assessment of each model, a leave-one-out or jackknife procedure was also used (Manel *et al.*

2001). In the process a single site is removed, the model reconstructed, and the probability of the presence or absence of the fish calculated; and then this process is repeated for all reference sites (SAS 1996). As a further independent test, we employed *k*-fold-partitioning (Fielding and Bell 1997). Huberty (1994) suggests using a ratio of $[1 + (p - 1)^{1/2}]^{-1}$, where p is the number of predictors, in this case it equates to retaining approximately 80% of the data for training. Thus, 20 sites were randomly selected and removed, then the remaining 98 sites were used to rebuild the discriminant model. The 20 removed sites were run through the discriminant model, and the percentage of sites correctly assigned to their presence/absence groups calculated. This process was repeated five times and the number of sites correctly assigned calculated to give 100 partitioned sites. For each of the species, matrices of confusion were derived, after (Fielding and Bell 1997), in which true presence and true absences were identified. These values were used to give measures of sensitivity (percentage of true presence correctly identified) and specificity (percentage of true absence correctly identified). To assess the global relationship between the number of sites each taxon was predicted to occur at and the number of sites they were observed at, these values were compared using linear regression.

Calculation of the expected number of taxa and O/E values

To identify the assemblages predicted to occur at sites the individual models (one for each taxon) were combined using the following procedure originally described by (Wright *et al.* 1984). The probabilities of the predicted taxa are summed to give the 'expected number of taxa' (*E*). The number of species actually captured at a site, providing they were predicted to occur is the 'observed number of taxa' (*O*). The ratio of the observed to the expected number of taxa (*O/E*) and taxonomic composition is the output from the model (see (Moss *et al.* 1987).

The model error range was estimated by comparing the number of taxa observed at reference sites to that predicted by the discriminant models and generating a distribution of reference site *O/E* ratios (Hawkins *et al.* 2000). This distribution of errors is then used to determine if the *O/E* ratio of a test site is outside that expected from model error alone (Clarke 2000, Hawkins *et al.* 2000).

Assessment of test sites

Environmental data from test sites were input to the models, and *O/E* ratios were calculated as described. Analysis of variance (ANOVA) was then used to determine if test site *O/E* ratios differed from those of reference sites. To assess whether particular individual taxa occurred more or less often than expected at test sites, the total number of sites where each taxon actually occurred was plotted against the expected number of sites. To assess the relationship between the complete observed and expected assemblages the simple matching coefficient (Krebs 1999) was calculated for the test and reference sites. The coefficient was calculated using the predicted and observed assemblages at each site and the difference in percent similarity for the two groups was tested using analysis of variance. Finally, the effects of elevation and distance from the coast on *O/E* ratios were tested using analysis of variance for both reference and test sites. For this analysis, the reference and test datasets were split into five $\log_{10}$ groups for both elevation and distance from the sea.

Results

Fish and decapod fauna

A total of 11967 fish and decapods from 23 species were caught at 200 sites during this study. Of these, 16 fish and the 2 crustacean species (a diadromous shrimp and a non-migratory crayfish) are native. Eleven fish species, including all non-native species, were found at less than 2% (4) of the sites and were not used in any of the analyses. These rare species were not included because most were introduced and their rarity made association of habitat variables problematical. At the 118 reference sites, ten fish taxa (including elvers) and two crustacean species were used in analyses after the removal of rare species. The most common was the longfin eel occurring at 88% of the sites, and rarest the torrentfish at 4% of the sites (Table 1.)

Table 1. Twelve fish and crustacean species found at more than 2% of the 200 potential reference sites and the number of 118 reference sites at which they were present (‡ indicates non-migratory species, * indicates decapod).

Scientific name	*Common name*	Code	present	absent
Anguilla australis	shortfin eel	ANGAUS	25	93
Anguilla dieffenbachii	longfin eel	ANGDIE	105	13
Anguilla spp.	elvers	ELVERS	7	111
Cheimarrichthys fosteri	torrentfish	CHEFOS	5	113
Galaxias fasciatus	banded kokopu	GALFAC	73	45
Galaxias maculatus	inanga	GALMAC	51	67
Gobiomorphus basalis	Crans bully ‡	GOBBAS	51	67
Gobiomorphus cotidianus	common bully	GOBCOT	23	95
Gobiomorphus huttoni	redfin bully	GOBHUT	42	76
**Paranephrops planifrons*	koura ‡	PARANE	72	46
**Paratya curvirostris*	shrimp	PARATA	69	49
Retropinna retropinna	common smelt	RETRET	7	111

Environmental variables

From the thirty-six environmental variables available 18 were considered unlikely to be affected by human influences and were retained for use in the discriminant models (Table 2). The test sites occurred over the full range of measured variables as revealed by the similar ranges and means of the variables for the test and reference sites (Table 2).

Validation of individual presence-absence models

The classification results for the individual species models reveal that for each species on average more than 79% of the reference sites were correctly classified (Table 3). The resubstitution and jackknife procedures revealed similar patterns of model accuracy for each species based on correct rates of prediction. The sensitivity and specificity rates showed that there was no bias towards false positive or false negative classifications for the calibration data (Table 3). The results for the validation data not used in the model (1/5 partitioning) from five separate random partitions each of 20 sites are also shown in Table 3. The percentage of the partitioned validation sites correctly assigned averaged 79%. As with the training data, inanga, the two crustaceans (koura and shrimp), and redfin bully had the highest classification error rates (between 65 and 70% correct) while the other species ranged from 75% up to 96%. The sensitivity and specificity results for the validation dataset reveal a lower mean sensitivity rate (Table 3) for the species that were relatively rare. Sensitivity, which is the percentage of sites where species presence was correctly predicted, was low for elvers, smelt, and torrentfish. This bias indicates that the models were poor when predicting the presence of rare species when using validation data. To assess the relationship between the number of sites where a taxon was predicted to be present and where they were observed the observed number of sites was plotted against the predicted number of sites using all the partitioned validation sites (Fig. 2). The observed number of sites was similar to the number of sites predicted over all species (R^2 = 0.98, predicted number of sites = 1.001 observed number of sites + 0.59, $P < 0.0001$).

Table 2. Range of measures observed among the 118 reference and 35 test sites for the variables used in predictive models. The coefficients ranking was calculated from the first five canonical coefficients for each species model (see table 5) (1 = most important to least important 9; SD = standard deviation).

	Coefficient		Reference sites			Test sites		
Variable	ranking	Code	Min	Mean	Max	Min	Mean	Max
Geographic Variables								
Altitude (m)	6=	Altit	1.0	47.9	220.0	2.0	62.1	200.0
Distance to coast (km)	4=	Dist	0.1	15.3	79.5	0.3	17.5	79.8
Longitude	7=	East	26397	26641	27282	26379	26662	27293
Latitude	6=	North	64419	65074	65611	64522	65082	65560
Channel variables								
Mean width (m)	3	Width	0.8	3.3	31.8	0.8	3.8	17.3
SD width	9=	Swidth	0.2	1.0	9.8	0.3	1.4	9.9
Mean depth (m)	4=	Depth	0.15	.42	1.96	0.02	0.47	1.42
SD Depth	6=	Sdepth	0.25	1.79	9.58	3.48	0.22	6.27
Median substrate	9=	Medsub	0.0	9.1	39.1	0	6.2	40.0
Leaf litter	8=	Litter	1.0	2.1	4.0	1.0	2.0	4.0
% Pool	1	Pool	0	75.3	100.0	0	59.4	100.0
% Riffle	4=	Riffle	0	12.7	70.0	0	24.3	100.0
% Run	2	Run	0	10.8	90.0	0	15.7	100.0
Cobble packing	7=	Pack	0	1.8	3.0	0	1.9	3.0
Water variables								
Temperature (° C)	6=	Temp	12.9	16.5	24.0	13.5	16.8	20.9
Conductivity ($\mu S/cm^{-1}$)	6=	Cond	100.00	216.07	974.00	99.4	222.2	897.0
Velocity (m/sec)	5=	Veloc	0	0.2	1.3	0	0.2	1.0
pH	8=	pH	7.1	8.1	8.9	6.3	8.0	8.6

Table 3. Number of sites correctly assigned by discriminant functions for the 118 reference sites and 100 validation (5 validation groups combined) using jackknife and resubstitution, and sensitivity (percentage of true presence's correctly identified) and specificity (percentage of true absences correctly identified). Results for validation sites are mean values for 5 separate (20 site) tests

	Reference sites (118)					Validation sites (n = 100)			
Species	Number of sites present	% correctly classified (jackknife)	% correctly classified (resubstitution)	Specificity	Sensitivity	Number of sites present	% correctly classified	Specificity	Sensitivity
Banded kokopu	73	70	80	82	79	64	75	66	80
Common bully	23	75	83	87	59	21	75	84	40
Cran's bully	51	80	86	87	86	36	81	88	71
Elvers	7	86	96	96	75	10	85	91	28
Inanga	51	65	80	82	76	43	64	69	58
Koura	72	72	81	79	81	60	68	68	68
Longfin eel	105	89	92	83	93	82	84	56	90
Shrimp	69	66	75	74	75	60	65	56	71
Redfin bully	42	66	79	81	74	38	69	72	62
Shortfin eel	25	81	86	85	90	27	80	80	82
Common smelt	7	92	97	98	83	3	93	98	20
Torrentfish	5	92	97	98	75	2	96	99	33
Mean		79	87	87	80		79	79	52

Table 4. The five largest standardised canonical coefficients for each species discriminant model. Positive coefficient indicates variable positively related to presence of species (codes from table 3).

Cran's bully	Pool	(1.58)	Riffle	(-1.04)	Run	(-0.72)	Cond	(0.63)	Temp	(0.42)
Elvers	Cond	(0.89)	Veloc	(0.81)	Pool	(0.70)	Riffle	(0.57)	Run	(0.54)
Redfin bully	Altit	(-0.55)	Dist	(-0.55)	North	(0.51)	Depth	(-0.50)	Sdepth	(0.40)
Longfin eel	Cond	(-1.04)	Pack	(0.58)	Altit	(-0.43)	East	(-0.42)	Veloc	(-0.39)
Banded kokopu	Width	(-0.80)	Dist	(-0.55)	North	(0.42)	Temp	(-0.35)	pH	(-0.29)
Koura	Pool	(1.44)	Sdepth	(-0.87)	Depth	(0.74)	Riffle	(0.73)	Run	(0.70)
Torrentfish	Pool	(-1.90)	Run	(-1.56)	North	(-0.67)	Sdepth	(-0.54)	Litter	(0.48)
Inanga	Pool	(-0.95)	Run	(-0.84)	Altit	(-0.80)	Width	(0.59)	Sdwidth	(-0.43)
Shrimp	Width	(0.91)	Dist	(-0.75)	Veloc	(-0.65)	Pool	(-0.58)	Sdwidth	(-0.34)
Shortfin eel	Dist	(0.90)	Pool	(0.61)	Width	(0.61)	Riffle	(0.51)	Run	(0.44)
Common bully	Veloc	(-0.56)	Dist	(-0.51)	Pool	(-0.47)	Pack	(0.37)	Riffle	(-0.29)
Common smelt	Pool	(-1.23)	North	(-0.99)	Run	(-0.82)	Temp	(0.65)	East	(0.65)

Table 5. Descriptive statistics of *O/E* ratio values and simple matching coefficient between observed and expected assemblages for the reference and test sites.

	O/E ratios Test sites	*O/E* ratios Reference sites	Simple matching coefficient Test sites	Simple matching coefficient Reference sites
Number of sites	35	118	35	118
Mean	0.662	1.012	0.74	0.87
Median	0.698	0.995	0.76	0.85
Standard Error	0.043	0.024	0.02	0.001
Standard deviation	0.252	0.258	0.11	0.01
Minimum	0.198	0.287	0.53	0.61
Maximum	1.161	2.008	1	1

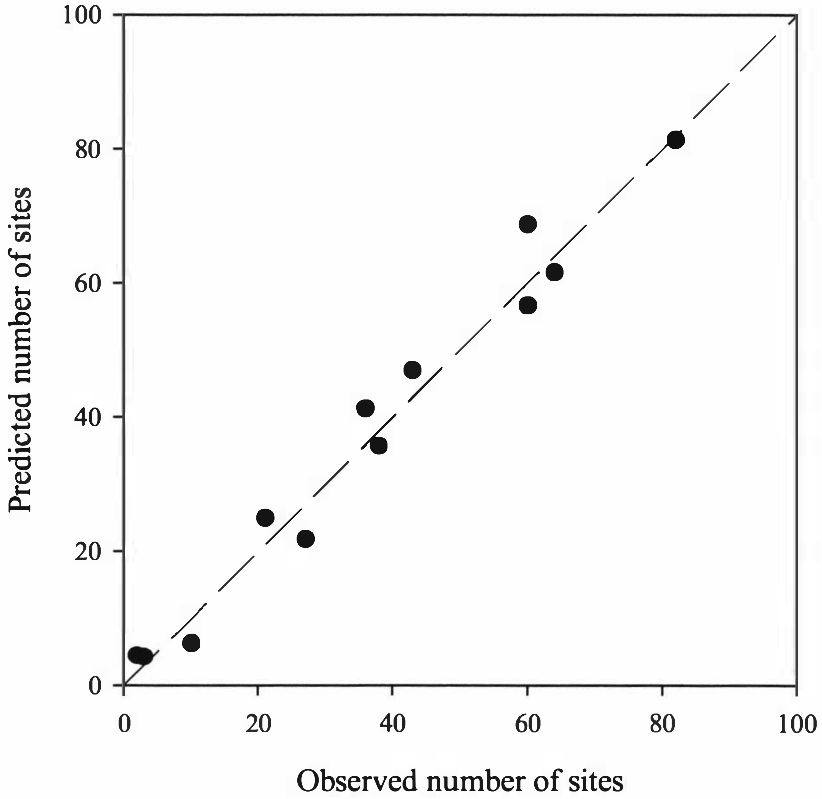

Fig. 2 Predicted vs. observed species occurrence for 12 fish and decapod taxa from 100 independent partitioned validation sites. The dashed diagonal line represents the line of perfect agreement.

Environmental variables associated with presence/absence groups

The values for the five highest canonical coefficients for the discriminant functions for each taxon are shown in Table 4. The only variable that did not appear in the leading five coefficients for all species was median substrate size (Table 4). All other variables were in the top 5 of at least one of the species models, indicating that there were different variables associated with individual discriminant models and therefore, species distributions (Table 4). The overall rankings averaged for each variable across all species are shown in Table 2. The percentage of pool was the highest-ranking variable followed by the percentage of run, mean width and then the distance from the site to the coast with no apparent association with either landscape scale, or reach scale variables (Table 2).

Distribution of reference O/E ratios

The number of taxa expected at each reference site was calculated and compared with observed to give the observed over expected (*O/E*) ratios. The distribution of *O/E* ratio values for the reference sites approximated a normal distribution centered on unity with a mean of 1.01 (Fig. 3 and Table 5).

Test site assessment

A number of potential impacts were identified at the 35 test sites (Table 6). When the total number of sites where each taxon was captured was plotted against the number of sites predicted, three taxa (non-migratory bullies, shortfin eels, and elvers) occurred at slightly more sites than predicted (Fig. 4). Common smelt was predicted and observed at one site and all other species were found at fewer sites than predicted. The two galaxiid species, inanga and banded kokopu, were at considerably fewer sites than predicted. Observed over expected values for the test sites were significantly lower than mean values for reference sites ($F_{1, 151} = 50$, $P < 0.0001$) (Table 5; Fig 3). The simple matching coefficient between the observed and predicted assemblages were more similar for reference sites (85% similarity) than for test sites (74% similarity) and this was significantly different for the two sets of sites ($F_{1, 151} = 318$, $P < 0.0001$).

Table 6. Site numbers, potential impacts, site elevations, predicted number of species, observed number of species and observed over expected ratios for the 35 test sites.

Potential impact	Altitude (m)	No. species predicted	No. species observed	*O/E* ratio
Pest fish	120	5.06	1	0.2
Access	10	4.92	1	0.2
Access	80	4.61	1	0.22
Geothermal	30	3.84	1	0.26
Forestry	5	5.95	2	0.34
Geothermal	65	2.7	1	0.37
Access	60	5.08	2	0.39
Access	20	4.48	2	0.45
Access	80	4.01	2	0.5
Access	50	5.89	3	0.51
Access	10	5.67	3	0.53
Access	60	3.34	2	0.6
Eutrophication	60	4.71	3	0.64
Access	60	4.67	3	0.64
Geothermal	20	4.63	3	0.65
Pest fish	35	4.61	3	0.65
Pest fish	2	6	4	0.67
Access	60	4.3	3	0.7
Access	140	2.81	2	0.71
Access	60	4.21	3	0.71
Access	80	4.14	3	0.72
Access	80	4.17	3	0.72
Sediment	100	3.95	3	0.76
Forestry	35	3.97	3	0.76
Intermittent flow	200	3.86	3	0.78
Pest fish	40	3.81	3	0.79
Forestry	120	4.96	4	0.81
Forestry	40	4.56	4	0.88
Access	20	5.39	5	0.93
Forestry	5	5.31	5	0.94
Forestry	180	4.05	4	0.99
Eutrophication	35	5.04	5	0.99
Access	110	3.97	4	1.01
Access	60	3.94	4	1.02
Tip leachate	40	3.45	4	1.16
Mean	62.06	4.46	2.91	0.66

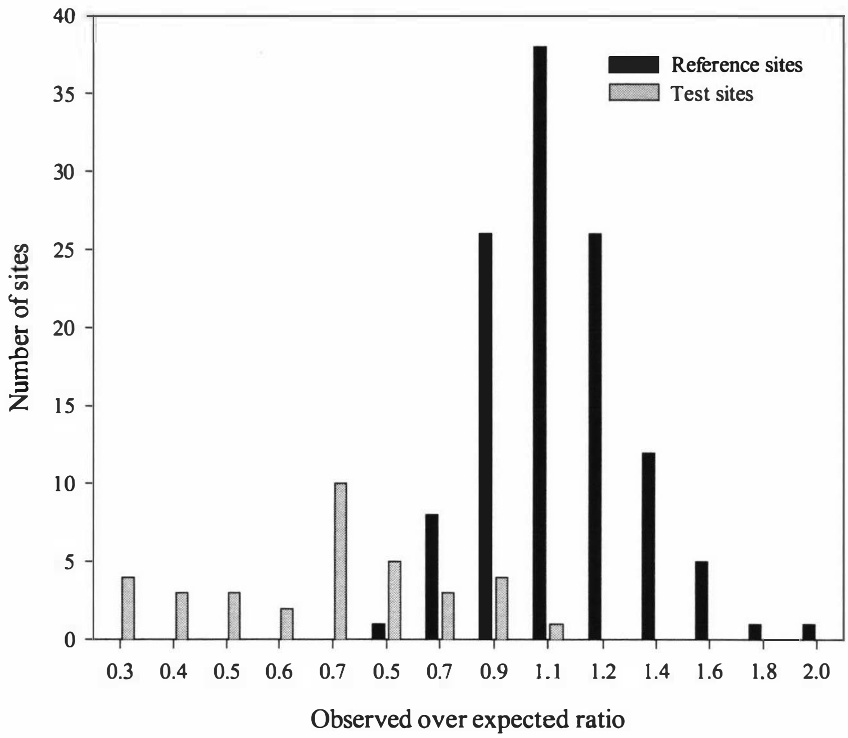

Fig. 3 Distribution of observed over expected ratio values. The *O/E* values for reference sites (black bars) and test sites (shaded bars) are shown.

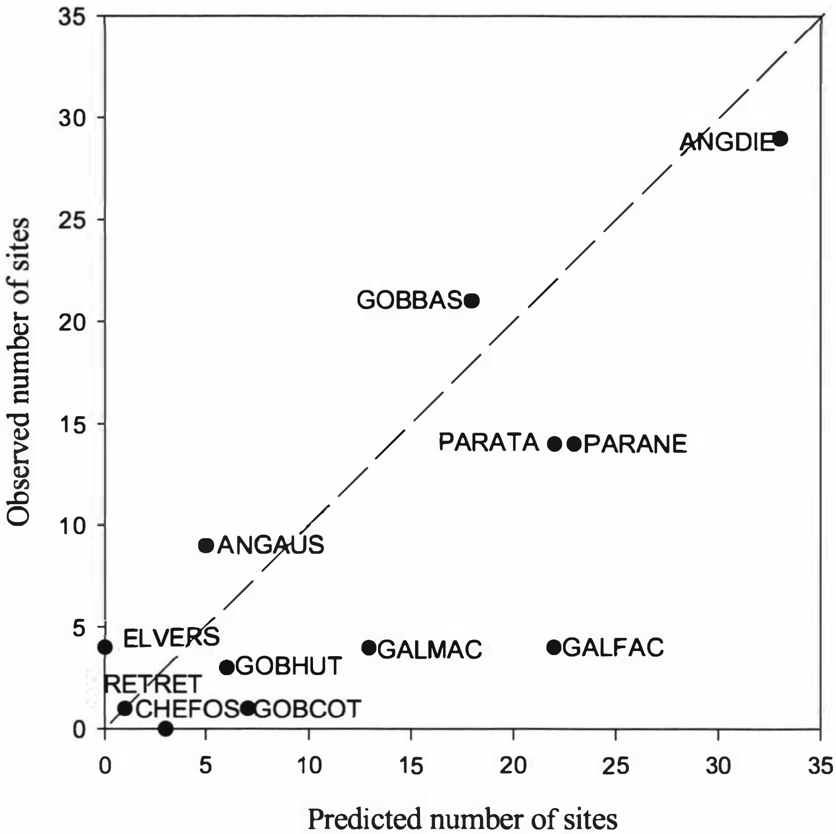

Fig. 4 The predicted vs. observed number of sites for 12 fish and decapod taxa for the 35 test sites. The dashed diagonal line represents the line of perfect agreement. (Species codes from Table 4)

Effects of elevation and distance from the coast

The analyses revealed that elevation and distance from the coast did not influence the mean *O/E* values for reference sites (Fig. 5) (altitude: $F_{1,116} = 0.0$, $P < 0.97$; distance from the coast: $F_{1,116} = 0.02$, $P < 0.88$). The mean *O/E* values for the test sites over the elevational range were not significantly different from each other ($F_{1,116} = 2.4$, $P < 0.13$), however mean test site *O/E* ratios were influenced by distance from the sea ($F_{1,33} = 5.37$, $P < 0.03$). These results indicate that the model predictions are not biased by differences in the variables that are normally associated with freshwater fish communities in New Zealand (Joy *et al.* 2000).

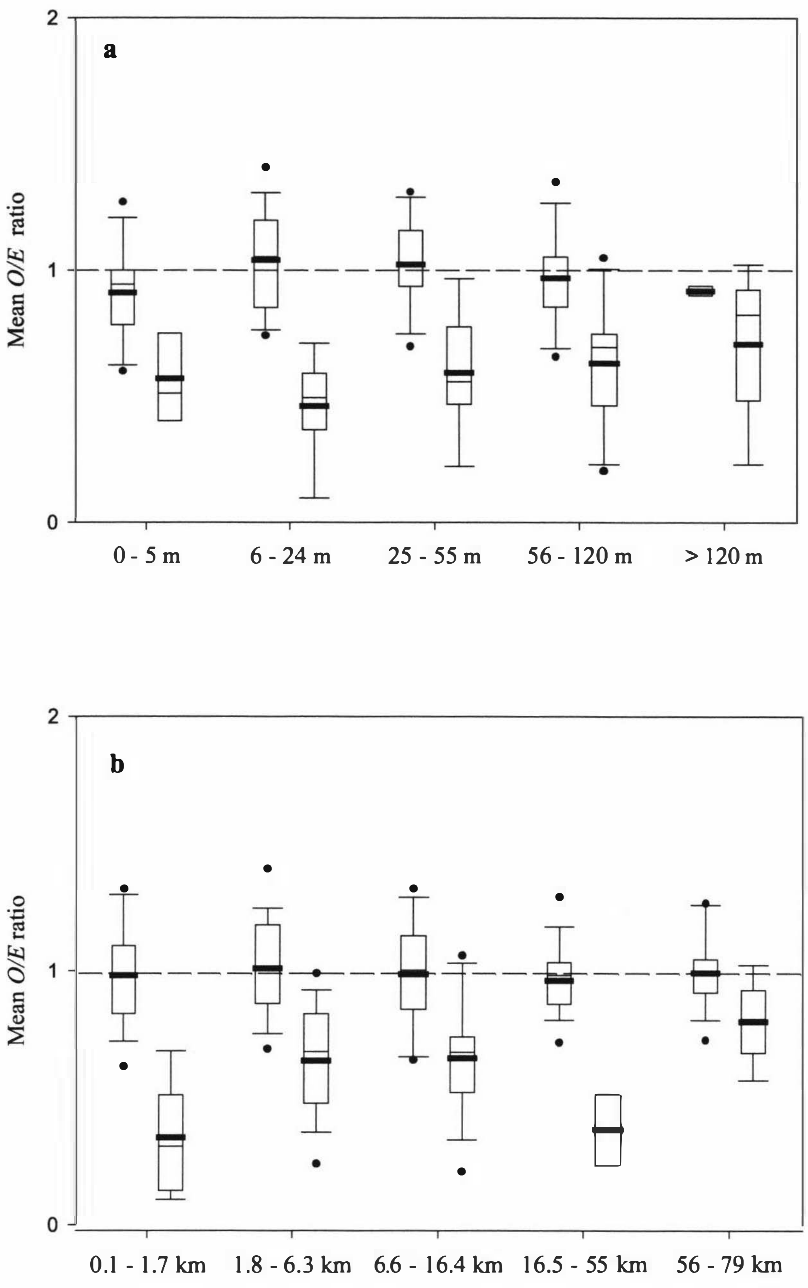

Fig. 5 Distribution of mean *O/E* values, both reference, and test, by log altitudinal groups (a) and log distance from the sea groups (b). For each pair, the reference data are on the left and test data on the right. Boxes contain interquartile range; circles denote 5th and 95th percentiles; fine centerline denotes median; heavy centerline denotes mean; whiskers extend to 1.5 × interquartile range.

Discussion

Individual models that relate occurrence or density of species to physical features are effectively habitat or resource selection functions (Boyce and McDonald 1999) and are commonly used in applied ecology (Guegan *et al.* 1998, Mastrorillo *et al.* 1998, Peeters and Gardeniers 1998, Manel *et al.* 1999, Brosse and Lek 2000, Oberdorff *et al.* 2001). These habitat selection functions yield probabilities of occurrence for species that are proportional to habitat use. Furthermore, as the models are multivariate, the influences of interactions with other species can be built into the models (Boyce and McDonald 1999). The habitat selection functions from reference sites therefore, give us the ability to relate the absence of taxa, predicted to be present, to possible habitat impairment.

The individual discriminant models showed a high level of classification accuracy for all species over the ranges of relative occurrences in the dataset ranging from torrentfish at 4% of the sites to longfin eel at 88%. Although precision was reduced when the discriminant functions were validated using independent data, the models remained relatively accurate for all but the rarest species (Table 3). This reduction in accuracy may be partly related to a reduction in training data for each of the five validation tests with only 100 sites available for training after five fold partitioning. For rare species, prediction is made more difficult with a lack of representation of the conditions found at the few sites where rare species are found and vice versa for common species. Another potential reason for a reduction in precision of discriminant models is the restriction placed on the models by using only the environmental variables not influenced by humans. If all the measured variables had been included in the model, the predictions would presumably have been more accurate but any predictions would not have been as useful for detecting impairment of the fish and decapod assemblages. Three species were consistently more difficult to predict than the others, redfin bully, and the two crustaceans koura and shrimp (Table 3). This difficulty in prediction may be because the important environmental variable(s) associated with their presence-absence was not measured, or were included in the variables associated with human influences not used in the models.

The strong elevational zonation of New Zealand fish species has traditionally nullified their use in biomonitoring (McDowall and Taylor 2000). In previous studies we have shown how predictive modeling (sensu RIVPACS) can be applied, to enable the use New Zealand fish to assess the impacts of hydroelectric dams (Joy and Death 2000) and a wider range of environmental degradation (Joy and Death 2002). However, a criticism of the RIVPACS approach is that it requires the classification of sites in to groups based on their biota's using multivariate techniques (Chessman 1999). This requires a somewhat arbitrary judgement about where to split groups of sites. Thus, a particular site is assigned to one group or another even if it could in part belong to both (although subsequent calculations nullify to a certain extent such splitting). The approach we have adopted here avoids this potential flaw by modeling species individually with respect to their presence or absence at a site. We believe this avoids some of the criticisms of the RIVPACS approach but still allows statistical modeling to be used accurate and objective bioassessment of fish, decapoda and potentially other organisms, in New Zealand and elsewhere.

In this study, we applied the discriminant models to presence-absence data and ignored species abundance. The importance of freshwater fish abundance as opposed to presence-absence in ecological assessment is equivocal. Pusey *et al.* (2000) found that presence-absence models did not detect flow variability differences that were evident when they considered species density, but conversely abundance models had less predictive precision than presence-absence models. In a North American river study, (Hoefs and Boyle 1992) found that fish assemblage differences associated with stream condition were related to presence-absence rather than abundance. In either case, predicting fish densities may be difficult because of temporal variation in abundance (Grossman *et al.* 1990) and the difficulty of assessing abundance in streams of varying size.

The estimates of species presence and absence were assessed to be a realistic representation of the actual species assemblage present in the sampled area. However, the precision of the sampling methods used when collecting the reference site data, on which the models were built, is subject to the inherent variability expected in any biological sampling. Notwithstanding this, the most important factor is that the same sampling methods are used when surveying test sites as was used when building models, thus any sampling bias is made equal across sites. Thus, the predictions are probabilities of capture and not necessarily presence.

Evaluation of models as an assessment tool

When making any statistical comparison there is a probability that the wrong conclusion will be reached. In this case, if the first decile of the reference *O/E* values were used as a cut-off level for impairment, any site with an *O/E* value less than 0.75 would be classed as impaired. Using this criterion, 22 of the 35 test sites would fail (Table 6). Thus, the probability of assessing an unimpaired site as impaired (type 1 error) was set at 10%. It can be assumed that the same percentage of test sites would be assessed as impaired even if they were not different from some of the reference sites. Using this assumption, the percentage of test sites failing (62%) would suggest that the model assessments conform to our *a priori* predictions of impairment. It must be remembered that these 35 test sites were initially selected as reference sites from geographic information so the proportion of test sites failed in this case is conservative. If test sites had initially been selected as impaired, we could expect even more deviation from reference conditions.

Conclusions

Our aims were to see if individual fish and decapod species could be predicted from site environmental characteristics and then to develop a method to use the predictions in biomonitoring. The results show that the assemblages could be predicted with relative precision using a restricted set of environmental variables. The method outlined to convert the predictions to *O/E* ratios for site assessment was based on the RIVPACS procedure (Wright 1995). However, our approach differs fundamentally in that we are modeling individual species and not assemblages. This predictive model meets the requirements of the ultimate goals of biomonitoring, that is, to provide an accurate assessment of the biological quality of freshwater ecosystems that can be easily understood by managers and policy makers (Karr 1991, Hawkins *et al.* 2000). We believe our results show that this predictive model has potential for use in assessing the biological quality and therefore water quality of New Zealand streams and elsewhere. In this context, biological quality is the extent to which a stream supports the fish and decapod assemblages expected in the absence of environmental stress. We anticipate that this model would be used to provide a target or benchmark of acceptable biological quality against which a site could be measured. The extent of the deviation from the benchmark assemblage is the relative level of impairment. The output from the model is not, however, an explanation of community response to impacts but rather a prediction of the expected assemblage given the environmental variables at a site. To elucidate an explanation of community response the individual species responses must be examined using more explicit field studies (e.g. (Joy and Death 2001).

This paper has concentrated on the use of predictive models to assess environmental impacts by considering degraded sites, but the same model also identifies sites where assemblages contain more species than predicted. These sites with high observed over expected ratios reveal sites where conditions support diverse fish assemblages. The identification of species rich sites is important for the identification of the habitat features associated with unusually diverse communities (see Wright *et al.* 2000). The results from the model can also be used to assist with the decision making process required when deciding on waterways for protection and those to be sacrificed when development pressures are high.

Acknowledgments

This work was funded by a grant from the sustainable management fund (New Zealand Ministry for the Environment grant No. 5099 "Nga Ika Waiora": stream health evaluation tool) and the Auckland Regional Council. Thanks to Allison Hewitt for assistance with all fieldwork and to Auckland Regional Council staff Brent Evans, Chris Hatton, David Glover, and John Maxted. For assistance with site access, information, and accommodation, we thank Conservation Department staff Chris Roberts, Thelma Wilson, Jo Ritchie, and Mike McGlynn.

References

Bailey, R. C., M. G. Kennedy, M. Z. Dervish, and R. M. Taylor. 1998. Biological assessment of freshwater ecosystems using a reference condition approach: comparing predicted and actual benthic invertebrate communities in Yukon streams. Freshwater Biology **39**:765-774.

Boyce, M. S., and L. L. McDonald. 1999. Relating populations to habitats using resource selection functions. Trends in Ecology and Evolution **14**:268-272.

Chessman, B. C. 1999. Predicting the macroinvertebrate faunas of rivers by multiple regression of biological and environmental differences. Freshwater Biology **41**:747-757.

Clarke, R. T. 2000. Uncertainty in estimates of biological quality based on RIVPACS. Pages 39-54 *in* J. F. Wright, D. W. Sutcliffe, and M. T. Furse, editors. Assessing the biological quality of freshwaters. RIVPACS and other techniques. Freshwater Biological Association, Ambleside, UK.

Fausch, K. D., J. R. Karr, and P. R. Yant. 1984. Regional application of an index of biotic integrity based on stream fish communities. Transactions of the American Fisheries Society **113**:39-55.

Fielding, A. H., and J. F. Bell. 1997. A review of methods for the assessemnt of prediction errors in conservation presence/absence models. Environmental Conservation **24**:38-49.

Grossman, G. D., J. F. Dowd, and M. Crawford. 1990. Assemblage stability in stream fishes: a review. Environmental Management **14**:661-671.

Hawkins, C. P., R. H. Norris, J. N. Hogue, and J. W. Feminella. 2000. Development and evaluation of predictive models for measuring the biological integrity of streams. Ecological Applications **10**:1456-1477.

Hoefs, N. J., and T. P. Boyle. 1992. Contribution of fish community metrics to the Index of Biotic Integrity in two Ozark rivers. Pages 283-304 *in* D. H. McKenzie, D. E. Hyatt, and V. J. McDonald, editors. Ecological Indicators. Elsevier, London.

Hughes, R. M., D. P. Larsen, and J. M. Omernik. 1986. Regional reference sites: a method for assessing stream potentials. Environmental Management **10**:629-635.

Information, N. Z. S. a. L. 1987. 1:50 000 Topographical Map Series. *in*. New Zealand Survey and Land Information NZSLI, Wellington.

Joy, M. K., and R. G. Death. 2001. Control of freshwater fish and crayfish community structure in Taranaki, New Zealand: dams, diadromy or habitat structure? Freshwater Biology **46**:417-429.

Joy, M. K., I. M. Henderson, and R. G. Death. 2000. Diadromy and longitudinal patterns of upstream penetration of freshwater fish in Taranaki, New Zealand. New Zealand Journal of Marine and Freshwater Research **34**:531-543.

Karr, J. R. 1981. Assessments of biotic integrity using fish communities. Fisheries **6**:21-27.

Manel, S., J. Dias, and S. J. Ormerod. 1999. Comparing discriminant analysis, neural networks and logistic regression for predicting species distributions: a case study with a Himilayan river bird. Ecological Modeling **120**:337-347.

McDowall, R. M. 1990. New Zealand Freshwater Fishes: A Natural History and Guide. Heinemann Reed, Auckland.

McDowall, R. M. 1998. Driven by diadromy: its role in the historical and ecological biogeography of the New Zealand freshwater fish fauna. Italian Journal of Zoology **65**:73-85.

Moss, D., M. T. Furse, J. F. Wright, and P. D. Armitage. 1987. The prediction of the macro-invertebrate fauna of unpolluted running-water sites in Great Britain using environmental data. Freshwater Biology **17**:41-52.

Pusey, B. J., M. J. Kennard, and A. H. Arthington. 2000. Discharge variability and the development of predictive models relating stream fish assemblage structure to habitat in northeastern Australia. Ecology of Freshwater Fish **9**:30-50.

Quinn, J. M., and C. W. Hickey. 1990. Magnitude of effects of substrate particle size, recent flooding, and catchment development on benthic invertebrates in 88 New Zealand rivers. New Zealand Journal of Marine and Freshwater Research **24**:411-427.

Reynoldson, T. B., R. C. Bailey, K. E. Day, and R. H. Norris. 1995. Biological guidelines for freshwater sediment based on BEnthic Assessment of SedimenT (the BEAST) using a multivariate approach for predicting biological state. Australian Journal of Ecology **20**:198-219.

Reynoldson, T. B., and J. F. Wright. 2000. The reference condition: problems and solutions. Pages 293-303 *in* J. F. Wright, D. W. Sutcliffe, and M. T. Furse, editors. Assessing the biological quality of freshwaters. RIVPACS and other techniques. Freshwater biological association, Ambleside, United Kingdom.

Rice, J., R. D. Ohmart, and B. W. Anderson. 1983. Habitat selection attributes of an avian community: a discriminant analysis investigation. Ecological monographs **53**:261-290.

Wiegleb, G. 1981. Application of discriminant analysis on the analysis of the correlation between macrophyte vegetation and water quality in running waters of Central Europe. Hydrobiologia **79**:91-100.

Williams, B. K. 1983. Some observations on the use of discriminant analysis in ecology. Ecology **64**:1283-1291.

Wolman, M. G. 1954. A method of sampling coarse river-bed material. Transactions, American Geophysical Union **35**:951-956.

Wright, J. F., D. Moss, P. D. Armitage, and M. T. Furse. 1984. A preliminary classification of running water-sites in Great Britian based on macro-invertebrate species and the prediction of community type using environmental data. Freshwater Biology **14**:221-256.

Wright, J.F. 1995. Development and use of a system for predicting the macroinvertebrate fauna in flowing waters. *Australian Journal of Ecology* 20: 181-197.

Wright, J. F., D. W. Sutcliffe, and M. T. Furse. 2000. Assessing the biological quality of fresh waters: RIVPACS and other techniques. Freshwater biological association, Ambelside, Cumbria, UK.

CHAPTER FIVE

Predictive modeling and spatial mapping of freshwater fish and decapod assemblages integrating GIS and neural networks

Joy, M. K., and R. G. Death. Submitted. Predictive modeling and spatial mapping of freshwater fish and decapod assemblages integrating GIS and neural networks. Journal of Applied Ecology.

Abstract

We used stream fish and decapod spatial occurrence data extracted from a national database combined with recent surveys and geospatial landuse data, geomorphologic, climatic, and spatial data in a geographical information system (GIS) to model fish occurrence in the Wellington Region, New Zealand. To predict the occurrence of each species at a site from a common set of predictor variables we used a multi-response, artificial neural network (ANN), to produce a single model which predicted the entire fish and decapod assemblage in one procedure. The predictions from the ANN using this landscape scale data proved very accurate and four other evaluation metrics independent of species abundance or probability thresholds, also confirmed the accuracy of the model. The important variables contributing to the predictions included the spatial and elevational position of the site reach, catchment area, the landuse and vegetation type proportions of the catchment, and catchment geology. The geospatial data available for the entire regional river network were then used to create a habitat-suitability map for all 18 species over the regional river network using GIS. This prediction map has many potential uses including; monitoring and predicting temporal changes in fish communities caused by human activities and shifts in climate, identifying areas in need of protection, biodiversity hotspots, and areas suitable for the reintroduction of endangered or rare species.

Keywords. prediction maps, artificial neural networks, New Zealand, freshwater fish, GIS, presence/absence, diadromous fish

Introduction

Effectively modeling the relationships between species and their environment has long been considered important in ecology and many methods have been utilised to produce these models (see Guisan and Zimmermann, 2000 for a review). The focus of these species-environment models has shifted recently from description and explanation to an emphasis on prediction accuracy (De'Ath, 2002). This movement towards predication accuracy has been driven by global environmental pressures such as climate change and species loss so that now accurate prediction is often seen as the principal objective of species-environment models (Peters, 1991; Franklin, 1998; Guegan *et al.*, 1998; Iverson and Prassad, 1998). The recognition that using statistical hypothesis tests as a method for addressing ecological questions can have drawbacks (Johnson, 1999) has also led to an increased emphasis on prediction rather than just description.

Species-environment models have been crucial in many fields of freshwater ecology including: species conservation (Armitage and Petts, 1992; Filipe *et al*, 2002; Olden, in press; Wright *et al*, 1998; Wright *et al*, 2000), assessment of the impacts of flow regulation (Joy and Death, 2000, 2001; Marchant and Hehir, 2002; Reyes-Gavilan *et al*, 1996), biological quality assessment (Brosse *et al*, 2001; Hoang *et al*, 2001; Joy and Death, In Press; Oberdorff *et al*, 2001; Oberdorff *et al*, 2002; Reynoldson and Rosenberg, 1996; Torgerson *et al*, 1999), and the prediction of aquatic macrophyte and diatom distribution (Chessman *et al*, 1999; Lehmann, 1998; Lehmann *et al*, 1997).

Artificial neural networks (ANNs) offer a promising alternative to traditional statistical approaches for predictive modeling when non-linear patterns exist. A recent comparison of traditional (e.g. logistic regression, linear discriminant analysis) and alternative techniques (classification trees and ANNs) for predicting species presence/absence using both simulated and empirical data showed that the accuracy of ANN predictions outperformed the alternatives particularly with non-linear data (Olden and Jackson, 2002a). ANNs also have advantages over traditional modeling methods because they are not dependant on particular functional relationships, need no assumptions regarding underling data distributions and no *a-priori* understanding of variable relationships (Olden and Jackson, 2001). Thus, this independence from assumptions make ANNs a powerful option for exploring complex potentially non-linear relationships such as the associations between stream fish and their environment.

ANNs have been used to solve a diverse array of questions in fields of study from economics to epidemiology, and have recently been applied in ecology because of their ability to recognise complex patterns (Guegan *et al*, 1998; Lek *et al*, 2000; Walley and Fontama, 1998). Even within aquatic research, the application of ANNs has been diverse. Examples range from water quality assessment (Brosse *et al*, 2001; Chon *et al*, 1996; Park *et al*, 2001; Scheiter *et al*, 1999; Walley and Fontama, 1998) to modeling phytoplankton production (Recknagel and Wilson, 2000; Scardi, 2001). Examples of fisheries applications using ANNs include the prediction of: fish production (Chen and Ware, 1999), fish species richness (Guegan *et al*, 1998; Mastrorillo *et al*, 1998; Oberdorff *et al*, 1995), fish presence/absence (Mastrorillo *et al*, 1997b; Olden and Jackson, 2001), and fish abundance (Brosse *et al*, 1999; Mastrorillo *et al*, 1997a).

Increasingly studies involving predictive modeling of species-occurrence combine the power of geographic information systems (GIS) with multivariate statistical tools to formalise the link between species and their habitat. When geospatial data in a GIS are combined with non-linear modeling techniques such as ANNs, the potential to create accurate predictive models is greatly enhanced. Furthermore, the capacity of predictive models using GIS to fill in the gaps between sample sites into unsampled areas has numerous potential benefits in conservation and resource management arenas. Aquatic resource managers are often faced with a lack of information on the distribution of the biological components of their area of responsibility. Maps of species distributions are often incomplete, outdated or non-existent. This means that decisions regarding resource-use are often made in the absence of even basic information, particularly when the potential impact is small in scale. These small-scale impacts tend not to be well investigated by regulatory authorities because of a perceived lack of importance and a concomitant lack of financial resourcing. However, even modest impacts can have cumulative effects and can thus become part of a large impact. By filling in the gaps on distribution maps, spatial predictive models will lead to better management decisions through improved knowledge of species distribution and ease of access to information.

We used fish and decapod spatial occurrence data extracted from a database New Zealand Freshwater Fish Database (McDowall and Richardson, 1983) combined with recent surveys and geospatial landuse, geomorphologic, climatic, and spatial data in a GIS to model fish occurrence in the Wellington Region, New Zealand. To predict the occurrence of each species at a site from a common set of predictor variables we used a multi-response, artificial neural network, to produce a single model to predict the entire fish and decapod assemblage in one procedure. The predictive model was then extended to fill in the gaps between the surveyed sites using a GIS river network to give a spatial map of species probability of occurrence for the entire region.

Data sources, study area and methods

The Wellington Region is a geopolitical region covering 8,130 km^2 of the southern North Island of New Zealand (41° S, 175° E). It contains the nation's capital city (Wellington) and a human population of 425,000. The region extends to the Waitohu River in the Northwest and the Mataikona and Ruamahanga River catchments in the Northeast. A mountain chain (the Tararua and Orongorongo Ranges) bisects the region running south to north. Approximately 22% of the region is covered by indigenous native forest; however, much of this area is above 300-m a.s.l. The region can be divided into three broad geographic areas, the western coastal lands, the rugged central axes and the remnant peneplain of the eastern Wairarapa. Landuse in the vicinity of Wellington Harbour at the southern end of the region is mainly urban-residential. The Wairarapa area to the east of the axial ranges is dominated by extensive pastoral dry-stock farming (38 % of the region) and exotic forestry (5 % of the region). The western coastal strip is predominantly sand country and has extensive residential areas.

Artificial neural networks

Of the many neural network types available (see Bishop, 1995) we chose a single hidden-layer feedforward multi layer perceptron trained using the backpropagation error algorithm (Rumelhart *et al*, 1986). This network consists of single predictor, hidden and output layers with a predictor neurone for each independent variable and an output neurone for each dependant variable. A single hidden layer was used because it reduces computation time and often produces similar results to networks with multiple hidden layers (Bishop, 1995; Kurkova, 1992).

We determined the optimal number of hidden neurones and number of training epochs iteratively by comparing the performances of different networks. To achieve this we compared networks with 20 to 120 (in intervals of 20) hidden neurones and varied the number of epochs from 50 to 250 (in 50 epoch intervals) and then selected the combination that produced the greatest predictive performance based on the evaluation procedures outlined below. We used crossvalidation (see the following paragraph for details of crossvalidation procedure) for this optimisation to ensure that the network was not 'overtrained'. Overtraining occurs when the network learns the training data extremely well but is not able to generalise well (Walley and Fontama, 1998). After optimisation, the network consisted of 69 predictor neurones representing each of the independent predictor variables. The hidden layer consisted of 70 neurones and there were 18 output nodes, one for each of the dependant variables (the 18 species being modelled) and the training was run for 100 epochs. The independent variables were converted to z – scores prior to training.

Model evaluation

To evaluate the performance of the predictive model we used a leave-one-out cross-validation method (jack-knife). This method involves excluding one observation, reconstructing the model and then predicting the response of the excluded observation. This process provides a nearly unbiased estimate of model performance (Olden and Jackson, 2000). To assess the overall classification success of the model we first derived matrices of confusion (Fielding and Bell, 1997). A matrix of confusion tabulates the observed and predicted presence/absence patterns to provide a summary of the number of correct and incorrect classifications from the model. Using these matrices, five metrics of prediction success were produced: (1) The overall classification accuracy of the model was measured as the percentage of sites where the model correctly predicted the presence/absence of each of the species. (2) The ability of the model to accurately predict species presence was assessed as model sensitivity (i.e. the percentage of site presences correctly predicted). (3) The ability of the model to correctly predict species absences was assessed and recorded as model specificity. (4) Cohen's Kappa coefficient of agreement (Titus *et al*, 1984) was used to examine if model performance differed from expectation based on chance alone. This is a relatively robust model evaluation method that is relatively independent of species frequency of occurrence (Manel *et al*, 2001).

Rather than simply following the conventional decision threshold of 0.5 (e.g. Oberdorff *et al,* 2001) for deciding on species presence we constructed receiver-operating characteristic (ROC) plots to estimate the predictive ability of the model over all decision thresholds and find the optimal threshold (Fielding and Bell, 1997; Zweig and Campbell, 1993). The ROC plot is obtained by plotting the true positive proportion on the y-axis against the false positive proportion on the x-axis as the decision threshold is varied over the entire range between 0 and 1. The optimum threshold is chosen to maximise overall classification performance of the model assuming equal costs of misclassification of species presence/absence. The area under the ROC curve (AUC) is an index of accuracy as it provides a single evaluation measure independent of any particular threshold. Confidence intervals were obtained for the AUC from 999 bootstraps from observed and predicted values for each taxon. We used the AUC from the ROC plots as the fifth evaluation method and to find the optimal probability threshold for each taxon (Hosmer and Lemeshow, 2000; Zweig and Campbell, 1993). Finally, we compared the predicted communities (using crossvalidation) to those observed at the 379 sites to measure the percentage of similarity between the observed and expected community using the simple matching coefficient, (Krebs, 1999).

Quantifying predictor variable contributions

To determine the relative importance of each predictor variable we used the method developed by Garson (1991) and later modified by (Goh, 1995) known as Garson's algorithm. This approach has been used in a number of ecological studies (e.g. Brosse *et al*, 2001; Gozlan *et al*, 1999; Olden and Jackson, 2002b). The method involves partitioning the hidden output connection weights of each hidden neurone into components associated with each output neurone. The result, expressed as a percentage gives the relative importance or distribution of output weights that can be attributed to each predictor variable.

However, Garson's algorithm does not provide the direction of the influence of the predictor variables on the output nodes (Olden and Jackson, 2002b). To elucidate the direction of relationship between predictions and predictor variables we used sensitivity analysis (Ozesmi and Ozesmi, 1999). This analysis involves varying one predictor variable over its entire range while holding all other variables at their mean value and examining the response of the output predictions. However, in this case where there are a large number of connections (6,300) the process becomes rather cumbersome. Therefore, we selected some variables of interest (generally associated with fish distribution) that also had high contributions from Garson's algorithm as examples and examined the direction of influence for these variables. All the procedures described above were run on Matlab version 6.1 using the Neural Networks toolbox.

Fish data

The data on freshwater fish presence and absence came from two sources, the New Zealand Freshwater Fish Database (NZFFD) (McDowall and Richardson, 1983) and recent surveys of the region (MKJ unpublished data). The data obtained from the NZFFD were for all sites sampled using electro-fishing since 1980 (259 sites). The remaining sites were sampled by single-pass electric fishing involving two people using a battery-powered electric fishing machine (EFM300; NIWA Instrument Systems), operated at 150-300 V depending on the water conductivity. Single pass electrofishing is a suitable survey method for assessing distribution of freshwater fish in New Zealand because all species are collected with equal probability on the first pass (Jowett and Richardson, 1996). Fish were collected in dip nets or stop nets held downstream and identified to species, counted and returned to the water. No attempt was made to use fish abundance/density data due to potential operator bias in data from NZFFD. Exact sampling details for the electro-fishing data from the NZFFD are not known but the general process described above is generally used in New Zealand freshwater fish surveys (all licensed operators have received the same training). Eighteen taxa were recorded at the 379 sites (Table 1). The survey sites gave a relatively broad coverage of the region apart from the eastern, lowland part of the region (Fig. 1).

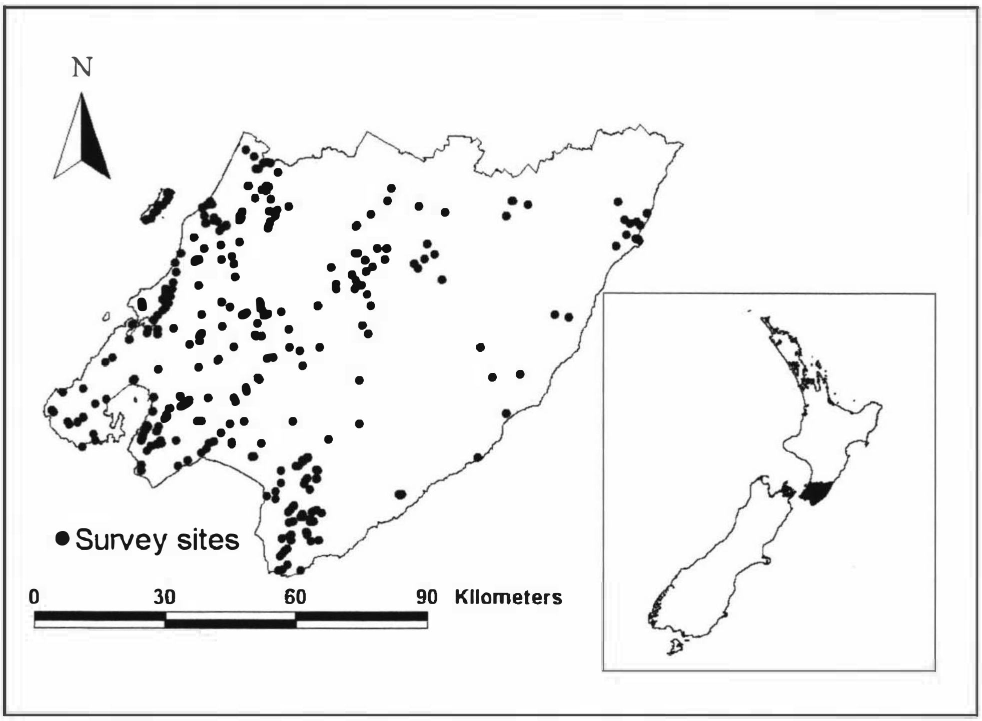

Fig. 1. Map of the Wellington Region, North Island New Zealand. Points denote survey sites.

Table 1. List of familial associations, species scientific and common names of the fish and decapod found at the 379 sites in the Wellington Region. (Taxa in bold non-migratory, asterisk denotes decapod)

Family	Scientific name	Common name
Anguillidae	*Anguilla australis*	Shortfin eel
	Anguilla dieffenbachii	Longfin eel
Pinguipedidae	*Cheimarrichthys fosteri*	Torrentfish
Galaxiidae	*Galaxias argentus*	Giant kokopu
	Galaxias brevipinnis	Koaro
	Galaxias divergens	**Dwarf galaxiid**
	Galaxias fasciatus	Banded kokopu
	Galaxias maculatus	Inanga
	Galaxias postvectis	Shortjaw kokopu
Geotriidae	*Geotria australis*	Lamprey
Eleotridae	*Gobiomorphus cotidianus*	Common bully

	Gobiomorphus gobioides	Giant bully
	Gobiomorphus hubbsi	Bluegill bully
	Gobiomorphus huttoni	Redfin bully
	Gobiomorphus breviceps* or *G. basalis	**Cran's or upland bully**
Parastacidae*	***Paranephrops planifrons***	Koura
Retropinnidae	*Retropinna retropinna*	Common smelt
Salmonidae	***Salmo trutta***	**Brown Trout**

GIS data

The GIS data on terrain, geology and land cover were obtained from the River Environment Classification (REC) (Snelder, 1998; Snelder, 2002). We used the raw data underlying the REC that describes in catchment proportions various environmental factors (e.g. geology, elevation, rainfall and vegetative cover), for individually numbered sections of the river network at a 1:50,000 mapping scale. The river network, stored as a set of GIS polygons (18,075 reaches for the region), consisted of adjoining sections (the reach units) derived from a 30m Digital Elevation Model (DEM). A uniquely identified node terminated each section (the minimum catchment area to define a section was 0.02 km^2). For each section of the region's river network (average length = 700m), each factor is described by the area of catchment occupied by various categories (e.g. geological categories include; greywacke, limestone, mudstone). The geology classifications for this dataset were sourced from the New Zealand Land Resources Inventory (NZLRI) (Newsome, 1990) which classifies geology into 55 categories at a scale of 1:50,000. Land cover catchment proportions came from the New Zealand Land Cover Database (NZLCDB) (Terralink International Limited). The NZLCDB defines land cover in 17 categories (e.g. indigenous forest, exotic forest, pastoral farming) and was derived from digitising satellite images at a mapping scale of 1:50,000.

For the climate variables, surfaces of annual mean precipitation, annual mean evapotranspiration and annual mean air temperature were estimated from thin plate splines fitted to meteorological station data as part of the REC process. Stream order was calculated from the network using the Strahler method (Strahler, 1964). All catchment data were associated with each corresponding sample site using the geoprocessing extension of ARCVIEW (ESRI Corporation, Redlands, California, USA). The steepest reach and the gradient of that reach for the waterway between the site and the coast were calculated using a 30-m spatial digital elevation model. The total distance from the site to the coast along the waterway was also calculated for each site by summing the reach lengths from the site to the sea. This process resulted in 69 variables for use as predictor variables (Table 2).

Table 2. Predictor variables used in neural network training from the raw data used in the REC classification (Snelder and Biggs 2002).

Variable		Minimum	Maximum	Mean
	Total catchment area (m^2)	22860	63400358	867587
	Estimated catchment area evaporation (mm)	637.1	740.7	704
	Estimated average annual flow (m^2/sec^{-1})	0.0027	23.6	1.2
	Average catchment rainfall (mm * m^2)	935.6	2926.9	1695
	Average catchment air temperature (C°)	7.6	12.9	10
	Average catchment slope (m/km)	0.6	61.9	36
	Average catchment elevation (m)	12	910	295
	Elevation at upstream end of site reach (m)	4	544	96
	Elevation at downstream end of site reach (m)	1	429	55
	Euclidean length of reach (m)	31	7071	723
	Length of reach (m)	31	8939	856
	Latitude of site reach	41.74	42.60	42
	Distance from site to sea (m)	72	169755	27506
	Maximum steepness downstream of site (m/km)	0.2	89	20.0
	Stream order (Strahler)	1	6	2.8
	Site reach identification number	381	18112	10092
	Reach slope (m/km)	0	65.7	7.7
	Lake	0	13.19	0.06
Geology	***Baserock***			
	Alluvium	0	75.23	0.47
	Argillite	0	99.72	0.17
	Argillite crushed	0	100	0.19
	Conglomerate or breccia	0	51.37	0.46
	Gravels	0	69.86	2.06
	Greywacke	0	100	60
	Limestone	0	11.93	0.04
	Loess	0	45.94	0.07
	Mudstone banded	0	34.85	0.07
	Mudstone betonic	0	42.06	0.02
	Mudstone jointed	0	35.81	0.05
	Mudstone massive	0	23.44	0.06
	Peat	0	23.66	0.07
	Sands	0	78.52	0.12
	Sandstone banded	0	6.33	0.03
	Sandstone massive	0	98.48	0.28

Unconsolidated	0	32.81	0.04
Surface rock			
Alluvium	0	99.99	1.48
Argillite crushed	0	99.99	0.19
Argillite	0	99.72	0.17
Conglomerate	0	46.08	0.23
Gravels	0	30.76	0.74
Greywacke	0	99.99	59.25
Limestone	0	14.69	0.043
Loess	0	86.73	1.48
Mudstone banded	0	29.55	0.069
Mudstone betonic	0	42.06	0.019
Mudstone jointed	0	35.81	0.053
Mudstone massive	0	21.92	0.059
Peat	0	23.66	0.075
Sands	0	78.52	0.12
Sandstone massive	0	98.48	0.18
Sandstone banded	0	6.33	0.03
Unconsolidated	0	59.13	0.03
Landcover			
Bare ground	0	0.12	0.0018
Coastal sands	0	0.04	0.0001
Coastal wetlands	0	0.0004	0.0000
Exotic forest	0	0.91	0.014
Horticultural	0	0.11	0.0003
Indigenous forest	0	1.0	0.19
Inland water	0	0.039	0.0004
Mines, dumps	0	0.06	0.0001
Pastoral	0	1.0	0.06
Scrub	0	1.0	0.14
Tussock	0	0.11	0.0013
Urban catchment	0	0.55	0.004
Urban open space	0	0.39	0.0009
Wetlands	0	0.01	0.0002

Extending predictions to the entire region

All predictor variables were available for the complete river network in the region and these data were entered into the completed ANN model and predictions produced for the entire river network. This resulted in a table of predictions for all 18 species at all the 18,075 numbered reaches; the probabilities were then mapped over the regional river network.

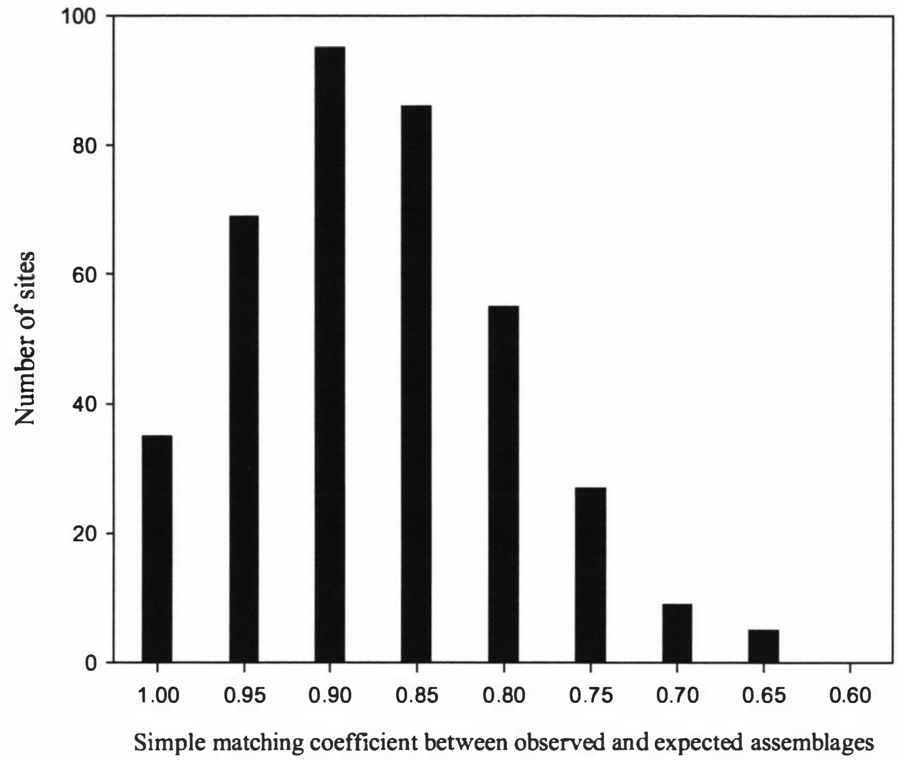

Fig. 2. Number of fish and decapod species correctly predicted using the ANN model at the 379 sites using crossvalidation.

Results

ANN evaluation

The structure of fish assemblages predicted by the neural network were very similar to the observed assemblages. There was on average an 86% similarity between the observed and generated crossvalidated assemblages using the simple matching coefficient. At more than 80 % of the sites the match between predicted and observed assemblages was greater than 90%, and at 90 % of the sites the match was 80% or greater (Fig. 2).

The ANN also exhibited very high levels of correct classification when considering the species individually. Using crossvalidation, the average correct classification rate for all species was 86% (Table 3). The least accurate predictions were for trout and redfin bullies at 71 and 74 % correctly classified, respectively. Sensitivity or the true positive prediction rate was moderate with an average of 51 %; the lowest value was for the rarest taxon, the giant bully. The average true negative rate (specificity) was higher at 87 %; the minimum was 16 % for the most common taxon, the longfin eel. The average Cohen's kappa value for all the species was 0.41. The average area under curve (AUC) value was also very high at 0.81 and all of the taxa had AUC values greater than 0.7. The optimal threshold values from the ROC analyses were all less than 0.5, once more reflecting the low prevalence of the majority of taxa.

Contribution of predictor variables

The results of the assessment of connection weights using Garson's algorithm are shown in Figure 3. The ranking reveals the relative importance of the contributions for each of the predictor variables. No particular variable class was more important than another with variables from geology landuse and longitudinal position all having high relative contributions. The first four variables relate to the longitudinal position of the sites along the waterway (reach length increases closer to the coast). The catchment area in scrub was the most important landcover variable followed by pastoral landcover, then exotic forestry. Loess was the most important geology variable, followed by baserock gravels. After the first 20 variables in the ranking the relative contribution was low and evenly spread among the remaining variables.

Table 3. Performance of neural network in predicting species presence/absence at 379 sites in the Wellington region based on leave-one-out crossvalidation. The reported values are the percent species occurrence (prevalence), the percent correct classification rate, sensitivity (true positive rate), specificity (true negative rate) and Cohen's kappa. The optimal decision threshold and area under curve based on receiver-operating characteristic analysis and the 95% confidence interval for the AUC was calculated from 999 bootstraps of the observed and predicted data.

Taxa	Prevalence %	Correct classification %	Sensitivity %	Specificity %	Cohen's kappa	Optimal threshold	Area under curve (AUC)	95 CI for AUC
Shortfin eel	33	79	61	87	0.51	0.53	0.84	0.79-0.88
Longfin eel	78	79	96	16	0.33	0.32	0.70	0.62-0.76
Torrentfish	9	91	25	97	0.38	0.66	0.79	0.71-0.88
Giant kokopu	7	86	45	90	0.27	0.94	0.84	0.68-0.85
Koaro	27	82	61	91	0.54	0.56	0.88	0.79-0.88
Dwarf galaxiid	7	94	43	98	0.48	0.56	0.86	0.62-0.74
Banded kokopu	11	86	52	90	0.39	0.24	0.80	0.73-0.87
Inanga	13	87	59	91	0.47	0.27	0.84	0.77-0.9
Shortjaw kokopu	7	94	44	98	0.49	0.52	0.84	0.75-0.92
Lamprey	4	97	36	99	0.49	0.70	0.80	0.66-0.92
Common bully	19	84	33	97	0.39	0.77	0.79	0.73-0.85
Giant bully	2	98	0	100	0.21	0.05	0.91	0.84-0.97
Bluegill bully	8	91	45	95	0.42	0.41	0.76	0.64-0.85
Redfin bully	46	71	79	63	0.43	0.37	0.75	0.7-0.79
Cran's or upland bully	19	89	53	97	0.58	0.55	0.80	0.74-0.88
Koura	24	78	46	88	0.37	0.45	0.72	0.65-0.78
Common smelt	19	91	77	91	0.26	0.05	0.91	0.84-0.97
Trout	40	74	72	75	0.46	0.42	0.76	0.71-0.81
Mean	21	86	51	87	0.41	0.46	0.81	

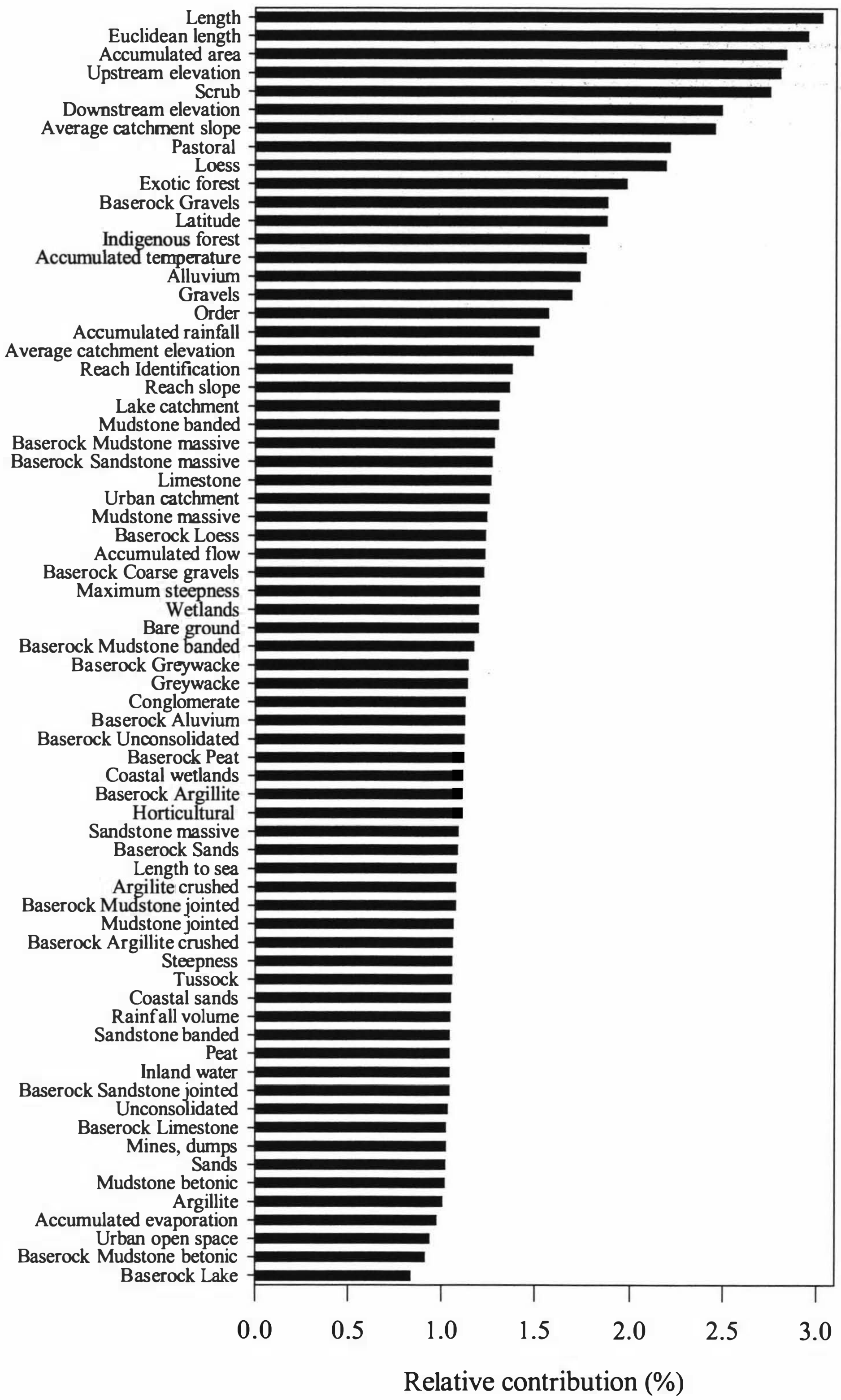

Fig. 3. Relative contribution of the predictor variables for predicting community composition averaged across all 18 species at the 379 sites in the Wellington region.

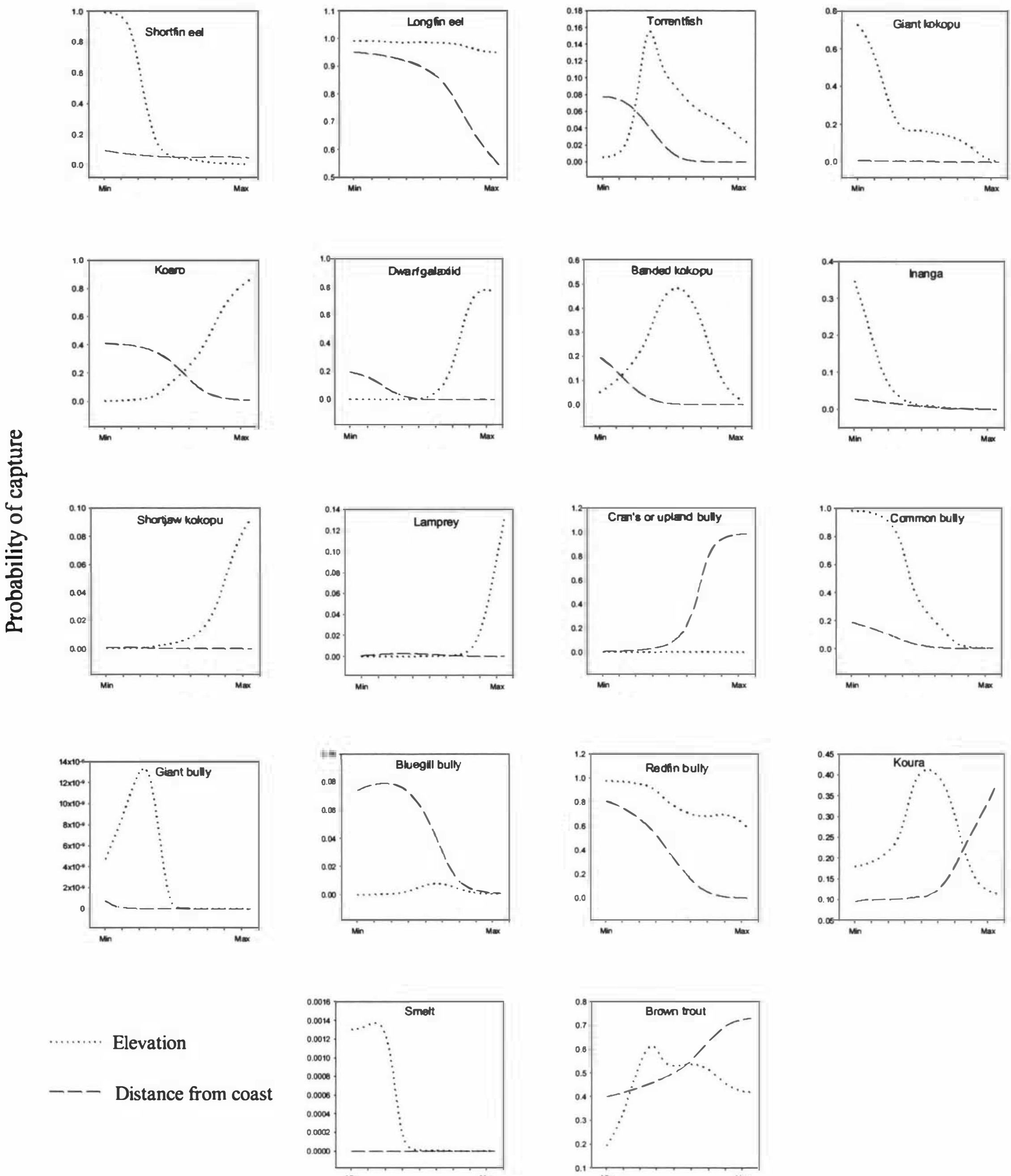

Fig. 4. Sensitivity analysis for distance from the coast (dashed line), and elevation (dotted line). In each plot (one for each taxon) all other variables were set to their mean values and the catchment proportions of pasture and scrub were varied over their full range. Note the y-axis scales differ depending on prevalence of the taxa.

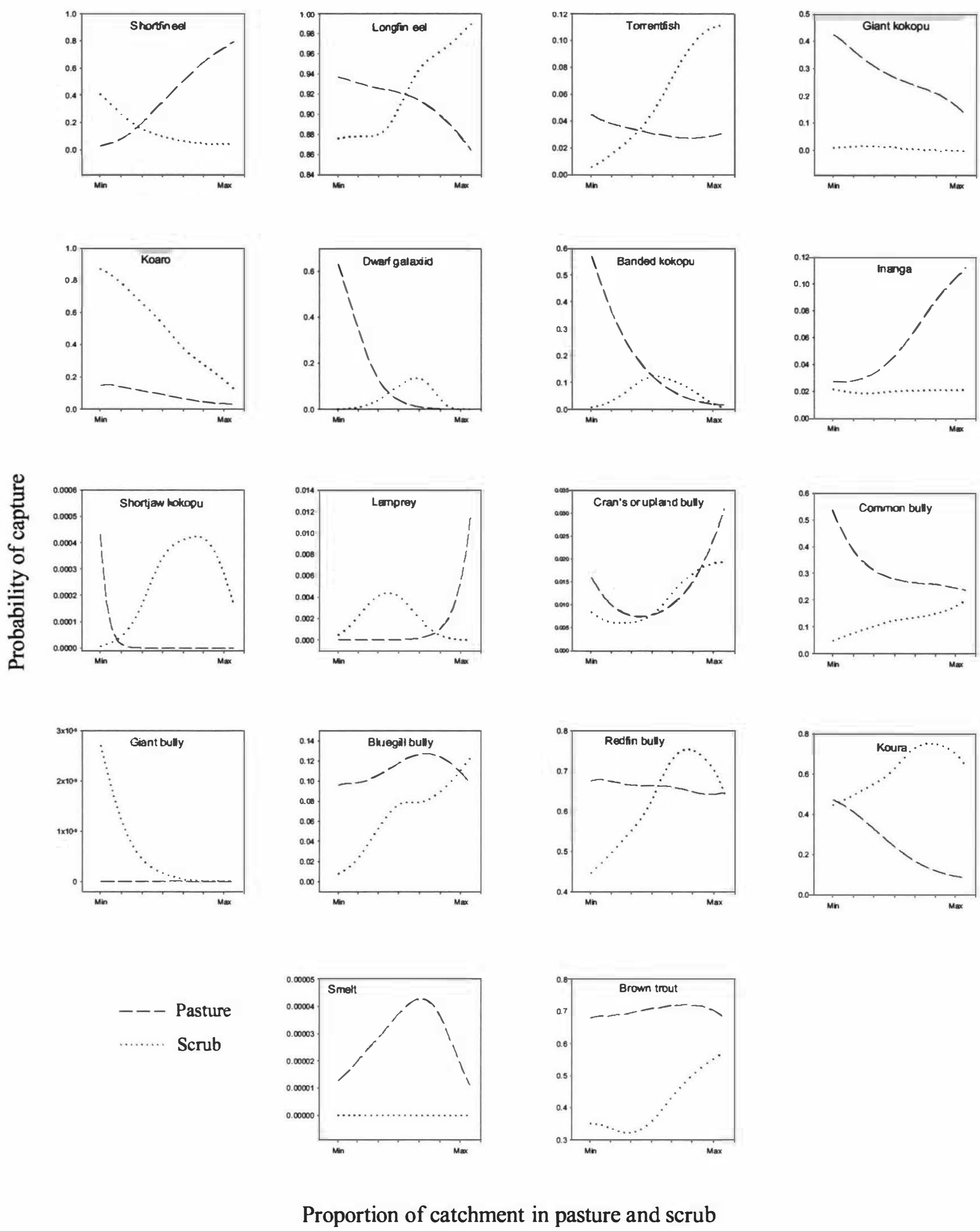

Fig. 5. Sensitivity analysis for the catchment proportions of pasture (dashed line), and scrub (dotted line). In each plot (one for each taxon) all other variables were set to their mean values and the catchment proportions of pasture and scrub were varied over their full range. Note the *y*-axis scales differ depending on prevalence of the taxa.

Sensitivity analyses

As an example of the use of sensitivity analysis to elucidate the direction of predictor variable effects on predictions, we selected 4 variables usually associated with freshwater fish distribution in New Zealand (Joy *et al*, 2000; McDowall, 1990). Distance from the coast and elevation revealed some general patterns with the non-migratory species (dwarf galaxiid, Cran's or upland bully, koura and brown trout) showing increasing probability of presence with increasing elevation or distance from the coast (Fig. 4). The migratory species generally showed the opposite pattern with probabilities of presence decreasing with increasing elevation or distance. Figure 5 shows the sensitivity analysis for catchment proportions of pasture and scrub. A number of species showed a similar pattern of increasing probability of presence with more scrub and decreasing probability of presence with increasing pasture.

Prediction map

To illustrate the use of the model we displayed the probabilities of occurrence using redfin bully as an example (Fig. 6). The map shows a mainly coastal distribution but extending farther inland on the southern and western coasts where indigenous forest occurs close to the coast. The large central-eastern area with no or low probability of occurrence is the part of the region dominated by pastoral farming. The inset map shows the southern Aorangi ranges and reveals the length of reaches and the level of detail of predictions. Maps can be produced for all species using a range of prediction categories.

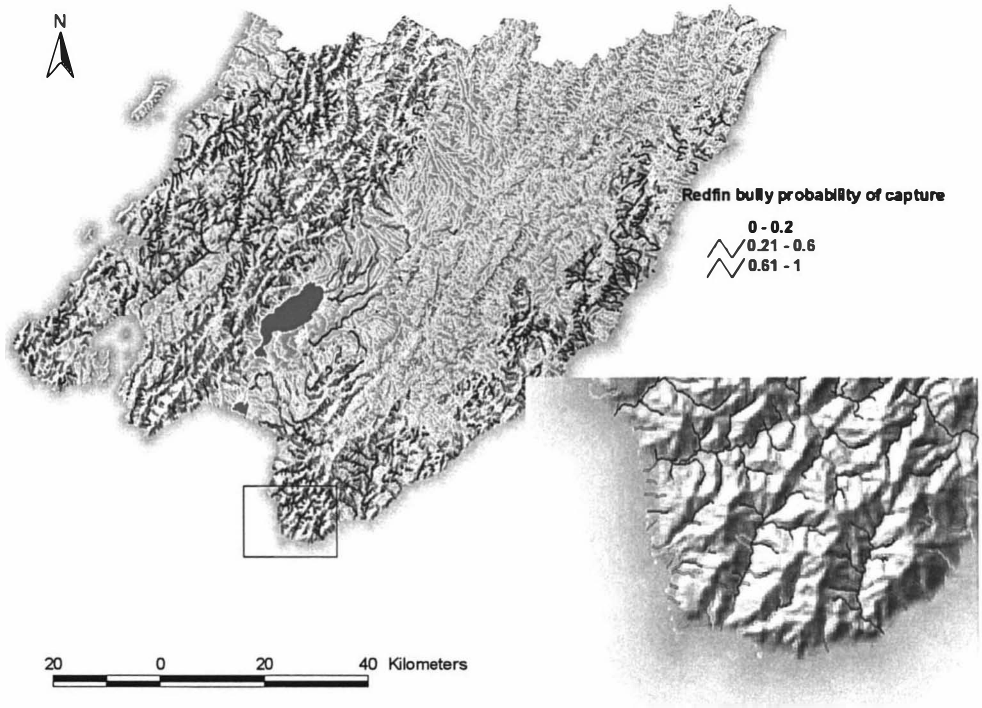

Fig. 6. Probability map of redfin bully (*Gobiomorphus huttoni*) in the Wellington Region (New Zealand). The inset shows the southern Aorangi Ranges.

Discussion

The ANN model described here succeeded in predicting fish and decapod communities with a high level of accuracy from geospatial landscape scale predictor variables. This result suggests that landscape scale patterns particularly geological, landcover, and spatial variables, at least at the reach scale, have considerable influence over New Zealand fish communities. Variables at this scale have also been found important as predictors of stream communities in temperate North American River systems (Milner *et al*, 1993; Paller, 1994; Poff and Allen, 1995). Furthermore, because these landscape scale characteristics are relatively constant over human time scales and can be remotely sensed, they allow species occurrence predictions to be mapped over an entire river network.

A number of studies have highlighted the importance of diadromy in structuring New Zealand fish communities because of the high proportion of diadromous species in the fauna (Joy and Death, 2001; Joy *et al*, 2000; McDowall, 1998a, b). Thus, large-scale variables such as elevation, distance to the sea, and the presence or absence of barriers to migration are likely to be even more important in controlling species' distribution in New Zealand than elsewhere. Undoubtedly, factors influencing fish and decapod distributions occur over a range of scales, but there is likely to be a hierarchy of effects so that the finer scale variables are constrained by the landscape scale factors (Poff, 1997; Levin, 1992).

The high level of accuracy of the neural network model presented here is likely to be related to a number of potential advantages that ANNs offer over conventional modeling techniques. These advantages include the ability to model non-linear relationships, handle mixed data types and the lack of a requirement for assumptions regarding data distributions, (Guegan *et al*, 1998; Olden and Jackson, 2002a). ANNs have another advantage; the ability to predict the entire community in one model (Olden in press) rather than by combining individual species models e.g. (Joy and Death, In Press; Oberdorff *et al*, 2001; Oberdorff *et al*, 2002) or groups representing fish communities (e.g. Joy and Death, 2002b). This also provides for the association of the relative importance of the influences of habitat variables on predicting the whole community, something that has not been possible with traditional statistical techniques.

GIS data

To our knowledge, this is the first regional scale application of a neural network predictive model to map the spatial occurrence of a complete riverine fish and decapod assemblage using geospatial data. The availability of a large GIS database containing information on catchment proportions of; geology, land use, climate etc. covering a complete regional river network presented us with a unique opportunity for producing a fish and decapod habitat suitability map. The geospatial predictor data we used were prepared as part of another project; a multi-scale river environment classification (Snelder and Biggs 2002). We used the raw data from the REC classification for our model to ensure we used the finest scale data (reach scale) available.

Model validation

Evaluation of the ANN model revealed high levels of predictive performance. Because the model was optimized using crossvalidation (jack-knifing) we are confident that it can accurately generalise the predictions over the region. Olden *et al*, (2002) demonstrated that a jack-knife approach to validation produced unbiased estimates of model performance as opposed to the highly biased resubstitution approach often used in predictive models. Considering all of the validation matrices (Table 3), the lowest model performance was for the prediction of the Giant bully. However, the low predictive power was not surprising, as this was the rarest taxon (occurring at only 2 % of the sites); the model has very little experience of giant bully preferred habitat.

Overall, the number of presences correctly predicted by the neural network (51 %) was lower than the number of absences (87 %) this is probably because most of the species were present at fewer than half of the sites (Table 3). Similar results have been found in other studies where the correct rate of predictions was higher for absences than presences with rare taxa (Manel *et al*, 2000; Manel *et al*, 1999; Manel *et al*, 2001; Olden and Jackson, 2002a; Olden *et al*, 2002). Thus, accounting for prevalence when validating predictive models is important (Manel *et al*, 2001; Olden and Jackson, 2002a; Olden *et al*, 2002). Manel *et al* (2001) found that Cohen's kappa was the most robust measure of model performance as it was only marginally affected by prevalence. The Cohen's kappa results for this model (Table 3) revealed the weakest predictions were for the rarest and most common species, the giant bully (0.21) and longfin eel (0.33), respectively.

Predictor variables

The relative contribution of the predictor variables based on the connection weights gave a general guide to their importance (Fig. 3). The two measures of reach length; both Euclidean and total had the highest ranking for contribution to predictions. The length of each reach is governed by distance between the entry point of one confluence and the next. Thus, these variables are proxies for and can therefore take into account facets of elevation, slope and stream size. The next two variables in the ranking were catchment area and elevation at the upstream end of the reach. These and many of the variables e.g. catchment area, stream order and accumulated flow are surrogates for different aspects of stream size and longitudinal network position. Hence they incorporate elevation and distance from the coast, trajectories that have been implicated in the distribution patterns of New Zealand freshwater fish in many studies (Joy and Death, 2001; Joy and Death, 2002a; Joy and Death, 2002b; McDowall, 1998a, b).

The first of the landcover/vegetation variables in the ranking, were scrub and pasture and these landcover variables have been associated with the distribution of fish species and the freshwater crayfish (koura) in other studies in New Zealand (e.g.Joy and Death, 2001; McDowall, 1997). Accumulated rainfall is a measure of the mean annual amount of rainfall in the catchment, which integrates climate with catchment landuse, and would influence water temperature and flow. Reach identification number is in effect a proxy for latitude and longitude as the reaches are numbered from west to east and then north to south. The first geological variable in the ranking was loess, which is wind blown material and is generally associated with stream sediment inputs. The next geology variable in the contribution ranking was baserock gravel, which is also likely to be related to water clarity. The rest of the variables had small relative contributions to overall predictions although it is probable they all had some influence in concert with other variables at some sites.

Sensitivity analyses

The plots of probability of capture against selected predictor variables (Fig. 4 and 5) revealed the response of taxa to predictor variables. However, these plots only reveal the general direction of the response as holding all other variables at their mean value may obscure many interaction effects. An example of this is the very low probabilities for some of the rare taxa in the plots. In a number of studies, the response curves have been shown to vary according to the value of the other variables not being varied. For example in a similar study (Olden and Jackson, 2002b) varied one variable through its full range while holding all others at their 20th, 40th, 60th and 80th percentiles and showed that while the general direction did not change there was often significant response variation over the percentile change. Thus, species with very low probabilities when the other variables are held at their mean do have high probabilities when realistic combinations of the other variables are used.

Predictive map

The development of predictive models that provide testable whole-community predictions have the potential to contribute significantly to restoration, and management of aquatic systems. The prediction map (Fig. 6) enables a visual representation of the probability of occurrence for each of the species and the associated habitat types. Each of the species can be mapped separately and the number of probability classes adjusted to suit requirements (e.g. two classes of presence/absence based on thresholds from table 3 or a range of probabilities based on the probability distribution). Furthermore, the variables of interest (e.g. landuse, geology, and elevation) can be overlaid over the predictive network to see how they relate to the occurrence of the species of interest. The maps can highlight areas where species are predicted to occur but are absent because of impacts, human induced or natural. An example is the Orongorongo River above a water supply weir where a number of diadromous fish species (redfin bully and banded kokopu) are predicted to occur but are absent as they are unable to scale the face of the dam. Similarly, there are a number of areas where the non-migratory dwarf galaxiids are predicted to occur but are not found (MKJ unpublished data). These are areas where they may have occurred in the past but have been locally extirpated by some impact such as an extreme flood event or drought and are then unable to recolonise. The individual species' predictive maps can be used by conservation agencies to prioritise survey areas by highlighting where rare or endangered species are predicted to occur but have not been surveyed. The maps can also be used to reveal where unexpected gaps appear in distributions or potential additional populations. The maps have been made available for water resource managers at the regulatory authority (The Wellington Regional Council) and the GIS format of the predictive map means that all all staff can have access to complete prediction maps and links lead users to details of ecology, research literature and habitat requirements for the taxa predicted. This ability to easily access relevant data will potentially improve resource allocation decision making.

Conclusion

The model we have presented here revealed a high predictive power given the scale of the predictor variables and number of species being modelled. However, further research is needed to survey sites to validate the model predictions away from areas where input data was available for model construction. The extension of the model predictions using GIS data to the entire river network opened up a range of potential uses. Examples of potential uses include; monitoring and predicting temporal changes caused by human activities and shifts in climate, the elucidation of areas in need of protection, biodiversity hotspots, and areas for the reintroduction of endangered or rare species. From a management perspective, the model provides easily accessible species distribution information in a simple format that can be linked to species specific biological information. Finally, this model conjugates the power of both GIS and ANNs to formalise the link between the species and their habitat.

Acknowledgments

This work was funded by a grant from the sustainable management fund (New Zealand Ministry for the Environment grant No. 5099 “Nga Ika Waiora”: stream health evaluation tool) and the Wellington Regional Council. Thanks to Allison Hewitt for assistance with all fieldwork. Thanks also to Murray McLea at the Wellington Regional Council for help with fieldwork logistics and funding. For comments on early drafts we thank Giles Foody and Julian Olden.

References

Armitage, P.D. and Petts, G.E. (1992) Biotic score and prediction to assess the effects of water abstractions on river macroinvertebrates for conservation purposes. *Aquatic Conservation: Marine and Freshwater Ecosystems*, **2**, 1-17.

Bishop, C.M. (1995) Neural networks for pattern recognition Oxford University Press Inc., New York, USA.

Brosse, S., Guegan, J.F., Tourenq, J.N., and Lek, S. (1999) The use of artificial neural networks to assess fish abundance and spatial occupancy in the littoral zone of a mesotrophic lake. *Ecological Modeling*, **120**, 299-311.

Brosse, S., Lek, S., and Townsend, C.R. (2001) Abundance, diversity, and structure of freshwater invertebrates and fish communities: an artificial neural network approach. *New Zealand Journal of Marine and Freshwater Research*, **35**, 135-145.

Chen, D.G. and Ware, D.M. (1999) A neural network model for forecasting fish stock recruitment. *Canadian Journal of Fisheries and Aquatic Science*, **56**, 2385-2396.

Chessman, B.C., Growns, I.O., and Plunket-Cole, N. (1999) Predicting diatom communities at genus level for the biological management of rivers. *Freshwater Biology*, **41**, 317-331.

Chon, T.S., Park, Y.S., Moon, K.H., and Cha, E.Y. (1996) Patternizing communities by using an artificial neural network. *Ecological Modeling*, **90**, 69-78.

De'Ath, G. (2002) Multivariate regression trees: a new technique for modeling species-environment relationships. *Ecology*, **83**, 1105-1117.

Fielding, A.H. and Bell, J.F. (1997) A review of methods for the assessemnt of prediction errors in conservation presence/absence models. *Environmental Conservation*, **24**, 38-49.

Filipe, A.F., Cowx, I.G., and Collares-Pereira, M.J. (2002) Spatial modeling of freshwater fish in semi-arid river systems: A tool for conservation. *River Research and Applications*, **18**, 123-136.

Franklin, J. (1998) Predicting the distribution of shrub species in southern California from climate and terrain-derived variables. *Journal of Vegetation Science*, **9**, 733-748.

Garson, G.D. (1991) Interpreting neural-network connection weights. *Artificial Intelligence Expert*, **6**, 47-51.

Goh, A.T.C. (1995) Back-propagation neural networks for modeling complex systems. *Artificial Intelligence Engineering*, **9**, 143-151.

Gozlan, R.E., Mastrorillo, S., Copp, G.H., and Lek, S. (1999) Predicting the structure and diversity of young-of-the-year fish assemblages in large rivers. *Freshwater Biology*, **41**, 809-820.

Guegan, J.F., Lek, S., and Oberdorff, T. (1998) Energy availability and habitat heterogeneity predict global riverine fish diversity. *Nature*, **391**, 382-384.

Guisan, A. and Zimmermann, N.E. (2000) Predictive habitat distribution models in ecology. *Ecological Modeling*, **135**, 147-186.

Hoang, H., Recknagel, F., Marshall, J., and Choy, S. (2001) Predictive modeling of macroinvertebrate assemblages for stream habitat assessments in Queensland (Australia). *Ecological Modeling*, **146**, 195-206.

Hosmer, D.W. and Lemeshow, S. (2000) Applied logistic regression, 2nd ed. A Wiley-Interscience publication, New York NY.

Iverson, L.R. and Prassad, A.M. (1998) Predicting abundance of 80 tree species following climate change in the eastern united states. *Ecological Monographs.*, **68**, 465-485.

Johnson, D.H. (1999) The insignificance of hypothesis testing. *Journal of Wildlife Management*, **63**, 763-772.

Jowett, I.G. and Richardson, J. (1996) Distribution and abundance of freshwater fish in New Zealand rivers. *New Zealand Journal of Marine and Freshwater Research*, **30**, 239-255.

Joy, M.K. and Death, R.G. (2000) Development and application of a predictive model of riverine fish community assemblages in the Taranaki region of the North Island, New Zealand. *New Zealand Journal of Marine and Freshwater Research*, **34**, 243-254.

Joy, M.K. and Death, R.G. (2001) Control of freshwater fish and crayfish community structure in Taranaki, New Zealand: dams, diadromy or habitat structure? *Freshwater Biology*, **46**, 417-429.

Joy, M.K. and Death, R.G. (2002a) A discriminant analysis investigation of reference site fish assemblages in the Manawatu-Wanganui region, North Island, New

Zealand. *Verhandlungen der Internationalen Vereinigung fur Theoretische und Angewandte Limnologie*, **28**, 319-322.

Joy, M.K. and Death, R.G. (2002b) Predictive modeling of freshwater fish as a biomonitoring tool in New Zealand. *Freshwater Biology*, **47**, 2261-2275.

Joy, M.K. and Death, R.G. (In Press) Neural network modeling of freshwater fish and macro-crustacean assemblages for biological assessment in New Zealand. In Modeling community structure in freshwater ecosystems (eds S. Lek, M. Scardi and S.E. Jorgensen). Springer Verlag.

Joy, M.K. and Death, R.G. (In Press) Assessing biological integrity using freshwater fish and decapod habitat selection functions. *Environmental Management*.

Joy, M.K., Henderson, I.M., and Death, R.G. (2000) Diadromy and longitudinal patterns of upstream penetration of freshwater fish in Taranaki, New Zealand. *New Zealand Journal of Marine and Freshwater Research*, **34**, 531-543.

Krebs, C.J. (1999) Ecological methodology, 2nd edn. Benjamin/Cummings, Menlo Park, California.

Kurkova, V. (1992) Kolmogrov's theorem and multilayer neural networks. *Neural Networks*, **5**, 501-506.

Landis, J.R. and Koch, G.G. (1977) The measurements of observer agreement for categorical data. *Biometrics*, **33**, 159-174.

Lehmann, A. (1998) GIS modeling of submerged macrophyte distribution using generalised additive models. *Plant Ecology*, **139**, 113-124.

Lehmann, A., Jaqet, J.M., and Lachavanne, J.B. (1997) A GIS approach of aquatic plant spatial heterogeneity in relation to sediment and depth gradient, Lake Geneva, Switzerland. *Aquatic Botany*, **58**, 347-361.

Lek, S., Giraudel, J.L., and Guegan, J. (2000). Neuronal Networks: algorithms and architectures for ecologists and evolutionary ecologists. In Artificial Neuronal Networks (eds S. Lek and J. Guegan), pp. 262. Springer-Verlag, Berlin.

Levin, S.A. (1992) The problem of scale in ecology. *Ecology*, **73**, 1943-1967.

Manel, S., Buckton, S.T., and Ormerod, S.J. (2000) Testing large-scale hypotheses using surveys: the effects of land use on the habitats, invertebrates and birds of Himalayan rivers. *Journal of Applied Ecology*, **37**, 756-770.

Manel, S., Dias, J.M., Buckton, S.T., and Ormerod, S.J. (1999) Alternative methods for predicting species distribution: an illustration with Himalayan river birds. *Journal of Applied Ecology*, **36**, 734-747.

Manel, S., Williams, H.C., and Ormerod, S.J. (2001) Evaluating presence-absence models in ecology: the need to account for prevalence. *Journal of Applied Ecology*, **38**, 921-931.

Marchant, R. and Hehir, G. (2002) The use of AUSRIVAS predictive models to assess the response of lotic macroinvertebrates to dams in south-east Australia. *Freshwater Biology*, **47**, 1033-1050.

Mastrorillo, S., Dauba, F., Oberdorff, T., Guegan, J.F., and Lek, S. (1998) Predicting local fish species richness in the Garonne River basin. *Comptes Rendus De L Academie Des Sciences Serie Iii-Sciences De La Vie-Life Sciences*, **321**, 423-428.

Mastrorillo, S., Lek, S., and Dauba, F. (1997a) Predicting the abundance of minnow Phoxinus phoxinus (Cyprinidae) in the River Ariege (France) using artificial neural networks. *Aquatic Living Resources*, **10**, 169-176.

Mastrorillo, S., Lek, S., Dauba, F., and Belaud, A. (1997b) The use of artificial neural networks to predict the presence of small-bodied fish in a river. *Freshwater Biology*, **38**, 237-246.

McDowall, R.M. (1990) New Zealand Freshwater Fishes: A Natural History and Guide. Heinemann Reed, Auckland.

McDowall, R.M. (1997) Indigenous vegetation type and the distribution of shortjawed kokopu, Galaxias postvectis (Teleostei: Galaxiidae), in New Zealand. *New Zealand Journal of Zoology*, **24**, 243-255.

McDowall, R.M. (1998a) Driven by diadromy: its role in the historical and ecological biogeography of the New Zealand freshwater fish fauna. *Italian Journal of Zoology*, **65**, 73-85.

McDowall, R.M. (1998b) Fighting the flow: downstream-upstream linkages in the ecology of diadromous fish faunas in West Coast New Zealand rivers. *Freshwater Biology*, **41**, 111-123.

McDowall, R.M. and Richardson, J. (1983). The New Zealand freshwater fish survey-a guide to input and output. Report. No. 12. Ministry of agriculture and Fisheries, Wellington.

Milner, N.J., Wyatt, R.J., and Scott, M.D. (1993) Variability in the distribution and abundance of stream salmonids, and associated use of habitat models. *Journal of Fish Biology*, **43**, 102-119.

Newsome, P.F.J. (1990) New Zealand Land Resources Inventory ARC/INFO Data Manual, Landcare Research, New Zealand.

Oberdorff, T., Guegan, J.F., and Hugueny, B. (1995) Global scale patterns of fish species richness in rivers. *Ecography*, **18**, 345-352.

Oberdorff, T., Pont, D., Hugueny, B., and Chessel, D. (2001) A probabilistic model characterizing fish assemblages of French rivers: a framework for environmental assessment. *Freshwater Biology*, **46**, 399-415.

Oberdorff, T., Pont, D., Hugueny, B., and Porcher, J. (2002) Development and validation of a fish-based index for the assessment of 'river health' in France. *Freshwater Biology*, **47**, 1720-1734.

Olden, J.D. and Jackson, D.A. (2000) Torturing data for the sake of generality: How valid are our regression models? *Ecoscience*, **7**, 501-510.

Olden, J.D. and Jackson, D.A. (2001) Fish-habitat relationships in lakes: Gaining predictive and explanatory insight by using artificial neural networks. *Transactions of the American Fisheries Society*, **130**, 878-897.

Olden, J.D. and Jackson, D.A. (2002a) A comparison of statistical approaches for modeling fish species distributions. *Freshwater Biology*, **47**, 1976-1995.

Olden, J.D. and Jackson, D.A. (2002b) Illuminating the "black box": a randomization approach for understanding variable contributions in artificial neural networks. *Ecological Modeling.*

Olden, J.D., Jackson, D.A., and Peres-Neto, P.R. (2002) Predictive models of fish species distributions: A note on proper validation and chance predictions. *Transaction of the American Fisheries Society*, **131**, 329-336.

Olden, J. D. 2003. A species-specific approach to modeling biological communities and its potential for conservation. Conservation Biology **17**:1-11.

Paller, M.H. (1994) Relationships between fish assemblages structure and stream order in South Carolina Coastal Plain Streams. *Transaction of the American Fisheries Society*, **123**, 150-161.

Park, Y.S., Kwak, I.S., Chon, T.S., Kim, J.K., and Jorgensen, S.E. (2001) Implementation of artificial neural networks in patterning and prediction of exergy in response to temporal dynamics of benthic macroinvertebrate communities in streams. *Ecological Modeling*, **146**, 143-157.

Peters, R. H. 1991. A Critique for Ecology. Cambridge University Press, Cambridge.

Poff, N.L. and Allen, J.D. (1995) Functional organisation of stream fish assemblages in relation to Hydrological variability. *Ecology*, **76**, 606-627.

Poff, N. L. 1997. Landscape filters and specific traits: toward mechanistic undrerstanding and prediction in stream ecology. Journal of the North American Benthological Society **16**:391-409.

Recknagel, F. and Wilson, H. (2000). Elucidation and prediction of aquatic ecosystems by artificial neuronal networks. In Artificial Neuronal Networks: Application to Ecology and Evolution (eds S. Lek and J.F. Guegan), pp. 262. Springer-Verlag, Berlin.

Reyes-Gavilan, F.G., Garrido, R., Nicieza, A.G., Toledo, M.M., and Brana, F. (1996) Fish community variation along physical gradients in short streams of Northern Spain and the disruptive effect of dams. *Hydrobiologia*, **321**, 155-163.

Reynoldson, T.B. and Rosenberg, D.M. (1996). Sampling strategies and practical considerations in building reference databases for the prediction of invertebrate community structure. In Study design and data analysis in benthic macroinvertebrate assessments of freshwater ecosystems using a reference site approach. (eds R.C. Baily, R.H. Norris and T.B. Reynoldson), pp. 1-31.

Rumelhart, D.E., Hinton, G.E., and Williams, R.J. (1986) Learning representations by back-propagating errors. *Nature*, **323**, 533-536.

Scardi, M. (2001) Advances in neural network modeling of phytoplankton primary production. *Ecological Modeling*, **146**, 33-45.

Scheiter, I.M., Borchardt, D., Wagner, R., Dapper, T., K., S., Schmidt, H., and Werner, H. (1999) Modeling water quality, bioindication and population dynamics in lotic ecosystems using neural networks. *Ecological Modeling*, **120**, 271-286.

Snelder, T., Biggs, B.J.F., Shankar, U., McDowall, B., Stephens, T., and Boothroyd, I.K.G. (1998) Development of a system of physically based habitat classification for water resources management of New Zealand rivers NIWA, Christchurch, New Zealand.

Snelder, T. and Biggs, B.J.F. (2002) Multi-Scale River Environment Classification for Water Resources Management. *Journal of the American Water Resources Association*, **38**, 1225-1240.

Strahler, A.N. (1964). Quantitative geomorphology of drainage basins and channel networks. In Handbook of Applied Hydrology (ed V.T. Chow), pp. 439-476. McGraw-Hill, New York, NY.

Terralink International Limited New Zealand Landcover Database (LCDB1) http://www.terralink.co.nz/metadata/ANZNZ3001800023_external.htm.

Titus, K., Mosher, J.A., and Williams, B.K. (1984) Chance-corrected classification for use in discriminant analysis. *The American Midland Naturalist*, **111**, 1-7.

Torgerson, C.E., Price, D.M., Li, H.W., and Mclntosh, B.A. (1999) Multiscale thermal refugia and stream habitat associations of Chinook salmon in north-eastern Oregon. *Ecological Applications*, **9**, 301-319.

Walley, W.J. and Fontama, V.N. (1998) Neural network predictors of average score per taxon and number of families at unpolluted river sites in Great Britian. *Water Research*, **32**, 613-622.

Wright, J.F., Furse, M.T., and Moss, D. (1998) River classification using invertebrates: RIVPACS applications. *Aquatic Conservation-Marine and Freshwater Ecosystems*, **8**, 617-631.

Wright, J.F., Gunn, R.J.M., Blackburn, J.H., Grieve, J.M., Winder, J.M., and Davy-Bowker, J. (2000) Macroinvertebrate frequency data for the RIVPACS lll sites in Northern Ireland and some comparisons with equivalent data for Great Britain. *Aquatic Conservation: Marine and Freshwater Ecosystems*, **10**, 371 - 389.

Zweig, M. and Campbell, G. (1993) Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clinical Chemistry*, **39**, 561-577.

CHAPTER SIX

Application of the index of biotic integrity methodology to New Zealand freshwater fish communities

.

Joy, M. K., and R. G. Death. Submitted. Application of the index of biotic integrity methodology to New Zealand freshwater fish communities. Environmental Management.

Abstract

An index of biotic integrity (IBI) was developed for freshwater fish in New Zealand streams. Data on freshwater fish occurrence for 5007 sites over the entire country were obtained from the New Zealand freshwater fish database for the period 1980 – 2002. Corresponding environmental descriptors for the stream catchments above or at each of these sites were obtained from a number of databases using a geographic information system. Of the 12 original North American IBI metrics, only six were adapted and applied because of differences between the fish faunas of New Zealand and the United States of America. A number of evaluation methods showed all six metrics contributed to the overall IBI scores with high levels of consistency. The IBI assessment of sites sampled at different times showed high levels of temporal concordance. The IBI scores differed significantly between some of the 15 geopolitical regions, the sampling method used and the year of survey. Overall, the results presented demonstrate the potential for New Zealand freshwater fish to be used to assess river condition at large spatial scales in New Zealand in the absence of specifically selected reference sites. This application demonstrates the effectiveness of the IBI approach even with a fauna of limited diversity, and limited ecological specialisation as in the New Zealand fish fauna.

Key words. Freshwater fishes; Index of biotic integrity; New Zealand; diadromy; biological indicators; reference sites.

Introduction

Using biological indicators to evaluate environmental quality has had a long history and continues to increase in popularity (Davis and Simon 1995). One way biological indicators can be used to evaluate riverine environmental quality is to measure the biological community of interest and then compare it with communities in similar representative unimpacted areas. These unimpacted or minimally impacted sites, referred to as reference sites, have been used to assess biotic attributes of freshwater sites using a variety of procedures. The reference site methodology can be applied using predictive multivariate techniques e.g. (Chessman *et al.* 1999, Simpson and Norris 2000, Joy and Death 2002) whereas others have applied it using multimetric techniques such as the Index of biotic integrity e.g. (Fausch *et al.* 1984, Angermeier and Karr 1986). While both are reference site approaches, they use different methods to associate test and reference sites, (Reynoldson *et al.* 1997). Multimetric approaches such as the IBI generally use an ecoregion or bioregion approach (Barbour *et al.* 1995, Omernik 1995) while the predictive approach uses multivariate biological classification (Wright *et al.* 2000, Joy and Death 2002).

An alternative to the *a priori* classification of sites as reference sites is the use of a large number of sites not selected as reference sites, but presumably including a number of minimally impacted sites that represent the best stream conditions in the region. The assumption is that as some of these sites are unimpacted they will provide an upper bound for expectations of biological structure (Simon and Lyons 1995). This is the technique used by (Karr 1981) and (Fausch *et al.* 1984) for the development of the index of biotic integrity (IBI) and is the approach we have taken by using a large number of sites from an existing database covering a full range of conditions.

Originally developed for fish assemblages in small streams in the northern United States (Karr *et al.* 1986) the IBI has since been modified for streams in Canada, Mexico, France, Australia, Africa, Belgium and India (Hughes *et al.* 1998). These IBI applications using fish have been based on the compositional, structural, and functional organisation of fish assemblages, and they generally incorporate multiple metrics of species richness, composition, reproductive function, abundance and fish condition (Fausch *et al.* 1990). The individual metrics and their scoring criteria represent an empirically based model of the characteristics of fish assemblages representing a range of conditions from unimpacted to degraded. Thus the metrics, theoretically at least, describe the changes that occur in fish assemblages, as an ecosystem becomes impaired.

McDowall and Taylor (2000) suggested that one reason why the IBI had not been applied in New Zealand was the overriding influence that diadromy has on fish distribution. The high proportion of migratory species means that differences in fish distribution may relate to factors other than habitat quality. McDowall and Taylor (2000) made the point that this makes application of the IBI as a measure of habitat quality problematical since the index depends on a strong relationship between population metrics and habitat quality. To alleviate this potential uncoupling of the association between metrics and habitat quality we accounted for the main variables driving New Zealand freshwater fish distributions (elevation and distance from the coast) by using these variables to plot maximum species richness expectations. The other unique aspects of the New Zealand freshwater-fish fauna that could potentially impact on the application of an IBI are the small number of species, comprising only 36 species from eight families and the presence of just a single trophic guild (McDowall 2000). Thus, the fauna is considerably less diverse and less ecologically specialised than the North American fauna where the IBI was developed (Karr 1981, McDowall and Taylor 2000).

In this application, we used metrics derived or adapted from previously published applications of the IBI. We modified or removed metrics to reflect the unique attributes of New Zealand stream fish assemblages but retained the general conceptual framework of the index and used a large number of sites over the entire country. The purpose of this study was to evaluate whether a fish index of biotic integrity could be applied to New Zealand streams by assessing the relationship between the IBI scores and landscape variables known to affect fish habitat quality. The next objective was to investigate the contributions of each of the metrics to the overall IBI score. Our third objective was to evaluate the IBI as a measure of stream condition by comparing IBI scores at sites with different landuse practices known to be associated with habitat and water quality degradation.

Methods and data sources

Environmental data

The GIS data on terrain, geology and landcover were obtained from the River Environment Classification (REC) (Snelder *et al.* 1998, Snelder and Biggs, 2002). We used the raw data underlying the REC, that describe in watershed attributes, various environmental factors (e.g. geology, elevation, rainfall and vegetative cover), for individually numbered sections of the river network at a 1:50 000 mapping scale. The river network, stored as a set of GIS lines (576 690 reaches for the country) consists of numbered adjoining sections (the reach units) derived from a 30m Digital Elevation Model (DEM). A uniquely identified node terminates each section (the minimum watershed area to define a section was 0.02 km^2). For each section of the region's river network (average length = 700 m), each factor is described by the area of watershed occupied by various categories (e.g. geological categories include; greywacke, limestone, mudstone). The geology classifications for this dataset were sourced from the New Zealand Land Resources Inventory (NZLRI) (Newsome 1990) which classifies geology into 55 categories at a scale of 1:50,000. Land cover watershed proportions came from the New Zealand Land Cover Database (NZLCDB) (Terralink International Limited). The NZLCDB defines land cover in 17 categories (e.g. indigenous forest, exotic forest, pastoral farming) and was derived from digitising satellite images at a mapping scale of 1:50,000. A number of geological and landuse variables were combined into groups based on their perceived relationship to stream biotas resulting in 51 variables (Appendix 1).

Table 1. Fish species from the New Zealand freshwater fish database used to construct Index of Biotic Integrity and their metric classes (see text for details on allocation of species to metrics).

Scientific name	Common name	Native	Benthic riffle	Benthic pool	Pelagic pool	Intoler ant
Aldrichetta forsteri	Yelloweye mullet	*			*	
Anguilla australis	Shortfin eel	*		*		
Anguilla dieffenbachii	Longfin eel	*	*	*		
Cheimarrichthys fosteri	Torrentfish	*	*			
Galaxias argenteus	Giant kokopu	*			*	*
Galaxias brevipinnis	Koaro	*	*	*		*
Galaxias divergens	Dwarf galaxias	*	*			*
Galaxias fasciatus	Banded kokopu	*		*	*	*
Galaxias maculatus	Inanga	*			*	
Galaxias spp.	Otago Galaxiids (non-migratory)	*	*			*
Galaxias paucispondylus	Alpine galaxias	*	*			*
Galaxias postvectis	Shortjaw kokopu	*		*	*	*
Galaxias prognathus	Longjaw galaxias	*	*			*
Galaxias vulgaris	Canterbury galaxias	*	*			*
Geotria australis	Lamprey	*		*		
Gobiomorphus basalis	Crans bully	*		*		
Gobiomorphus breviceps	Upland bully	*		*		
Gobiomorphus cotidianus	Common bully	*		*		
Gobiomorphus gobioides	Giant bully	*		*		*
Gobiomorphus hubbsi	Bluegill bully	*	*			*
Gobiomorphus huttoni	Redfin bully	*	*			*
Mugil cephalus	Grey mullet	*			*	
Neochanna apoda	Brown mudfish	*		*		*
Neochanna burrowsius	Canterbury mudfish	*		*		*
Neochanna diversus	Black mudfish	*		*		*
Retropinna retropinna	Common smelt	*			*	
Rhombosolea retiaria	Black flounder	*		*		*
Stokellia anisodon	Stokell's smelt	*				
Ameiurus nebulosus	Catfish			*		
Carassius auratus	Goldfish					
Ctenopharyngodon idella	Grass carp					
Gambusia affinis	Mosquitofish					
Oncorhynchus mykiss	Rainbow trout					
Oncorhynchus nerka	Sockeye salmon					
Oncorhynchus tshawytscha	Chinook salmon					
Perca fluviatilis	Perch					
Poecilia reticulata	Guppy					
Salvelinus fontinalis	Brook char					
Salmo trutta	Brown trout					
Scardinius erythrophthalmus	Rudd					
Tinca tinca	Tench					

Fish data

The biological data was derived from the New Zealand Freshwater Fish Database (NZFFD) (McDowall and Richardson 1983, Richardson 1989). This is an archive of data on distributions of fish in New Zealand freshwaters containing at present c. 18 000 sites. We used fish and decapod data from 1980 onwards that included: the presence or absence of 44 species (Table 1), the fishing method used, and whether the surveyor knew of a barrier downstream from the site. There were a number of different fishing methods listed in the database and these were classified into 2 broad groups; sites that were electrofished and sites that were sampled using traps or nets. For this study, we used a subset of this database. The selection was made based on the unambiguous association of a site with a its correct numbered stream reach using a geographic information system (ArcView 1999) (see below). Sites from the freshwater fish database were associated with the river network using the geoprocessing extension of ArcView (1999). Initially when sites were associated with their corresponding river reaches there were a number of cases where there was potential for a site to be associated with the wrong numbered reach particularly near stream junctions. This could occur where sample site co-ordinates were not accurately recorded or stream reaches were plotted imprecisely. To remove this potential for collating incorrect data, buffers lines were drawn around sites at 50 m, 100 m and 150 m diameters and stream reaches contacting the buffers were recorded. Sites that encountered a single unique stream reach at all three buffer lines were retained for analysis, as their association with the correct numbered stream reach was unambiguous. This process resulted in 5007 sites for IBI development (Fig. 1).

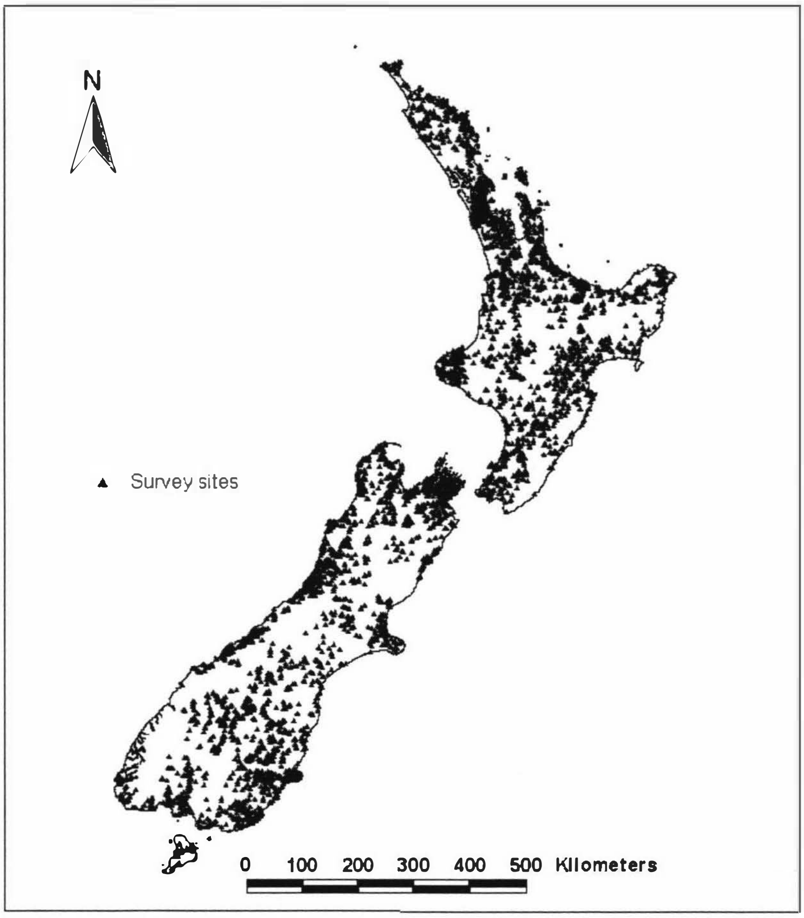

Fig. 1. Map of New Zealand showing the position of the 5007 study sites from the New Zealand freshwater fish database used for this application of the Index of Biotic Integrity.

Applying the IBI to New Zealand freshwater fish

For most IBI applications (e.g. in North America, France and Australia), catchment area is used to plot maximum species richness lines, as maximum species richness is a function of catchment area in these countries (Karr 1981, Fausch *et al.* 1990, Harris and Silveira 1999). However, examination of the data revealed little or no association between catchment area and maximum species richness for New Zealand freshwater fish. This is likely to be a result of the high proportion of diadromy in the New Zealand fish fauna (McDowall 1998a, b, Joy *et al.* 2000, Joy and Death 2001). Thus, species richness in New Zealand is a function of elevation and distance from the coast, not catchment area. However neither elevation nor distance from the coast can be used individually because of differences in river length and elevation so both were combined into a single synthetic variable using principal components analysis (PCA). Elevation and distance from the coast were combined using PCA to create a single variable, which is a combination of the two variables. The first principal component explained 97% of the variation of the two variables (elevation and distance from the coast) and this was used to plot the species richness metrics.

Candidate metrics

We identified candidate metrics of fish assemblage integrity from a large suite used in IBI applications elsewhere (Karr and Chu 2000). The IBI metrics typically fit into four broad groupings; taxonomic richness, habitat guilds, trophic guilds and individual abundance and health. However, because the New Zealand native freshwater fish fauna all belong to a single trophic guild (McDowall 1990) and show little evidence of fish disease in wild populations (Duignan *et al.* 2003) metrics dealing with trophic guild and individual health were not used. Furthermore, as fish abundance data from the NZFFD were considered unreliable, given the range of sampling methods and operators any metrics involving abundance were also discarded. Consequently, 6 non-exclusive metrics were selected: (1) the number of native species, (2) the number of riffle dwelling species, (3) the number of benthic pool species, (4) the number of native pelagic species, (5) the number of intolerant or sensitive species, and (6) the proportion of alien species. The assumptions underlying each of the metrics are given in Table 2.

Table 2. Assumptions underlying the metrics used in the Index of Biotic Integrity applied to New Zealand streams in response to declining environmental conditions (modified after (Fausch *et al.*, 1990). See Table 3 for details of metrics.

Effect on fish assemblage of declining environmental conditions		Metric
(i)	Reduction in the Number of native species	1 and 6
(ii)	Reduction in the number intolerant (sensitive) species	5
(iii)	Reduction in the number of species in specific habitats	2,3 and 4
(iv)	Increase in the number of exotic species	6

The first stage of the application of an IBI to a new region or country is the assignment of the fish fauna into groups for the metrics (Tables 1 and 2). These assignments were determined after consulting the literature (see individual metrics below), expert opinion, and from personal observations.

Taxonomic richness

Metric 1 is the number of native species, an attribute of freshwater biotas commonly used in biological assessment. We used native species richness, as opposed to total species richness as non-native species may prefer degraded habitats and thus increase species richness. The assumption underpinning the use of the species richness metric is that environmental degradation will change diverse communities containing many species to simple assemblages dominated by a few species (Fausch *et al.* 1990).

Habitat Guilds

Metric 2, the number of benthic riffle species is used as an indicator of degradation in riffle zones in rivers. Metric 3 is the number of benthic pool species and metric 4 is the number of native pelagic pool species. These metrics were used to make the index sensitive to changes in stream geomorphology resulting from the effects of channelisation and dams on habitats required by fish in these guilds. Only native pelagic pool species were included because many of the alien species indicative of degradation found in New Zealand are pelagic.

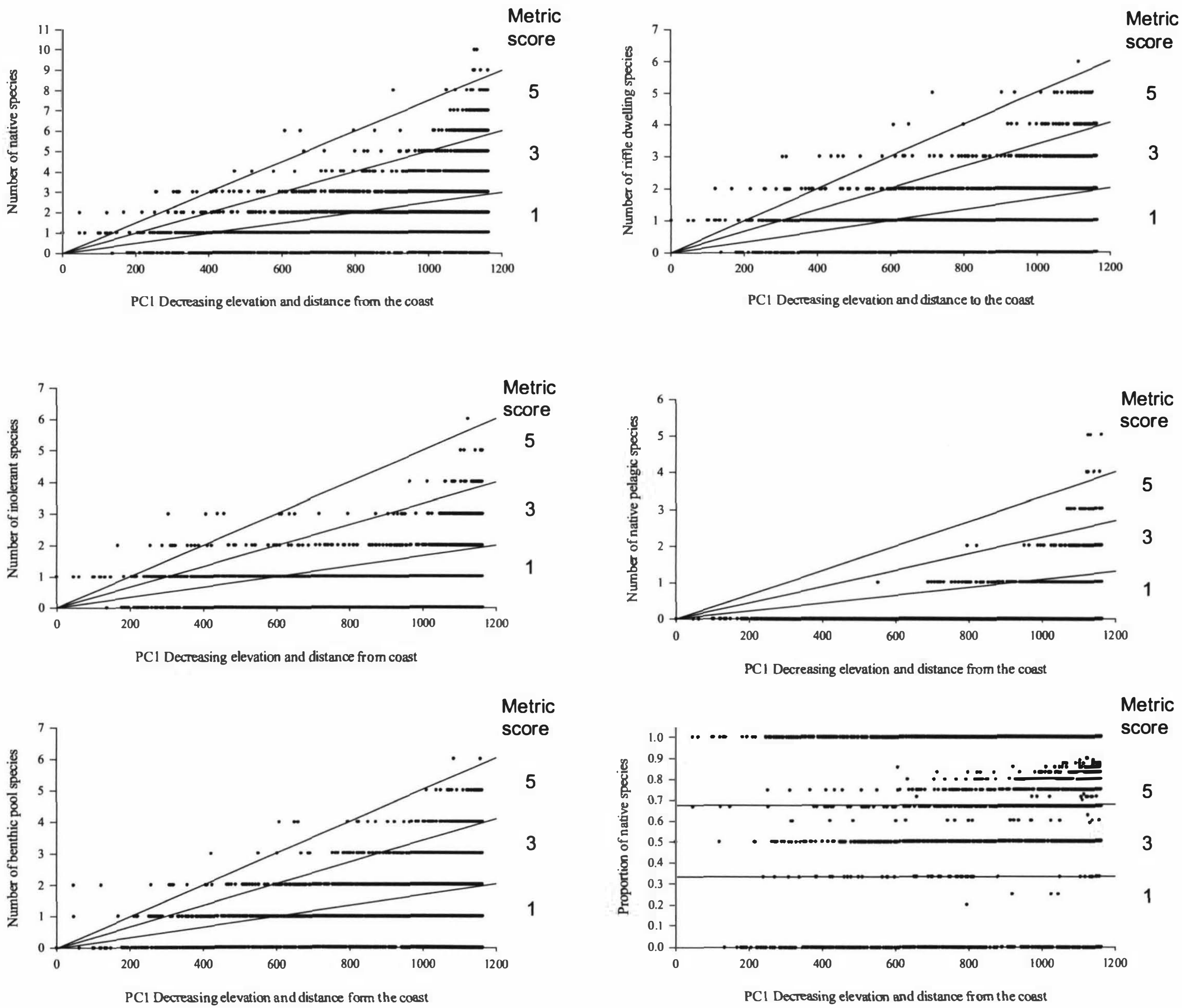

Fig. 2 The procedure for establishing the maximum species richness line (MRSL) for setting the six metrics (see text for details on fitting lines).

Tolerant species

Metric 5 is the number of intolerant species and makes use of limited information on the tolerence of New Zealand freshwater fish to different environmental variables. Species were selected based on their tolerence to impacts such as migration barriers (McDowall 1990, Joy and Death 2000, 2001) and water quality variables such as temperature, sediment and ammonia (Richardson *et al.* 1994, Richardson 1997, Richardson *et al.* 2001).

Invasive species

Metric 6 is the proportion of native to alien species and measures the extent to which the fish assemblage has been invaded by introduced species. The presence of non-native species reflects biological pollution, and generally, these species in New Zealand are more tolerant of degradation of habitat and water quality than the native species and thus, they may indicate degraded conditions.

Calculation of IBI metrics

The IBI was calculated for each sample site following the methods outlined in (Karr *et al.* 1986). The five measures of species composition were scored from Maximum Species Richness Lines (MSRL) (Karr *et al.* 1986). This line indicates the potential species richness for all sites plotted against elevation and distance from the coast. The MSRL was fitted by eye to include about 95% of the sites and this line reflects the number of species expected in relatively undisturbed sites of various elevations and distances from the sea (Fig. 2). The area under the MSRL was trisected to determine metric ratings. A site scored the maximum points (5) when the species richness measure was 0.67 MSRL or higher, 3 points for a site with species richness between 0.33 and 0.67 of the MSRL, and 1 point for less than 0.33 MSRL. For metric 6, the proportion of native species, the position of the site along the stream was not taken into account. Sites with the proportion of native species > 0.67 scored 5, sites between 0.33 and 0.67 received a score of 3 and sites with a proportion of native species less than 0.33 scored the minimum score of 1 (Table 3) (Fig. 2).

Table 3 Criteria used for scoring Index of Biotic Integrity metrics adapted for this study from (Karr *et al.* 1986).

Metric	Scoring criteria		
	5	3	1
(1) Number of native species	> 67 % MSRL	33 – 67 % MSRL	< 33 % MSRL
(2) Number of riffle dwelling species	> 67 % MSRL	33 – 67 % MSRL	< 33 % MSRL
(3) Number of benthic pool species	> 67 % MSRL	33 – 67 % MSRL	< 33 % MSRL
(4) Number of pelagic species	> 67 % MSRL	33 – 67 % MSRL	< 33 % MSRL
(5) Number of intolerant species	> 67 % MSRL	33 – 67 % MSRL	< 33 % MSRL
(6) The proportion of native species	> 67 %	33 % - 67 %	< 33 %

To calculate the total IBI, the scores for the six metrics were summed to give the IBI score for each sampling site (maximum possible was 30 and minimum 6). The ranges of qualitative assessments given for the IBI scores were based on those provided by (Karr *et al.* 1986): excellent (26 – 30), good (22 – 25), fair (17 – 21), poor (12 – 16), and very poor (6 – 11) (the IBI is undefined at sites where no fish are caught and are designated as 'no fish').

Influence of sampling method, date, region and migratory barriers on IBI scores

A four-way factorial analysis of variance (ANOVA) was used to assess the differences in IBI scores over the 15 New Zealand geopolitical regions, 23 years of sampling, 2 sampling methods (electrofishing and nets/traps) and whether or not a barrier was listed on the database as present downstream from the sample site. To determine the differences in mean values for regional IBIs we used Duncan's multiple range test.

Performance of metrics

The performance of metrics was examined using four techniques (Angermeier and Karr 1986). First, to investigate the structure of the metric's correlation matrix we used principal components analysis. The six individual metric scores were analysed to assess the variance accounted for by each principal component. Second, to assess internal metric consistency we calculated Cronbach's alpha (Cronbach 1951), which is a positive function of the average correlation between items in a combined index. The alpha values were then calculated for each metric by sequentially removing that metric and calculating the alpha coefficient for the remaining metrics, which we then compared with the overall alpha. Third, the relationships between each metric and the IBI score were examined using the Kendall's correlation coefficient. Finally, the overall IBI score was compared with the IBI scores after the removal of each metric (IBI-metric). This was derived by sequentially omitting each metric score from the IBI score, for all six metrics. This relationship was again analysed using the non-parametric Kendall's tau.

Validating the IBI

To test the temporal validity of the index we compared sites surveyed at different times on the same river reach. Within the database, there were 369 numbered river reaches where two samples had been taken at different times. We compared the total IBI scores for two samples from the 369 reaches by tabulating the absolute difference between IBI scores.

Associating IBI scores with environmental variables

Fausch *et al.* (1984) suggested that IBI scores ought to be related to measures of physical and chemical habitat quality. To assess the relationship between IBI scores and the available catchment landuse variables we used two procedures. First, we used canonical discriminant function analysis (DFA). DFA is a multivariate procedure that generates functions of variables that best distinguish among previously defined groups of samples by maximising separation between those groups (Williams 1983). We used DFA to find discriminant functions of the variables to best discriminate between the five IBI output classification classes for each site (excellent, good, fair poor and very poor). Then to visualise the influence of these variables on the IBI output classifications the sites were plotted in discriminant space and the environmental variables were plotted in this space using vector arrows. To achieve this, the biplot arrows were drawn from the centroid to the co-ordinates of the pooled within-class standardised canonical coefficients for each of the variables strongly discriminating between the output classification classes.

To investigate the relationships between IBI scores and landuse variables we trisected the sites into groups based on their catchment proportions in each different landuse (e.g. urban, farming, native). We then compared the mean IBI scores for these three categories using analysis of variance (ANOVA) in SAS (2000).

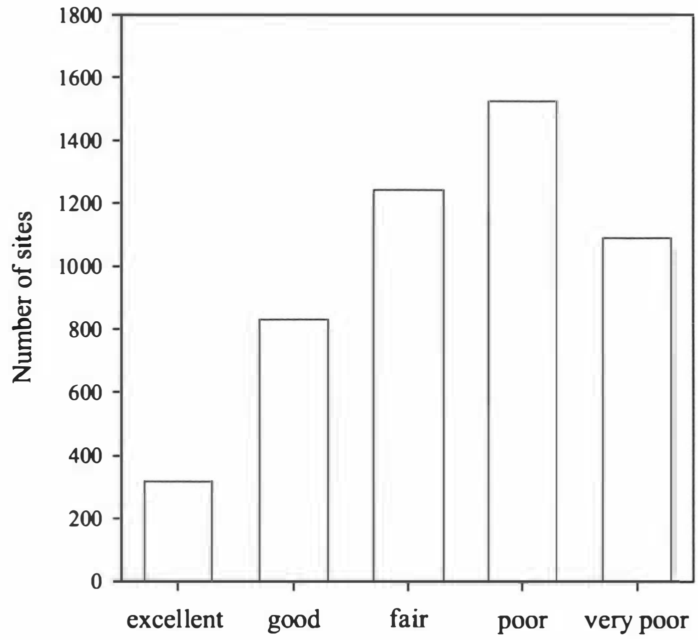

Fig. 3 The distribution of IBI qualitative classes for all 5007 sites throughout New Zealand the classes were calculated from IBI scores excellent (26 – 30), good (22 – 25), fair (17 – 21), poor (12 – 16), and very poor (6 – 11).

Results

Overall results

The IBI varied from a minimum of 6 (very poor) to a maximum value of 30 (excellent) with an overall mean score of 16.3 (SE = 0.08). Most sites were not classed as excellent, or good (Fig. 3). The scores differed between the 15 regions, 23 sample years, 2 sampling methods used, and the presence of a barrier (Table 4). There was a difference in IBI scores between years and a significant interaction between region and year and between regions and sampling method. Regionally mean IBI scores ranged from 18.8 in Gisborne to 13.4 in Otago (Table 5). The mean IBI score was higher (16.6) for the 3904 electro-fished sites than for the 1103 sites sampled by nets and traps (14.9) (Table 4). The mean IBI score (16.5) was higher for the 3918 sites listed in the database as having no downstream barrier than the 1089 sites listed as having a barrier (mean IBI score 15.4).

Table 4 ANOVA results for differences in index of biotic integrity scores for region, year of sample, fishing method used and presence of a downstream blockage at 5007 fish sampling sites in New Zealand from the NZFFD.

Source	d.f.	F-value	P-value
Region	14	3.15	< 0.0001
Block	1	2.04	0.1529
Region * Block	14	2.95	0.0002
Year	22	2.65	< 0.0001
Region * Year	254	1.39	< 0.0001
Region * Block * Year	135	1.55	< 0.0001
Method	1	10.81	0.001
Method * Region	14	2.13	0.0081
Method * Block	1	2.96	0.0854
Method * Region * Block	9	1.79	0.0647
Method * Year	22	1.43	0.0894
Method * Region * Year	107	1.32	0.0155
Method * Block * Year	16	1.99	0.0105
Method * Region * Block * Year	6	0.25	0.9583
Error	4354		

Table 5. Mean IBI scores and number of sample occasions for the 15 New Zealand geopolitical regions. Regions with the same letter not significantly different using Duncan's multiple range test ($P < 0.05$).

Region	Mean IBI score	Number of sites	Duncan grouping
Gisborne	18.78	79	a
Hawke Bay	18.28	247	b a
Taranaki	18.01	181	b a
Wellington	17.89	133	b a
Manawatu	17.72	320	b a c
Canterbury	17.66	415	b a c
Marlborough	17.56	252	b c
Tasman	16.73	208	d c
Northland	16.65	233	d c
Westcoast	16.42	728	d
Waikato	16.33	524	d
Bay of Plenty	16.21	256	d
Southland	16.19	232	d
Auckland	14.60	445	e
Otago	13.44	754	f

Performance of metrics

The first two principal components of the correlation matrix of IBI scores accounted for 69% of the variance among the six metrics (Fig. 4). The plot of the two components shows the relationships between the metrics. Component 1 showed the contrast between metrics 1, 2 and 3 (number of native, riffle and benthic pool species)

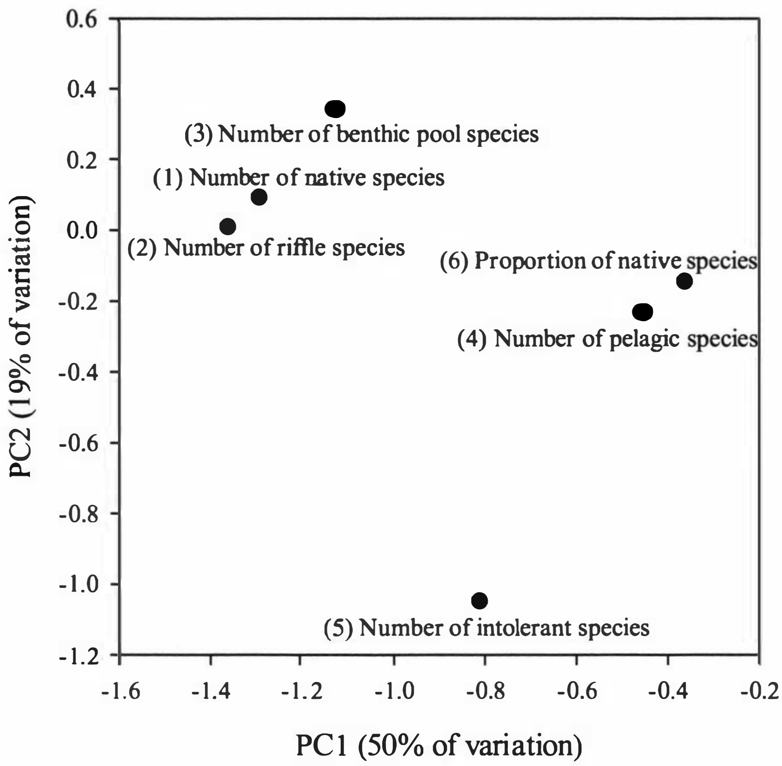

Fig. 4. Principal components plot showing the relationships between IBI metric scores from 5007 New Zealand sites.

and metrics 4 and 6 at the top right (proportion of native and number of pelagic species). This separated the metrics relating to benthic species richness from those related to pelagic and non-native species. The second component revealed the contrast between metric 5 (number of intolerant species) and the other five metrics.

Coefficients of concordance (Cronbach's alpha and Kendal's tau) between the IBI and metric scores and within the metric scores indicated that all the metrics were significantly positively correlated (Table 6). The first column, the correlation between the total IBI score and the IBI after each metric was sequentially removed (IBI-metric) revealed high correlations for metrics 1,4 and 6. Higher correlations between the IBI and the omission scores indicate that metric scores are invariant or highly correlated. For example removing metrics 4 and 6 individually had little influence on the overall IBI score as revealed by the high correlation between the IBI-metric and the IBI before and after their removal. The correlation between the metric and the total IBI (column 2) showed relatively high correlations for the first three metrics and low values for the second three and again metrics 4 and 6 exhibited the lowest concordance. The overall Cronbach's alpha score of 0.76 indicated an acceptable level of internal consistency. All metric alpha scores were lower or equal to the overall score after their individual omission revealing that no metrics had a negative effect and all metrics apart from metric 4 added to the overall IBI score.

Table 6. Coefficients of concordance (Kendall's tau) and (Cronbach's alpha) were calculated between the Index of Biotic Integrity and each of the metrics. In the first column Kendall's tau was calculated for the total IBI score and the IBI after that metric was removed (IBI-metric) (high correlations indicate redundancy). Column 2 has the coefficient of concordance between the metric and the total IBI. Column 3 has Cronbach's alpha calculated for the five metrics after each metric was sequentially omitted. The coefficients were all significant ($P < 0.0001$)

No.	Metric	Kendall's tau (IBI-metric)	Kendall's tau	Cronbach's alpha
1	Number of native species	0.94	0.77	0.69
2	Number of riffle dwelling species	0.91	0.67	0.71
3	Number of benthic pool species	0.91	0.62	0.72
4	Number of native pelagic species	0.96	0.30	0.76
5	Number of intolerant species	0.90	0.45	0.75
6	The proportion of native species	0.95	0.38	0.75
Overall Cronbach's alpha				0.76

Validating the IBI

The temporal consistency of IBI assessments taken in different years at 369 sites on the same reach was high (Fig. 5). At 50 % of the sites the IBI score for the two sample dates differed by less than 3 points, and at a further 30 % the scores differed by 4 – 6 points. The time between samples varied considerably, the mean number of years between sample dates for these multiple sampled reaches was 5.8 years, (range 22-yr. S.E. 0.27-yr.). The mean time between sample dates appeared to increase with increasing difference between scores (Fig. 5).

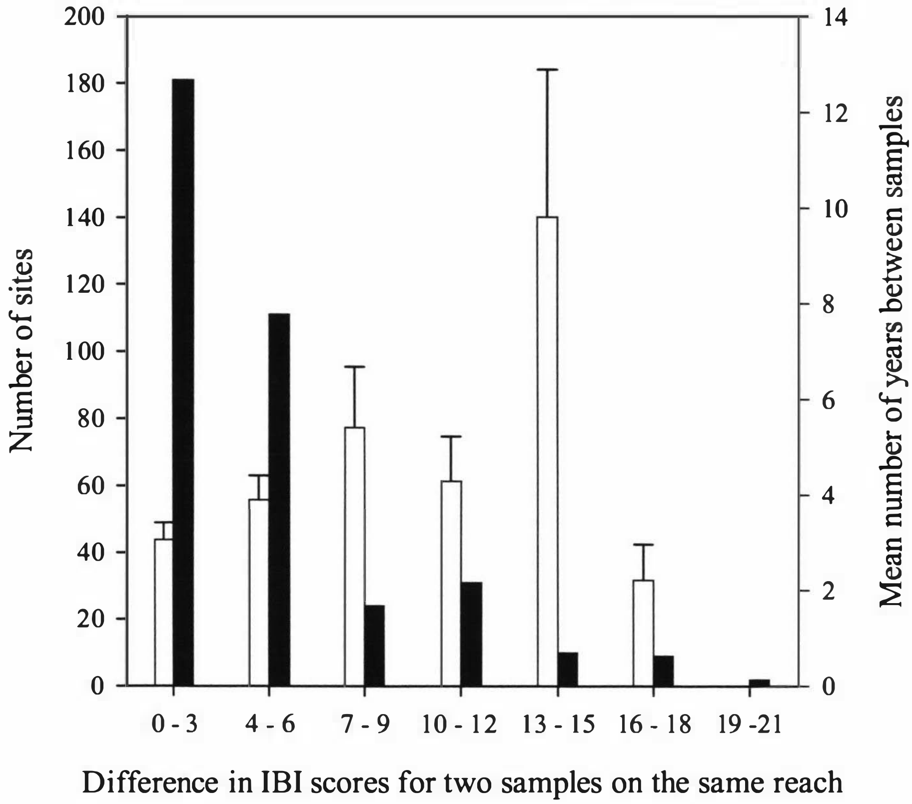

Fig. 5. Distribution of absolute differences between pairs of sample sites on the same numbered reach (filled bars) and mean number of years between samples (open bars) within each group of sites. (The mean number of years between sample dates for the pairs of samples was 5.8 years, range 22-yr. SE 0.27-yr.).

Association with landscape variables

Discriminant function analysis was used to discriminate between sites with scores in the five quantitative IBI output classes (excellent, good, fair poor and very poor). The sites with high IBI scores were associated with high proportions of the catchment in native forest, greywacke geology, and high elevation and distance from the coast (Fig. 6). Conversely, the sites with low IBI scores were associated with high proportions of catchment in farming, and urban landuse, sand and schist geology, and higher mean air temperatures.

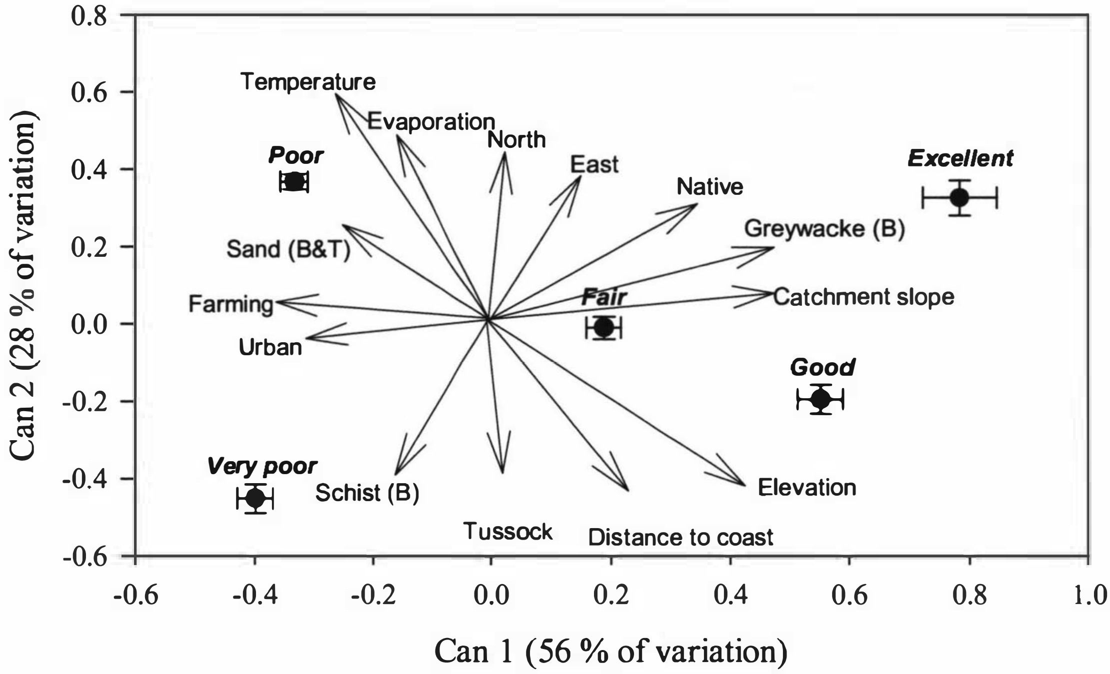

Fig. 6 Position of the mean co-ordinates and standard errors of the five output classes from the IBI classification of 57 New Zealand fish survey sites in discriminant space. The position of the mean co-ordinates and biplot arrows were obtained using linear discriminant analysis on 51 environmental variables (see Appendix 1 for details). The biplot arrows were drawn from the centroid to the co-ordinates of the pooled within-class standardised canonical coefficients. Only variables with coefficients > 0.3 were plotted.

IBI scores were found to differ between sites split into groups with low (< 33 %) medium (33 – 67 %), and high (> 67 %) proportions of catchment in different land use categories (Fig. 7). Mean IBI score increased at sites with increasing catchment proportions of native forest ($F_{2, 5004} = 58.6$ $P < 0.0001$). Conversely, the mean IBI score reduced with increasing proportions of catchments in farming and urban landuse ($F_{2, 5004} = 20.6$ $P < 0.0001$, and $F_{2, 5004} = 25.8$ $P < 0.0001$ respectively). Mean IBI score was significantly higher at sites with a medium proportion of catchment (33 - 67 %) in scrub than at sites with higher and lower proportions ($F_{2, 5004} = 9.1$ $P < 0.0001$). Mean IBI scores were not significantly different among different catchment proportions of exotic forest ($F_{2, 5004} = 0.57$ $P < 0.56$) (not shown).

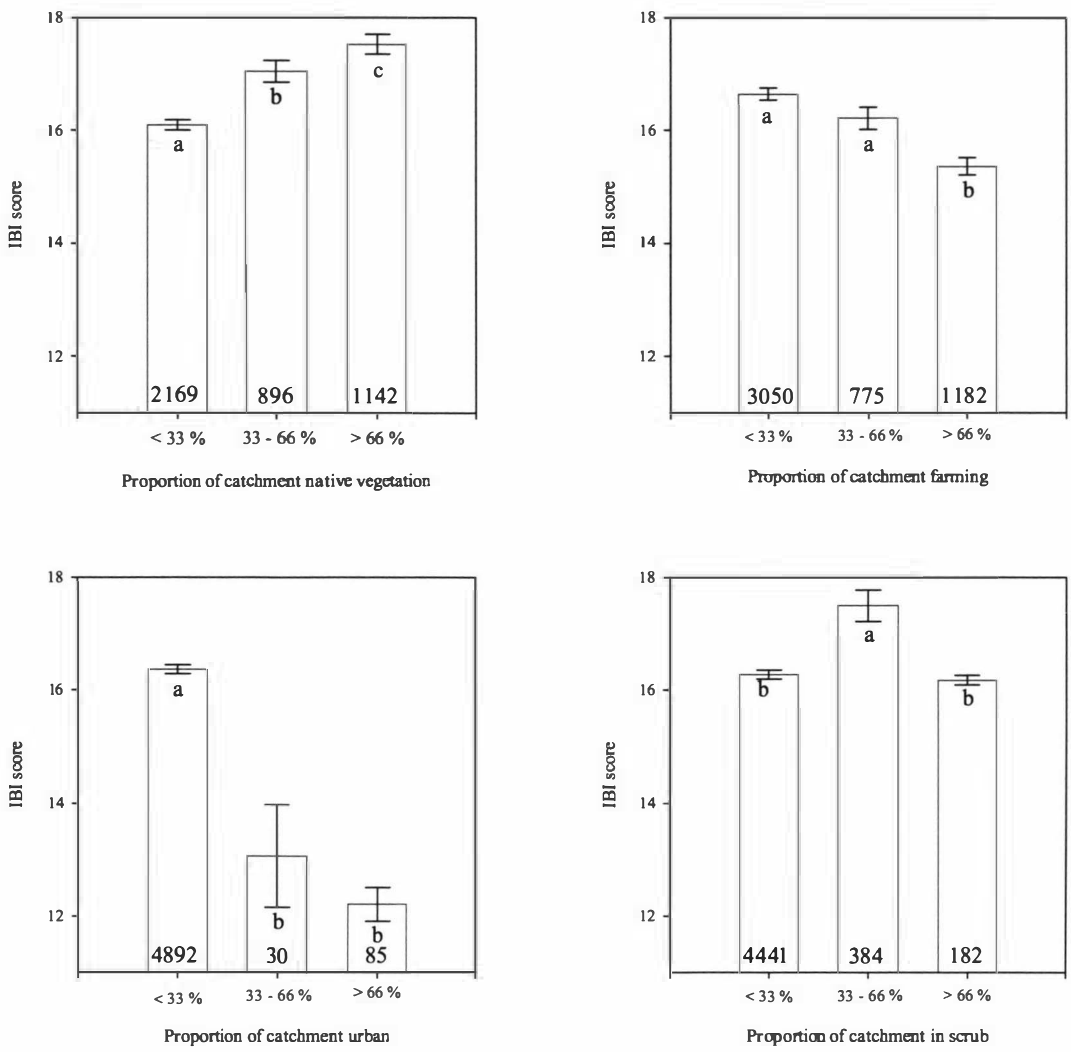

Fig. 7 Mean index of biotic integrity (IBI) scores and standard errors for catchment proportions low (< 33 %) medium (33 – 67 %), and high (>67 %) of different land use categories for 57 sites in New Zealand from the New Zealand freshwater fish database. Bars with different letters significantly different ($P < 0.05$); numbers in bars denote number of sites in each category.

Discussion

Our results illustrate the strong influence of upstream landuse on New Zealand fish communities. High IBI scores were associated with high proportions of catchment in native forest, whereas high levels of agriculture were associated with low IBI scores. The effects of urbanisation were particularly noticeable with a significant drop in IBI scores when the catchment in urban landuse exceeded 33%. These findings are consistent with studies in other countries using the IBI (Wang *et al.* 1997, Schleiger 2000). New Zealand studies have also found farming and urban landuse impact on fish communities (Hanchet 1990, Jowett *et al.* 1998, Rowe *et al.* 1999, Rowe *et al.* 2000, Allibone *et al.* 2001, Richardson and Jowett 2002). The sites with low IBI scores were also those with other non-chemical measures generally associated with environmental degradation including increased temperature, lower elevation and shorter distance from the coast. This gives us confidence that the index is assessing important aspects of biotic integrity despite the difficulties imposed by the large variation in spatial scale, limited species richness and limited specialisation of the fauna.

Many more sites were classed as poor than were classed as excellent and the distribution of IBI scores is skewed with more low scores than high. As the sites were not selected based on a statistical design, we cannot make quantitative statements about the relative number of impaired sites. Nevertheless, it would seem probable that such a large database of sites would be representative of conditions in New Zealand. Our results therefore suggest that biotic integrity is impaired in many New Zealand rivers.

There were significant differences in mean IBI scores across the country and this is likely to reflect the different farming intensity and landuse practises in these regions. The Gisborne region had the highest mean IBI score but only 79 sites were available in this region. The Auckland region has many small streams (> 90% are second order or less) and has considerably fewer fish species than in other regions (Joy and Death In press-b) and this may cause a reduction in IBI scores independent of environmental quality. However, as Auckland contains New Zealand's largest city, and 33% of the national population, the low IBI scores may equally be related to poor environmental quality. There are no obvious patterns to the regional ranking but there were unexpected results with the region commonly thought to be the least developed, the West Coast appearing relatively low on the list. Geology may also be related to some of the regional variation in IBI scores. Catchment proportions of schist were associated with low IBI scores and this rock type predominates is the Southland and Otago regions. Greywacke geology was associated with high IBI scores and proportions of this rock type vary regionally.

An important aspect for evaluating an assessment method is the repeatability of results over time. In this application, more than 70% of repeat samples were within six IBI points of each other even over relatively long time scales. New Zealand freshwater fish lifecycles extend from one year to decades but relatively few are short lived (McDowall 1990), thus comparatively long temporal scales are covered when using fish in environmental assessment when compared to invertebrates.

Sampling efficiency

Fish catchability is a potential source of error when measuring fish communities because it varies both spatially and temporally. Different sampling methods can also introduce error into assessments and may have done so in this study as revealed by the difference IBI scores for the two methods (Table 4). The mean IBI score was higher for electrofishing and this was the most commonly method used. However, this difference could also be the result of the sampling of different habitats using different methods. (Cowx 1991) noted that the environmental sources of variability in electrofishing catchability electrofishing included water depth, conductivity, turbidity, velocity and temperature.

However, by not using abundance data we have minimised some of the potential for error from variation in catchability. While electrofishing increases the chances of catching all species there are many sites where electrofishing is not suitable e.g., where depth and or conductivity are high and these are the sites where the other methods such as trapping and netting are important. Thus, while the IBI was higher at electrofished sites this may be because surveyors at degraded sites are more inclined to use sampling methods other than electrofishing. The method and efficiency of assessment of fish assemblages is very important to an IBI assessment, but not necessarily when setting the ranges for metrics using a large database such as this study. While the levels of fishing efficiency in the NZFFD are variable, this does not negate their usefulness as there will undoubtedly be many sites that are sampled efficiently and these set the upper bounds for the metrics. When making regional comparisons (Table 5) it is assumed that the likelihood of errors is equal for all regions and thus not biased toward any particular region. However, a number of steps can be taken to minimise sampling error for IBI assessments. Sampling should only take place in late summer, as this is when maximum species diversity occurs in New Zealand streams because all diadromous species are present in fresh water (McDowall 1990). The use of a combination of fishing methods will increase the chances of capturing all species present at a site. Furthermore, the full range of habitat types present at a site should be sampled, because although riffle habitats may contain higher fish densities than pools (Jowett *et al.* 1996) other habitats contain different species and these are assessed by specific metrics.

Conclusions

This initial application of the IBI although applied to a low-diversity unspecialised, highly mobile fauna and using a reduced number of metrics has shown that the index can produce repeatable results that can distinguish differing levels of biotic integrity. The index allows for comparison of stream biotic conditions across a large spatial and temporal range of river types. However, further knowledge is needed about the accuracy of the IBI in comparison to other river assessment systems. Other river assessment approaches including predictive reference sites models using fish (Joy and Death 2002, In Press-b) and invertebrates (Wright *et al.* 2000, Joy and Death 2003) as well as single index approaches such as the macroinvertebrate community index (MCI) (Stark 1993) may provide these opportunities for comparison. This application of the IBI parallels the current shift in interest from physical or chemical attributes of rivers to ecological quality assessment and management. This represents a movement towards ecosystem level assessment by taking into account natural physicochemical, geographic, and climatic factors as well as those resulting from human activities. This is the direction given by environmental legislation in New Zealand with the Resource Management Act (1991) with its emphasis on sustainable management of whole ecosystems.

Acknowledgments

This work was funded by a grant from the sustainable management fund (New Zealand Ministry for the Environment grant No. 5099 "Nga Ika Waiora": stream health evaluation tool). Thanks to Bob McDowall for comments on an early draft.

References

Allibone, R., J. Horrox, and S. M. Parkyn. 2001. Stream classification and instream objectives for Auckland's urban streams. ARC00257, NIWA, Hamilton.

Angermeier, P. L., and J. R. Karr. 1986. Applying an index of biotic integrity based on stream fish communities: considerations in sampling and interpretation. North American Journal Fisheres Management **6**:418-429.

ArcView. 1999. Getting to know Arc View GIS. Environmental Systems Research Institute, Inc., Redlands California.

Barbour, M. T., J. B. Stribling, and J. R. Karr. 1995. Multimetric approach for establishing biocriteria and measuring biological condition. Pages 63-77 *in* W. S. Davis and T. P. Simon, editors. Biological assessment and criteria. Tools for water and resource planning and decision making. Lewis publishers, Boca Raton, Florida.

Chessman, B. C., I. O. Growns, and N. Plunket-Cole. 1999. Predicting diatom communities at genus level for the biological management of rivers. Freshwater Biology **41**:317-331.

Cowx, I. G. 1991. Catch Effort Sampling Strategies. Fishing News Books, Oxford.

Cronbach, L. J. 1951. Coefficient alpha and the internal structure of tests. Psychometrika **16**:297-334.

Davis, W. S., and T. P. Simon. 1995. Biological assessment and criteria: Building on the past. Pages 15-29 *in* W. S. Davis and T. P. Simon, editors. Biological Assessment and Criteria: tools for water resource planning and decision making. Lewis publishers, Boca Raton.

Duignan, P. J., P. M. Hine, M. K. Joy, N. Gibbs, and G. W. Jones. 2003. Freshwater fish health survey. Ministry of Agriculture and Forestry Biosecurity Authority, Palmerston North.

Fausch, K. D., J. R. Karr, and P. R. Yant. 1984. Regional Application of an Index of Biotic Integrity Based on Stream Fish Communities. Transactions of the American Fisheries Society **113**:39-55.

Fausch, K. D., J. Lyons, J. R. Karr, and P. L. Angermeier. 1990. Fish communities as indicators of environmental degradation. American Fisheries Society Symposium **8**:123-144.

Hanchet, S. M. 1990. Effect of land use on the distribution and abundance of native fish in tributaries of the Waikato River in the Hakarimata Range, North

Island, New Zealand. New Zealand Journal of Marine and Freshwater Research **24**:159-171.

Harris, J. H., and R. Silveira. 1999. Large-scale assessments of river health using an index of biotic integrity with low diversity fish communities. Freshwater Biology **41**:235-252.

Hughes, R. M., M. Kaufman, A. T. Herlihy, T. M. Kincaid, L. Reynolds, and D. P. Larsen. 1998. A process for developing and evaluating indices of fish assemblage integrity. Canadian Journal of Fish and Aquatic Science **55**:1618-1631.

Jowett, I. G., J. W. Hayes, N. Deans, and G. A. Eldon. 1998. Comparison of fish communities and abundance in unmodified streams of Kahurangi National Park with other areas of New Zealand. New Zealand Journal of Marine and Freshwater Research **32**:307-332.

Jowett, I. G., J. Richardson, and R. M. McDowall. 1996. Relative effects of in-stream habitat and land use on fish distribution and abundance in tributaries of the Grey River, New Zealand. New Zealand Journal of Marine and Freshwater Research **30**:463-475.

Joy, M. K., and R. G. Death. 2000. Development and application of a predictive model of riverine fish community assemblages in the Taranaki region of the North Island, New Zealand. New Zealand Journal of Marine and Freshwater Research **34**:243-254.

Joy, M. K., and R. G. Death. 2001. Control of freshwater fish and crayfish community structure in Taranaki, New Zealand: dams, diadromy or habitat structure? Freshwater Biology **46**:417-429.

Joy, M. K., and R. G. Death. 2002. Predictive modelling of freshwater fish as a biomonitoring tool in New Zealand. Freshwater Biology **47**:2261-2275.

Joy, M. K., and R. G. Death. 2003. Biological assessment of rivers in the Manawatu-Wanganui region of New Zealand using a predictive macroinvertebrate model. New Zealand Journal of Marine and Freshwater Research **37**:367-379.

Joy, M. K., and R. G. Death. In Press-b. Neural network modelling of freshwater fish and macro-crustacean assemblages for biological assessment in New Zealand. *in* S. Lek, M. Scardi, and S. E. Jorgensen, editors. Modelling community structure in freshwater ecosystems. Springer Verlag.

Joy, M. K., and R. G. Death. In Press-a. Assessing biological integrity using freshwater fish and decapod habitat selection functions. Environmental Management.

Joy, M. K., I. M. Henderson, and R. G. Death. 2000. Diadromy and longitudinal patterns of upstream penetration of freshwater fish in Taranaki, New Zealand. New Zealand Journal of Marine and Freshwater Research **34**:531-543.

Karr, J. R. 1981. Assessments of biotic integrity using fish communities. Fisheries **6**:21-27.

Karr, J. R., and E. W. Chu. 2000. Sustaining living rivers. Hydrobiologia **422/423**:1-14.

Karr, J. R., K. D. Fausch, P. L. Angermeier, P. R. Yant, and I. J. Schlosser. 1986. Assessing biological integrity in running waters: A method and its rationale. Illinois Natural History Survey Special Publication **5**:28pp.

McDowall, R. M. 1990. New Zealand Freshwater Fishes: A Natural History and Guide. Heinemann Reed, Auckland.

McDowall, R. M. 1998a. Driven by diadromy: its role in the historical and ecological biogeography of the New Zealand freshwater fish fauna. Italian Journal of Zoology **65**:73-85.

McDowall, R. M. 1998b. Fighting the flow: downstream-upstream linkages in the ecology of diadromous fish faunas in West Coast New Zealand rivers. Freshwater Biology **41**:111-123.

McDowall, R. M. 2000. The Reed field guide to New Zealand freshwater fishes. Reed Publishing, Auckland.

McDowall, R. M., and J. Richardson. 1983. The New Zealand freshwater fish survey-a guide to input and output. Fisheries Research Division Information leaflet 12, Ministry of agriculture and Fisheries, Wellington.

McDowall, R. M., and M. J. Taylor. 2000. Environmental indicators of habitat quality in a migratory freshwater fish fauna. Environmental Management **25**:357-374.

Newsome, P. F. J. 1990. New Zealand Land Resources Inventory ARC/INFO Data Manual. Landcare Research, New Zealand.

Omernik, J. M. 1995. Ecoregions: a spatial framework for environmental management. *in* W. S. Davis and T. P. Simon., editors. Biological

Assessmant and Criteria: Tools for water resource planning and decision making. Lewis publishers, Boca Raton.

Reynoldson, T. B., R. H. Norris, V. H. Resh, K. E. Day, and D. M. Rosenberg. 1997. The reference condition: a comparison of multimetric and multivariate approaches to assess water-quality impairment using benthic macroinvertebrates. Journal of the North American Benthological Society **16**:833-852.

Richardson, J. 1989. The all-new freshwater fish database. Freshwater Catch **41**:20-21.

Richardson, J. 1997. Acute ammonia toxicity for eight New Zealand indigenous freshwater species. New Zealand Journal of Marine and Freshwater Research **31**:185-190.

Richardson, J., J. A. T. Boubee, and D. W. West. 1994. Thermal tolerance and preference of some native New Zealand freshwater fish. New Zealand Journal of Marine and Freshwater Research **28**:399-407.

Richardson, J., and I. Jowett. 2002. Effects of sediment on fish communities in East Cape streams, North Island New Zealand. New Zealand Journal of Marine and Freshwater Research **36**:431-442.

Richardson, J., D. K. Rowe, and J. P. Smith. 2001. Effects of turbidity on the migration of juvenile banded kokopu (Galaxias fasciatus) in a natural stream. New Zealand Journal of Marine and Freshwater Research **35**:191-196.

Rowe, D. K., B. L. Chisnall, T. L. Dean, and J. Richardson. 1999. Effects of land use on native fish communities in east coast streams of the North Island of New Zealand. New Zealand Journal of Marine and Freshwater Research **33**:141-151.

Rowe, D. K., M. Hicks, and J. Richardson. 2000. Reduced abundance of banded kokopu (Galaxias fasciatus) and other native fish in turbid rivers of the North Island New Zealand. New Zealand Journal of Marine and Freshwater Research **34**:547-558.

SAS. 2000. SAS/STAT Version 8.2. SAS Institute Inc, Cary, North Carolina.

Schleiger, S. L. 2000. Use of an index of biotic integrity to detect effects of land uses on stream fish communities in west-central Georgia. Transactions of the American Fisheries Society **129**:1118-1133.

Simon, T. P., and J. Lyons. 1995. Application of the index of biotic integrity in freshwater ecosystems. Pages 245-262 *in* W. S. Davis and T. P. Simon, editors. Biological Assessment and Criteria: Tools for water resource planning and descision making. Lewis, Boca Raton Florida.

Simpson, J., and R. H. Norris. 2000. Biological assessment of water quality: development of AUSRIVAS models and outputs. Pages 125-142 *in* J. F. Wright, D. W. Sutcliffe, and M. T. Furse, editors. Assessing the biological quality of freshwaters. RIVPACS and other techniques. Freshwater Biological Association, Ambleside, UK.

Snelder, T., and B. J. F. Biggs. 2002. Multi-Scale River Environment Classification for Water Resources Management. Journal of the American Water Resources Association **38**:1225-1240.

Snelder, T., B. J. F. Biggs, U. Shankar, B. McDowall, T. Stephens, and I. K. G. Boothroyd. 1998. Development of a system of physically based habitat classification for water resources management of New Zealand rivers. NIWA, Christchurch.

Stark, J. D. 1993. A macroinvertebrate community index of water quality for stony streams. New Zealand Journal of Marine and Freshwater Research **27**:463-478.

Terralink International Limited. New Zealand Landcover Database (LCDB1). http://www.terralink.co.nz/metadata/ANZNZ3001800023_external.htm

Wang, L., J. Lyons, P. Kanehl, and R. Gatti. 1997. Influences of watershed landuse on habitat quality and biotic integrity in Wisconsin Streams. Fisheries **22**:6-12.

Williams, B. K. 1983. Some observations on the use of discriminant analysis in ecology. Ecology **64**:1283-1291.

Wright, J. F., D. W. Sutcliffe, and M. T. Furse. 2000. Assessing the biological quality of fresh waters: rivpacs and other techniques. Freshwater biological association, Ambelside, Cumbria, UK. 373 pp.

Appendix 1. Landscape variables used in discriminant function analysis from the raw data used in the River Environment Classification (REC) (Snelder and Biggs, 2002).

Variable	min	max	mean
Total catchment area (km^2)	228600	16543.540	73.029
Estimated catchment area evaporation (mm)	486.3	849.5	699.5
Estimated average annual flow (m2/sec^{-1})	0	505.8	2.4
Average catchment rainfall (mm * m^2)	349.38	8233.1	1713.1
Average catchment air temperature (C°)	3.2	15.7	10.7
Average catchment slope (m/km)	0	89.9	19.0
Average catchment elevation (m)	0	1569.0	398.1
Distance from site to sea (km)	0	448.0	66.3
Stream order (Strahler)	1	8.000	2.706
Landcover catchment (proportion)			
Bare ground	0	0.819	0.019
Coastal sands	0	0.378	0.001
Exotic forest	0	1.0	0.081
Indigenous forest	0	1.0	0.322
Pastoral	0	1.0	0.321
Scrub	0	1.0	0.119
Tussock	0	1.0	0.107
Urban catchment	0	1.0	0.022
Top rock geology (proportion)			
Peat	0	1.0	0.009
Loess	0	1.0	0.048
Windblown material	0	1.0	0.005
Alluvium	0	1.0	0.143
Sandstone	0	1.0	0.092
Other	0	1.0	0.062
Mudstone	0	1.0	0.028
Calcareous	0	1.0	0.007
Schist	0	1.0	0.245
Plutonic and Felsic	0	0.9	0.000

Ultramafic	0	1.0	0.000
Volcanic (hard)	0	1.0	0.073
Volcanic (soft)	0	1.0	0.146
Greywacke	0	1.0	0.095
Baserock geology			
Peat	0	1.0	0.006
Loess	0	0.8	0.004
Windblown material	0	0.9	0.004
Alluvium	0	1.0	0.149
Sandstone	0	1.0	0.103
Other	0	1.0	0.082
Mudstone	0	1.0	0.035
Calcareous	0	0.9	0.007
Plutonic and Felsic	0	1.0	0.036
Ultramafic	0	0.9	0.002
Volcanic (hard)	0	1.0	0.056
Volcanic (soft)	0	1.0	0.122
Greywacke	0	1.0	0.232
Schist	0	1.0	0.134
Year	1980	2002	1992
Altitude (m)	0	1460	181
Distance from coast (km)	0	448	66

Synthesis

Synthesis

The aim of this PhD was to improve and extend the limited range and depth of bioassessment tools available to resource managers in New Zealand. This objective was achieved by the development of multimetric, multivariate, and multi-biotic-group predictive models. However, trying to authenticate a new bioassessment model and show an improvement over an existing model is difficult. Comparison with existing models is speculative because a lack of similarity between outputs does not indicate whether the new model is an improvement or not. Another approach is to look at the relationship between scores and physicochemical variables but this can also be problematic and may become circular by pre-supposing relationships between the biota and variables. Given this inability to prove or disprove the authenticity of new bioassessment models the only alternative is to place absolute importance on the presence of sound theoretical, statistical and ecological underpinnings. For models to have sound theoretical, statistical and ecological foundations the use of prediction is paramount. The importance of prediction in ecological modeling has been propounded by many authors e.g. (Peters 1991, Harris 1994, Pace 2001, Olden et al. 2002). However, ultimately the regulatory authorities and other resource managers will make the final decision with their choice of biomonitoring tools.

To improve and enhance bioassessment in New Zealand I introduced a number of new approaches outlined below:

1. By using fish an important and well known component of riverine biology previously ignored has been incorporated into bioassessment. Apart from the obvious advantages from expanding the range of biotic groups assessed, fish at or near the apex of stream food webs integrate many stream processes.
2. The inclusion of prediction into modeling is coherent with Peter's (1991) call for a "more rigorously scientific, more informative and more useful ecology". The inclusion of prediction will enhance bioassessment as hypotheses can be tested by evaluating the ability of the model to make correct predictions.
3. The use of multiple variables has a number of advantages over a single index approach. While single metrics are attractive because they produce a single

score that is readily comparable to a target value and are a traditional approach familiar to many managers they have a number of disadvantages. Multivariate models use all available information, make no assumptions regarding either individual species responses or the impacts being assessed and can assess any impacts not just the single impact used when assigning scores to metrics. Thus multivariate and multimetric models maximize objectivity.

4. The inclusion of Artificial Neural Networks (ANNs) in bioassessment has a number of advantages over traditional statistical approaches. These advantages include the ability to model non-linear relationships, furthermore there are no underlying assumptions about variable distributions and finally ANNs have the unique ability to predict multiple outputs.

The models produced in this thesis are being used in day to day resource management by the regional councils involved in funding this research. Other regions have shown interest and models are being produced at present for the Hawkes Bay Regional Council.

References

Harris, G. P. 1994. Pattern, process and prediction in aquatic ecology. A limnological view of some general ecological problems. Freshwater Biology **32**:143-160.

Karr, J. R. 1991. Biological Integrity - a Long-Neglected Aspect of Water- Resource Management. Ecological Applications **1**:66-84.

Metcalfe-Smith, J. L. 1996. Biological water-quality assessments of rivers: use of macroinvertebrate communities. Pages 17-43 *in* G. E. Petts and P. Calow, editors. River restoration. Blackwell, Oxford.

Olden, J. D., D. A. Jackson, and P. R. Peres-Neto. 2002. Predictive models of fish species distributions: A note on proper validation and chance predictions. Transaction of the American Fisheries Society **131**:329-336.

Pace, M. L. 2001. Prediction and the aquatic sciences. Canadian Journal of Fisheries and Aquatic Sciences **58**:63-72.

Peters, R. H. 1991. A Critique for Ecology. Cambridge University Press, Cambridge.