

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

MICHAEL K. JOY

RUSSELL G. DEATH

Institute of Natural Resources—Ecology  
Massey University  
Private Bag 11 222  
Palmerston North, New Zealand  
email: mikejoy@clear.net.nz

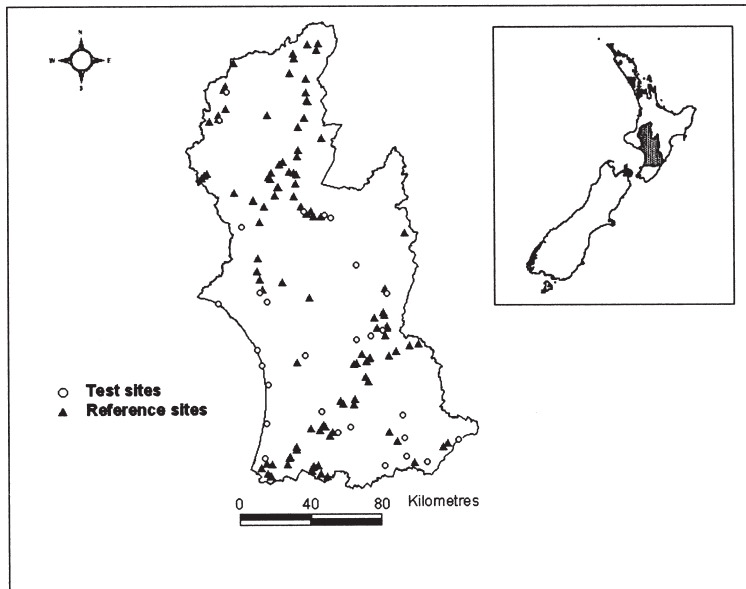
**Abstract** This study presents a river invertebrate and classification system (RIVPACS) type bioassessment methodology for the Manawatu-Wanganui region of New Zealand. Aquatic macroinvertebrates and related physico-chemical data were collected at 127 sites, with minimal human impacts (reference sites) 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 modelling 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 modelling approach combines information on environmental variables and macroinvertebrate assemblages in a predictive format (Winterbourn 1999). Therefore, the application of a predictive modelling 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 were used



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

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 & 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 modelling approach has not been applied in New

Zealand except with fish and macro-crustaceans (Joy & Death 2000, 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 & 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<sup>2</sup>) 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′–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.

### 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 which 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  $\mu$ m 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  $\mu$ m 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 species) using existing keys (e.g., Winterbourn & 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.

#### *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 & 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 & 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

Modelling 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 & Mefford 1999). This process classifies both samples and species simultaneously, based on dividing reciprocal averaging ordination space. The TWINSPAN analysis was achieved using five

**Table 1** Environmental variables estimated at reference and test sites and details on collection and measurement. Variables in bold were not used as predictors but were used to assess test sites. (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 c. 100 m of the reach sampled (Joy & 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 fished; 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 & 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 observer's evaluation of predetermined criteria. Three totals relate to three regions of the stream channel: upper banks, lower banks, and stream bottom; n, percentage of backwater, pool, run, riffle, or rapid was visually estimated over 50 m 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.)

Variable	Units	
<b>Water</b>		
pH		a
Alkalinity	mg litre <sup>-1</sup> Ca CO <sub>3</sub>	b
Temperature	°C	c
Conductivity	µs cm <sup>-1</sup>	d
Velocity	m s <sup>-1</sup>	e
Nitrate	mg litre <sup>-1</sup> NO <sub>3</sub> -N	f
Reactive phosphorus	mg litre <sup>-1</sup> PO <sub>4</sub> <sup>3-</sup>	f
<b>Physical attributes</b>		
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 <sup>-1</sup>	g
Reach slope (over surveyed reach)	m km <sup>-1</sup>	h
Width (mean and SD)	m	i
Depth (mean and SD)	mm	i
Embeddedness	j	
Ecoregion	k	
Stream order	Strahler	h
Average catchment elevation	m	h
Total catchment rainfall	annual rainfall × catchment area	h
Total catchment area	km <sup>2</sup>	h
Mean and median substrate size	cm	l
<b>Proportion of reach surveyed composed of:</b>		
% still		n
% backwater		n
% pool		n
% run		n
% riffle		n
% rapid		n
% undercut		n
% debris jam		n
<b>Pfankuch channel stability index score</b>		
upper		m
lower		m
bottom		m

## Variable

**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
<b>Catchment area landcover</b>	
% primarily pastoral	h
% primarily horticultural	h
% indigenous forest	h
% planted forest	h
% scrub	h
% tussock	h
% coastal sands	h
% bare ground	h

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  $\text{Log}_{10}(x + 1)$ . 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. Cross-validation was used to check whether sites were allocated to their correct groups. The cross-validation process (also known as jack-knife or leave-one-out validation) involves leaving out each site in turn, then rebuilding 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, 2001; Olden et al. 2002). A site was considered to be correctly classified if the probability of belonging to the correct group was higher than it was 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 & 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 & 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

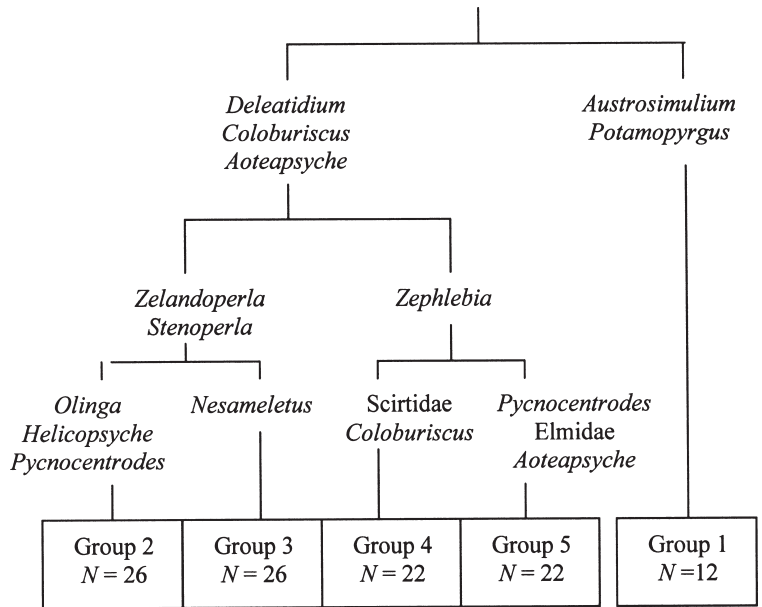
### Faunal characteristics of the groups

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* sp. as the indicator taxa and were distinct from another group containing the stoneflies *Zelandoperla* sp. 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 *Pycnocentroides* sp. The group 4 indicator taxa were the beetle family Scirtidae and the mayfly *Coloburiscus*. The indicator taxa for group 5 sites were *Pycnocentroides*, *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.

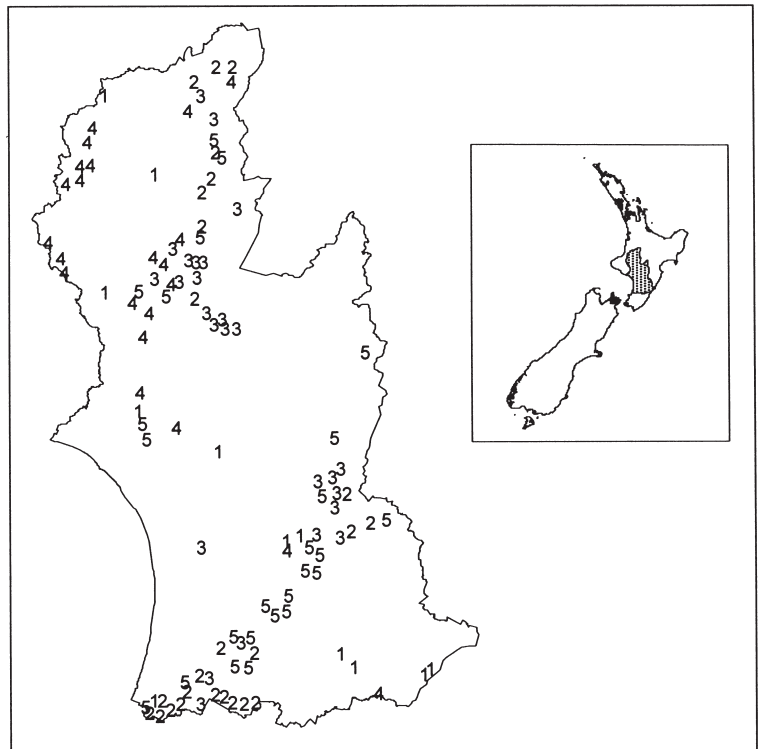
### Geographic patterns

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

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



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



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)

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

Variable	<i>F</i> 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	2.8	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	131.68 ± 15.67
Distance to coast	2.35	110.67 ± 22.95	156.42 ± 21.70	200.65 ± 11.78	195.95 ± 17.19	
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)	2.0	85 ± 0.16	0.51 ± 0.07	0.40 ± 0.04	1.02 ± 0.11	0.55 ± 0.06

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

**Table 3** Number of taxa observed (*O*) and expected (*E*) 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).

Site no.	Sampled 1997 (Polglase 2000)			Sampled 2000 (this study)		
	No. of taxa expected	Observed no. of taxa	<i>O/E</i>	No. of taxa expected	Observed no. of taxa	<i>O/E</i>
7	5.82	6	1.03	5.82	5	0.86
18	4.28	6	1.40	4.28	7	1.63
19*	5.76	5	0.86	5.76	6	1.04
20	7.86	6	<b>0.76</b>	7.86	5	<b>0.64</b>
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	1.20	5.85	7	1.20
90	5.78	5	0.87	5.78	6	1.04
138†	7.02	5	<b>0.71</b>	7.02	4	<b>0.57</b>
139*	5.98	6	1.00	5.98	6	1.00
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	<i>O/E</i>
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 Papaitonga, 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 Apati	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	0.70
Porewa Stream, near Marton	Sheep/beef farming/sewage	0.70
Mangatainoka River, near Pahiatua	Dairy/sheep/beef farming	0.70
Mongotai Stream, near Wanganui	Exotic forest	0.70
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

coastal sand, stream alkalinity, and stream width all increased with reducing *O/E* values whereas 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.

## DISCUSSION

The results presented here are the first published application of the RIVPACS/AUSRIVAS methodology using macroinvertebrates in New Zealand. This application in the Manawatu-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 land use 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

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

Variable	Spearman <i>r</i>	<i>P</i> 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

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 & 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, demonstrating 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 modelling 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. 20th ed. Washington DC, United States, American Public Health Association, American Water Works Association, Water Pollution Control Federation.
- Bailey, R. C.; Kennedy, M. G.; Dervish, M. Z.; Taylor, R. M. 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.; Furse, M. T.; Wright, J. F.; 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.
- Furse, M. T.; Moss, D.; Wright, J. F.; 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.
- Harding, J. S. 1994: Lotic Ecoregions of New Zealand. Unpublished PhD thesis, University of Canterbury, Christchurch, New Zealand.

- Harding, J. S. 1999: Changes in agricultural intensity and river health along a river continuum gradient. *Freshwater Biology* 42: 345–357.
- Hawkins, C. P.; Norris, R. H.; Hogue, J. N.; Feminella, J. W. 2000: Development and evaluation of predictive models for measuring the biological integrity of streams. *Ecological Applications* 10: 1456–1477.
- Hickey, C. W.; Clements, W. H. 1998: Effects of heavy metals on benthic macroinvertebrate communities in New Zealand streams. *Environmental Toxicology and Chemistry* 17: 2338–2346.
- Hughes, R. M.; Larsen, D. P.; Omernik, J. M. 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. In: Davis W. S.; Simon T. P. ed. Biological assessment and criteria: tools for water resource planning and decision making. Boca Raton, Lewis Publishers. Pp. 31–47.
- Joy, M. K.; 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.; Death, R. G. 2002: Predictive modelling of freshwater fish as a biomonitoring tool in New Zealand. *Freshwater Biology* 47: 2261–2275.
- McCune, B.; Mefford, M. J. 1999: PC-ORD. Multivariate analysis of ecological data. Glenden Beach, Oregon, United States, MjM Software Design.
- Manel, S.; Dias, J.; Ormerod, S. J. 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.
- Manel, S.; Williams, H. C.; 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. 1989: A sub-sampler for samples of benthic invertebrates. *Bulletin of the Australian Society for Limnology* 19: 49–52.
- Marchant, R. 2002: Do rare species have any place in multivariate analysis for bioassessment? *Journal of the North American Benthological Society* 21: 311–313.
- Marchant, R.; Hirst, A.; Norris, R. H.; Butcher, R.; Metzeling, L.; Tiller, D. 1997: Classification and prediction of macroinvertebrate assemblages from running waters in Victoria, Australia. *Journal of the North American Benthological Society* 16: 664–681.
- Moss, D.; Furse, M. T.; Wright, J. F.; 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.
- Olden, J. D.; Jackson, D. A.; 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.
- Omernik, J. M. 1995: Ecoregions: a spatial framework for environmental management. In: Davis W. S.; Simon T. P. ed. Biological assessment and criteria: tools for water resource planning and decision making. Boca Raton, Lewis Publishers.
- Pfankuch, D. J. 1975: Stream reach inventory and channel stability evaluation. United States Department of Agriculture Forest service, Region 1, Missoula Montana, United States.
- Polglase, M. A. 2000: The development of the Waterway self assessment form, a stream management tool for landowners. Unpublished MSc thesis, Massey University, Palmerston North, New Zealand.
- Quinn, J. M.; 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.; Wright, J. F. 2000: The reference condition: problems and solutions. In: Wright, J. F.; Sutcliffe, D. W.; Furse, M. T. ed. Assessing the biological quality of freshwaters. RIVPACS and other techniques. Ambleside, United Kingdom. Freshwater Biological Association. Pp. 293–303.
- Reynoldson T. B.; Bailey R. C.; Day K. E.; 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 2000: SAS/STAT Version 8.2. Cary, North Carolina, SAS Institute Inc.

- Simpson, J.; Norris, R. H. 2000: Biological assessment of water quality: development of AUSRIVAS models and outputs. *In: Wright, J. F.; Sutcliffe D. W.; Furse, M. T. ed. Assessing the biological quality of freshwaters. RIVPACS and other techniques.* Ambleside, United Kingdom. Freshwater Biological Association. Pp. 125–142.
- Smith, M. J.; Kay, W. R.; Edward, D. H. D.; Papas, P. J.; Richardson, K. S. J.; Simpson, J. C.; Pinder, A. M.; Cale, D. J.; Horwitz, P. H. J.; Davis, J. A.; Yung, F. H.; Norris, R. H.; Halse, S. A. 1999: AusRivAS: using macroinvertebrates to assess ecological condition of rivers in Western Australia. *Freshwater Biology* 41: 269–282.
- Snelder, T.; Biggs, B. J. F. 2002: Multi-scale river environment classification for water resources management. *Journal of the American Water Resources Association* 38: 1225–1240.
- Snelder, T.; Guest, P. 2000: The ‘river ecosystem management framework’ and the use of river environment classification as a tool for planning. NIWA Client Report CHC00/81. Christchurch, New Zealand.
- Snelder, T.; Biggs, B. J. F.; Shankar, U.; McDowall, B.; Stephens, T.; Boothroyd, I. K. G. 1998: Development of a system of physically based habitat classification for water resources management of New Zealand rivers. NIWA Client Report CHC98/66. Christchurch, New Zealand.
- 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.; Trihadiningrum, Y.; Hart, B. T.; Davies, P. E.; Humphrey, C.; Norris, R. H.; Simpson, J.; Thurtell, L. 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.; Flack, L. K.; Norris, R. H.; Simpson, J.; Waddell, N. 1999: Assessment of river condition at a large spatial scale using predictive models. *Freshwater Biology* 41: 283–298.
- Winterbourn, M. J. 1999: Recommendations of the New Zealand Macroinvertebrate Working Group: monitoring for sustainable river ecosystem management and the role of macroinvertebrates. *In: Winterbourn M. J. ed. The use of macroinvertebrates in water management.* Wellington, New Zealand, Ministry for the Environment.
- Winterbourn, M. J.; Gregson, K. L. D. 1989: Guide to the aquatic insects of New Zealand. Revised edition. *Bulletin of the Entomological Society of New Zealand No. 9.* 95 p.
- 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. 2000: An introduction to RIVPACS. *In: Wright, J. F.; Sutcliffe D. W.; Furse M. T. ed. Assessing the biological quality of freshwaters. RIVPACS and other techniques.* Freshwater Biological Association, Ambleside, United Kingdom.
- Wright, J. F.; Moss, D.; Armitage, P. D.; Furse, M. T. 1984: A preliminary classification of running-water sites in Great Britain based on macroinvertebrate species and the prediction of community type using environmental data. *Freshwater Biology* 14: 221–256.
- Wright, J. F.; Furse, M. T.; Armitage, P. D. 1993: RIVPACS—a technique for evaluating the biological quality of rivers in the UK. *European Water Pollution Control* 3: 15–25.
- Wright, J. F.; Sutcliffe, D. W.; Furse M. T. 2000: Assessing the biological quality of fresh waters: RIVPACS and other techniques. Freshwater Biological Association, Ambleside, Cumbria, United Kingdom.

