BioOne COMPLETE

Evaluating AUSRIVAS Predictive Model Performance for Detecting Simulated Eutrophication Effects on Invertebrate Assemblages

Authors: Nichols, Susan J., Reynoldson, Trefor B., and Harrison, Evan T.

Source: Freshwater Science, 33(4) : 1212-1224

Published By: Society for Freshwater Science

URL: https://doi.org/10.1086/678946

BioOne Complete (complete.BioOne.org) is a full-text database of 200 subscribed and open-access titles in the biological, ecological, and environmental sciences published by nonprofit societies, associations, museums, institutions, and presses.

Your use of this PDF, the BioOne Complete website, and all posted and associated content indicates your acceptance of BioOne's Terms of Use, available at www.bioone.org/terms-of-use.

Usage of BioOne Complete content is strictly limited to personal, educational, and non - commercial use. Commercial inquiries or rights and permissions requests should be directed to the individual publisher as copyright holder.

BioOne sees sustainable scholarly publishing as an inherently collaborative enterprise connecting authors, nonprofit publishers, academic institutions, research libraries, and research funders in the common goal of maximizing access to critical research.

Downloaded From: https://bioone.org/journals/Freshwater-Science on 28 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

Evaluating AUSRIVAS predictive model performance for detecting simulated eutrophication effects on invertebrate assemblages

Susan J. Nichols[1,2], Trefor B. Reynoldson[1,3], and Evan T. Harrison[1,4]

[1]Institute for Applied Ecology, University of Canberra, Australian Capital Territory 2601, Australia

Abstract: Confidence in any bioassessment method is related to its ability to detect ecological improvement or impairment. We evaluated Australian River Assessment (AUSRIVAS)-style predictive models built using reference-site data sets from the Australian Capital Territory (ACT), the Yukon Territory (YT; Canada), and the Laurentian Great Lakes (GL; North America) area. We evaluated model performance as ability to correctly assign reference condition with independent reference-site data. Evaluating model ability to detect human disturbance is generally more problematic because the actual condition of test sites is usually unknown. Independent reference-site data underwent simulated impairment by varying the proportions of sensitive, intermediate, and tolerant taxa to simulate degrees of eutrophication. Model performance was related to differences in data sets, such as number and distribution of invertebrate taxa. Sensitive taxa tended to have lower expected probabilities of occurrence than more-tolerant taxa, but the distribution of taxa grouped by tolerance categories also differed by data set. Thus, the models differed in ability to detect the simulated impairment. The ACT model performed best with respect to Type 1 error rates (0%) and the GL model the worst (38%). The YT model performed best (10% error) for detecting moderate impairment, and the ACT model detected all severely impaired sites. AUSRIVAS did not assign most mildly impaired sites to below-reference condition, but a reduction in observed/expected values for some of the mildly impaired sites was observed. Models did not detect mild impairment that simply changed taxon abundances because presence–absence data were used for models. However, in comparison with other models described in this special issue (that did use abundance data), the AUSRIVAS model performance was comparable or better for detecting the simulated moderate and severe impairments.

Key words: reference condition approach, AUSRIVAS, simulated impairment, bioassessment, predictive modeling, macroinvertebrates

The Australian River Assessment System (AUSRIVAS) has been Australia's national standard method for biological assessment of river health for over a decade (Davies 2000, Simpson and Norris 2000, eWater CRC 2012). AUSRIVAS consists of a standardized invertebrate sampling method, predictive models, and software for assessing river health (Simpson and Norris 2000) that uses the reference-condition approach (Reynoldson et al. 1997). Adoption of AUSRIVAS bioassessment by water and environment agencies was rapid with implementation into state policy and regulatory frameworks and a variety of environmental management settings, by government, community, and industry (Davies 2000). AUSRIVAS has been used for targeted impact assessment (e.g., Marchant and Hehir 2002, Sloane and Norris 2003, Nichols et al. 2006, Growns et al. 2009, White et al. 2012), state/regional assessments of river condition (e.g., Turak et al. 1999, ACT Government 2006, Rose et al. 2008, Norris and Nichols 2011), community-based river assessment programs (e.g. WaterWatch; Davies 2007), and very broad-scale assessment at multijurisdictional and national levels (Turak et al. 1999, Norris et al. 2001a, b, EPA 2004, Davies et al. 2010, Harrison et al. 2011). A major strength of national systems like AUSRIVAS, River Invertebrate Prediction and Assessment System (RIVPACS) and Canadian Aquatic Biomonitoring Network (CABIN) is the broad-scale bioassessment and biomonitoring opportunities such programs allow (Rosenberg et al. 2000, Wright et al. 2000, Norris et al. 2001a, 2007). For example, AUSRIVAS data were the only data with national coverage used to report in-stream biological condition for Australia's 2011 State

E-mail addresses: [2]sue.nichols@canberra.edu.au; [3]Present address: Canada Creek, Nova Scotia B0P 1V0, Canada, trefor.reynoldson@gmail.com; [4]evan.harrison@canberra.edu.au

DOI: 10.1086/678946. Received 13 April 2013; Accepted 08 May 2014; Published online 2 October 2014.
Freshwater Science. 2014. 33(4):1212–1224. © 2014 by The Society for Freshwater Science.

Downloaded From: https://bioone.org/journals/Freshwater-Science on 28 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

of the Environment report (Harrison et al. 2011). Thus, AUSRIVAS has national significance for monitoring and assessing river condition in Australia.

In a review of alternatives to the River Invertebrate Prediction and Classification system (RIVPACS)-style predictive models, Johnson (2000) concluded that it was a robust approach for predicting assemblage structure and found no compelling reason justifying a change to other techniques. The AUSRIVAS method has produced models that work well in many of Australia's varied environments, and they have proved useful for river assessment (for further examples see Marchant and Hehir 2002, Hose et al. 2004, Metzeling et al. 2006, Nichols et al. 2010). However, since Johnson's review, other modeling methods have been used more extensively (Linke et al. 2005, Van Sickle et al. 2006, Chessman 2009, Webb and King 2009, Aroviita et al. 2010, Feio and Poquet 2011), and investigators have identified some limitations of the AUSRIVAS approach. For example, model performance is poor where reference sites are problematic or lacking (Chessman et al. 2010).

Implementation of national-scale water reforms and statutory water planning (Tomlinson and Davis 2010, Connell 2011, EU 2012) will necessitate evaluation of interventions designed to improve river conditions. Renewed pertinence of adequate assessment tools and continued advances in river assessment methods have prompted interest in development of new and improved tools for assessing ecological effects of human activities. Given the large initial investment in the AUSRIVAS approach, the utility of the method, and almost 20 y of experience since its inception, an appraisal seems timely and was one motivation prompting this special series of papers.

User confidence in any bioassessment modeling method is related to the method's ability to detect ecological improvement or impairment. Updating of predictive models that are in widespread use or introduction of alternative modeling options should involve careful evaluation and comparison of their performance. Evaluations of model performance generally are based on how well models predict group membership of reference sites and how well models predict the taxa found at new reference sites (Coysh et al. 2000, Hawkins et al. 2000). Such validation usually involves a data set of reference sites that are independent of those used to create the predictive model. However, evaluating the ability of models to detect human disturbance is more problematic than validating with reference sites because the biological condition of test sites is usually unknown. One approach is to use simulated impairments to determine the sensitivity of a method for detecting impairment (Cao and Hawkins 2005, Bailey et al. 2012). Evaluating both Type 1 and Type 2 error rates provides a better indication of model performance.

We used independent reference sites and simulated impairment (Bailey et al. 2014) to evaluate AUSRIVAS-style models built from reference-site data collected in Australia, the Yukon Territory (Canada), and the Laurentian Great Lakes (GL) area of North America to compare model performance for 3 very different environments. Independent reference-site data were artificially impaired to simulate 3 degrees of eutrophication. Evaluating model performance in this way allowed us to test the ability of models to detect known impairment. The results provided by the standard AUSRIVAS method used in our study form the basis for comparison with other modeling methods presented in this special series.

METHODS

Data sets

Authors of all papers in this special series analyzed the same data sets (described in full by Bailey et al. 2014). The reference-site data (invertebrate and environmental data) were collected from wadeable streams in the Australian Capital Territory (ACT) region (i.e., the upper Murrumbidgee River catchment), the Yukon Territory (YT), and from near-shore sites in the Laurentian Great Lakes (GL; North America). Each region had 2 reference-site data sets, 1 for model training and another independent data set consisting of 20 sites for model validation (D0). The invertebrate data from the validation sites were artificially impaired to simulate the effects of 3 degrees (D1 = mild, D2 = moderate, and D3 = severe) of eutrophication by varying the proportions of sensitive, intermediate, and tolerant taxa (Bailey et al. 2014). Impairment was simulated at each site for each level by altering the abundance of taxa or by removing some taxa. The simulated impairment was applied to randomly selected taxa within tolerance categories (e.g., sensitive, intermediate, tolerant) that were based on region-specific tolerance scores (Barbour et al. 1999) for the YT data, Hilsenhoff tolerance values (Hilsenhoff 1988) for the GL data, and Stream Invertebrate Grade Number (SIGNAL) values (Chessman 2003) for the ACT data.

AUSRIVAS modeling methods

We developed a standard AUSRIVAS model (Smith et al. 1999, Simpson and Norris 2000) using the reference-site training data for each data set. AUSRIVAS developers adapted the modeling approach originally described by the authors of the RIVPACS models (Wright et al. 1984, Moss et al. 1987, Wright 1995). In accordance with AUSRIVAS methods, we excluded sites with <6 taxa and taxa that occurred at <10% of sites in the training data. For each model, we grouped reference sites based on the similarity of their invertebrate assemblages using Unweighted Pair Group Method with Arithmetic Mean (UPGMA) cluster analysis on presence–absence data (Belbin 1993). We then selected a subset of the environmental variables (predictor variables) by using stepwise discriminant function analysis to determine which environmental variables best discriminated among reference-site groups (Smith et al.

Downloaded From: https://bioone.org/journals/Freshwater-Science on 28 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

1999, Simpson and Norris 2000). We used the discrimination cross-validation procedure as an internal model check regarding error rates for assigning training sites to correct groups (Smith et al. 1999). AUSRIVAS models predict the taxa expected at a site by summing the individual probabilities of occurrence for all the taxa predicted to have ≥50% probability of occurrence (Simpson and Norris 2000), resulting in a site-specific expected taxa list.

AUSRIVAS bands of biological condition

How much the observed invertebrate assemblage (O) deviates from that expected (E) is a measure of the severity of environmental impairment. AUSRIVAS assigns O/E scores to quality bands that represent different levels of biological condition (Coysh et al. 2000). Sites with O/E scores in band A are similar to reference condition, whereas sites with O/E values in band B or lower are considered impaired (Table 1). The distributions of the training reference-site O/E scores were used to determine the width of the quality bands. Thus, band widths are specific to each model (Table 1).

Model validation and performance evaluation

Outside model experience We used standard AUSRIVAS methods to assess whether validation sites were within the environmental scope of the reference data set for each model. We calculated the Mahalanobis distance of each site to each canonical variate (as per Clarke et al. 1996). We then used a χ^2 test to determine whether each site was within the 99% confidence interval of the centroid of ≥1 reference-site group. If a site's environmental characteristics differed significantly from the training data set (which could indicate underrepresentation of that site type in the training data set) then that site would have no appropriate reference group for comparison. At that stage, the site may be identified as 'outside the experience of the model', and the model predictions and site assessments treated as suspect (Coysh et al. 2000).

Model validation We assessed the ability of each model to correctly assign a new reference-site O/E to band A (Table 1) with the validation data set. Assuming the validation sites were truly in reference condition, <10% of the sites should mistakenly fall below AUSRIVAS band A (Coysh et al. 2000). A failure rate >10% would indicate that the model had a greater than expected Type 1 error rate (sites failed that should have passed). Thus, we based the Type 1 error rate on % validation sites with O/E values <10th percentile of the training-data O/E distribution.

Model performance for detecting impairment We used the simulated impairment validation data sets to assess Type 2 error rates (sites passed that should have failed). We tested the ability of models to detect the 3 degrees of simulated impairment as the percentage of sites assessed as below band A. Thus, we based the Type 2 error rate on % simulated impairment sites with O/E values >10th percentile of the training-data O/E distribution. We used the other AUSRIVAS quality bands to assess the ability of the models to detect a disturbance gradient.

Table 1. Australian River Assessment System (AUSRIVAS) bands of biological condition for the Australian Capital Territory (ACT; upper Murrumbidgee River catchment in ACT region of Australia), the Yukon Territory (YT), and the Laurentian Great Lakes (GL; North America) models, showing observed/expected (O/E) taxa range, band descriptions, and interpretations (Coysh et al. 2000).

Band	O/E value	Band description	O/E interpretations
X	ACT > 1.15 YT > 1.28 GL > 1.18	More biologically diverse than reference; O/E > 90th percentile of reference sites used to create the model	More taxa found than expected; site is potentially rich in biodiversity or may have mild organic enrichment that initially could increase the number of taxa because of increased food resources resulting from the increase in nutrients; test site requires further consideration before drawing conclusions
A	ACT 0.85–1.15 YT 72–1.28 GL 0.82–1.18	Similar to reference; O/E within range of central 80% of reference sites used to create the model	Most/all of the expected taxa found; water quality or habitat condition similar to reference sites
B	ACT 53–84 YT 0.15–71 GL 0.43–0.81	Significantly impaired; O/E < 10th percentile of reference sites used to create the model; band B width = band A width	Fewer families than expected; water quality or habitat quality possibly impaired, resulting in loss of expected taxa
C	ACT 0.21–52 YT 0–0.14 GL –0.08–0.44	Severely impaired; O/E value < band B; band C width = band A width	Many fewer families than expected; poor water quality or habitat quality resulting in loss of expected invertebrate richness
D	ACT 0–0.20 YT No band D GL 0–0.07	Extremely impaired; O/E value < band C to 0	Few expected families and only hardy and pollution-tolerant taxa remain; extremely poor water quality or habitat quality resulting in severe impairment

Downloaded From: https://bioone.org/journals/Freshwater-Science on 28 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

Low-probability taxa E for a site is the sum of the site-specific expected probabilities of the individual taxa with ≥0.50 probability. We tested whether sensitive taxa (as defined by Bailey et al. 2014), on average, had lower expected probabilities of occurrence than more-tolerant taxa. If so, excluding sensitive taxa with low probability of occurrence might obscure the simulated impairment that removed selected sensitive taxa. For each model, we evaluated the effect of the tolerance category of low-probability taxa on the model and the model's ability to detect the 3 levels of simulated impairment. Taxa with 0 (to 3 decimal places) expected probabilities were excluded from the analyses because some taxa are naturally restricted to particular stream types and, therefore, are naturally excluded from a proportion of the sites (as per Clarke and Murphy 2006). Including many 0 values would have distorted the frequency distributions. We calculated the expected probability of occurrence of each taxon across all the validation sites and compared the distributions of the probability values within each tolerance category (box plots).

RESULTS

Model summary

We created 1 model from each of the 3 data sets (Tables 2, 3). The number of taxa removed from the data set because they were considered rare and, therefore, excluded from the models was high (ACT: 46%, YT: 52%, GL: 59%). The ACT model used all available training sites, but sites were removed from the YT (n = 17) and GL (n = 40) data sets because they had too few taxa for modeling purposes (≤5). The cross-validation error for the YT model (44%) was greater than usually desired for an AUSRIVAS model (Table 2). The YT model also produced the widest range of O/E values for the training sites and had the widest quality bands of all models (Table 1). A wide band A equates to a wide range of accepted reference condition.

The data sets differed in the total number of taxa and total taxa used for modeling (Table 2). Taxon richness was ≤6 at 50% of the GL and 17% of the YT sites, whereas no ACT sites had <10 taxa and 50% had >18 taxa. The YT and GL data sets had 8 and 14 fewer taxa, respectively, than the ACT data set. The models developed with these data sets also varied in the number of taxa expected to occur at sites (using the standard 0.5 probability cutoff) (Table 2). The ACT model predicted 8 to 10 more taxa than the low estimates predicted by the YT and GL models (Table 2), a greater percentage than the difference in total taxa. This result indicates that factors other than the difference in total number of taxa present in the data sets are required to explain differences among models regarding predicted taxa.

Model validation and performance for detecting simulated impairment

Some validation sites appeared dissimilar to training sites used for model development based on ordination of the biological data (Fig. 1A–C), particularly in the YT and GL models where some validation sites and the removed low-richness training sites shared a similar ordination space (Fig. 1B, C). However, for all models, no validation sites were outside the model experience regarding their environmental character.

Type 1 errors The ACT model correctly assigned all validation sites to band A and had the lowest Type 1 error rate (Table 4). The Type 1 error rates for the YT and GL models were >10%. The GL model had the highest Type 1 error rate (Table 4).

Type 2 errors The ACT model detected all severely impaired (D3) sites, which were allocated to AUSRIVAS band C (severely impaired) or near the boundary of bands C and B (Table 4, Fig. 2A). Most (80%) of ACT sites with moderate (D2) levels of impairment were allocated to AUSRIVAS band B (significantly impaired) (Table 4, Fig. 2A). Except for 2 sites, the mildly impaired (D1) ACT sites did not fall below band A (Fig. 2A). However, some D1 sites had lower O/E values within band A (Fig. 2A) than did the original unimpaired validation sites (D0). The ACT

Table 2. Summary details for Australian River Assessment (AUSRIVAS) predictive models developed for the Australian Capital Territory (ACT; upper Murrumbidgee River catchment in ACT region of Australia), the Yukon Territory (YT), and the Laurentian Great Lakes (GL; North America). O/E = observed/expected taxa.

	Initial data set		Model data set		Model details					
Model	Sites	Taxa	Sites	Taxa	Groups (and sites per group)	Cross-validation error	Band width	Upper value of impaired O/E range	Predicted taxa range	Expected taxa range
ACT	87	67	87	36	3 (45, 25, 17)	0.28	0.31	0.84	13–19	10.3–13.9
YT	118	59	101	28	4 (27, 25, 31, 18)	0.44	0.56	0.71	5–12	3.6–9.0
GL	124	54	84	22	3 (21, 24, 39)	0.31	0.36	0.81	5–9	4.3–7.2

Downloaded From: https://bioone.org/journals/Freshwater-Science on 28 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

Downloaded From: https://bioone.org/journals/Freshwater-Science on 28 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

Table 3. Mean (SD) predictor variables and number of taxa per group for Australian River Assessment (AUSRIVAS) predictive models developed for the Australian Capital Territory (ACT; upper Murrumbidgee River catchment in ACT region of Australia), the Yukon Territory (YT), and the Laurentian Great Lakes (GL; North America). EC = ecological condition, UCA = upstream catchment area, EC = electrical conductivity, nonprod = nonproductive, snow = total annual snowfall, rif = riffle, boul = boulders.

Groups	Statistic	Taxa	Predictor variables						Summary of group characteristics
ACT			Altitude (m)	UCA (km^2)	EC (µS/cm)	Riffle depth (cm)	% rif boul	Velocity : depth	
Group 1	Mean	18 (2.7)	740 (189.9)	499 (1178.5)	96 (82.8)	18 (8.4)	18 (11.7)	14.7 (3.0)	Streams with high proportion of boulder-sized substrate, mid-range EC values
	Range	11–24	380–1095	12–5709	24–459	4–40	0–50	10–20	
Group 2	Mean	19 (2.5)	982 (207.4)	80 (142.5)	52.58 (45.87)	18.28 (8.78)	11.4 (10.56)	13.4 (2.6)	High altitude sites with small catchment area, low EC values, greater taxon richness
	Range	15–25	620–1360	3.9–572	17–245	3–36	0–40	10–18	
Group 3	Mean	14 (2.7)	665 (153.7)	1539 (2405.8)	222.2 (162.8)	19.94 (9)	11.76 (11.45)	14 (3.4)	Larger rivers, greater EC values, substrate of smaller particle size
	Range	10–20	380–910	5.5–6589	41.7–509	6–39	0–30	7–19	
YT			% metamorphic	% sedimentary	% alpine	% nonprod forest	Snow (mm)	Wetted width (m)	
Group 1	Mean	10 (3.3)	15 (36.2)	50 (48.3)	15 (25.3)	20 (24.2)	133 (21.6)	7.3 (6)	Metamorphic/sedimentary geology, higher taxon richness
	Range	6–20	0–100	0–100	0–99.7	0–88.5	103–167	1–25	
Group 2	Mean	8 (1.58)	16 (37.4)	66 (44.2)	18 (25.9)	37 (32.6)	122 (18.5)	4 (2.3)	Small streams, metamorphic/ sedimentary geology, less snow, wide sediment size range, lowest taxon richness
	Range	6–12	0–100	0–100	0–93.4	0–95.6	98–167	1.3–11	
Group 3	Mean	13 (3.23)	0 (0)	75 (41.6)	30 (19.9)	37 (25.6)	148 (19.8)	7.3 (3.7)	Catchments with high % alpine and % nonproductive forest, greatest snowfall, greatest taxon richness
	Range	8–20	0–0	0–100	0–73.7	0–67.9	106–167	2–14.6	
Group 4	Mean	10 (2.17)	0 (0)	94 (13)	35 (40.3)	14 (21.4)	125 (14.4)	6.04 (6.3)	High % alpine and low % nonpro ductive forest, sedimentary bedrock geology
	Range	6–14	0–0	51–100	0–99	0–61	106–167	1.3–30	
GL			Latitude (dec °)	% clay	Water depth (m)	% sediment MgO			
Group 1	Mean	7 (1.4)	45.5 (1.8)	46.3 (20.9)	20 (10.2)	2.7 (1.4)			Most lakes in this group, large % clay substrates, low taxon richness
	Range	6–10	42–49	6.7–77	8–51	1–8			
Group 2	Mean	8 (2.6)	46.8 (1.9)	20.6 (22.1)	36 (22.9)	3.2 (2.1)			Upper oligotrophic lakes (Huron, Superior), sediment with high % MgO, deeper water
	Range	6–14	43.8–48.9	0–83.8	6.8–87	0.6–8.4			
Group 3	Mean	11(2.3)	44.3 (1.4)	20 (18.7)	7 (6.6)	2.1 (1.1)			Only group with lower lake (Erie) sites (more mesotrophic), shallow water depth, high taxon richness
	Range	6–17	42.3–48.9	0–76.1	0.5–27.5	0.95–7.3			

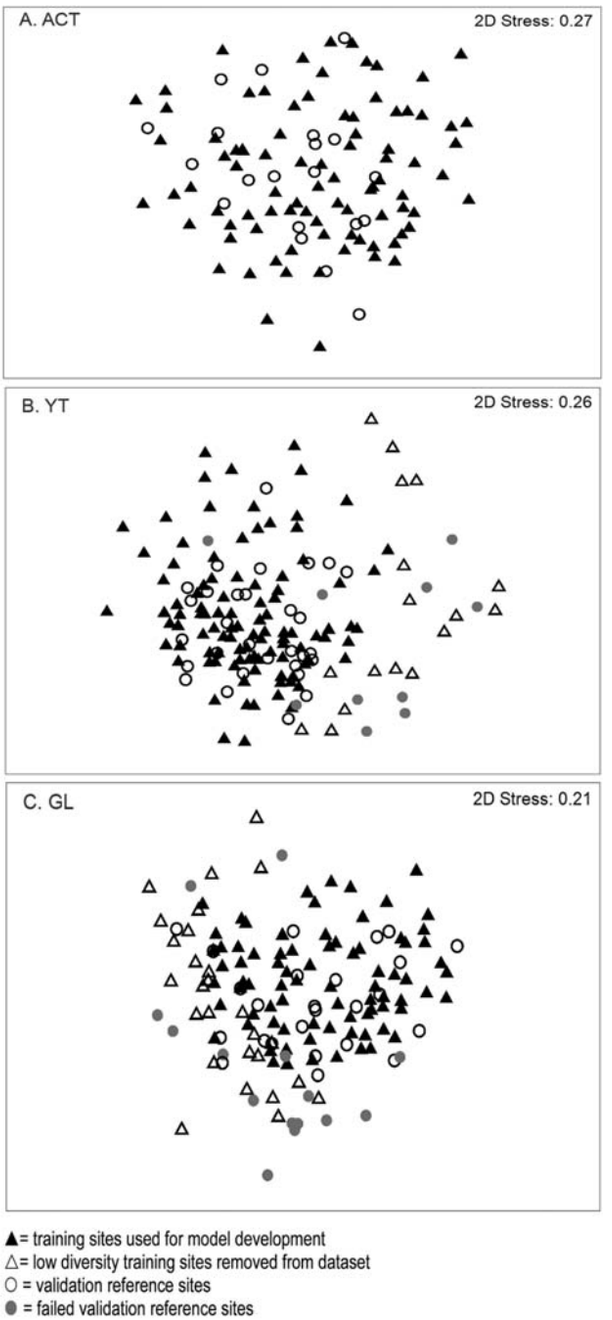


Figure 1. Two-dimensional ordinations of the invertebrate presence–absence data for Australian Capital Territory (ACT) (A), Yukon Territory (YT) (B), and Laurentian Great Lakes (GL) (C).

model produced O/E values that distinguished best between D2 and D3 sites (Fig. 2A).

O/E values produced by all models were distributed along a gradient, but the gradients produced by the YT and GL models did not always correspond to the simulated impairment levels (Fig. 2A–C). The YT and GL models generally did not distinguish D1 sites from the original D0 sites (Fig. 2B, C) because the data did not differ. Between 46 and 59% of the total taxa in the original data sets were not used for model creation because they occurred at <10% of reference sites, and many of those excluded taxa were also involved in the simulated impairment process that was applied to data sets prior to model development (ACT: 45%, YT: 42%, GL: 53%). Thus, this situation contributed to nondetection of mild simulated impairment.

Low-probability taxa For ACT, the median probability of occurrence for predicted taxa in the sensitive category was 0.34, which is lower than the medians for taxa in the tolerant (0.65) and intermediate (0.46) categories. Compared with the other models, the ACT model had the most taxa above the 0.5 probability cut-off value (Fig. 3A, Table 5). For YT, taxa in the intermediate category had the greatest median value (0.47), and most probability values >0.5 were for taxa in the intermediate category (Table 5, Fig. 3B). For GL, most taxa with probability values >0.5 were in the tolerant category, but the median values for probabilities in all tolerance categories were <0.5 indicating the presence of many low-probability taxa (Fig. 3C, Table 5). This difference between models in taxon probabilities in tolerance categories (combined with fewer taxa overall) contributed to differential ability to detect simulated impairment among models and explains the low number of expected taxa for the YT and GL models (Table 2).

DISCUSSION

We built an AUSRIVAS-style predictive model for reference-site data sets from very different environments (ACT, YT, GL; Tables 2, 3) and used the simulated impairment data sets to evaluate the ability of each model to detect impairment. The data sets differed in total number

Table 4. Type 1 and Type 2 error rates for the Australian Capital Territory (ACT), Yukon Territory (YT), and Laurentian Great Lakes (GL) Australian River Assessment System (AUSRIVAS) predictive models. Each category had 20 sites artificially impaired to simulate degrees of eutrophication impact (D0 = validation sites with no impairment, D1 = mild, D2 = moderate, and D3 = severe).

	Type 1 error rate (% validation sites assessed as impaired)	Type 2 error rate (% simulated impairment sites assessed as equal to reference (%))		
Model	D0	D1	D2	D3
ACT	0	90	20	0
YT	25	75	10	12.5
GL	37.5	57.5	32.5	12.5

Downloaded From: https://bioone.org/journals/Freshwater-Science on 28 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

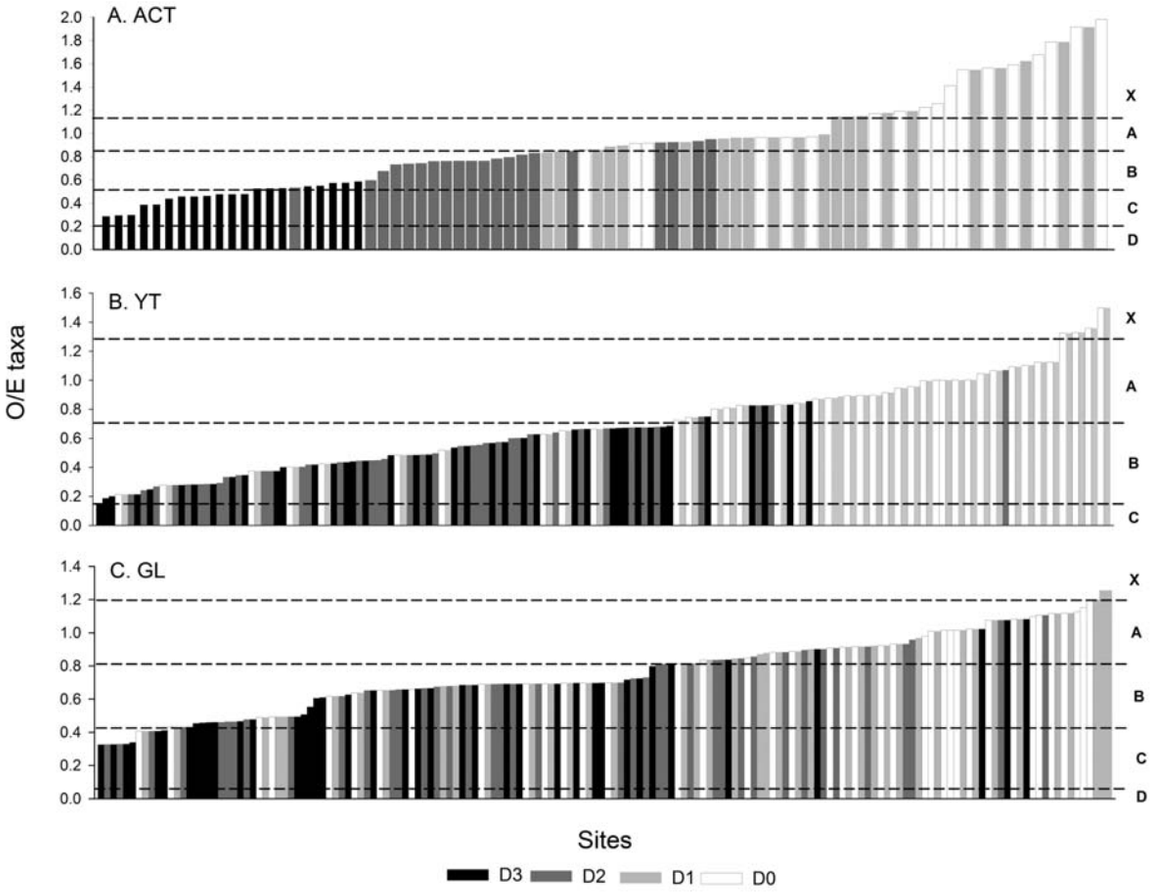


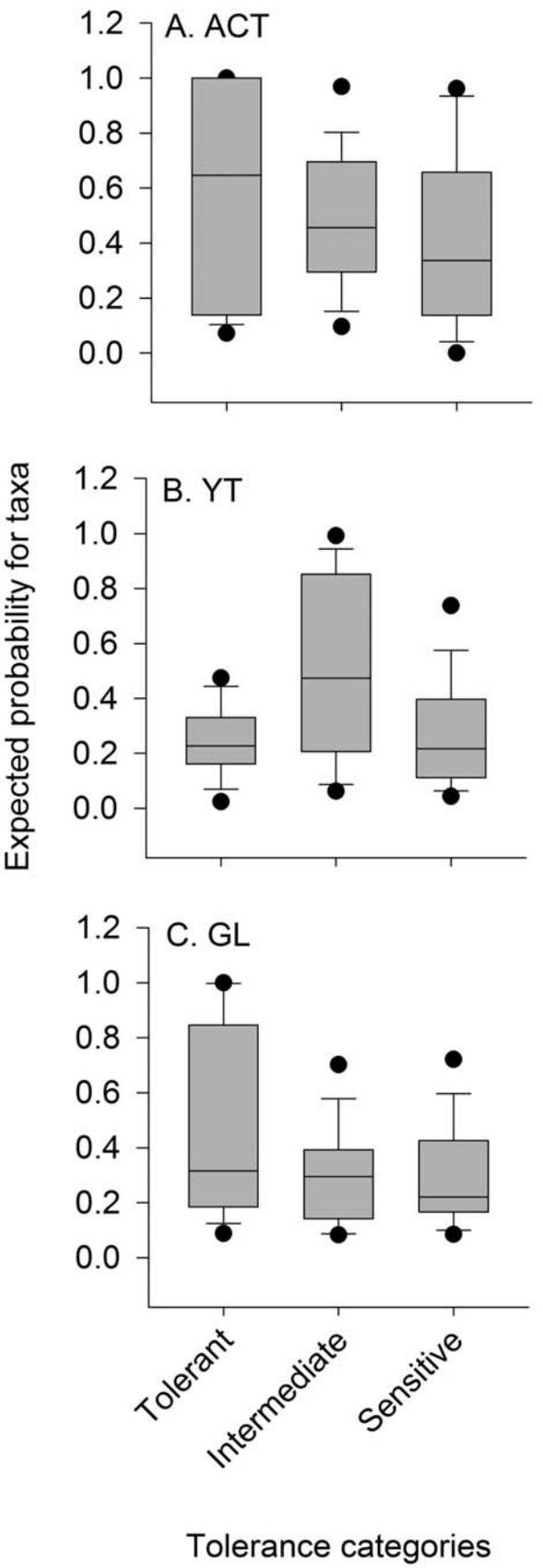
Figure 2. Distribution of observed/expected (O/E) taxa values for the validation sites (D0) and the 3 degrees of simulated impairment (D1 = mild, D2 = moderate, D3 = severe) for the Australian Capital Territory (ACT) (A), Yukon Territory (YT) (B), and Laurentian Great Lakes (GL) (C) data sets. The horizontal dotted lines represent the Australian River Assessment System (AUSRIVAS) bands of biological condition. Band A is equivalent to reference condition, and lower bands represent increasingly impaired condition. Band X is more diverse than reference.

of taxa (Table 2) and in the distribution of taxa (Bailey et al. 2014). These major differences and inherent characteristics of each data set influenced Type 1 and Type 2 error rates of the models.

The Type 1 error for the ACT model was 0%, but the YT and GL models had greater-than-expected Type 1 error rates. The YT and GL data sets contained sites that were similar in terms of measured environmental variables but that differed in their invertebrate assemblages, a combination that makes modeling difficult. The failed sites were within the environmental scope of the models (i.e., not outside model experience based on the predictor variables used) but biologically, many were dissimilar to the training sites used in the models (Fig. 1B, C). The invertebrate assemblages of these failed validation sites were similar to those of the unused, low-richness sites that were considered to have too few taxa for modeling. These low-richness sites may constitute a particular site type that was consequently not represented in the model. If such sites could be characterized (e.g., if they were all sites from harsh glacial environments in the Yukon Territory or oligotrophic systems in the Laurentian Great Lakes region), the model limitations could be characterized and subsequent model users could be advised that assessment of these types of sites will underestimate the O/E value. Knowledge of the model's limitations would enable users to identify particular site types that a model will not adequately match to reference sites. Users could then select an alternative assessment method or biological group more suitable for assessing those sites. Knowledge that test sites were being compared with an appropriate set of reference sites would provide users with greater confidence in the site assessments provided by the predictive models.

Models with wide biological-quality bands may have lower probabilities of misbanding than models with narrow bands (Barmuta et al. 2003). However, wide bands mean wide ranges of acceptable reference condition and, possibly, less sensitivity to impairment because the im-

Downloaded From: https://bioone.org/journals/Freshwater-Science on 28 Mar 2024
Terms of Use: https://bioone.org/terms-of-use



paired condition is more likely to fall within the range of acceptable condition. Regardless of the cause of wide bands (an inadequate set of reference sites or a naturally wide range of reference condition), such a model may have low power to detect impairment. However, the GL model was least able to detect impairment even though the YT model had the greatest band widths.

Other potential sources of error in estimates of the expected taxa and reference condition include an inadequate set of reference sites or insufficient environmental predictor variables to distinguish among reference-site groups (Ostermiller and Hawkins 2004, Clarke and Hering 2006, Bailey et al. 2012). New spatial tools are becoming increasingly available, particularly geographic information system (GIS) tools and an array of catchment-scale map layers describing attributes, such as geology, landuse, vegetation type, and climate (Frazier et al. 2012). GIS layers and remotely sensed data offer alternative approaches to defining reference sites (Yates and Bailey 2010) and are sources of potential predictor variables (Armanini et al. 2012). The predictor variables used in our study (Table 3) were a selected subset of the available data set, but variables that more completely characterize the factors controlling invertebrate distribution might improve the models (Ostermiller and Hawkins 2004).

Simulating mild impairment involved decreasing the abundance of sensitive taxa, increasing the abundance of tolerant taxa, and removing 2 randomly selected sensitive taxa (Bailey et al. 2014). With a few exceptions for the ACT model, the AUSRIVAS models did not perform well in detecting such mild impairment. Three factors contributed to the nondetection of the mild level of simulated impairment. First, the AUSRIVAS observed taxa list will not change if the taxa selected for simulated removal were not used for model development and, thus, were not included in the list of expected taxa. The standard AUSRIVAS modeling procedure is to remove (rare) taxa that occur at <10% of reference sites in the training data set. For all models, the number of taxa removed before model creation was high (Table 2). Second, where the artificially impacted taxa had <0.5 probability of occurrence at the site, they would not contribute to the O/E score. Third, we developed the models using presence–absence data, and thus, the models will not detect impacts within the data sets that are manifested only by a change in abundance. Thus, the

Figure 3. Distribution of the expected probability of occurrence for taxa across all validation sites for the Australian Capital Territory (ACT) (A), Yukon Territory (YT) (B), and Laurentian Great Lakes (GL) models. Taxa are grouped by their tolerance category (see Bailey et al. 2014 for assignments of taxa to categories). Lines in boxes are medians, box ends are quartiles, whiskers are 90th and 10th percentiles, and circles are 5th and 95th percentiles.

Downloaded From: https://bioone.org/journals/Freshwater-Science on 28 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

Table 5. Number of taxa probability values > 0.5 by taxon tolerance category for the Australian Capital Territory (ACT), Yukon Territory (YT), and Laurentian Great Lakes (GL) Australian River Assessment System (AUSRIVAS) predictive models (number of sites used for each model in brackets).

Model	Tolerant	Intermediate	Sensitive	Total
ACT (sites = 87)	40	177	96	313
Taxa per category	5	18	13	36
YT (sites = 101)	7	187	64	258
Taxa per category	6	10	12	28
GL (sites = 84)	199	29	20	248
Taxa per category	14	5	3	22

taxon richness (the basis for the O/E score) for most of the mildly disturbed sites was similar to that of the original validation data set and the taxa observed (O) (i.e., the taxa captured from the list of predicted taxa) differed little, or not at all, between the validation sites and those that were mildly impaired for all models (Table 4, Fig. 2A–C).

Consequently, the models had large Type 2 error rates regarding mildly disturbed sites (particularly the ACT model; Table 4). The Type 2 errors for the D1 sites were largely an inverse reflection of the Type 1 error rates. The difference among models regarding the Type 2 errors for mildly impaired sites is related to the random nature of taxa removed to simulate the impairment and to the differential effects that missing taxa have in relation to the number of taxa expected (which differed by model; Table 2). For example, removing 2 taxa from a YT site at which 3.6 taxa (56% of expected taxa) are expected will have greater effect on the O/E value than removing 2 taxa from an ACT site at which 13.9 taxa (14.4% of expected taxa) are expected.

If the simulated impairment data sets accurately represented eutrophication disturbance, then the AUSRIVAS model better detected such disturbance in the ACT region than did the models developed for the GL or YT regions. The ACT model most accurately displayed the gradient from moderate to severe impairment (Fig. 2A). The other 2 models displayed a gradient of O/E values, but the impaired and validation sites were more randomly distributed along that gradient. Often the different levels of impairment at specific sites in the GL and YT data sets were not distinguishable (Fig. 2B, C). The ability to detect the simulated impairment depended on whether the simulated disturbance was severe enough to remove taxa and whether those same taxa were used for modeling. The ACT model was created with a data set that had more taxa per site and more uniformly distributed taxa than in the other data sets, so a greater proportion of the biological data were used for model development. Thus, the taxa that underwent simulated impairment had a greater chance of being used in the ACT model than in the YT and GL models, which increased the probability of detecting the ACT impairment.

Regardless of whether abundance or presence–absence data (as for AUSRIVAS) are used for modeling, detection of eutrophication or any other disturbance in the real world will depend on the invertebrate sampling and processing methods. As sites become increasingly stressed, more of the sensitive taxa will disappear from the sites and the samples (Cao and Hawkins 2005). The sampling and subsampling methods will influence the proportion of locally rare taxa observed in a sample (Clarke and Murphy 2006). For example, sampling methods that collect the maximum number of different taxa regardless of their abundance may cause the model to have trouble detecting a mild disturbance that simply changes the relative abundances of taxa (Nichols and Norris 2006), whereas a sampling method that collects taxa relative to their abundance at the site (Nichols et al. 2000, Nichols and Norris 2006, Environment Canada 2012) could enable the model to detect a change in abundance before the impact removes taxa from the site, even when relying on taxon richness measures for assessment. Data sets collected with different sampling methods at the same stressed site can give the appearance of different responses to the disturbance simply because of the sampling or subsampling method (Ostermiller and Hawkins 2004). When simulated impairments were applied to an existing data set, the assumption made was that the sampling methods had not influenced the assemblage structure of the data set. Clearly, this assumption was not correct. Nonetheless, a simulated impairment data set is the only way to evaluate method performance regarding Type 2 errors. The accuracy of the representation of the impairment is less important than knowing the level to which the data set was impaired. Moreover, we evaluated only the O/E taxa modeling method, which is only 1 component of the AUSRIVAS bioassessment protocol, which includes standardized sampling methods and other outputs and indices to aid interpretation of the O/E result.

AUSRIVAS models from the different regions varied in the number of taxa predicted and, thus, expected (Table 2). In models with low numbers of expected taxa, the O/E score is vulnerable to the chance omission of observed taxa at a site (Barmuta et al. 2003). Such chance omissions could result in misbanding a site and failing it when it should have passed (Type 1 error). Marchant (2002) suggested that O/E scores calculated from <20 expected taxa may be too variable and unreliable to use. In reality, the argument regarding the chance omission of observed taxa should be viewed relative to the probability of missing or misidentifying taxa at a site (Barmuta et al. 2003). If taxa at the low-richness sites also have a low probability of being misidentified or missed during sampling then the problem may not be great (Barmuta et al. 2003). Replicated sampling at naturally low-richness sites used for modeling (such as

Downloaded From: https://bioone.org/journals/Freshwater-Science on 28 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

those from harsh environments, e.g., some YT sites) may help to ensure that the reference condition is not underestimated for initial model development (Barmuta et al. 2003). Such replicated sampling could reduce the chance of Type 2 errors (sites passing that should fail) by providing a more reliable estimate of reference condition at low-richness sites.

We used the standard AUSRIVAS method in our study so that we could compare our results with those produced by other methods presented in this special series. Thus, we used the standard (although somewhat arbitrary) AUSRIVAS probability cut-off of 0.5, which excludes low-probability taxa from the expected taxa list. AUSRIVAS uses a 0.5 probability cut-off because taxa with ≤0.5 probability of occurrence have an equal or greater probability of *not* being observed at a site. Decreasing the cut-off value to <0.5 may increase the number of expected taxa. However, any new taxa will add increasingly slowly to the count of expected taxa and will not necessarily strengthen confidence in the O/E scores (see Marchant 2002, Clarke and Murphy 2006). Excluding taxa with a low probability of occurrence will result in less variable O/E estimates (Clarke and Murphy 2006), but the optimal cut-off value does not have universal consensus (Hawkins et al. 2000, Marchant 2002, Ostermiller and Hawkins 2004, Clarke and Murphy 2006, Van Sickle et al. 2007). Clark and Murphy (2006) found the marginally best cut-off value to be 0.2, but the power of detecting impacts was similar up to 0.5. Van Sickle et al. (2007) found that excluding taxa with <0.5 probability increased ability to detect impairment. Including low-probability taxa in the O/E calculations assumes they are reliable and not simply absent by chance from new assessment sites. Marchant (2002) concluded that low-probability taxa play no useful role in predictive models, such as AUSRIVAS. The results of such studies caused AUSRIVAS developers to use the 0.5 cut-off. Moreover, O/E cut-off thresholds must be standardized when comparing site assessments in multijurisdictional bioassessment programs, or they should be treated as different indices (Clarke and Murphy 2006). Nevertheless, by comparing the performance differences among models from the 3 regions, we found that the distribution of low-probability taxa contributed to whether a particular data set could produce an adequate predictive model for the detection of impairment. Compared with the other 2 data sets, the GL data set had the most severely skewed invertebrate frequency distribution (i.e., more of the low-probability taxa) (Bailey et al. 2014), and therefore, was more vulnerable to the effects of excluding taxa with <0.5 probability of occurrence.

Taxa assigned to the sensitive category tended to have lower expected probabilities of occurrence than more-tolerant taxa (Fig. 3A–C). Other investigators found similar patterns in the expected probabilities of sensitive taxa for RIVPACS-style predictions (Clarke and Murphy 2006). The sensitive taxa tended to be less widespread among reference sites particularly for the YT and GL data sets and, thus, had considerably lower average expected probabilities (Fig. 3B, C). Thus, use of the 0.5 probability threshold excluded more sensitive taxa than taxa in other categories and contributed to nondetection of impairment by the YT and GL models.

AUSRIVAS and most other RIVPACS-style predictive models use discriminant function analysis (DFA), which requires identification of reference-site groups (Van Sickle et al. 2006). However, in most reference data sets with many sites, invertebrate data are not characterized by discrete community assemblages (Hawkins and Vinson 2000). Rather, the data structure displays sites along a continuum of ≥1 taxonomic gradients. Each taxon's array of environmental requirements and habitat preferences determine the gradients evident in the invertebrate data sets (Resh et al. 1994, Menezes et al. 2010). The spatial scale of sampling also influences the underlying structure revealed by analysis, such as classification and ordination (Marchant et al. 1999), and gradients may become more obvious as the size of the reference data set (or the density and spatial scale of reference-site coverage) increases (Turak et al. 1999). The ACT data were collected from a relatively small area (12,000 km^2) compared with both the YT (840,000 km^2) and GL (244,160 km^2) data sets (Bailey et al. 2014). Thus, the density of ACT reference-site coverage also was greater. Our results indicated that the ACT model performed best, and the spatial ordination and density of reference sites may have contributed to this outcome.

Classifying discrete groups of sites is a requirement of DFA rather than a representation of the reality of the invertebrate assemblages. Other modeling approaches may explicitly acknowledge the continuum in taxon distributions and avoid the use of classification groups by using the ordination space of reference sites as the basis for predicting site-specific invertebrate assemblages (Linke et al. 2005). However, AUSRIVAS does not base the probability of taxon occurrence on just 1 classification group that is most similar to a site, unlike some other methods, e.g., Benthic Assessment of Sediment (BEAST; Reynoldson et al. 2001). Rather, AUSRIVAS uses the weighted probabilities of the site membership to all of the groups, in a sense accounting for the assemblage continuum. The use of weighted probabilities of the site membership to all classification groups may moderate the effects of misclassification errors associated with large cross-validation errors, as for the YT model (Table 2).

Conclusion

The ability of our models to detect the simulated impairment depended on whether the simulated disturbance was severe enough to remove taxa from the data set and on whether the removed taxa had been used for modeling. Rare taxa, which have a patchy distribution in the

Downloaded From: https://bioone.org/journals/Freshwater-Science on 28 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

data sets, were removed before developing the AUSRIVAS models. To further improve predictive performance, we removed sites with naturally low richness (from the YT and GL data sets). Thus, this low-richness site type was not represented in the models, a limitation for the particular model. Other methods or biota may be better for assessing the condition of sites with naturally low invertebrate richness. Thus, effectiveness and performance of the models was related to differences in the total number of invertebrate taxa and to the distribution of taxa in the data sets. In short, data sets with highly skewed taxon distributions are difficult to model.

Careful evaluation of model performance should consider both Type 1 and Type 2 errors because confidence in the assessment is related to the model's ability to detect impairment. Use of a simulated impairment data set is the only way to evaluate Type 2 errors in model performance. The YT model was the best for detecting moderate impairment (10% error), and the ACT model detected all severely impaired sites. All models detected a gradient in O/E scores, but the ACT model best distinguished between moderate and severe simulated impairment. Moreover, AUSRIVAS was able to assign site O/E scores to a band of biological quality, thereby indicating the level of impairment. AUSRIVAS did not assign most mildly impaired sites to below reference condition, but a reduction in O/E values within band A was observed for some mildly impaired sites. AUSRIVAS did not detect simulated mild impairment that simply changed taxon abundance in a data set because presence–absence data were used for model development. Nevertheless, in comparison with other models described in this special issue (that did use abundance data), the AUSRIVAS model performance was comparable or better for detecting the simulated moderate and severe impairments.

ACKNOWLEDGEMENTS

The data from the upper Murrumbidgee River catchment were provided by SN, EH, and the Environment and Sustainable Development Directorate, ACT Government, Canberra, Australia. GL data were provided by Lee Grapentine and Environment Canada. John Bailey, Yukon Government, Fisheries and Oceans Canada, and University of Western Ontario (UWO) provided the YT data set. We also acknowledge 2 anonymous referees who generously contributed their time and effort. Their recommendations greatly enhanced the value of this manuscript. Last, we acknowledge the major contributions of the late Richard Norris who initiated much of the AUSRIVAS work in Australia.

LITERATURE CITED

ACT (Australia Capital Territory) Government. 2006. Environmental flow guidelines 2006. Environment ACT, Canberra, Australia. (Available from: http://www.legislation.act.gov.au/di/2006-13/default.asp)

Armanini, D. G., W. A. Monk, L. Carter, D. Cote, and D. J. Baird. 2012. Towards generalised reference condition models for environmental assessment: a case study on rivers in Atlantic Canada. Environmental Monitoring and Assessment 185: 6247–6259.

Aroviita, J., H. Mykrä, and H. Hämäläinen. 2010. River bioassessment and the preservation of threatened species: towards acceptable biological quality criteria. Ecological Indicators 10: 789–795.

Bailey, R. C., S. Linke, A. G. Yates. 2014. Bioassessment of freshwater ecosystems using the Reference Condition Approach: comparing established and new methods with common data sets. Freshwater Science 33:1204–1211.

Bailey, R. C., G. Scrimgeour, D. Coté, D. Kehler, S. Linke, and Y. Cao. 2012. Bioassessment of stream ecosystems enduring a decade of simulated degradation: lessons for the real world. Canadian Journal of Fisheries and Aquatic Sciences 69:784–796.

Barbour, M. T., J. Gerritsen, B. D. Snyder, and J. B. Stribling. 1999. Rapid Bioassessment Protocols for use in streams and wadeable rivers: periphyton, benthic macroinvertebrates and fish. 2nd edition. EPA 841-B-99-002. Office of Water, US Environmental Protection Agency, Washington, DC.

Barmuta, L. A., L. Emmerson, and M. P. Otahal. 2003. The sensitivity of AusRivAS to variations of input values, low natural diversity, and temporal variation. Final report. University of Tasmania, Hobart, Australia. (Available from: http://secure.environment.gov.au/water/publications/environmental/rivers/nrhp/pubs/errors-2.pdf)

Belbin, L. 1993. PATN technical reference. Division of Wildlife and Ecology, Commonwealth Scientific and Industrial Research Organisation, Canberra, Australia.

Cao, Y., and C. P. Hawkins. 2005. Simulating biological impairment to evaluate the accuracy of ecological indicators. Journal of Applied Ecology 42:954–965.

Chessman, B. 2003. SIGNAL 2—a scoring system for macroinvertebrate ('water bugs') in Australian rivers. Monitoring River Heath Initiative Technical Report 31. Commonwealth of Australia, Canberra, Australia. (Available from: http://www.environment.gov.au/system/files/resources/a9ad51d4-a8a2-4e21-994d-c6381f7445ee/files/signal.pdf)

Chessman, B. C. 2009. Climatic changes and 13-year trends in stream macroinvertebrate assemblages in New South Wales, Australia. Global Change Biology 15:2791–2802.

Chessman, B. C., H. A. Jones, N. K. Searle, I. O. Growns, and M. R. Pearson. 2010. Assessing effects of flow alteration on macroinvertebrate assemblages in Australian dryland rivers. Freshwater Biology 55:1780–1800.

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 the expected fauna. Journal of Applied Statistics 23:311–332.

Clarke, R. T., and D. Hering. 2006. Errors and uncertainty in bioassessment methods—major results and conclusions from the STAR project and their application using STARBUGS. Hydrobiologia 566:433–439.

Clarke, R. T., and J. F. Murphy. 2006. Effects of locally rare taxa on the precision and sensitivity of RIVPACS bioassessment of freshwaters. Freshwater Biology 51:1924–1940.

Downloaded From: https://bioone.org/journals/Freshwater-Science on 28 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

Connell, D. 2011. Water reform and the federal system in the Murray–Darling Basin. Water Resources Management 25: 3993–4003.

Coysh, J., S. Nichols, G. Ransom, J. Simpson, R. Norris, L. Barmuta, and B. Chessman. 2000. AUSRIVAS macroinvertebrate bioassessment: predictive modelling manual. CRC for Freshwater Ecology, Canberra, Australia. (Available from: http://ausrivas.ewater.com.au/index.php/manuals-a-datasheets)

Davies, P. E. 2000. Development of a national river bioassessment system (AUSRIVAS) in Australia. Pages 113–124 *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.

Davies, P. E. 2007. AusRivAS: its utility, possible future governance and funding arrangements. Report to Department of the Environment and Water Resources. Technical Report. Department of the Environment and Water Resources, Canberra, and Freshwater Systems, Hobart, Australia. (Available from: http://eprints.utas.edu.au/3180/1/AUSRIVAS_scoping_paper_25_Oct_2007.pdf)

Davies, P. E., J. H. Harris, T. J. Hillman, and K. F. Walker. 2010. The Sustainable Rivers Audit: assessing river ecosystem health in the Murray–Darling Basin, Australia. Marine and Freshwater Research 61:764–777.

Environment Canada. 2012. Canadian Aquatic Biomonitoring Network field manual: wadeable streams. Science and Technology Branch, Environment Canada, Dartmouth, Nova Scotia. (Available from: http://www.ec.gc.ca/Publications/C183563B-CF3E-42E3-9A9E-F7CC856219E1/CABINFieldManual_EN_2012.pdf)

EPA (Environmental Protection Agency). 2004. Biological objectives for rivers and streams—ecosystem protection. Publication 793.2. Environmental Protection Agency. Victoria, Melbourne, Australia.

EU (European Union). 2012. The European Union Water Framework Directive: an approach to integrated river basin management for Europe. European Commission, Brussels, Belgium. (Available from: http://ec.europa.eu/environment/water/water-framework/index_en.html)

eWater CRC (eWater Cooperative Research Centre). 2012. AUSRIVAS: Australian River Assessment System. eWater, Canberra, Australia. (Available from: http://ausrivas.ewater.com.au/)

Feio, M. J., and J. M. Poquet. 2011. Predictive models for freshwater biological assessment: statistical approaches, biological elements and the Iberian Peninsula experience: a review. International Review of Hydrobiology 96:321–346.

Frazier, P., D. Ryder, E. McIntyre, and M. Stewart. 2012. Understanding riverine habitat inundation patterns: remote sensing tools and techniques. Wetlands 32:225–237.

Growns, I., I. Reinfelds, S. Williams, and G. Coade. 2009. Longitudinal effects of a water supply reservoir (Tallowa Dam) on downstream water quality, substrate and riffle macroinvertebrate assemblages in the Shoalhaven River, Australia. Marine and Freshwater Research 60:594–606.

Harrison, E. T., S. Nichols, B. Gruber, F. Dyer, A. Tschierschke, and R. Norris. 2011. AUSRIVAS: Australia's in-stream biological health 2003–2010. 2011 State of the Environment report. Report prepared for The Australian Government, Department of Sustainability, Environment, Water, Population and Communities, Canberra, Australia. (Available from: http://www.environment.gov.au/system/files/pages/ba3942af-f815-43d9-a0f3-dd26c19d83cd/files/soe2011-supplementary-water-ausrivas.pdf)

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.

Hawkins, C. P., and M. R. Vinson. 2000. Weak correspondence between landscape classifications and stream invertebrate assemblages: implications for bioassessment. Journal of the North American Benthological Society 19:501–517.

Hilsenhoff, W. L. 1988. Rapid field assessment of organic pollution with a family-level biotic index. Journal of the North American Benthological Society 7:65–68.

Hose, G., E. Turak, and N. Waddell. 2004. Reproducibility of AUSRIVAS rapid bioassessments using macroinvertebrates. Journal of the North American Benthological Society 23: 126–139.

Johnson, R. K. 2000. RIVPACS and alternative statistical modelling techniques: accuracy and soundness of principles. Pages 323–332 *in* J. F. Wright, D. W. Sutcliffe and M. T. Furse (editors). Assessing the biological quality of fresh waters: RIVPACS and other techniques. Freshwater Biological Association, Ambleside, UK.

Linke, S., R. H. Norris, D. P. Faith, and D. Stockwell. 2005. ANNA: a new prediction method for bioassessment programs. Freshwater Biology 50:147–158.

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., and G. Hehir. 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.

Marchant, R., A. Hirst, R. Norris, and L. Metzeling. 1999. Classification of macroinvertebrate communities across drainage basins in Victoria, Australia: consequences of sampling on a broad spatial scale for predictive modelling. Freshwater Biology 41:253–268.

Menezes, S., D. J. Baird, and A. M. V. M. Soares. 2010. Beyond taxonomy: a review of macroinvertebrate trait-based community descriptors as tools for freshwater biomonitoring. Journal of Applied Ecology 47:711–719.

Metzeling, L., D. Tiller, P. Newall, F. Wells, and J. Reed. 2006. Biological objectives for the protection of rivers and streams in Victoria, Australia. Hydrobiologia 572:287–299.

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.

Nichols, S. J., and R. H. Norris. 2006. River condition assessment may depend on the sub-sampling method: field live-sort versus laboratory sub-sampling of invertebrates for bioassessment. Hydrobiologia 572:195–213.

Nichols, S. J., R. Norris, W. Maher, and M. Thoms. 2006. Ecological effects of serial impoundment on the Cotter River, Australia. Hydrobiologia 572:255–273.

Downloaded From: https://bioone.org/journals/Freshwater-Science on 28 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

Nichols, S. J., W. A. Robinson, and R. H. Norris. 2010. Using the reference condition maintains the integrity of a bioassessment program in a changing climate. Journal of the North American Benthological Society 29:1459–1471.

Nichols, S. J., P. Sloane, J. Coysh, C. Williams, and R. Norris. 2000. Australian Capital Territory, Australian River Assessment System (AUSRIVAS) sampling and processing manual. Cooperative Research Centre for Freshwater Ecology, University of Canberra, Canberra, Australia. (Available from: http://ausrivas.ewater.com.au/index.php/resources/category/5-act-sampling-and-datasheets?download=9).

Norris, R. H., S. Linke, I. Prosser, W. J. Young, P. Liston, N. Bauer, N. Sloane, F. Dyer, and M. Thoms. 2007. Very-broad-scale assessment of human impacts on river condition. Freshwater Biology 52:959–976.

Norris, R. H., P. Liston, N. Davies, J. Coysh, F. Dyer, S. Linke, I. Prosser, and B. Young. 2001a. Snapshot of the Murray–Darling Basin river condition. Murray–Darling Basin Commission, Canberra, Australia.

Norris, R., and S. Nichols. 2011. Environmental flows: achieving ecological outcomes in variable environments. Pages 331–349 *in* Q. Grafton and K. Hussey (editors). Water resources planning and management. Cambridge University Press, Cambridge, UK.

Norris, R. H., I. Prosser, B. Young, P. Liston, N. Bauer, N. Davies, F. Dyer, S. Linke, and M. Thoms. 2001b. The Assessment of River Condition (ARC). An audit of the ecological condition of Australian Rivers. Final report submitted to the National Land and Water Resources Audit Office, Canberra, Australia. (Available from: http://piku.org.au/reprints/2001_Norris_etal_The_assessment_of_river_condition.pdf)

Ostermiller, J. D., and C. P. Hawkins. 2004. Effects of sampling error on bioassessments of stream ecosystems: application to RIVPACS-type models. Journal of the North American Benthological Society 23:363–382.

Resh, V. H., A. G. Hildrew, B. Statzner, and C. R. Townsend. 1994. Theoretical habitat templets, species traits, and species richness: a synthesis of long-term ecological research on the Upper Rhône River in the context of concurrently developed ecological theory. Freshwater Biology 31:539–554.

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.

Reynoldson, T. B., D. M. Rosenberg, and V. H. Resh. 2001. Comparison of models predicting invertebrate assemblages for biomonitoring in the Fraser River catchment, British Columbia. Canadian Journal of Fisheries and Aquatic Sciences 58:1395–1410.

Rose, P., L. Metzeling, and S. Catzikiris. 2008. Can macroinvertebrate rapid bioassessment methods be used to assess river health during drought in south eastern Australian streams? Freshwater Biology 53:2626–2638.

Rosenberg, D. M., T. B. Reynoldson, and V. H. Resh. 2000. Establishing reference conditions in the Fraser River catchment, British Columbia, Canada, using the BEAST (Benthic Assessment of SedimenT) predictive model. Pages 181–194 *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.

Simpson, J. C., and R. H. Norris. 2000. Biological assessment of river 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.

Sloane, P. I. W., and R. H. Norris. 2003. Relationship of AUSRIVAS-based macroinvertebrate predictive model outputs to a metal pollution gradient. Journal of the North American Benthological Society 22:457–471.

Smith, M. J., W. R. Kay, D. H. D. Edward, P. J. Papas, K. S. 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.

Tomlinson, M., and R. Davis. 2010. Integrating aquatic science and policy for improved water management in Australia. Marine and Freshwater Research 61:808–813.

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.

Van Sickle, J., D. D. Huff, and C. P. Hawkins. 2006. Selecting discriminant function models for predicting the expected richness of aquatic macroinvertebrates. Freshwater Biology 51: 359–372.

Van Sickle, J., D. P. Larsen, and C. P. Hawkins. 2007. Exclusion of rare taxa affects performance of the O/E index in bioassessments. Journal of the North American Benthological Society 26:319–331.

Webb, J. A., and E. L. King. 2009. A Bayesian hierarchical trend analysis finds strong evidence for large-scale temporal declines in stream ecological condition around Melbourne, Australia. Ecography 32:215–225.

White, H. L., S. J. Nichols, W. A. Robinson, and R. H. Norris. 2012. More for less: a study of environmental flows during drought in two Australian rivers. Freshwater Biology 57:858–873.

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. 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, Ambleside, UK.

Yates, A. G., and R. C. Bailey. 2010. Selecting objectively defined reference sites for stream bioassessment programs. Environmental Monitoring and Assessment 170:129–140.

Downloaded From: https://bioone.org/journals/Freshwater-Science on 28 Mar 2024
Terms of Use: https://bioone.org/terms-of-use