Bioassessment of freshwater ecosystems using the Reference Condition Approach: comparing established and new methods with common data sets

Robert C. Bailey[1,4], Simon Linke[2,5], and Adam G. Yates[3,6]

[1]Cape Breton University, P.O. Box 5300, 1250 Grand Lake Road, Sydney, Nova Scotia, Canada B1P 6L2
[2]Australian Rivers Institute, Griffith University, Kessels Road, Nathan, 4111 Queensland, Australia
[3]Canadian Rivers Institute and Department of Geography, Western University, 1151 Richmond Street, London, Ontario, Canada N6A 5C2

Abstract: Although used in many jurisdictions around the world, analytical approaches of the Reference Condition Approach (RCA) to bioassessment of freshwater ecosystems have evolved quite slowly over the past 2 decades. For this special series of papers in *Freshwater Science*, researchers analyzed 3 data sets that included both benthic macroinvertebrate and environmental data from a number of reference sites. Australian Capital Territory (ACT) reference sites (n_{total} = 107) were wadeable streams in the upper Murrumbidgee River catchment, Australian Capital Territory, Australia. Yukon Territory (YT) reference sites were wadeable streams (n_{total} = 158) in the Yukon Territory, Canada, part of the Yukon River basin. Great Lakes (GL) sites (n_{total} = 164) were all nearshore (<20 m) lentic sites in the North American Great Lakes. For each data set, sites were divided into model-building (training) and model-testing (validation) groups. Each validation site was further subjected to 3 levels of simulated degradation based on the sensitivity of the biota to eutrophication. The analytical approaches ranged from standard or slight modifications of methods used in national programs (Australian River Assessment [AUSRIVAS], Canadian Aquatic Biomonitoring Network [CABIN]), to improved matching of sites to be assessed and appropriate reference sites, and Bayesian and machine-learning modeling. In comparing Type 1 error rates (proportion of validation sites deemed not in reference condition) and power (proportion of simulated impairment sites deemed not in reference condition), we found no obvious pattern among the 3 data sets or approaches. Approaches commonly used in RCA programs would benefit from incorporating newer methods that better match reference and test-site environments and build better predictive models.
Key words: bioassessment, reference condition approach, AUSRIVAS, predictive models, BEAST, simpacts

The Reference Condition Approach (RCA) to bioassessment, originally developed >25 y ago and still implemented as the River Invertebrate Prediction and Classification System (RIVPACS) program in the UK (Wright 1995), has been used globally in regional and national programs for more than a decade. Substantial work has been done on the statistical techniques used to analyze data in RCA bioassessment programs since publication of a primer 10 y ago (Bailey et al. 2004). A special session held at the annual meeting of the Society for Freshwater Science in Louisville, Kentucky (2012) was devoted to using both long-established and new approaches to RCA analysis with common data sets. Authors presented their analyses of data from 3 RCA programs, and a selection of these presentations were further refined and are presented in this special issue of *Freshwater Science*.

RCA DATA SETS

The 7 research groups contributing to this special issue analyzed 3 data sets that include macroinvertebrate taxon abundances and environmental descriptors from reference sites (relatively unexposed to human activity) in: 1) Australian Capital Territory (ACT), 2) Yukon Territory (YT) in northwestern Canada in the Yukon River basin, and 3) Laurentian Great Lakes (GL) in North America.

The ACT sites (n_{total} = 107) were wadeable streams in the upper Murrumbidgee River catchment, Australia, which

E-mail addresses: [4]robert_bailey@cbu.ca; [5]simon.linke@gmail.com; [6]adam.yates@uwo.ca

DOI: 10.1086/678771. Received 07 July 2014; Accepted 11 July 2014; Published online 6 October 2014.
Freshwater Science. 2014. 33(4):1204–1211.

covers an area of 12,000 km^2. The climate is temperate, characterized by cool winters (average daily maximum temperature 11°C) and warm summers (average daily maximum temperature 27°C) (Parsons et al. 2003). Across the catchment, rainfall is relatively uniform throughout the year, and snowfall is common on the western ranges >1200 m asl. The southern part of the catchment receives 500–700 mm, the northern part receives 600–700 mm, and the western ranges receive 800–1200 mm of rain/y (Davies et al. 2000). However, temperature and rainfall are variable over time, and severe droughts and floods occur periodically (Parsons et al. 2003).

YT reference sites were all wadeable streams (n_{total} = 158). The Yukon River basin is the 6$^{\text{th}}$ largest in North America and covers ~840,000 km^2 (Bailey 2005). Roughly ⅔ of the basin is in Alaska. Most of the Canadian portion (90%) is in Yukon Territory, but many of the headwater lakes and streams are in northern British Columbia. The river extends ~3200 km from the headwaters to its outflow to the Bering Sea in Alaska. Terrestrial ecoregions inside the Yukon Territory include the Yukon Interior Dry Forests in the headwaters area, Interior Alpine Tundra in the middle portion of the basin and Lowland Taiga in the most northerly portion, including the Porcupine River subbasin.

The GL sites (n_{total} = 164) were all nearshore (<20m), lentic sites in areas of the North American Great Lakes that were relatively unexposed to human activity. The Laurentian Great Lakes are the largest freshwater lake network in the world, with a total lake area of 244,106 km^2 and shoreline length of 17,017 km (USEPA and Canada 1995). The watershed spans a diverse range of climates, geology, soils, and ecoregions. Northern parts are characterized by a cold climate, granitic bedrock, thin acidic soil layers, and coniferous forests. Southern regions are comparatively much warmer, underlain by sedimentary bedrock with glacial sediment deposits, deeper soil layers, and historically deciduous forests (USEPA and Canada 1995).

Each data set consisted of reference sites divided into *training* and *validation* groups. The training sites were used to build models that related the biota at a reference site to its environment. The validation sites were subjected to the model to determine whether they had the biota expected if the site were in reference condition. Thus, the proportion of validation sites deemed not in reference condition (i.e., failed the assessment) is a robust measure of the true Type 1 error rate (α) of the RCA (a reference site mistakenly deemed degraded) because it is independent of whatever decision rule was used to deem the site reference (pass) or nonreference (fail). It is robust because it was not based on sites that had been used to build the predictive model. The ACT validation sites (n_{val} = 20) were randomly selected from the full range of reference sites, leaving n_{train} = 87 training sites for the ACT analysis. YT (n_{val} = 40) and GL (n_{val} = 40) validation sites were selected using Principal Components Analysis (PCA) of the environmental descriptors, leaving n_{train} = 118 sites for the YT analysis and n_{train} = 124 sites for the GL analysis. The validation sites were randomly selected from within the 25$^{\text{th}}$ to 75$^{\text{th}}$ percentile range of the first 3 PC scores of the training sites.

SIMULATED IMPAIRMENT OF DATA SETS

Each validation site (D0 = no degradation) was artificially impaired to simulate 3 levels of degradation by eutrophication (D1 = mild degradation, D2 = moderate degradation, D3 = severe degradation) (Table 1). Hilsenhoff's (1988) Family Biotic Index (FBI) was used to designate YT and GL taxa as sensitive (FBI = 1–4), insensitive (FBI = 5–6), or tolerant (FBI = 7–10). Stream Invertebrate Grade Number (SIGNAL [SG]; Chessman 2003) was used to designate each ACT taxon as sensitive (SG = 7–10), insensitive (SG = 4–6), or tolerant (SG = 1–3). Taxon tolerance levels were used in an algorithm to simulate impairment of validation sites (Table 1).

Each simulated impairment (D1, D2, D3) validation site had the same values for environmental descriptors as its corresponding reference (D0) validation site and, thus, had the same predicted biota when subjected to the RCA model. Thus, the proportion of simulated impairment validation sites deemed to be in reference condition was a measure of the Type 2 error rate (β) of the RCA analysis (a degraded site mistakenly deemed in reference condition). The converse of the Type 2 error rate (1 − β) is the power of the bioassessment to detect a given deviation from reference condition.

Table 1. Simulating eutrophication degradation (D).

Impairment level	Algorithm
D0: Reference	Unchanged validation sites
D1: Mild	Sensitive taxa: reduce abundance 25% and eliminate 10% of taxa
	Intermediate taxa: unchanged
	Tolerant taxa: increase abundance 75%
D2: Moderate	Sensitive taxa: reduce abundance 75% and eliminate 50% of taxa
	Intermediate taxa: reduce abundance 50% and eliminate 20% of taxa
	Tolerant taxa: reduce abundance 25% and eliminate 10% of taxa
D3: Severe	Sensitive taxa: eliminate all taxa
	Intermediate taxa: reduce abundance 75% and eliminate 50% of taxa
	Tolerant taxa: reduce abundance 50% and eliminate 20% of taxa

The biota descriptors in each data set (Table 2), show a gradient of taxon richness from ACT to YT to GL, with a related, variable effect of the simulated impairment. Each program had similar categories of environmental descriptors available as potential predictors (Tables 3–5), but they differed in some notable ways, particularly with respect to descriptors that might be influenced by the human activity of interest (Bailey et al. 2004). The scale of environmental descriptors varied from almost all site scale (ACT, GL) to almost all landscape scale (YT).

ANALYTICAL APPROACHES

A total of 7 approaches were used to analyze the 3 data sets:

Nichols et al. (2014) used the standard Australian River Assessment System (AUSRIVAS) approach that has been implemented in various contexts in Australia for almost 20 y (Nichols et al. 2010). This method probably is the closest of all analytical methods used in this collection to the original RIVPACS approach to RCA (Wright 1995). In this method, training reference sites are grouped using classification analysis based on the similarity of their invertebrate communities, and a stepwise discriminant function analysis (DFA) is used to build a model that predicts the faunally defined group from the environment of the reference site. The proportion of occurrence (p_{occur}) of taxa that occur at ≥50% of sites ($p_{\mathrm{occur}} > 0.5$) in the predicted group are summed to an expected taxon richness (E), which is compared to the observed taxon richness (O) at the site. The ratio of the number of taxa observed at the site to those expected given the predicted reference group (O/E) is compared to the distribution of the O/E values for the training reference sites.

Strachan and Reynoldson (2014) used the Benthic Assessment of Sediment (BEAST) approach that has been implemented since the mid 1990s in the Canadian Aquatic Biomonitoring Network (CABIN) program (Reynoldson et al. 2001). Like AUSRIVAS, it classifies reference-training sites based on their invertebrate community composition and then uses a stepwise DFA to predict group membership from the reference sites' environmental descriptors. It differs from AUSRIVAS by doing a low-dimension ordination (based on the invertebrate community) of each test site and its predicted reference group before making a reference or nonreference decision that depends on whether the test site is inside a normal distribution confidence ellipse around the predicted reference group in ordination space. Reynoldson et al. (2014) modified the BEAST approach by building sequential stepwise-DFA models corresponding to the structure of the hierarchical classification of the reference training sites. For example, if a faunally based, hierarchical classification resulted in 3 groups of reference sites (A, B1, B2), a DFA model would be built to distinguish A from B1 and B2 sites, and a 2nd model would be built to distinguish B1 from B2 sites.

Many developments in RCA bioassessment analysis have been concerned with improving the so-called *matching* of the site being assessed with appropriate reference sites. Sarrazin-Delay et al. (2014) used Assessment by Nearest Neighbour Analysis (ANNA; Linke et al. 2005) and Redundancy Analysis (RDA; Legendre and Legendre 1998) to

Table 2. Sampling times and biological characteristics of the Australian Capital Territory (ACT), Yukon Territory (YT), and Great Lakes (GL) data sets. Min = minimum, max = maximum, D1–D3 = simulated impairment reference sites, where D1 = mildly impaired, D2 = moderately impaired, D3 = severely impaired.

Data set	Sampling period	Ecosystem	Sampling method	No. taxa	No. training sites	No. validation sites	Median site richness (min–max)	Median site abundance (min–max)
ACT	Spring 1994–1995	Stream riffle	Kick net from 10 m of riffle	67	87	20	18 (9–26)	Proportions only
						D1: 20	17.5 (9–25)	
						D2: 20	14 (8–19)	
						D3: 20	8 (5–13)	
YT	Summer 2006–2007	Stream	Traveling (3-min) kick net	59	118	40	10 (1–22)	263 (8–8825)
						D1: 40	9 (2–17)	222 (10–8863)
						D2: 40	5 (2–11)	58 (5–1131)
						D3: 40	4 (1–10)	45 (3–2156)
GL	Autumn 1991–2010	Nearshore lake	Box corer, mini-box corer, or Ponar grab	54	124	40	7 (3–24)	8957 (241–256,031)
						D1: 40	6 (3–22)	17,860 (377–390,847)
						D2: 40	6 (2–19)	5896 (166–162,440)
						D3: 40	4 (2–14)	3287 (60–105,850)

Table 3. Environmental descriptors in the Australian Capital Territory data set (Nichols et al. 2000).

Category	Description	Range
Geographic	Latitude (°)	35.1°S–39.7°S
	Longitude (°)	148.5°E–149.5°E
	Stream order	1–7
	Catchment area (km^2)	3.9–6589
	Distance from source (km)	3–222
	Altitude (m)	380–1360
Channel morphology	Bankfull width (m)	1–80
	Bank height (m)	0.3–8
	Reach riffle habitat (%)	5–90
Water chemistry	Conductivity (μS/cm)	17–509
	Alkalinity (mg/L)	8–370
Substrate/particle size (reach)	% bedrock	0–50
	% boulder (>256 mm)	0–40
	% cobble (64–256 mm)	0–60
	% pebble (16–64 mm)	5–40
	% gravel (2–16 mm)	3–35
Substrate/particle size (riffle)	% bedrock	0–60
	% boulder (>256 mm)	0–50
	% cobble (64–256 mm)	0–60
	% pebble (16–64 mm)	5–50
	% gravel (2–16 mm)	0–35
Depth/velocity	Water depth (cm)	3–40
	Water velocity (m/s)	0–1.26
Habitat assessment	Bottom substrate/available cover	5–20
	Pool/riffle, run/bend ratio	3–15

achieve matching without initially classifying the training sites based on their biota (as in the BEAST and AUSRIVAS methods). First, biota in each training and validation site, including the simulated impairment sites, were described using the summary metrics, abundance, diversity, and composition. Then each site was compared to others with similar environments, where site similarity was either defined in PCA (ANNA) or biota-correlated (redundancy analysis [RDA]) multivariate space of the environmental descriptors. Last, Test Site Analysis (TSA; Bowman and Somers 2006) was used to combine results from each metric in an assessment of each site.

Chessman (2014) used Limiting Environmental Difference Analysis (LEDA; Chessman et al. 2008) in a similar effort to define a better comparator group for assessing a site. In LEDA, assemblage composition at a site is compared only to that of sites with similar values for what are considered 'limiting' environmental features (e.g., depth in lentic systems). This approach is somewhat similar to ANNA (Linke et al. 2005), so as in the approach taken by Sarrazin-Delay et al. (2014), reference sites are not grouped a priori based on biota.

Bayesian modeling approaches are increasingly common in ecology and environmental science (Clark 2005), but they are rarely used in RCA bioassessment. Webb et al. (2014) developed Bayesian models to predict summary metrics of invertebrate diversity and composition from continuous and categorical environmental descriptors. They also incorporated the faunally defined classification groups determined by Strachan and Reynolds (2014) as categorical predictors in their models.

Feio et al. (2014) applied a variety of machine-learning tools in their analysis of the 3 data sets. Each method used the training sites to model the occurrence of individual taxa based on their environment, and then, similar to AUSRIVAS, summed the probability of occurrence

Table 4. Environmental descriptors in the Yukon Territory data set (Reynoldson et al. 2007).

Category	Description	Range
Geographic	Latitude (°)	60.2°N–67.6°N
	Longitude (°)	141.0°W–132.8°W
	Stream order	1–5
	Catchment area (km^2)	2.0–1856.0
	Catchment perimeter (km)	9.2–359.2
	Stream length in catchment (m)	204–642,137
	Stream density (m/km^2)	51.0–701.2
Land cover in upstream catchment	Alpine (proportion of catchment)	0–1.0
	Forested (proportion)	0–1.0
	Lake (proportion)	0–0.177
	Nonproductive forested (proportion)	0–0.922
	River (proportion)	0–0.486
	Unforested (proportion)	0–0.605
	Urban (proportion)	0–0.002
	Wetland (proportion)	0–0.597
Catchment bedrock geology	% metamorphic	0–100
	% sedimentary	0–100
	% sedimentary/volcanic	0–67.7
	% ultramafic	0–1.6
	% ultramafic/metamorphic	0–54.6
	% unclassified	0–100
	% unconsolidated	0–100
	% volcanic	0–100
Catchment long-term climate	Rainfall in January (mm)	0–2.0
	Rainfall in June (mm)	20.2–67.7
	Total annual rainfall (mm)	114.0–267.2
	Snowfall in January (mm)	10.5–32.4
	Snowfall in June (mm)	0–1.3
	Annual snowfall (mm)	97.7–166.9
	January total precipitation (mm)	6.9–31.6
	June total precipitation (mm)	20.4–64.3
	Mean maximum temperature in January (°C)	−27.2 to −13.6
	Mean minimum temperature in January (°C)	−37.0 to −23.0
	Mean temperature in January (°C)	−31.6 to −18.0
	Mean maximum temperature in June (°C)	16.8–20.5
	Mean minimum temperature in June (°C)	2.6–6.7
	Mean temperature in June (°C)	10.3–13.5
Stream channel at site	Average depth (cm)	6.6–200.0
	Mean velocity (m/s)	0.010–1.360
	Wetted width (m)	1.0–30.0
Water	pH	4.25–8.87
	Temperature (°C)	3.50–18.50

Table 5. Physicochemical descriptors in the Great Lakes data set (Reynoldson and Day 1998).

Category	Description	Range
Geographic	Latitude (°)	42.0°N–48.9°N
	Longitude (°)	88.9°W–74.7°W
General limnology (overlying water)	Water depth (m)	0.52–90.1
	Alkalinity (mg/L)	38.5–99.2
	Dissolved O_2(mg/L)	5–12.82
	pH	6.6–9.53
	Temperature (°C)	4.21–21.2
Nutrient concentrations	Total Kjeldahl N (water, mg/L)	0.042–0.599
	NO_2+NO_3-N (water, mg/L)	0.0025–0.392
	Total P (water, mg/L)	0.0027–0.037
	Total N (sediment, mg/L)	140–9900
	Total P (sediment, mg/L)	20–7180
	Total organic C (sediment, %)	0.05–7.78
	Loss on ignition (sediment, %)	0.59–56.3
Sediment/particle size	% clay	0–83.8
	% gravel	0–36.19
	% sand	0–99.79
	% silt	0–95.75
	Particle size: 25th percentile (μm)	2.2–3654.4
	Particle size: 75th percentile (μm)	0–405.19
	Particle size: mean (μm)	2–1167.4
Major element oxide concentration (sediment)	% Al_2O_3	3.26–16.1
	% CaO	0.7–32.15
	% Fe_2O_3	1.06–21.7
	% K_2O	0.5–5.25
	% MgO	0.39–9.3
	% MnO_2	0.02–1.74
	% Na_2O	0.005–4.68
	% P_4O_6	0.015–4.89
	% SiO_2	19.84–81.17

(p_{occur}) for each taxon that had p_{occur} > 0.5. This sum is the expected taxon richness if the site is in reference condition, and it is compared to the observed richness using the O/E ratio as described for AUSRIVAS.

OUTCOMES

The final results of these approaches were strikingly different (Table 6). Analyses based on BEAST (Reynoldson et al. 2014, Strachan and Reynoldson 2014) had the highest Type 1 error rates and consistently low power (Table 6). Other approaches did quite well (i.e., low Type 1 error rate and increasing power with greater degradation) with the ACT data set (relatively high taxon richness, proportional abundance only), but only Chessman (2014), Webb et al. (2014), and Feio et al. (2014) also had good results for the YT and GL data sets. Given the diversity among the data sets in geographic location and scale, nature of the ecosystems, environmental descriptors, biota, and the variation in approaches among the research groups, simple explanations and generalizations for the results are difficult and risky. The authors of each paper suggest why their approach might be useful and its constraints. One thing does seem clear from this special series: regional and national bioassessment programs, like AUSRIVAS, CABIN, and RIVPACS, should move to implementing the analytical methods developed over the last decade to achieve the most useful results for their stakeholders.

ACKNOWLEDGEMENTS

This group of papers was inspired by our collective work with our colleague and friend, the late Richard Norris, so we dedicate

Table 6. Proportion of validation (D0) and simulated impairment (D1, D2, D3) reference sites deemed to be not in reference condition by the assessment. Sites in D0 were not degraded so p(D0) represents Type 1 error (concluding a site is in reference condition when it is not). Other proportions (p[D1], p[D2], p[D3]) represent the power of the assessment to detect known impairments of increasing severity. ANNA = Assessment by Nearest Neighbour Analysis, RDA = Redundancy Analysis, D1 = mild, D2 = moderate, and D3 = severe impairment.

	ACT				Yukon				Great Lakes			
Author	p(D0)	p(D1)	p(D2)	p(D3)	p(D0)	p(D1)	p(D2)	p(D3)	p(D0)	p(D1)	p(D2)	p(D3)
Nichols et al.	0.00	0.10	0.80	1.00	0.25	0.25	0.90	0.88	0.38	0.42	0.68	0.88
Strachan and Reynoldson	0.75	0.80	0.80	0.95	0.30	0.35	0.40	0.45	0.30	0.35	0.40	0.45
Reynoldson et al.	0.70	0.80	0.85	0.90	0.45	0.43	0.62	0.78	0.48	0.48	0.38	0.43
Sarrazin-Delay et al.												
ANNA	0.05	0.15	0.65	1.00	0.10	0.08	0.15	0.30	0.12	0.12	0.08	0.10
RDA	0.10	0.20	0.70	1.00	0.15	0.15	0.13	0.34	0.08	0.05	0.08	0.05
Chessman	0.25	0.30	0.80	1.00	0.17	0.18	0.60	0.73	0.07	0.13	0.30	0.53
Webb et al.	0.05	0.10	0.30	1.00	0.20	0.22	0.65	0.75	0.08	0.42	0.67	0.87
Feio et al.	0.15	0.35	0.80	1.00	0.12	0.25	0.68	0.78	0.15	0.55	0.63	0.70

it to him knowing that he would have enjoyed reviewing and critiquing it. We thank Susan Nichols, Evan Harrison, and the Environment and Sustainable Development Directorate, ACT Government, Canberra, Australia for the ACT (Murrumbidgee River) data set; John Bailey and the Yukon Government DFO and UWO for the Yukon data set; and Lee Grapentine and Environment Canada for the Great Lakes data set. Editor Pamela Silver was a great help in assembling this special issue and in keeping the guest editors moving.

LITERATURE CITED

Bailey, R. C. 2005. Yukon River Basin. Pages 775–801 *in* A. C. Benke and C. E. Cushing (editors). Rivers of North America. Elsevier Academic Press, Oxford, UK.

Bailey, R. C., R. H. Norris, and T. B. Reynoldson. 2004. Bioassessment of freshwater ecosystems using the Reference Condition Approach. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Bowman, M. F., and K. M. Somers. 2006. Evaluating a novel Test Site Analysis (TSA) bioassessment approach. Journal of the North American Benthological Society 25:712–727.

Chessman, B. 2003. SIGNAL 2—a scoring system for macroinvertebrates ('Water Bugs') in Australian rivers. Technical Report No. 31. Monitoring River Health Initiative, Commonwealth of Australia, Canberra, Australia.

Chessman, B. 2014. Predicting reference assemblages for freshwater bioassessment with limiting environmental difference analysis. Freshwater Science 33:1261–1271.

Chessman, B. C., M. Muschal, and M. J. Royal. 2008. Comparing apples with apples: use of limiting environmental differences to match reference and stressor-exposure sites for bioassessment of streams. River Research and Applications 24:103–117.

Clark, J. S. 2005. Why environmental scientists are becoming Bayesians. Ecology Letters 8:2–14.

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

Feio, M. J., C. Viana-Ferreira, and C. Costa. 2014. Testing a multiple machine learning tool (HYDRA) for the bioassessment of fresh waters. Freshwater Science 33:1286–1296.

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.

Legendre, P., and L. Legendre. 1998. Numerical ecology. 2nd English edition. Elsevier, Amsterdam, The Netherlands.

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.

Nichols, S. J., T. B. Reynoldson, and E. T. Harrison. 2014. Evaluating AUSRIVAS predictive model performance for detecting simulated eutrophication effects on invertebrate assemblages. Freshwater Science 33:1212–1225.

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)

Parsons, M., M. C. Thoms, and R. H. Norris. 2003. Scales of macroinvertebrate distribution in relation to the hierarchical organization of river systems. Journal of the North American Benthological Society 22:105–122.

Reynoldson, T. B., and K. E. Day. 1998. Biological guidelines for the assessment of sediment quality in the Laurentian Great Lakes. NWRI Report No. 98-232. National Water Research Institute, Environment Canada, Burlington, Ontario, Canada.

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.

Reynoldson, T. B., C. Logan, T. Pascoe, S. Thompson, S. Sylvestre, C. Mackinlay, and H. McDermott. 2007. CABIN field manual for streams. Environment Canada. (Available from: http://www.ec.gc.ca/Publications/C183563B-CF3E-42E3-9A9E-F7CC856219E1/CABINFieldManual_EN_2012.pdf)

Reynoldson, T. B., S. A. Strachan, and J. L. Bailey. 2014. A tiered method for discriminant function analysis models for the Reference Condition Approach: model performance and assessment. Freshwater Science 33:1238–1248.

Sarrazin-Delay, C., K. M. Somers, and J. L. Bailey. 2014. Using Test Site Analysis and two nearest neighbor models, ANNA and RDA, to assess benthic communities with simulated impacts. Freshwater Science 33:1249–1260.

Strachan, S., and T. B. Reynoldson. 2014. Performance of the standard CABIN method: comparison of BEAST models and error rates to detect simulated degradation from multiple data sets. Freshwater Science 33:1225–1238.

USEPA and Canada (US Environmental Protection Agency and Government of Canada). 1995. The Great Lakes: an environmental atlas and resource book. 3rd edition. EPA 905-B-95-001. Great Lakes National Program Office, US Environmental Protection Agency, Washington, DC.

Webb, J. A., E. L. King, T. B. Reynoldson, and M. Padgham. 2014. Bayesian reference-condition models achieve comparable or superior performance to existing standard techniques. Freshwater Science 33:1272–1285.

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.