Interpreting the outputs from AUSRIVAS

Published by Land and Water Resources Research and Development Corporation

GPO Box 2182
Canberra ACT 2601
Telephone: (02) 6257 3379 and facsimile (02) 6257 3420
E-mail: public@lwrrdc.gov.au
WebSite: www.lwrrdc.gov.au

Publication data 'Interpreting the outputs from AUSRIVAS' by Dr Leon A. Barmuta, Dr Bruce C. Chessman and Professor Barry T. Hart. Occasional Paper 02/98.

Authors

Dr Leon A. Barmuta
Department of Zoology and the Centre for Biodiversity and Evolutionary Biology
University of Tasmania
GPO Box 252-05
Hobart, Tasmania 7001
Telephone: (03) 6226 2785
Facsimile: (03) 6226 2745
Email: Leon.Barmuta@utas.edu.au

Dr Bruce C. Chessman
NSW Department of Land and Water Conservation
PO Box 3720
Parramatta, NSW 2124
Telephone: (02) 9895 7154
Facsimile: (02) 9895 7867
Email: bchessman@dlwc.nsw.gov.au

Professor Barry T. Hart
Water Studies Centre
CRC for Freshwater Ecology
Monash University
Caulfield East, Victoria 3145
Telephone: (03) 9903 2326
Facsimile: (03) 9571 3646
Email: Barry.T.Hart@sci.monash.edu.au

ISSN 1320-0992
ISBN 0 642 26715 4

Design by WhizzbangArt
Printed by CPN Printing

November 1998

contents

acknowledgments

In the early stages of this project we benefited from discussions and correspondence with Dr David Stockwell (Symbiotik Pty Ltd) and Dr Dan Faith (CSIRO DWE) about alternative analytical procedures, and Dr R.T. Clarke of the Institute of Freshwater Ecology, U.K. who generously shared his insights into the problems of estimating errors for biological indices. Dr R. Sweeting (National Rivers Authority) and Dr L. Ruse (Thames Water) gave 'top down' and 'bottom up' views respectively of the past and future uses of RIVPACS in the U.K. Our thanks to all these people.

Associate Professor Richard Norris (CRC for Freshwater Ecology) and his team have collaborated closely with us during this project and we are especially indebted to Mr Justen Simpson for assisting with data analyses and his input into discussions about banding, and to Mr Paul Blackman for his programming expertise. This project would not have been completed without their dedicated hard work in analysing the agency data as it became available.

Four State and Territory agencies made their data and analyses available to us for our analyses. We would like to thank Dr Stuart Halse (CALM, WA), Mr. Leon Metzeling (EPA, Victoria), Mr David Oldmeadow and Dr David Fuller (DPIF, Tasmania) and Mr Peter Liston (Department of Environment, Land and Planning, ACT) for their prompt attention to our queries and for their comments on and discussion of our work.

Our own colleagues and technical and office staff were essential to this project, and we acknowledge the contributions of Dr Jane Growns (formerly AWT EnSight), Ms Pam Dickinson (Monash University), Mr Edwin Chester and Ms Sarah Tilley (University of Tasmania).

We are grateful to the personnel of all the State and Territory agencies who helped to organise and participated in the series of regional workshops in 1996. Their insights and suggestions, together with the contributions by others from universities, community groups and the water industry, have helped make this project successful.

summary

The Australian River Assessment Scheme (AUSRIVAS) is a series of procedures and associated software for the rapid assessment of river conditions or 'health' using macroinvertebrate communities. AUSRIVAS includes a set of predictive computer models relevant to particular Australian States and Territories, seasons (spring, autumn or spring and autumn combined) and habitats for river macroinvertebrates (eg. riffles and pool edges). The models attempt to predict the macroinvertebrate families in a specified subsample from a standard kick or sweep net sample, assuming natural or near-natural conditions at the sampling site. The predictions are based on physical, chemical and vegetative features of the site. The predicted fauna can then be compared with the fauna actually observed to infer departures of the macroinvertebrate community from its natural state.

In this project we developed methods for interpretation of the raw output from the AUSRIVAS models, which consists of the probabilities of occurrence of the various families of macroinvertebrates. We used a series of structured workshops with river scientists, managers and interested community representatives, and a telephone conference, to facilitate the development of outputs tailored to various end users and uses. The latter include environmental summary reporting (eg. State of the Environment), agency regulatory activities, catchment planning and management, community-based assessment, and education and training.

The interpretation is based on two ratio indices. One incorporates the number of families expected to occur in a standard subsample and the number of those actually observed (*O/E* FAMILIES). The other is based on the modelled and observed values of the SIGNAL biotic index (*O/E* SIGNAL); this ratio incorporates the tolerances of different macroinvertebrate families to common types of water pollution. Detailed guidelines are provided for combining values of these indices in each habitat, to generate bands reflecting varying degrees of departure of macroinvertebrate communities from model predictions. Band A represents conditions similar to those at most sites that are minimally disturbed by human activities (reference sites). Bands B, C and D comprise sites with progressive reductions in numbers of families, or in the average pollution sensitivity of the families. Band X includes sites with more families than predicted, which may be unexpectedly biodiverse either because of natural factors or mild nutrient enrichment as a consequence of human activities.

The present models, indices and banding scheme are seen as in need of considerable further testing and improvement. Specific areas for research and development include the following:

- More standardised rules for the selection of reference sites on which the models are based, including clear ecological and management-based criteria for the degree of human disturbance that is acceptable at a reference site (see Section 3.1).
- Quantification and, where possible, reduction of the sources of variability and statistical error in the derivation of both the predicted and the observed macroinvertebrate community composition at a site (see Section 3.4).
- A more rigorous approach to the selection and use of predictor variables in the models, especially those variables whose values at test sites can be affected by human activities (see Section 3.2).
- Further investigation of the merits of alternative statistical and numerical methods of prediction, in particular detailed comparisons of DDRAM and e-ball with the present method (see Section 3.5).
- Testing of the merit of combining predictions from different habitats to obtain an overall prediction for a site using weighted probabilities (see Section 3.3).
- Further development of SIGNAL or other approaches using tolerance information to provide diagnosis of causes of impairment of macroinvertebrate communities at test sites (see Section 3.7).

- The inclusion and reporting of non-predictor environmental variables for diagnostic purposes (see Section 3.8).
- Investigation of alternative banding schemes including schemes based on defined effect sizes (degrees of family loss) that jeopardise attainment of policy objectives or can be shown to be detrimental to ecosystem functioning (see Section 3.4).
- A close, formal liaison between agencies responsible for environmental audits to ensure integration of AUSRIVAS into State of the Environment reporting procedures in a scientifically valid manner (see Section 3.9).

how to use this document

for the end-user

This document provides for end users of AUSRIVAS a guide to the outputs from the present modelling procedure employed for the rapid bioassessment protocol for macroinvertebrates.

The flow chart on the following page provides different options for the sequence that could be followed when reading this report when depending on the primary interest of the reader. This flow chart provides an overview only. Some of the sections of this document refer to more detailed or technical information in the appendices.

guide to the achievement criteria

This milestone report has six achievement criteria. These are itemised below in italics and are cross-referenced to the relevant sections of this report.

a) Protocol for the interpretation of bioassessment outputs from the NRHP using metrics.

This is in Section 1.

b) Documentation of the indices and rules used in NRHP bioassessment outputs, and inclusion of this material in the final software package resulting from the statistical support project led by Associate Professor Richard Norris.

The indices and rules are fully documented in Section 2. Supporting material for the software has been included in the documentation from that project, and is supplied in Appendix A.

c) Results of the review of the protocol and delivery system by an expert panel.

Final proposals for outputs and their preferred form were circulated to all lead agencies in February 1997, and a telephone conference was held in early March to review these proposals. This is summarised in Section 4.2.

d) A summary of communication and integration activities, including the results of the regional workshops.

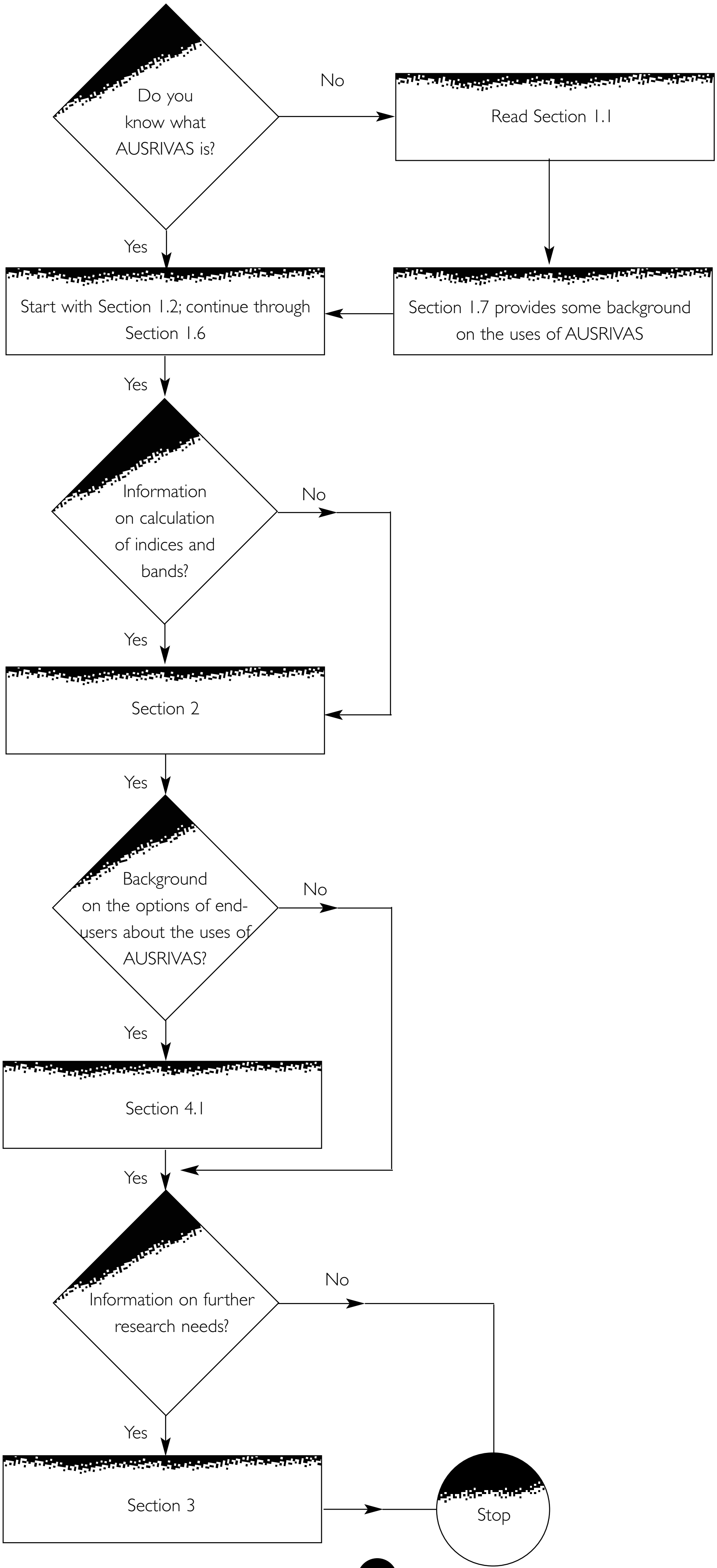
See Section 4.

e) Listing of NRHP lead agencies to which two copies of the final milestone report should be sent.

See Section 5.

f) One copy of (d) as an attachment to the final report.

This is in Section 4.


Do you know what AUSRIVAS is?
No
Read Section 1.1
Yes
Start with Section 1.2; continue through Section 1.6
Section 1.7 provides some background on the uses of AUSRIVAS
Yes
Information on calculation of indices and bands?
No
Yes
Section 2
Yes
Background on the options of end-users about the uses of AUSRIVAS?
No
Yes
Section 4.1
Yes
Information on further research needs?
No
Yes
Section 3
Stop

1. **procedure** for the interpretation of AUSRIVAS outputs

1.1 **overview** and background

AUSRIVAS stands for the **Aus**tralian **Riv**er **A**ssessment **S**cheme and is being developed as part of the Monitoring River Health Initiative (MRHI) of the National River Health Program (NRHP). AUSRIVAS is designed to be a rapid biological assessment tool that ultimately may use different components of the in-stream biota to assess the 'ecological health' of segments of rivers and streams (Schofield & Davies 1996). Most of the research for the MRHI has concentrated on using macroinvertebrates, although pilot projects using fish, benthic algae and stream metabolism are being carried out. This report focuses entirely on interpretation of the outputs from the macroinvertebrate program of AUSRIVAS.

Assessing the ecological health of a river site at first seems intuitively simple. However, like the term 'water quality', this can mean different things to different people. Schofield and Davies (1996) interpret 'health' as the similarity of a river to its natural or pristine state. For most Australian rivers this natural state is difficult to determine because no pristine examples remain. Aboriginal firing regimes altered the vegetation across large tracts of Australia, and in some regions this probably affected runoff patterns, river hydrology and water quality. European and Asian settlement has also affected runoff and hydrology, first through land clearance and agriculture, and later with dams, flow diversions and irrigations schemes. In addition rivers have received pollutants, either from point-sources such as sewage treatment plants, or diffuse sources such as agricultural runoff, and instream habitats have been altered by channel works and changes to the flow regime of the river. Therefore, to discuss ecological health meaningfully, we need to clarify some terms.

In this document we define a **human disturbance** as a human intervention with the aquatic environment. A human disturbance may have a measurable effect on the biota, which we term an impact; not all human disturbances result in a measurable impact. Impacts include such things as changes to the diversity of species or to the populations sizes of plants and animals. Not all impacts will be detrimental. Those that are detrimental result in **impairment** of the biota of the river. Impairment depends on the social context within which an impact is judged detrimental; thus, value judgements play an important role in setting that context. Detection of impact requires that a comparison be made with conditions where the human disturbance of concern is absent. This is similar to the notion of a 'control' in designed experiments. A site on a river that fulfils this criterion is called a **reference site** because it provides a benchmark against which other sites can be compared. Note that in this framework a reference site is not necessarily a pristine site, completely undisturbed by humans. In many parts of Australia, reference sites have had to be chosen to represent the least disturbed condition within a region. Sites that we are seeking to assess are termed test sites. A hypothetical example that explains these issues further is given in Box 1 overleaf.

Box 1. A hypothetical example explaining the terms **human disturbance, impact, impairment, reference site** and **test site.** These terms are used in explaining and reporting the outputs of AUSRIVAS.

Suppose we have been set the task of assessing a small lodge development that discharges highly treated wastewater into an otherwise pristine stream. Suppose further there are several other, similar streams close by which have no human intervention; these will serve as appropriate **reference sites** with which we can compare the **test sites** downstream of the discharge. These streams are naturally poor in nutrients, have low productivity and a low diversity of aquatic life as a result. The **human disturbance** in this example is the discharge from the lodge.

Although the waste is highly treated, suppose the effluent still raises slightly the nutrient concentrations in the receiving water. Whether that increase in nutrients results in an **impact** on the aquatic biota is the task set for the investigating biologist to detect. Suppose the biologist finds that the diversity of plants and animals in the stream has increased downstream of the discharge. This might be because the elevated nutrients have increased productivity and permitted some more species to establish, while the nutrients have not increased to levels that have eliminated the resident species. Whether this impact is judged to be an **impairment** depends on the values encoded in the management goals for this area. If the area is being managed as a wilderness area then this impact is an impairment. By contrast, if the management goal was to increase the population size and range of an endangered species, then the increased productivity from the discharge might be regarded as beneficial instead. Impairment includes value judgements about the size and nature of the impact. The goal of bioassessment is to quantify impacts.

The task of quantifying impacts is a comparative exercise, where impact is judged relative to some benchmark or target. Rapid bioassessment seeks to do this cheaply and quickly (Resh & Jackson 1993), and there are two major types of procedure that have been implemented over a wide area for invertebrate benthos. The United States Environmental Protection Agency (US EPA) has used a battery of Rapid Biological Assessment Protocols (Plafkin *et al.* 1989), and different regions can choose components from these protocols that best suit their local conditions. This approach requires each region to 'calibrate' the protocols as outlined in Box 2. By contrast, the British RIVPACS (**R**iver **I**n**v**ertebrate **P**rediction **a**nd **C**lassification **S**cheme) (Wright *et al.* 1984; Wright *et al.* 1988; Wright, Armitage & Furse 1989) attempts to represent benchmark conditions across the entire country so that a model predicting faunal composition can be built using environmental variables that are unlikely to be affected by human disturbance. The same environmental variables are measured for sites that are to be assessed and the model is used to predict a 'target' faunal composition of the sites. The fauna actually occurring at the assessment sites is then sampled, recorded and compared with the predictions of the model. If the observed fauna differs from that predicted by the model, then the site is judged to be impacted. In practice, three indices are calculated to compare the differences between the observed and predicted fauna and a set of rules is used to combine these indices into a final assessment (Wright *et al.* 1991; National Rivers Authority 1994).

AUSRIVAS follows the RIVPACS procedure, although there are some important differences. The chief difference is that, wherever possible, two habitats are sampled and separate models are constructed for each, whereas for RIVPACS, only one sample is collected from a site in an attempt to integrate across the different habitat types present (Furse *et al.* 1981). (Parsons & Norris 1996) explain why separate habitats need separate models. In most States the habitats are pool edges (mostly slow-water, depositional environments) and riffles (fast-water environments). In some States the habitats are macrophyte (where vascular plants dominate) and channel (open water).

Box 2. An outline of the Rapid Bioassessment Protocols used by the US EPA.

The United States Environmental Protection Agency (US EPA) advocates a battery of Rapid Biological Assessment Protocols (Plafkin *et al.* 1989), many of which have been developed from earlier quantitative and qualitative methods employed at smaller spatial scales in Europe and North America. The US EPA procedures rely on dividing a region into biomes and then identifying sites of high quality within the biomes to act as reference sites against which test sites can be compared. A series of test sites with a range of known disturbances is also selected, and these test and reference sites are then sampled contemporaneously. Many different indices (called 'metrics') are computed from the resulting data, and those indices that respond to the known disturbances are used in subsequent assessments. A set of rules is then used to synthesise these different indices into a final assessment (Plafkin *et al.* 1989), and there is provision for combining assessments made on different components of the biota (eg. fish, macroinvertebrates, periphyton) to derive a summary grading for a given test site (US Environmental Protection Agency 1996). Some of these indices may be redundant, but biomes differ in terms of which indices are redundant, and which are most sensitive to disturbance (Plafkin *et al.* 1989; Barbour *et al.* 1992; Kerans & Karr 1994; Barbour *et al.* 1995). Although the synthesis of these indices is potentially complex, a potential strength of this approach is the high level of flexibility, allowing the investigators to choose the best indices for a particular set of environmental conditions (US Environmental Protection Agency 1996).

In summary, AUSRIVAS is a rapid procedure to quantify impact on the in-stream biota. At present, this is achieved by predicting the occurrence of families of macroinvertebrates at test sites from environmental variables and a large database of high-quality reference sites. The raw output from this procedure is a list of the families of invertebrates expected in a standard sample from the site, the probability of occurrence of each family in that sample and a tally of which of those families did occur in an actual sample. This project examined the best procedures for reporting this information for a wide range of end-users.

1.2 **outline** of the outputs

The remaining parts of Section 1 describe the recommended procedure for interpreting and reporting the outputs for AUSRIVAS.

Optimally, when a test site is assessed, samples are taken from two designated habitats in two designated seasons. The habitats used and the seasons sampled depend on the region for which the model has been developed. The data are then compared with the model developed for combined seasons for each habitat. Thus there are two model outputs, one for each habitat. For each of these outputs (one for each habitat), two summary indices are computed (Section 1.3), and the site can be assigned to a category of impact which compares it with the expectation of the reference conditions called a band (Section 1.4). Comparing the allocation of the test site to a band on each of the indices provides some diagnostic information and some possible interpretations about the nature of impairment at the test site (Section 1.5). Rules for combining the bands for each of the habitats for summary reporting are described in Section 1.6, while Section 1.7 outlines the different uses (and hence reporting requirements) for AUSRIVAS that were identified by potential end-users during a series of regional workshops conducted during the development of AUSRIVAS in 1996. These workshops are summarised in Section 4.

If only a single habitat can be sampled or only one sampling occasion can be used, the reporting procedure is modified as described in Section 1.6.

In addition, the AUSRIVAS computer package provides detailed output from which the indices were calculated. This output can be consulted by experienced biological personnel in concert with the program's documentation to provide additional diagnostic information for narrative background for report preparation. An output sheet for a chosen test site has also been developed to assist agency biologists in interpreting the output. This ouput will be included in the AUSRIVAS software, and an example of the output is given in Appendix E.

1.3 **the** indices

Two complementary indices are used to summarise outputs from the analysis of macroinvertebrate survey data using the AUSRIVAS predictive models. These are:

1. *O/E* FAMILIES, which is a ratio relating the number of families of macroinvertebrates recorded in a sample to the number of families expected in that sample according to the predictions of the model for least-disturbed conditions; and
2. *O/E* SIGNAL which is the ratio of the observed SIGNAL value for a sample to the expected SIGNAL value. SIGNAL (Stream Invertebrate Grade Number Average Level) is a biotic index that uses a grade assigned to each family according to its sensitivity to pollution (Chessman 1995). A grade of 10 represents a high sensitivity to pollution, while a grade of 1 represents a high tolerance of pollution. The sum of the grades of all families present in a sample is divided by the number of families to give an average sensitivity (SIGNAL value) for the sample.

The rationale for choosing these indices is given in Section 2.1, which includes a worked example.

Note that the 'O' (observed) portion of these two ratios refers only to families that were expected to occur in the sample and were recorded in the sample. Other families that were observed in the sample but were not expected are not included in these indices. Such 'unexpected families' may provide additional diagnostic information for interpretation by expert biologists, but there is no simple method of incorporating these families into an index.[BC1] Unexpected families occurring in a sample are listed in the raw outputs of AUSRIVAS.

Both indices have a minimum value of 0 if none of the families expected in a sample is found. Any value less than 1 indicates either that some of the expected families were not collected (*O/E* FAMILIES), or that the suite of families collected is, on average, more tolerant of pollution than expected (*O/E* SIGNAL). In either case, impairment of the macroinvertebrate community could reasonably be inferred. However, in practice a value less than 1 could result from chance variations in the families captured, or inevitable statistical uncertainty in the model predictions. Therefore, impairment is inferred only if values are sufficiently lower than 1 for these alternative explanations to be highly unlikely.

Index values of 1 indicate a perfect match between the families expected and those found (*O/E* FAMILIES) or that the families collected are exactly as pollution-sensitive, on average, as those expected (*O/E* SIGNAL). A value exceeding 1 indicates that more families were found than were expected (*O/E* FAMILIES), or that those found were on average more sensitive than expected (*O/E* SIGNAL). There are three potential explanations for index values greater than 1. Firstly, chance factors may result in more families being collected than expected. Differences between operators in macroinvertebrate sampling efficiency fall into this category. Secondly, the site may have unusual microhabitat or other ecological features that permit the coexistence of more families than expected. Thirdly, the site may be disturbed by human activities in such a way as to allow additional families to establish. An example of such a disturbance is a mild increase in nutrient levels, which can sometimes result in macroinvertebrate communities that are more abundant or diverse than those that would occur naturally. As previously discussed, deciding whether this type of impact should be considered as impairment involves a value judgment.

1.4 **the** banding scheme

For reporting purposes, the values of each index are divided into categories or bands. The width of the bands is based on the distribution of index values for the reference sites in a particular model. The width of the 'reference' band, labelled 'A' in Table 1 below, is centred on the value 1 and includes the central 80% of the reference sites. A test site whose index value exceeds the upper bound of these values (ie. the index value is greater than the 90th percentile of the reference sites) is judged to be richer than the reference condition and is allocated to band X. A site whose index value falls below the lower bound (ie. the index value is smaller than the 10th percentile of the reference sites) is judged to have a deficit of expected families and/or a lower SIGNAL value than expected and is

Table 1 Division of the indices into bands or categories for reporting. The names of the bands refer to the relationship of the index value at a test site to the reference condition (band A). Under comments, for each index, an explanation of the band is stated first, followed by possible interpretations as dot-points.

BAND LABEL	BAND NAME	COMMENTS O/E FAMILIES	COMMENTS O/E SIGNAL
X	Richer than reference	Very high occurrence of expected families • Biodiversity 'hot-spot'; or • Mild nutrient enrichment	Appreciably greater SIGNAL value than expected. • Site of exceptionally high water quality; or • Impact unrelated to organic/nutrient/sediment pollution
A	Similar to reference	Families found similar to those expected • Water and habitat quality roughly equivalent to those at reference sites; or • Impact on water or habitat quality of a type that does not result in loss of families	SIGNAL value similar to expected value • Water quality roughly equivalent to that at reference sites; or • Impact on water quality of a type that does not result in loss of families sensitive to organic/nutrient/sediment pollution
B	Poorer than reference	Several expected families not found • Impairment of either water quality or habitat quality or both	Appreciably lower SIGNAL value than expected • Impairment of water quality
C	Much poorer than reference	Many expected families not found • Substantial impairment of water and/or habitat quality	Much lower SIGNAL value than expected • Substantial impairment of water quality
D	Far poorer than reference	Very few of the expected families found • Severe impairment of water and/or habitat quality	Very low SIGNAL value • Severe impairment of water quality

allocated to one of the lower bands (B, C or D) according to the *O/E* value. The widths of bands B and C are the same as for band A, the reference band. Band D, which extends down to zero, may be wider or narrower than these bands depending on the variability in the index values of the reference sites in the model. In most cases, sites falling in band D on either index will be severely deficient in families expected at the site.

In many cases the values of both *O/E* indices will allocate a site to the same band. Occasionally, a site may be allocated to one band on the basis of *O/E* FAMILIES and to another band (either higher or lower) based on the value of *O/E* SIGNAL. These mis-matches can provide valuable diagnostic information, as illustrated in the next section.

1.5 **use** of AUSRIVAS indices for diagnosis

The different, but complementary emphases of *O/E* FAMILIES and *O/E* SIGNAL can provide some diagnostic information about the sample collected from a test site. Figure 1 below explains how the index values might arise. For example, a value of *O/E* SIGNAL that is richer than reference, shows that the SIGNAL value is greater than expected. If the sample also has a value of *O/E* FAMILIES which is poorer than reference (upper left corner of Figure 1), then there are two ways that this could have arisen; either nearly all of the sensitive families were found in the sample or fewer of the tolerant families that were expected were found or some combination of these two explanations.

Figure 2 outlines the possible causes for each different combination of index values relative to the reference band. To continue with the previous example, where a sample is allocated to band X based on *O/E* SIGNAL and band B (or lower) on *O/E* FAMILIES, a potential cause is an impact on the invertebrate fauna related to causes other than a disturbance that changes water quality. This interpretation arises because the families sensitive to water pollution (as indicated by SIGNAL) have been found in the sample while there has been an overall reduction in the total number of families expected.

The possible causes summarised in Figure 2 are only potential explanations. Further diagnostic information needs to be sought from environmental data collected at the same time, and expert biologists should consult the summary lists of families expected and observed for the sample to interpret these data more fully. Also the steps outlined in the next section need to be followed to ensure that an assessment of the site is justified based on the data available.

1.6 **use** of AUSRIVAS for reporting

Assessment of a test site will often be based on data collected from two habitats in two seasons. In this case there will be eight index values to be synthesised for final reporting. The following procedure is to be used for reporting such data.

1.6.1 **choosing** between single-season and combined models

The procedure first requires a choice of the most appropriate seasonal model for computing the indices. Usually, data from both seasons for a particular habitat will be used in the combined seasons model for that habitat. If resources or time only allow one sampling occasion, then the seasonally more appropriate model will be used.

Figure 1. Meaning of situations where the two indices place sites in different bands. *O/E* SIGNAL is the observed v. expected score of a site based on the SIGNAL biotic index; *O/E* FAMILIES is the observed v. expected number of families for a site. The vertical and horizontal dashed lines indicate the upper and lower bounds for band A, which includes 80% of the reference (= 'low disturbance') sites for the relevant model.

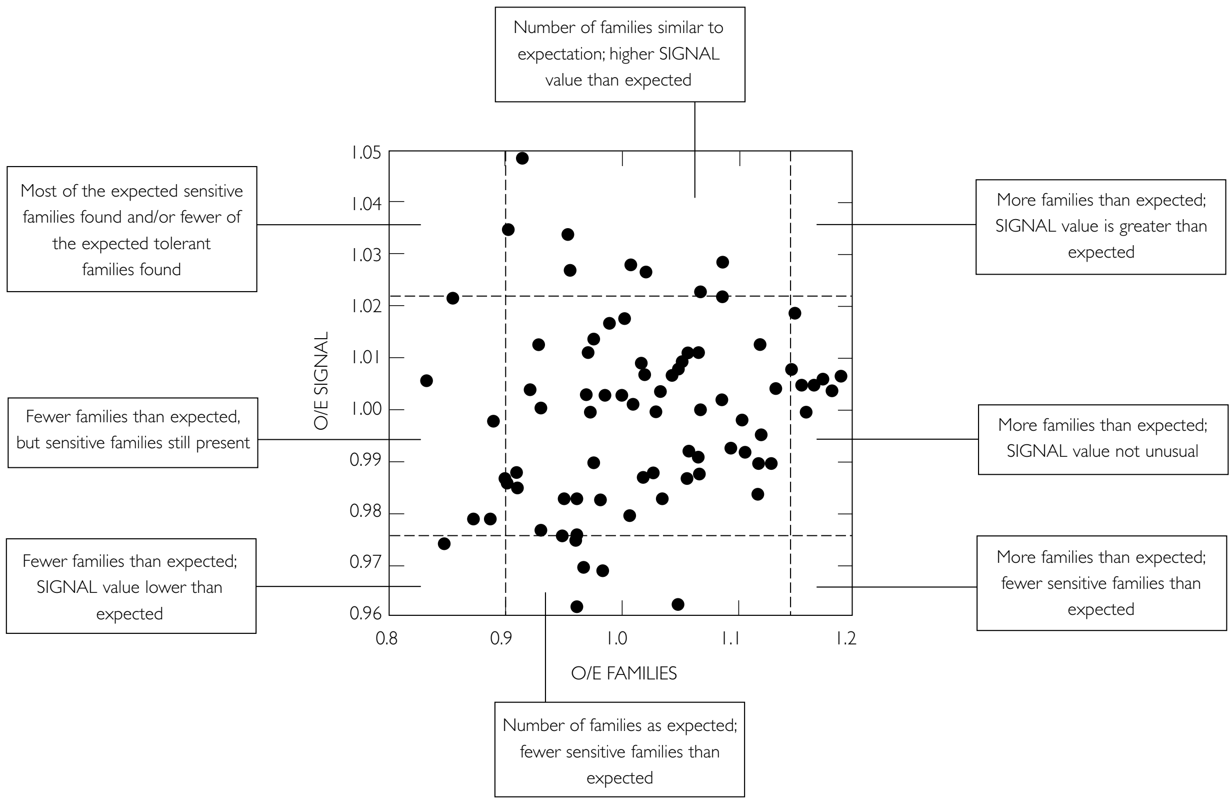
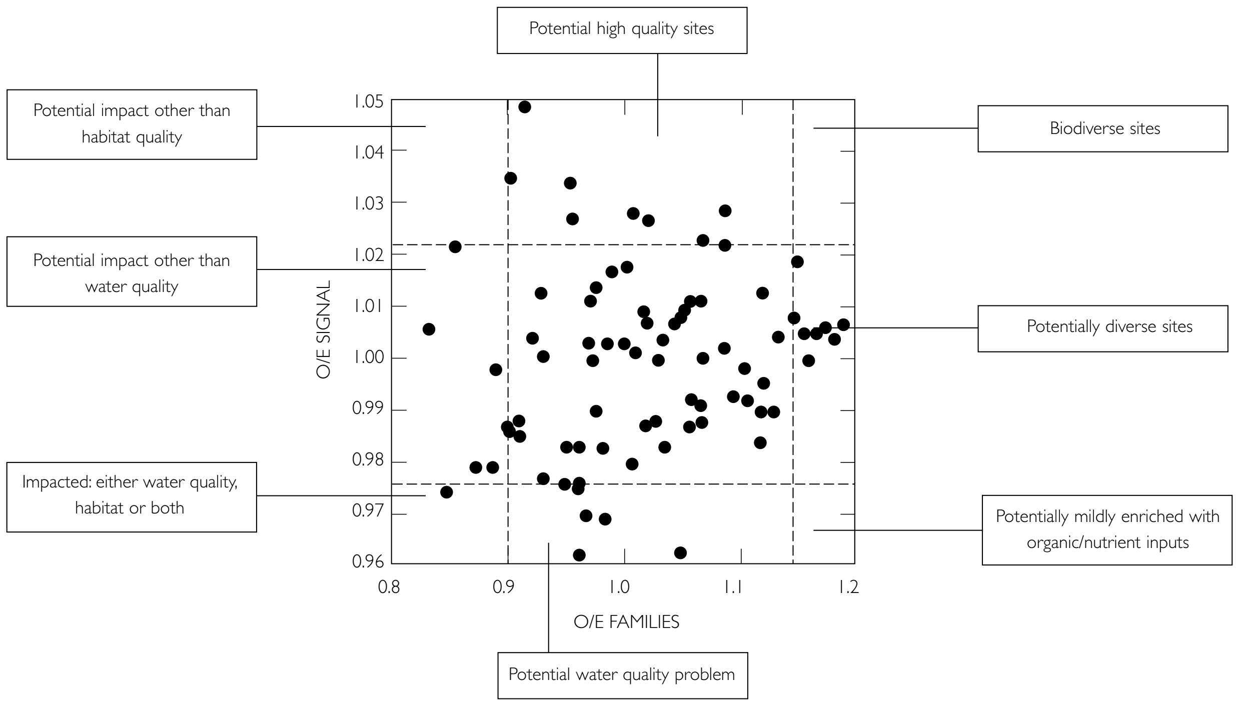


Figure 2. Possible causes for placement of sites in bands for the two indices. *O/E* SIGNAL is the observed v. expected score of a site based on the SIGNAL biotic index; *O/E* Families is the observed v. expected number of families for a site. The vertical and horizontal dashed lines indicate the upper and lower bounds for band A, which contains 80% of the reference (= 'low disturbance') sites for the relevant model.



In some circumstances, it may be of interest whether the status of a site has changed between sampling occasions. For example, a human disturbance such as a waste spillage or river restoration may have occurred between the sampling occasions. In that case the appropriate single-season model will be applied to each sampling occasion.

1.6.2 **reconciling** different band allocations from the same habitat

If the two indices allocate a site to different bands the following steps must be followed.

1. Check that the data were entered correctly. Check both environmental and biotic information. If any typographic errors are found, correct them and re-run the analysis.
2. Determine whether any sampling problems could have affected the results. Review the field data sheets and confer with those who conducted the sampling and identification, and review any factors that might have affected the quality of the data. For example, poor light, severe weather conditions or accidents while sampling or sorting can bias the families that are found or identified. If this process reveals problems with the data collection other than typographic, data-entry errors, then proceed to Step 5. Otherwise, proceed to Step 3.
3. Determine whether there are any natural factors that might have differentially affected one index compared with the other. The following circumstances may influence index values and allocation to a band.
   a. Low overall abundance because of natural, but unusual habitat conditions such as domination of the substratum by bedrock. This can lead to low values of *O/E* FAMILIES.
   b. The proximity of tributaries (even small tributaries) to the test site. It is possible that macroinvertebrates may be washed in from these tributaries and be recorded as part of the observed fauna for the site when, in reality, they belong to families that are unlikely to be resident at this site. This can lead to overestimation of either index. If contamination by fauna from tributaries is suspected, then proceed to step 4.

c. Recent, unusual, natural flow events (spates or low flows) that might affect invertebrate community composition or abundance. Typically, high flows reduce the numbers of invertebrates encountered, and so may reduce *O/E* FAMILIES to a greater extent than *O/E* SIGNAL, although very high flows that disturb most of the river bed will probably result in low values for both indices. Severe low flows can depress both indices. If natural flow events are likely to have depressed index values, then proceed to step 4. (Note that this does not apply to artificial high or low flows resulting from factors such as irrigation releases from dams or water abstraction respectively which are part of a human disturbance that is being assessed.)

4. If steps 1–3 indicate no problems with the original data, then allocate to the band that is farther from band A except where the more extreme band has obviously resulted from natural factors. In this case, allocate the site to the band that is most indicative of the level of human impact. In the unlikely event that the alternative bands are band B and band X, allocate to band B, because this is the more precautionary approach.
5. If the preceding steps indicate problematic data, then there should be **no *post hoc* alteration or 'correction' of the faunal or environmental data beyond typographic, data-entry errors.** The integrity of the data is paramount, and alterations such as deletions of families whose presence is 'explained away' cannot be tolerated.

   The options for further action in order of preference are as follows.

   a. Re-sampling and re-assessment. This is the obvious choice if time and resources permit re-sampling. If this is not possible, then
   b. Draw a conclusion of 'no reliable assessment possible'. This is the most conservative approach. Diagnostic information can still be presented, explained and qualified, but no allocation should be made to a band. The reasons that no reliable assessment could be made should be made explicit.

1.6.3 **combining** assessments from different habitats

In some circumstances only one habitat will be assessed, either because other habitats are not present or because the investigation is targeted at only one habitat (such as the habitat deemed most susceptible to the disturbance of concern, for example). In such cases, the fact that only one habitat has been employed in the assessment should be made explicit.

In many cases two bandings will be available for a given index and test site: one for each habitat. Where the bandings from both habitats allocate the site to the same band, then that is the final band allocation for the site. Where there is a mis-match in the band allocation from the two habitats, then allocate the site to the band that is farther from band A. In the rare event that the alternative bands are band B and band X, allocate to band B.

Allocation to Band X should result in further assessment to determine whether the site is richer than reference because of naturally high biodiversity or an impact such as mild nutrient enrichment.

1.6.4 **assessments** of multiple sites

Assessments of several sites simultaneously should use a consistent basis for comparison; mixing assessments based on different seasonal models or mixtures of single- and two-habitat data should be discouraged.

1.7 **different** uses of AUSRIVAS and different reporting requirements

During the regional workshops, it became clear that AUSRIVAS is going to be used for a number of different purposes, and that these purposes potentially had differing reporting requirements (See Section 4.1). Although different emphases could be used for the various requirements, subsequent consultations and the nature of the data collected for AUSRIVAS has constrained and modified the consensus recommendations from the regional workshops.

From the workshop discussions, five broad categories of use for AUSRIVAS were identified. Each of these uses is outlined below, together with a commentary on the use of the AUSRIVAS outputs for each use.

1.7.1 **environmental** reporting

This use encompasses State of the Environment (SoE) reporting, State of the Rivers reporting as used by some States and Territories and Environmental Audits. The goal of this use is to provide a broad-scale assessment of river 'health' for different regions within a State or Territory. Most environmental reporting procedures also imply that the process will be repeated at regular intervals in the future, suggesting that documenting trends in, as well as spatial patterns of, environmental health are also important.

Of all the uses for AUSRIVAS, environmental reporting is the category which requires the most consultation to determine the final output components which best suit the goals of summary reporting. Some options are discussed further in Section 3.9, but the procedure stipulated in Section 1.2 to 1.6 suffices for the First National Assessment of River Health (FNARH) that commenced in 1997.

1.7.2 **regulatory** uses

This includes monitoring for compliance with licence conditions, water quality guidelines, and the forms of environmental impact assessment where rapid biological assessment is sensitive enough to be used.

For these uses, the procedure stipulated in Section 1.2 to 1.6 suffices, although the lack of a method of quantifying the error or confidence in band allocation hampers the use of AUSRIVAS for subtle impacts for the time being. However, gross departures from expectation (eg. a decline of >1 band) are probably detected by AUSRIVAS reliably, and this use of AUSRIVAS is appropriate, especially in conjunction with other (eg. physicochemical) information when assessing either compliance or impact.

Access to the raw outputs from the AUSRIVAS software is an essential adjunct for agency biologists who are preparing reports or submissions for regulatory applications of AUSRIVAS. Detailed expertise about the local or regional behaviour of the expected or observed families can be bought to bear when preparing background material. The prospect of further development of automated procedures for documenting this knowledge is discussed in Section 3.7.

1.7.3 **catchment** planning and management

An assessment of the current status of a river as part of the planning or catchment management process is a clear example of where rapid biological assessment can be a cost-effective tool for managers. As such the procedure stipulated in Section 1.2 to 1.6 provides an adequate level of summary reporting for these purposes.

In cases where ecosystem protection is a key environmental value, targets or goals can be established in terms of rapid biological assessment *O/E* values. For example, in the most recent State Environment Protection Policy for the Yarra catchment, the Victorian EPA has established objectives for different river segments in terms of numbers of key macroinvertebrate families present and SIGNAL index score (Victorian Environmental Protection Authority 1995).

1.7.4 **community-based** monitoring

The involvement of community groups in baseline monitoring, monitoring the performance of restoration or rehabilitation works and in catchment management and planning requires the summary reporting procedure stipulated in Section 1.2 to 1.6 as a minimum reporting requirement. Setting aside the issues associated with training for community representatives if they are to be accredited to collect samples and identify invertebrates, there is a clear need for agency biologists to provide some narrative background about the fauna for those members of the group who are interested to develop their understanding of AUSRIVAS further.

Once confidence intervals for the index values can be calculated, reporting of actual index values rather than just band allocations may be useful, especially if the activity being monitored is likely to take some time to lead to large improvements in the fauna. Smaller changes in the fauna would be documented by reporting index values so that slight improvements that might provide valuable encouragement for continued community effort are documented. Reporting confidence limits of index values are, however, essential to prevent chance fluctuations being mis-interpreted as a trend.

An alternative suggestion of sub-dividing the bands (eg. B+, B–) is not recommended because increasing the number of bands increases the probability of mis-banding resulting from inherent (and as yet unquantified) natural variation in index values.

1.7.5 **education** and training

The educational uses of AUSRIVAS remain to be explored fully. The full range of outputs need to be available for training agency staff and others who will either contribute data to AUSRIVAS or use it for their own assessments, monitoring and reporting. For users with a more general educational interest, access not only to the on-line documentation with AUSRIVAS but also to the supporting research and published papers is also essential. Whether these supporting documents are made available on-line as well as through conventional sources such as libraries should be contingent on the demand for this information being demonstrated.

2. **rationale** for the outputs

2.1 **the** indices

The two indices employed in the outputs are *O/E* FAMILIES and *O/E* SIGNAL. (Their equivalents in the British RIVPACS system are *O/E* and *O/E* ASPT respectively.) Small values indicate impact, while values >1 may indicate exceptional sites worthy of special conservation management (Moss *et al.* 1987) or sites that may have more families than expected because of mild nutrient enrichment. *O/E* FAMILIES is easily understood as an index: it is calculated as the ratio of the number of expected families (those with a modelled probability of occurrence greater than 0.5) that are actually observed in a sample divided by the sum of the probabilities of occurrence of the expected families. The diagnostic or interpretive value of the output can be increased by incorporating a measure of the sensitivity to pollution of the families involved. In Australia, Chessman (1995) and Chessman *et al.* (1997) have developed SIGNAL, a biotic index that assigns a grade number between 1 and 10 to each family depending on its sensitivity to pollution (see Appendix F). *O/E* SIGNAL (the ratio of observed and expected SIGNAL values) has the additional advantage that unlike *O/E* FAMILIES it is robust to inter-operator variations in sampling technique and effort.

These two indices were selected for several reasons. In tests on trial data indices based on *O/E* ratios performed better on most criteria used to select indices than did several alternative types of index (Barmuta, Chessman & Hart 1996). In addition, Stockwell and Faith (1996) demonstrated that the accuracy of this group of indices was not sensitive to decreases in the predictive accuracy of the model. They attributed this feature to the division of the observed tally of families by the sum of the weighted probabilities (Stockwell & Faith 1996, p.11).

Barmuta *et al.* (1996) also suggested that another index, *O/E* SIGNS (Stream Invertebrate Grade Number Sum), was suitable as well. This index is equivalent to the *O/E* BMWP of RIVPACS, and is the ratio of the sum of the SIGNAL grades of the expected families that were observed divided by the sum of the expected SIGNAL grades of each family multiplied by the weighted probability of its occurrence. *O/E* SIGNAL differs from *O/E* SIGNS in that the observed and expected sums are divided by the number of families involved in the summation where the sums of the values have not been divided by the number of families involved. (Effectively *O/E* SIGNS = *O/E* FAMILIES x *O/E* SIGNAL: Clarke *et al.* 1996). As Clarke *et al.* (1996) found for *O/E* BMWP, *O/E* SIGNS is redundant because it is highly correlated with *O/E* FAMILIES (r>0.9).

Furthermore, Barmuta *et al.* (1996) qualified their support for the use of *O/E* SIGNAL because it seemed less sensitive than the other *O/E* indices. However, this qualification was based on the very small and limited sets of data available for testing and the preliminary nature of SIGNAL grades at that time. Furthermore, their assessment of *O/E* SIGNAL did not take into account the lower variability of this index amongst reference sites compared with *O/E* FAMILIES. Given the present information available, using these two indices together in assimilating the outputs of AUSRIVAS is justified because both perform well and are complementary in their emphases (see Section 2.1.2).

2.1.1 **computation** of indices

O/E FAMILIES

Within the AUSRIVAS program, *O/E* FAMILIES is calculated as follows for each habitat (after Moss *et al.* 1987):

1. After determining whether a test site falls within the range of environmental variables for which an AUSRIVAS model can be used validly, the environmental data from the site are used to estimate the probability that the site belongs to each of the groups of reference sites. The probability that a test site belongs to a group is a function of the proximity of that site to the centroid of a group and the size of the group (Institute of Freshwater Ecology 1991, p. 17).
2. The probability of occurrence of family i in a standard sample from the test site is calculated by summing the products of the probability that the site belongs to group *j* (for *j* = 1, …, n, where n is the number of groups) and the proportion of members of group *j* at which family i occurs in a standard sample.
3. To calculate the expected number of families, only those with ≥50 % probability of occurring in a standard sample from the test site are included. We term this procedure selecting a **threshold probability** because families with probabilities less than this value are excluded from the analysis. This exclusion is justified on evidence that the occurrence of such low-probability families is haphazard, and could make results too 'noisy' (Moss *et al.* 1987; Institute of Freshwater Ecology 1991).
4. *E,* the number of families expected to be found in the sample is then calculated as the sum of the individual probabilities of each family above the threshold probability.
5. To count the number of observed families in the sample, *O*, the families found are tallied with the important proviso that only those families that are predicted to occur at above the threshold probability identified in step 3 are included. Thus the observed number of families, *O*, is really the observed number of expected families.

   In some circumstances, an impacted site could have some of its expected families replaced by more tolerant, but unexpected, families. If all of the families observed in a sample were included in the tally of O, such a site could retain a value of *O* similar to *E*. This would give an *O/E* of close to unity, creating the false impression that there was no impact (Wright, Furse & Armitage 1994).

O/E SIGNAL

The SIGNAL grades are incorporated into the *O/E* index as follows:

1. As described above, the expected number of families in a sample is the sum of the weighted probabilities of occurrence, p_i, where all families included in the calculation are at or above the threshold probability (Step 3 above). This quantity is denoted as E_T, ie.

   $$E_T = \sum_i p_i$$

   The SIGNAL grade for each family is incorporated by multiplying the probability of each family's occurrence by its SIGNAL grade, ie.:

   $$E_S = \sum_i (p_i S_i)$$

   where S_i is the SIGNAL grade for family

   The expected value of SIGNAL, E_{SIGNAL}, is given by:

   $$E_{SIGNAL} = \frac{E_S}{E_t}$$

   (*cf.* calculation for *O/E* ASPT in the British RIVPACS: Institute of Freshwater Ecology 1991, p. 17).

2. The observed value of SIGNAL, OSIGNAL, is obtained by summing the SIGNAL grades for all of the observed families with expected probabilities of occurrence at or above the threshold probability, and dividing by the number of these families, n_f, ie.:

$$O_{SIGNAL} = \frac{\sum_i (S_i)}{n_f}$$

A worked example

Worked example of calculations for O/E families and O/E SIGNAL

FAMILY	SIGNAL GRADE (S_i)	PROBABILITY OF OCCURRENCE (P_i)	E_s (= P_i/S_i)	FAMILIES OBSERVED AT SITE (O) (0 = absent, 1 = present)	OBSERVED SIGNAL GRADES (S_i x O)
Aeshnidae	6	0.4	2.4	0	0
Baetidae	***5***	***0.8***	***4.0***	***1***	***5***
Chironomidae	***1***	***0.9***	***0.9***	***1***	***1***
Dugesiidae	3	0.0	0.0	1	3
Ephydidea	2	0.0	0.0	1	2
Glossosomatidae	***8***	***0.6***	***4.8***	***0***	***0***
Hydridea	4	0.1	0.4	0	0
Isostictidea	***7***	***0.6***	***4.2***	***1***	***7***
Janiridae	5	0.3	1.5	0	0
Kokiriidae	***10***	***0.8***	***8.0***	***0***	***0***
	Sum	3.7	21.9	3	13

Expected number of families is the sum of the probabilities > = threshold probability (ie. values in bold italics in column above.)

Expected SIGNAL = 21.9/3.7 = 5.92

Observed SIGNAL = 13/3 = 4.33

These families are not included in the observed sums because their probabilities of occurrence are < threshold probability

O/E Families	= 3.00/3.70	0.81
O/E SIGNAL	= 4.33/5.92	0.73

Families with a probability of occurrence > = 0.5 are in bold italic type

2.1.2 **why** use these two indices?

The reason for using two indices is that they have slightly different emphases, which can provide additional diagnostic information for the end user. In most situations, both indices are likely to agree. *O/E* FAMILIES should be sensitive to a wide variety of disturbances provided these result in the loss of families of macroinvertebrates from the habitats sampled at a site. Thus this index should detect not only loss of families due to deteriorated water quality, but also loss because of physical habitat degradation. A potential disadvantage of *O/E* FAMILIES, however, is that number of families observed depends on sampling effort. Although sampling effort is standardised in AUSRIVAS, and is carried out by trained operators, it is not a quantitative technique and is subject to minor differences in style between operators. Extremely inclement weather or over-zealous sampling could also result in under- or over-estimation of the observed number of families.

O/E SIGNAL weights the families by their sensitivity to water pollution. Accordingly, *O/E* SIGNAL can detect situations where water pollution has resulted in the loss of only a few, but very sensitive, families. This is illustrated in the hypothetical worked example in the last section. In addition, *O/E* SIGNAL averages the contributions of the various expected and observed families to the final value of the index. Thus this index is potentially less sensitive to variations in sampling effort than *O/E* FAMILIES. However, there are two reasons for caution in using *O/E* SIGNAL.

Firstly, the range of disturbances for which SIGNAL grades have been derived is presently limited. Most of the data used to derive the SIGNAL grades have come from rivers where the pollutants are organic effluents or diffuse agricultural and urban runoff (eg. treated sewage and sediment and nutrient-laden stormwater). The database is also biased towards south-eastern Australia. Secondly, not all human disturbances impinge on water quality *per se*. For example, increased bedload may alter the habitat of the benthos and result in the loss of some families; not all of the families that disappear will be 'high-grade' (ie. pollution-sensitive) families. As a result, *O/E* SIGNAL may stay relatively high because the remaining families are still sensitive to pollution, but *O/E* FAMILIES will indicate a net loss of families.

The potential advantages and disadvantages of the two indices are summarised in the following table.

	O/E FAMILIES	**O/E SIGNAL**
Advantages	Simple to understand. Should reflect a wide variety of disturbances including habitat degradation as well as diminished water quality.	Takes sensitivity of taxa to pollution into account; therefore emphasises water quality effects on fauna. Because of averaging, this index is less sensitive to variations in sampling effort
Disadvantages	Can be sensitive to sampling effort. Higher variability amongst reference sites compared with *O/E* SIGNAL.	Some feel grades need more testing with wider range of known disturbances. Some situations where *O/E* SIGNAL remains close to reference whereas *O/E* FAMILIES shows substantial impact

2.1.3 **background** on the development of SIGNAL

In many parts of the world biotic indices are calculated by summing or averaging pollution-sensitivity values assigned to species, genera, families and higher taxa of stream macroinvertebrates. Indices of this type include the Biological Monitoring Working Party (BMWP) score system used in Great Britain (Armitage *et al.* 1983), Stark's (1985) New Zealand Macroinvertebrate Community Index (MCI), Hilsenhoff's (1987; 1988) Wisconsin Biotic Index (BI) and Family Biotic Index (FBI), the Spanish Biological Monitoring Water Quality (BMWQ) score system (Camargo 1993b) and the South African Score System (SASS) of Chutter (1994).

Such indices are increasingly popular because they are responsive to different types of anthropogenic disturbance (Pinder & Farr 1987b; Barton & Metcalfe-Smith 1992; Camargo 1993a; Resh & Jackson 1993; Growns *et al.* 1995), robust to variations in sample size (Armitage *et al.* 1983; Pinder & Farr 1987b; Stark 1993; Growns *et al.* 1997), and have low variability both within a site and over time (Jones *et al.* 1981; Barton & Metcalfe-Smith 1992; Hannaford & Resh 1995).

Chessman (1995) designed SIGNAL (Stream Invertebrate Grade Number – Average Level) as a provisional Australian index of this type. For the initial version (SIGNAL-95), sensitivity values (grade numbers) were set for common macroinvertebrate families using information from the few published Australian studies that contained sufficiently detailed data on the responses of river macroinvertebrate communities to specific anthropogenic disturbances. These disturbances included the discharge of biodegradable organic pollutants including sewage and sugar mill wastewater (Campbell 1978; Arthington *et al.* 1982; Watson, Arthington & Conrick 1982; Pearson & Penridge 1987; Cosser 1988), trace metal contamination resulting from mining activities (Nicholas & Thomas 1978; Norris, Lake & Swain 1982; Norris 1986; Mackey 1988), sedimentation caused by land clearing and urban development (Hogg & Norris 1991) and the disposal of power station and pulp mill effluents (Marchant *et al.* 1984; Chessman & Robinson 1987).

SIGNAL-95 was tested in the Nepean River, Blue Mountains and Sydney regions of New South Wales (Growns *et al.* 1995; Growns *et al.* 1997), where it showed a high degree of sensitivity to sewage pollution and salinity and little response to gradients in natural factors such as stream size and elevation. Subsequently, Chessman *et al.* (1997) developed a modified version of SIGNAL for the Hunter River basin, New South Wales, using an iterative algorithm to assign basin-specific grade numbers objectively. The modified index (SIGNAL-HU97) was highly correlated with water turbidity and electrical conductivity, as well as with altitude and the condition of the stream banks and bed. Later, Chessman (unpublished manuscript) used similar iterative algorithms to assign grade numbers to a wide range of families using data obtained during the Monitoring River Health Initiative from all Australian states and territories.

A particular difficulty in the construction of SIGNAL and similar indices is that individual taxa may not be equally sensitive to all types of anthropogenic disturbance. For example, laboratory studies show that particular species of aquatic macroinvertebrates vary quite widely in their tolerances of specific pollutants (Chapman, Farrell & Brinkhurst 1982; Slooff 1983; Ewell *et al.* 1986). Similarly, the abundances of particular macroinvertebrate taxa can differ greatly among streams affected by flow alteration and various types of land-use activities and wastewater discharges (Yoder & Rankin 1995). In these circumstances, it is difficult to assign representative sensitivity values to individual taxa. The sensitivity grade numbers for macroinvertebrate families used in SIGNAL-95 represented a subjective compromise between responses to a wide range of types of disturbance.

Chessman and McEvoy (in prep.) tested the possibility of developing a suite of biotic indices targeted to specific types of disturbances. Each index would be constructed using separate grade numbers reflecting sensitivities to the particular disturbance concerned. By calculating values of each index for a study site, it would then be possible to obtain not only an indication of the degree of impact, but also a diagnosis of the particular type of impact. They sampled rivers affected by three well-studied types of disturbance: large dams, pollution by municipal sewage effluent, and pollution by trace metals originating from historic mining, and used the algorithm of Chessman *et al.* (1997) to derive grade numbers for each river and for combinations of all rivers subjected to a particular disturbance type. They also compared grade numbers for different disturbances and tested the ability of grades derived from one river to interpret impacts on other rivers.

It appears from this study that the development of diagnostic indices is a realistic possibility, at least for some types of disturbance. However, differences in sensitivities to particular pollutants among species within a family may necessitate the derivation of grade numbers at the species or possibly the genus level. Because of the sporadic occurrence of most macroinvertebrate taxa, and differences in responses between river systems, large amounts of data will be required to calculate reliable grades. There is also uncertainty about the number of different types of disturbance for which separate grading systems will need to be developed. It is probable that for disturbances such as organic pollution, a high degree of generality exists. However, for other disturbances such as flow alteration, it is likely to be necessary to differentiate a number of types of disturbance regimes. Even then, a lack of generality in response (Castella *et al.* 1995) may prevent the development of adequate diagnostic systems, and other approaches may be required.

Chessman has developed a set of interim SIGNAL grades for use in AUSRIVAS (Appendix D of this report). These include generalised grades which should be appropriate for pollution by nutrients, organic wastes and fine sediment (which in practice are often associated with one another as a result of diffuse agricultural and urban runoff). A set of alternative grades is also provided for situations where contamination by trace metals is suspected. Refinement of SIGNAL grades is expected to continue as more data become available through the National River Health Program.

2.2 **the** banding scheme

2.2.1 **why** divide the indices into bands?

Banding refers to the procedure of dividing a continuous measure into a small number of categories for the purposes of reporting and comparison across space and time. Inevitably this results in some loss of information, and is problematic in borderline cases. (To use an analogy, most of us can recall the pain we felt as students when our 'nearly distinction' score on a continuous scale was converted to a mere 'credit' grade when the teacher divided those scores into bands.)

Banding has to be both scientifically credible and easily interpreted by the end-user (Institute of Freshwater Ecology 1991, p. 21). The issues that need to be addressed in any attempt at banding are: the number of measures that will be banded and how these separate measures will be amalgamated, the number of bands for each measure, and the thresholds for each band.

2.2.2 **criteria** for division of indices into bands

Two options for banding schemes were considered during selection of the final scheme. These were banding based on explicit criteria and banding based on variability of reference sites.

Banding on explicit criteria would require river ecologists to define index levels that correspond to mild, moderate and severe impairment of the macroinvertebrate community. For example, a loss of 10% of expected families (*O/E* FAMILIES of 0.9) might be considered to represent mild impairment, a 25% loss could be defined as moderate impairment, and a 50% loss as severe impairment. Statistically, this is akin to defining the effect size of interest in conventional hypothesis testing.

The advantage of this approach is that levels of impairment are unambiguously defined by an explicit value judgment, and are easy to understand. These levels can be readily linked to water quality guidelines that specify acceptable levels or targets for biological indicators of water quality. The band boundaries are fixed and do not automatically change as the AUSRIVAS models develop over time.

However, agency biologists were generally uncomfortable with this approach, believing that experience with AUSRIVAS and the responses of macroinvertebrate communities across the country was insufficient to reach consensus on such boundaries at the present time. There was also a view that such an approach could be too inflexible.

No matter which approach is used, there is variability amongst reference sites, and that the statistical nature of the final values of the indices needs to be incorporated into the banding procedure. To elaborate: even if all observations and measurements could be made without error, not every reference site would have a value of *O/E* FAMILIES or *O/E* SIGNAL equal to exactly 1. This is because most of the reference sites are not completely undisturbed by humans, and so the *O/E* indices for the reference sites will vary around 1. In short, observations are being compared with an average expectation, and there is variability inherent in that expectation (Clarke *et al.* 1996). This variation is in addition the variation inherent in deriving the observed parts of the indices (Discussed further in Sections 2.2.3 and 3.4).

The alternative approach of banding on variability of reference sites was therefore adopted. In this approach the boundary between band 'A' (the 'reference band') and lower bands is set at a certain percentile of *O/E* values for reference sites in a particular model. This percentile is set at a low value so that very few sites that are in equivalent condition to the reference sites will be banded below 'A'. (For statistical reasons, it is impossible to prevent some sites that are equivalent to reference being banded as 'B' or lower.) The advantage of this approach is that it takes account of inherent uncertainties and variability in predictions of the expected fauna, and automatically adjusts band boundaries according to the variability of *O/E* values for a particular model.

While this approach is simple to implement, it has the disadvantage of being arbitrary; ie. there is no particular reason for choosing a certain percentile of reference site *O/E* values at which to set the boundary between bands 'A' and 'B'. Thresholds as varied as the 25th percentile, the 10th percentile and two standard deviations below the mean have been used elsewhere (Wright *et al.* 1991); (Hannaford & Resh 1995); (Wright *et al.* 1995). Nor is there any generally agreed principal on how to set band boundaries below band 'B'. These have sometimes been set according to the width of band 'A' and sometimes to set to divide the range from band 'A' to zero into equal intervals.

A further disadvantage of this approach is that a test site can change bands when the underlying predictive model changes, even if the fauna of the site remains constant. For example, if an improved model allows more accurate predictions of reference site fauna, the boundary between bands 'A' and 'B' will move upward, and a site previously in band 'A' can move to band 'B'.

Bands set on the basis of variability of reference sites *must not be labelled with names such as 'unimpaired' or 'moderately impaired',* because the degree of biological change that constitutes a particular level of impairment is *not defined in this approach.* Instead we have used terms such as 'similar to reference' or 'much poorer than reference.' As noted in Section 1.1, impairment is a value judgement that depends on the context of the assessment.

The advantages and disadvantages of the two approaches are summarised in the following table.

	BANDING ON EXPLICIT CRITERIA	**BANDING ON VARIABILITY OF REFERENCE SITES**
Advantages	Simple to understand	Simple to implement
	Levels of impairment are explicitly defined	Does not require value judgments about levels of impairment
	Band boundaries fixed, providing consistent targets	Automatically adjusts to individual models
	Links readily to water quality guidelines	
Disadvantages	Current knowledge insufficient to reach consensus on band boundaries	Band boundaries are arbitrary
		Easily misinterpreted
		Band boundaries may change as models improve, creating 'moving goalposts'

The final number of bands for summary reporting was a compromise between the need for fine resolution for some reporting purposes (eg. Section 1.7.4), the variability inherent in reference site data, and the necessity to minimise the number of bands to reduce the probability of mis-banding sites (Clarke *et al.* 1996). The five bands designated in Section 1.4 satisfy these criteria and result in a nationally consistent reporting scheme. The numerical values of the band boundaries depend on the inherent variability amongst reference sites of each model and thus allow the necessary flexibility to account for regional differences in the variation of the reference data set. Percentile values were used to designate band boundaries rather than multiples of the standard deviation because some reference data sets showed slightly skewed distributions of the values of the indices. The 10th and 90th percentile values of the distribution of the reference sites was used to define the width of the reference band, A, so that the band boundaries err on the side of test sites being judged different from reference condition. This was felt to reinforce the precautionary approach pursued in the reporting procedure and also makes some allowance for the lack of sensitivity that is probably inherent in a rapid bioassessment approach that uses family-level presence/absence information. The analyses used to make these decisions are presented in Appendix D.

2.2.3 **errors** in estimates of index values and bands

The term 'error' in this context refers to the variation inherent in estimating either the observed or expected component of either index. This variation results from the probabilistic nature of sampling. Replicate samples taken from the same habitat in the same site on the same day will have slightly different numbers of families. Similarly replicated measurements of environmental variables used for predicting the fauna will vary. These variations will occur even with highly trained personnel operating with extreme rigour. As a result, a test site has a chance of being misallocated to a band, with the probability of misallocation rising as one approaches the boundaries between the bands (Clarke *et al.* 1996).

What is needed, therefore, is some estimate of the natural variation around the value of an index for a given site so that a confidence interval can be constructed for the value of the index. This is not easy to calculate because replicated samplings from the same site are not part of the Rapid Bioassessment Protocol because of the motive to keep the procedure as rapid as possible. Consequently, there are several sources of potential error which need to be quantified and combined so that confidence intervals and the probability of misallocation to a band can be estimated. The sources of error in estimating the value of an index are as follows (see also Clarke *et al.* 1996):

1. Variation in the capture of families or sampling variation. Although the sampling regime is standardised as part of the River Bioassessment Manual (Anonymous 1994), not every family present within a given habitat at a given site will be captured; some rare families in particular will inevitably be missed. Different operators will vary in their sampling style and efficiency, and different suites of organisms will be collected depending on which 10 m transect of riffle, pool edge etc. is selected for sampling.
2. Errors resulting from sorting samples and identification of the families, and data recording and transcription errors. Even experienced operators will miss or misidentify organisms in samples.
3. Variation in the values of the environmental variables. Variables measured from maps (eg. altitude, distance from source) or in the field (eg. river width, composition of the substratum) may differ between operators or between successive measurements by the same operator. This can lead to errors in generating the 'expected' portion of either index by affecting the values the site takes on the discriminant functions used to derive the weighted probabilities of the families for a site. In addition, for models based on data from combined seasons, average values of the environmental variables are used, and there are errors inherent in averaging these values which will also affect the weighted probabilities of the families for a site.

4. Natural temporal variation. There will be intra- and inter-annual differences in the fauna of a site resulting from natural, climatic changes such as droughts and floods. El Niño – Southern Oscillation events, in particular, have profound effects on streamflows in southern Australia. Predictions made using data obtained under one climatic regime may be invalid under another regime.

Clarke (1996) also identify two further sources of error in estimating the expected families for a site, which relate more to the choice of variables and analytical methods rather than the four classes of 'sampling error' mentioned thus far. As such these two further sources of error are regarded as part of the definition of the 'expected' portion of output indices rather than sampling errors (R.T. Clarke, Institute of Freshwater Ecology, pers. comm., (Clarke *et al.* 1996). These other two sources are:

5. Omission of other environmental variables that might better predict faunal composition of a site.
6. Use of sub-optimal numerical procedures in analysis. There may be methods that make better use of the data resulting in more accurate predictions of the weighted probabilities of the families for a site. Stockwell and Faith (1996) have investigated alternative analytical procedures for AUSRIVAS, and (Wright *et al.* 1995) compare a variety of methods for the British RIVPACS data. Stockwell and Faith's (1996) recommendations are discussed further in Section 3.5.

Sources 5 & 6 are, therefore, part of the definition of expectations; if the resulting values are not sensible, then the variables and methods employed are inadequate for defining the expected faunal composition.

At present, it is not possible to quantify these sources of error. Some of them could be assessed by collecting replicated samples from the same site, habitat and time. Although some replicated samples have been collected by some State or Territory agencies, they have mostly not yet been processed or analysed. Quantification of errors was outside the ambit of this project. Suggestions for further research and development on this aspect of the outputs are outlined in Section 3.4.

2.3 **rules** for combining indices

Both *O/E* FAMILIES and *O/E* SIGNAL can vary from 0 to greater than 1 (the observed values of both indices are set to zero if no expected families are present). However, simply averaging the two indices is a poor method of combining them. This is because *O/E* SIGNAL is less variable amongst reference sites than *O/E* FAMILIES, and so the two indices are not strictly commensurate. Accordingly, a set of rules based on banding the two indices separately is most appropriate.

There are many different ways that the indices could be combined (Institute of Freshwater Ecology 1991), and several options were canvassed in the regional workshops and discussed by the expert panel (Section 4.2). For the time being the most precautionary approach was favoured by the potential end-users so that test sites were allocated to the band that was farthest from reference conditions based on the values of either index. In some cases this may mean that 'borderline' sites may be banded lower than they should be. However, there seemed to be general consensus that this risk was justifiable on three grounds. Firstly, the procedure is a rapid bioassessment procedure and is, therefore, less sensitive than more time-consuming and expensive quantitative procedures. Secondly, the margin of error or confidence intervals for the indices cannot be estimated yet, so the most precautionary approach was also the most defensible. Thirdly, adopting the 'worst case' was the simplest rule and the most publicly accountable.

This rule for combining indices differs from the British 5M system that was chosen by the National Rivers Authority (1994) for further investigation and testing. Under the 5M set of rules, *O/E* indices are calculated for the number of taxa (equivalent to *O/E* FAMILIES of AUSRIVAS), ASPT (equivalent to *O/E* SIGNAL of AUSRIVAS) and BMWP (equivalent to *O/E* SIGNAL but without averaging the SIGNAL values across the observed or expected families; we previously called this *O/E* SIGNS (Barmuta, Chessman & Hart 1996). The site allocated to the band indicated by *O/E* ASPT when this value is the lowest of the three, or to the band of the middle ranking index where this is not the case. The justifications for giving primacy to the *O/E* ASPT are that ASPT is less sensitive to differences in sampling effort and that *O/E* ASPT is better predicted by RIVPACS than *O/E* based on the number of taxa (Clarke *et al.* 1996).

At present, most end-users felt that giving similar primacy to *O/E* SIGNAL was premature given the current state of development of SIGNAL. This was also borne out by some of our own analyses of State and Territory data where a small number of test sites that were clearly impacted were banded lower on *O/E* FAMILIES than on *O/E* SIGNAL.

A further complication for AUSRIVAS compared with the British RIVPACS, is that there are usually two habitats sampled, where as the British procedure attempts to 'integrate' all habitats within the one sampling episode. Again, assigning the site to the band farther from reference conditions is the most precautionary approach adopted when combining bands from two habitats, although there was some debate about whether one habitat could be regarded as more sensitive compared with the other. For example, some researchers felt that riffle habitats harboured more sensitive families than edge habitats. This was countered by observations that some disturbances may affect edge habitats more than riffle habitats; for example, increased sediment inputs will deposit preferentially in slow edge habitats and affect the fauna there, but may have little effect in rapidly-flowing riffles. In summary, habitats may differ in their susceptibility to different types of impacts, and the most precautionary approach is warranted when combining assessments across habitats.

3. **future** developments & research needs

As part of the regional workshops we received extensive feedback on concerns that managers and technical staff had about AUSRIVAS. Only some of these concerns could be addressed by the first rounds of sampling and analysis of the MRHI. Indeed, this project concerning the outputs of AUSRIVAS was limited in its scope by the lack of large data sets collected in a fashion commensurate with the MRHI, and the unfinished status of concurrent research projects that bear on the quality and interpretation of AUSRIVAS outputs.

The following sections outline the key issues raised in all the regional workshops that require further research and development for AUSRIVAS to become a flexible and reliable tool for the wide variety of uses identified by the end-users. Some of these issues are not strictly about the formulation or presentation of outputs, but they nevertheless affect the quality of the outputs from AUSRIVAS and thus the trustworthiness of the procedure.

3.1 **selection** of reference sites

Reference sites for AUSRIVAS models are required to represent 'least disturbed' conditions. Reference sites also need to be chosen in such a way that they represent adequately the full range of river types in a region for which an AUSRIVAS model is to be developed (for example, large and small, upland and lowland, bedrock-confined and alluvial, naturally intermittent and naturally perennial) (see (Norris 1994) for a fuller discussion).

The reference sites in the current models have been chosen somewhat subjectively by agency staff, and often on the basis of limited information. This was necessary because of time and resource constraints, and the lack of prior studies of many river systems. During the process of model development, the reference data sets were censored by removal of those sites that were poorly predicted in preliminary model runs. The assumption underlying deletion of these sites is that despite initially being assumed to be relatively undisturbed, the fact that they have low *O/E* ratios suggests that their macroinvertebrate communities are nevertheless impaired. However, it is possible that in some cases they are inaccurately predicted simply because they represent unusual river types that are poorly represented in the reference data set.

There is a need for a more objective and standardised procedure for selection of reference sites. One option would be to undertake a detailed catchment analysis to identify and quantify disturbances and a screening process to delete catchment segments that do not satisfy minimum disturbance criteria.

To this end, the database being developed by Environment Australia as part of the Wild Rivers Project warrants further investigation (Stein *et al.* in prep.). This database has gathered data relating to sizes of settlements, land clearance, agriculture and forestry, mining leases and other disturbances across the entire country. Disturbances are weighted according to their type and distance from the river, and composite indices computed for river segments and entire rivers. The initial use of this database by the Australian Heritage Commission is to screen for rivers and river segments that are potentially free of disturbances due to European settlement, and thus identify rivers that are in near-pristine condition. Extension of the use of this database to rank reference sites would require further work to ensure that the weightings for the different disturbance types were appropriate for assessing river health and were commensurate across catchments and States and Territories. For example, information on sewage treatment plants merely records their presence and does not rank them based on the type of treatment employed or the performance of the plant. For assessing whether a river is in 'near-pristine' condition, the presence of any sewage treatment plant is sufficient information, whereas for ranking reference sites for biological assessment of river health, the level of treatment and the performance of the plant are important pieces of information.

One potential difficulty in implementing this approach is that the relationships between many types of human disturbances and the responses of river macroinvertebrate communities are poorly understood. The screening process might therefore eliminate sites exposed to disturbances that actually have minimal impact on the macroinvertebrate fauna. Another difficulty is that insufficient information on disturbance types and levels is available for many catchments, and some disturbances could pass undetected.

In some regions, the selection of reference sites is particularly difficult because all significant rivers have been substantially disturbed across large geographic areas. Examples of such areas are the wheat belt in the south-west of Western Australia, and the highly regulated floodplains of the Murray-Darling Basin in western New South Wales. It is arguable whether the RIVPACS/AUSRIVAS approach is applicable to such regions. Biological targets for such areas may need to be set not by predictive modelling, but by other means.

3.2 **selection** of predictor variables

In addition to assessments of macroinvertebrate communities, various physical, chemical and biological habitat variables are measured at AUSRIVAS reference and test sites. Each variables is measured for one two reasons: either as a potential predictor variable for possible use in the AUSRIVAS models, or as an explanatory variable to assist in diagnosis of the reasons for impairment of the macroinvertebrate community at a test site. The nature of the variables used for predicting the fauna at test sites was queried closely by some participants in the regional workshops.

Such environmental variables range from those that are entirely unaffected by human activity (eg. altitude) to those that are greatly modified by human activity, at least at some test sites (eg. phosphorus concentration in the water). The former are appropriate as predictor variables since the values of such variables measured at test sites, which will be input to the models to generate predictions, will be natural. The latter are generally inappropriate as predictor variables since their values at test sites will often not be the natural values for those sites.

Nevertheless, some of the current AUSRIVAS models use at least some human-influenced variables as predictors. For example, the New South Wales models incorporate alkalinity (affected by acid mine drainage, disturbance of acid sulphate soils, urban runoff and many other activities) and water depth in riffles (affected by flow regulation). The ACT models include the percentage cover of riparian trees less that 10 m high (affected by clearing and grazing). If unnatural values of these variables are input to the models, unnatural faunas will be predicted.

Four solutions to this dilemma are possible. The models could be reconstructed to avoid using human-influenced variables, but this may reduce predictive power. A variation on this theme is to only include human-influenced variables if they are unlikely to be affected by human disturbances within the region to which that model applies. (As a hypothetical example pH is potentially affected by human activities but might be included in the set of predictors in a region where such disturbances as acid mine drainage or acidic deposition were highly unlikely). Alternatively, these variables could continue to be used in the models, but the values input could be not those actually measured at the test sites, but those assumed to apply at the test sites under natural conditions. This could only be done in situations where pre-settlement values could reasonably be inferred, for example by using historical data or extrapolating backwards from current trends. Finally, the present approach could be continued, but with an explicit recognition that natural targets are not being sought for test sites. However, this would represent a major shift in philosophical perspective from the currently stated basis of AUSRIVAS (Davies and Schofield 1995).

Whichever approach is pursued, it needs to be clearly documented, and the inclusion of human-influenced variables should be identified in at least the raw outputs of future versions of the AUSRIVAS software.

3.3 **combining** habitat assessments

At present assessments from different habitats are combined using a set of rules (Section 1.6). An alternative which we did not have time to investigate fully would be to derive a probability of occurrence for each family for the entire site using basic probability theory.

Let $p1$ be the probability of finding a family in habitat 1 and $p2$ be the probability of finding the family in habitat 2, then the probability of finding the family in the entire site is given by:

$$p_1 \cup p_2 - p_1 \cap p_2 = (p_1 + p_2) - (p_1 \times p_2)$$

At present, the AUSRIVAS software keeps the family lists separate for the two habitat models, so incorporation of this method of combining habitat assessments should be assessed, preferably with the full range of data from FNARH, and compared with the present procedure prior to any re-writing of the AUSRIVAS software.

3.4 **estimating** errors in the indices

There were four sources of sampling error described in Section 2.2.3. To recapitulate, these are:

1. Variation in the capture of families or sampling variation.
2. Errors resulting from sorting samples and identification of the families. Research concurrent with this project is nearing completion (C. Humphrey, *eriss,* & A. Storey, University of Western Australia, pers. comm.) that should quantify the error rates of live- and laboratory-sorting procedures allowed by the River Bioassessment Manual (Anonymous 1994).
3. Errors in the values of the environmental variables.
4. Natural temporal variation. There will be intra- and inter-annual differences in the fauna of a site resulting from natural, climatic changes such as droughts, floods and, in some regions, wildfire.

All of these sources of error can potentially affect the value of both indices for a given site and result in misallocating the site to a band as a result. The necessity of describing a confidence interval for the index values was highlighted in all the regional workshops, and should be regarded as a high priority for further research and development for AUSRIVAS.

The first three sources of error itemised above need to be quantified using procedures outlined by Clarke *et al.* (1996) and described more fully by Furse *et al.* (1995). Specifically, replicated, empirical data need to be collected across a variety of environmental types and a variety of environmental qualities in order to quantify the magnitude of the sampling errors from sources 1–3 inclusive. In addition, the degree of tolerance of the models to errors in measuring the environmental variables can be estimated as described by Clarke *et al.* (1996) and Clarke *et al.* (1994).

Once the size of these errors has been estimated they can be combined using Monte Carlo simulation methods to generate confidence intervals for index values. If the distribution of the indice is assumed to be known, then analytical error estimates can be derived as well (Clarke *et al.* 1996).

The effects of natural temporal variation on values of the indices remains more problematic. Long-run data sets using procedures identical to those employed in the River Bioassessment Manual do not exist yet, and Humphrey (1997) has been using long-term, quantitative data sets from several environmental types in Australia to start to investigate this issue. The first two sampling occasions of the MRHI encompassed a severe drought that affected much of eastern Australia, especially central and south-western Queensland and western New South Wales. Comparisons of the models derived from the drought and non-drought years will need to be made, and decisions made about whether data from climatically extreme years should be included in the models that provide the expected probabilities of occurrence of the fauna. An alternative may be to exclude data collected in such extreme conditions and use it instead to develop separate models that are more applicable to tests carried out in comparably extreme conditions. A further alternative may be to exclude data from extreme years in model development and also not to attempt assessments of test sites under those climatic conditions. While conservative this last approach limits the usefulness of AUSRIVAS, and would make using this procedure difficult for Environmental Audits or State of the Environment Reporting if the reporting cycle coincided with these extreme conditions.

3.5 **integration** of alternative analytical methods

A prominent point discussed at all the regional workshops was the complexity of the numerical methods used to build the models for AUSRIVAS. Participants had difficulty in understanding fully how the procedure works, and many expressed doubts about being able to explain it to lay audiences. Stockwell and Faith (1996) investigated an number of alternative and novel methods for analysing these data, and their recommended methods are simpler, more direct and easier to explain. They also performed as well as or better than RIVPACS-type procedures (on which AUSRIVAS is presently based), although the procedures were compared on a limited number of small data sets which were all that were available at the time they conducted their research.

Stockwell and Faith's (1996) preferred procedures were the Distance Dissimilarity Regression Analysis Method (DDRAM) to build a model to judge whether a test site differed from the appropriate reference conditions, followed by a procedure called e-ball to provide diagnostic information on expected taxa for a test site that could then be used for an *O/E* index.

DDRAM uses linear regression of the biotic dissimilarities (measured by the Bray-Curtis measure) against environmental distances (measured by the Euclidean metric) for all combinations of the reference sites. The environmental data for a test site is then used to predict the biotic dissimilarity of the test site from the reference sites using the regression. If the observed dissimilarity of the test site falls outside the range of variation predicted by the regression then the site is judged to be impacted (Stockwell & Faith 1996). (The Z statistic is used to provide a formal hypothesis test of deviation from the regression, and hence impact.) Because the intermediate steps (eg. construction of a list of weighted probabilities of occurrence of each family) used in RIVPACS-type models are not used, e-ball provides diagnostic information as follows. The environmental data for a test site are used to position the test site amongst the reference sites and all the sites within a given radius are identified. These are the sites most environmentally similar to the test site. The faunal compositions of the reference sites within this radius are averaged to provide and expected list of families for the test site. The number of these expected families observed at the test site is then compared with the expected number of families, thus providing *O/E* ratios as for AUSRIVAS and RIVPACS.

When comparing DDRAM, e-ball and RIVPACS-type models, Stockwell and Faith (1996) found broad agreement between the methods, but found that some test sites differed in whether they were shown as being impacted by the two approaches. They recommended further testing of DDRAM and e-ball alongside the present RIVPACS-like approach adopted by AUSRIVAS, preferably with a known range of human disturbances included in the test-site database so that the different emphases of the two procedures could be compared. We strongly support this suggestion. The advantages of DDRAM and e-ball over the RIVPACS-type method from the perspective of reporting rapid bioassessments are:

- The method is simple to explain.
- Impact can be judged by a formal hypothesis test.
- The quality of the model is easily assessed from regression diagnostics
- E-ball could be developed so that the reference sites that are most similar environmentally to the test site are identified explicitly. This was identified in several regional workshops as being a desirable additional output for diagnosis, and for giving users a 'feel' for whether the predictions were biologically reasonable.

Stockwell and Faith (1996) also document the relative technical and statistical merits of DDRAM and RIVPACS-type methods, and outline procedures for removing environmental variables that reduce the quality of the model.

When comparing DDRAM and the present implementation of AUSRIVAS, the following issues need to be addressed:

1. A range of well-documented test sites with known disturbances needs to be included in the analysis so that the biases or emphases of the two models can be compared directly.
2. The responses and robustness of the two approaches to the errors inherent in rapid bioassessment.
3. The nature of the statistical test in DDRAM needs to be examined. For example, one issue is whether the *Z*-test is most appropriate given the small size of some of the reference data bases used to create a model.
4. Several options for e-ball deserve further investigation. For example, the size of the radius used and whether the reference sites within the radius should be weighed relative to their proximity to the test site are obvious issues that need to be explored more fully.

3.6 **calibration** of outputs using a variety of test sites

Testing of the outputs in this project was limited by the number and variety of test sites available for use with AUSRIVAS models. Although the indices and procedures recommended here worked well with the data sets available, further testing is desirable to calibrate the outputs and to test the robustness and sensitivity of AUSRIVAS.

The process of calibrating and further testing the outputs is, however, complicated and potentially circular. The most obvious approach is to examine relationships between water quality variables and output indices, and this has been used frequently in the past (reviewed in Barbour *et al.* 1992; Kerans & Karr 1994; Barbour *et al.* 1995). The disadvantage of relying solely on this method is that the biota may be responding to other variables which have not been measured. Human disturbances to rivers usually involve mixtures of impacts, so it is difficult to establish a simple relationship between 'severity of disturbance' and an *O/E* index. In other words, the metaphor of a dose-response relationship between disturbance and an index is an oversimplification. Besides, one of the primary motivations for using the biota in the first place is to provide information on the nature of the suspected impacts in addition to that already available from physicochemical measures (Cairns & Pratt 1993).

Thus there is no equivalent in biological assessment of the analytical chemist's procedure of 'spiking' a solution with a known concentration of a chemical to determine the accuracy of a procedure (Diamond, Barbour & Stribling 1996; US Environmental Protection Agency 1996). The alternative approach is to calibrate the performance of output measures against test sites subjected to a carefully selected range of disturbance types and severities. For the British RIVPACS, this has been attempted across the entire country to evaluate the performance of different indices (Wright *et al.* 1995), while the US Environmental Protection Agency (1996) advocates a performance-based methods system (PBMS), which incorporates a similar procedure within each region to select both replicated reference sites and range of test sites with known disturbances and impacts (Diamond, Barbour & Stribling 1996). Similar procedures have been used by other researchers seeking to validate different output indices for rapid bioassessments (Barbour *et al.* 1992; Kerans & Karr 1994; Barbour *et al.* 1995).

In Australia, as part of FNARH, the staff of State and Territory agencies have been asked to nominate sets or putatively 'highly disturbed', 'moderately disturbed' and 'least disturbed' sites, which provides the opportunity to further test the present outputs of AUSRIVAS, and compare the performance of the alternative analytical procedures of Stockwell and Faith (1996) (see Section 3.5). Inclusion of some replicated samplings in both test sites and reference sites within FNARH will also address issues of reproducibility and variability in model outputs.

3.7 **further** development of SIGNAL, variants of SIGNAL and other diagnostic tools

Further refinement of SIGNAL is generally impractical at present because of a lack of necessary raw data from a wide variety of impacts with differing, well-documented degrees of severity. However, development of a provisional metals diagnostic version is expected to be completed, using data from a variety of published and unpublished sources, before the end of 1997. It is likely that the First National Assessment of River Health will generate data that can be used to develop a variety of diagnostic versions, provided that a large number of sites with known types of disturbances is sampled.

In a related project, Chessman and McEvoy (in prep.) attempted to develop two other versions of SIGNAL, one related to sewage impacts and the other to the effects of dams and impoundments. However, the impacts of dams and impoundments are varied, so the faunal response will differ depending on the type of impact imposed by a particular dam and its discharge regime. For example, a deep impoundment that stratifies during summer will have a different impact if it discharges anoxic, cold water from the hypolimnion compared with a discharge from the warmer, oxygenated epilimnion. Accordingly, further diagnostic versions of SIGNAL will need to be oriented towards specific impacts that result from human disturbances rather than the disturbances themselves.

In developing further variants of SIGNAL, two issues relating to reporting need to be addressed. Firstly, the initial SIGNAL variants developed by Chessman and McEvoy were found to be highly correlated with each other and SIGNAL itself. Thus Barmuta *et al.* (1996) found that sites in Queenstown, Tasmania that were heavily impacted by heavy metal contamination scored low on *O/E* FAMILIES, *O/E* SIGNAL and *O/E* SIGNAL for metals and also *O/E* SIGNAL for sewage and dam impacts as well – despite the lack of human disturbances related to either sewage or dams for these sites. Naïve use of the variants of SIGNAL would lead to incorrect diagnoses of what the impacts were at these sites; this would become problematic in sites subjected to a variety of human disturbances where the task might be to evaluate the relative contributions of the different disturbances to the impacts at the site. The second issue relates to having too many diagnostic indices, which was a concern voiced by some of the participants in the regional workshops. Whether all index values have to be reported and how they are summarised were viewed as complications to a process that needs to be transparent and readily understood by the public so that they have confidence in the summary produced.

Nevertheless, there was strong support in most of the workshops for additional supporting or diagnostic information to eventually be output from AUSRIVAS. Only SoE reporting is satisfied by a simple summary banding; most other applications require additional information to infer which disturbances are responsible for the impact detected by AUSRIVAS. There are two approaches, which are not mutually exclusive, which could be pursued to make progress on this issue. The first is to continue development of impact-specific indices which are only employed for diagnosis by expert biologists, with *O/E* FAMILIES and *O/E* SIGNAL being used as the two summary indices on which bands are based for summary reporting. The second approach suggested by one group in one of the regional workshops was to develop a rules-based expert system to provide verbal diagnostic support for agency biologists. An example of such a rule might be 'if Leptophlebiidae were expected at the test site with a high probability (>0.7), but were absent with Conidia being found instead then suspect impact involving increased deposition of fine sediments'.

Such an expert system could be used to accommodate regional and local differences in faunal responses, and shows considerable promise to help formalise verbal interpretations of AUSRIVAS outputs. If this approach was combined with links to refereed and grey literature that documented particular faunal responses, then the system could form a valuable training and educational tool for less experienced staff and interested members of the public. Formulation of the rules would be a major exercise, however, and would require the participation of active biologists to ensure the scientific rigour and consensus of the interpretations. Furthermore, not all families are well enough known to permit formulation of rules for all of them; however, the process could be open-ended and added to as empirical knowledge improves. Indeed, the system may provide a formal framework to guide the acquisition and documentation of faunal responses in the future.

3.8 **diagnostic** environmental variables

At present, AUSRIVAS uses environmental variables only for building predictive models. There is scope for using the environmental information for further interpretation and diagnosis of AUSRIVAS outputs, which could include variables that are likely to be influenced by human disturbances.

Initially this could be a print-out of measured values to be used by expert scientists for interpretation and report preparation. Formalising the reporting of these variables would require further work, because regional and seasonal differences in physical and chemical variables preclude simple, nation-wide targets or standards to be set. The procedures followed for the Index of Stream Condition in Victoria (ISC) (Department of Natural Resources and Environment, ID&A Pty Ltd & Hydrology 1997) provide one example for formalised reporting of such data, and the US Environmental Protection Agency (1996) also have procedures for summarising physicochemical information, and further procedures for combining biological and physicochemical data into one summary measure. Both these approaches rely on a regionalisation of rivers and streams, so that targets and standards are set at appropriate levels.

Note, however, that these procedures are principally tools for summarisation and reporting. Department of Natural Resources and Environment *et al.* (1997), for example, emphasise that the primary uses of their ISC are for benchmarking, setting objectives and assessment of long term effectiveness of management intervention, and that not all of the information required by managers will be represented in the index. Thus, formalisation of diagnostic tools for use by agency scientists and biolgists for physical and chemical information will probably require more than a simple adaptation of these methods.

3.9 **use** in State of the Environment reporting and Environmental Audits

AUSRIVAS has been proposed as a 'state'or 'condition' variable for the purposes of State of the Environment Reporting (SoE) and Environmental Audits (P. Fairweather, CSIRO Division of Land and Water, pers. comm.). From our discussions during the regional workshops in 1996, it was clear that a wide variety of approaches to SoE reporting had been adopted by different States and Territories, and there was little consistency. There was general enthusiasm for using invertebrates and AUSRIVAS as a means of achieving consistent reporting for SoE, but there was also a misapprehension by some that the reference sites for AUSRIVAS had been selected specifically with the next round of SoE reporting in mind. This is not the case. Reference sites were selected on the basis that they represented the least disturbed or impacted conditions for a region, whereas SoE will require a network of sites encompassing a range of conditions from pristine to strongly impaired.

To use AUSRIVAS for SoE reporting or broad-scale Environmental Audits, three issues need to be addressed.

1. **Site selection.** At present, reference sites and test sites for the for the development of AUSRIVAS and for FNARH have not been selected for nation-wide SoE as a primary goal. Most States and Territories have oriented site selection towards areas most likely to require rapid bioassessment, so that the coverage of a given State or Territory is usually uneven. Likewise, test sites for the FNARH have been selected to help validate and further develop AUSRIVAS.

   It is likely, therefore, that a dedicated network of sites geared specifically towards SoE reporting would need to be developed, although some sites already within the AUSRIVAS network would probably be included. At present, SoE sites would need to be selected in regions where good predictive models already exist because predictions of faunal composition are not possible outside the environmental 'envelope' prescribed by the reference sites for a particular model.

   One option for SoE sites would be to locate them as fixed sites at the downstream end of catchments subjected to a variety of human disturbances. If the goal of environmental management is to improve river health through improved catchment management, then the focus would be on documenting trends in the AUSRIVAS outputs at these sites over time. An alternative method for locating SoE sites is to use some form of stratified sampling design so that sites from a wide range of different degrees of human disturbance were included in the SoE set. The hope would be that over time, heavily impacted sites would become closer to reference sites, and that sites rated as similar to reference did not deteriorate.

   The final issue about site selection for SoE reporting relates to the nature of the reference sites themselves, especially in regions where the least disturbed reference sites are already subjected to one or more human disturbances so that the fauna at the reference sites is probably already depauperate. There two concerns voiced in the regional workshops. First, that test sites within such regions could easily become as good as reference sites just by the occurrence of a very small number of families. Thus heavily disturbed areas may appear to recover more quickly compared with test sites in regions with a richer expected fauna. Second, the reference sites within such regions may themselves change, either improving as catchment management improves or deteriorating further in the face of added disturbances. In either case, the reference condition is a 'moving target', and this is likely to be disproportionately important in regions with a depauperate expected fauna.

2. **Temporal variability and frequency of sampling.** Droughts and floods are prominent features of Australian river systems, and rigid timetables for SoE reporting may coincide with, for example, El Niño droughts. Even if an

entire re-sampling of all the reference sites was carried out simultaneously with the SoE test sites, it is unclear whether reference sites themselves would remain faunally distinct from test sites under severe, natural droughts or floods.

In addition, if the goal of SoE reporting is to document changes in river health resulting from changes in management, then the reporting cycle needs to be geared to the expected speed of response of the river. For example, five years may be sufficient to expect changes from mitigation of nutrients from sewage treatment plants, but not long enough for a long-term program to reduce dryland salinity.

3. **Reporting unit and band number and widths.** There was some debate within the regional workshops about whether SoE reporting was better served by reporting the numbers of sites within each of the bands or the number of river kilometres within each band. Expanding the size of the reporting unit to kilometres of river requires more research about the spatial representativeness of the sites at they are sampled under AUSRIVAS. There was also a consensus amongst workshop participants that SoE reporting would be better served by a small number of bands, which could be achieved either by amalgamating some of the bands below reference or by using different boundary values for the (fewer) bands below band A. The justification for a smaller number of bands was simplicity of reporting.

Overall, however, the success or failure of the use of AUSRIVAS for SoE or Environmental Audits rests with the strategy adopted for selecting sites. At present there seems to be little consistency amongst the States and Territories about how this is to be achieved, and there is a clear and urgent need for SoE efforts to be co-ordinated if such reporting is going to be commensurate and meaningful.

4. **communication** and integration activities

4.1 **the** regional workshops

A series of five regional workshops were held in August and September 1996 to provide an opportunity for resource managers to comment on the way in which data being collected as part of the National River Health Program might best be interpreted and displayed. This section summarises the outputs from these workshops.

4.1.1 **purpose**

The objectives of the workshops were to obtain the views of river managers on:

- the uses of the outputs from the National River Bioassessment Program
- methods and formats for environmental quality ratings using rapid biological assessment
- desirable formats and media for presenting raw output, indices and environmental quality ratings.

4.1.2 **workshop** format

Full day workshops were run by the project team (Leon Barmuta, Bruce Chessman, Barry Hart) in five states as follows:

- Adelaide (26 August, 1996 – South Australian and NT participants)
- Perth (27 August, 1996 – Western Australian participants)
- Brisbane (29 August, 1996 – Queensland participants)
- Sydney (30 August, 1996 – NSW and ACT participants)
- Melbourne (2 September, 1996 – Victorian participants).

Subsequently, Leon Barmuta consulted separately with Tasmanian participants in September 1996.

A full list of participants is provided in Appendix B. All participants were provided with a background issues paper prior to the workshop. A copy of this is provided in Appendix C.

The format adopted for each workshop is summarised below. BTH denotes Professor Barry Hart, LB denotes Dr Leon Barmuta, and BC denotes Dr Bruce Chessman.

Morning

- Project background and objectives (BTH)
- Background information session 1: RIVPACS, *O/E* ratios, possible uses, banding, errors, case studies (LB/BC)
- Workshop 1 – participants divided in 3–4 groups, each to discuss the following three issues:
    1. Uses of biological indices
       How will these biological indices be used in Australia (eg. environmental reporting, regulatory activities, catchment management performance)?

2. Banding

   How many bands are appropriate? Where and how should bands be set? Should we have different bands for different uses? Should bands be nationally consistent? Should bands have descriptive labels?

3. Borderline sites

   How should we deal with borderline sites (eg. treat conservatively, apply 'benefit of doubt' or more investigation)? How to balance 'false alarm' and 'false security'?

- Plenary – groups report back

Afternoon

Background information session 2: details on indices (eg. *O/E*, *O/E* SIGNS, *O/E* SIGNAL, others) (LB/BC); delivery and outputs (eg. simple text, GIS-based outputs) (BTH)

Workshops – break into three–four groups, each to discuss the following three issues:

1. Indices

   Usefulness of the different indices (eg. *O/E*, *O/E* SIGNS, *O/E* SIGNAL, Dissimilarity Index). Is there a need for different grades for different regions? Is there a need to develop impact-specific versions of these indices?

2. Changes and modifications to the indices

   How are we to cope with improvements and changes to the indices that will occur with time?

3. Outputs

   Should these be user related? What options are available and what would people prefer (eg. simple text, geographically referenced, other)? What access should there be to the data (WWW, restrictions)?

- Plenary – groups report back
- General discussion
- Summarise Workshop outcomes (BTH)

4.1.3 **key** workshop outputs

A summary of the key points to emerge from the workshops is provided below under the six headings that were the focus of the discussions.

1. Possible uses of indices

The main question addressed was: How will these biological indices be used in Australia?

A wide range of uses of the information were identified. These are summarised below. It was pointed out that different sampling and reporting techniques may be needed depending upon the intended use of the information.

- Environmental reporting
  - useful in range of environmental reporting areas, eg. State of Environment, State of Rivers, Condition of the Catchment
  - useful to determine trends over time in the condition of particular river systems
  - will allow comparisons between sites
  - will be used as core SoE indicator in SA; several other states showed interest in using this approach for SoE and environmental audits

- Regulatory activities
  - information useful for assessing compliance monitoring, for environmental impact assessment and assessment of particular pollution cases leading to prosecution
  - particularly useful for these purposes if backed up with relevant chemical data
  - may not stand up to legal challenge because confidence intervals or standard errors are not yet available for index values
  - some groups saw difficulties in using the AUSRIVAS information for impact assessment, because to do this effectively will need 'before/after' and 'control/impact' data; also methods for assessing natural variation in the outputs need to be developed
- Catchment planning & management
  - useful for assessing condition of rivers in particular catchment, but will be one of a number of indicators (that will also include physico-chemical water quality, habitat assessment, etc)
  - in Victoria will be part of the Index of River Condition involving geomorphology, hydrology, riparian vegetation and water quality
  - may be used in conjunction with economic costs for making decisions on river maintenance, restoration and remediation
  - information useful in establishing rehabilitation targets
  - will have a role in catchment planning (eg. setting benchmarks and targets, identifying priorities for problem areas)
  - will be useful in evaluating the effectiveness of management initiatives, such as environmental flows, habitat restoration and water quality improvement
- Water quality guidelines
  - will contribute to the national ANZECC water quality guidelines
  - useful for indicating when quality degrading, but not what is causing the degradation
  - need more research on the tolerance of particular animals
- Community monitoring of river 'health'
  - provides a useful 'standardised' technique for assessing river health that is transparent and understandable by the community
  - will also require input from experts and adoption of good quality control procedures
- Water allocations and environmental water requirements
  - may be useful as a 'standardised' rapid biological assessment technique for following the changes in the biota as a result of modifications to the flow
- Education & training
  - will be useful in increasing community awareness of ecological degradation
  - will play a role in increasing general knowledge about the biological quality of rivers

2. Indices

The groups addressed the following questions: What is the usefulness of the different indices (eg. O/E, O/E SIGNS, O/E SIGNAL, dissimilarity Index). Is there a need for different grades for different regions? Is there a need to develop impact-specific versions of these indices?

- Summary of comments:
  - agreement that all three rapid indices are useful (eg. *O/E*, *O/E* SIGNS, *O/E* SIGNAL). A number of groups saw advantage in having available a range of indices.
  - two groups argued that multivariate analysis was a more robust technique for identifying environmental relationships, and for linking to biodiversity.
  - the general lack of comment on the desirability of other methods was probably related to a lack of knowledge of their advantages and disadvantages. Some groups felt that feeding groups and community composition indices may be useful in the future.
  - one group argued that grades need to be consistent within catchments, but if families differ in sensitivity between regions they should have different scores
  - all groups agreed that impact-specific indices would be useful and should be developed as the information becomes available, although two groups were concerned about too many indices being used for summary reporting. One suggestion was to reserve the used of other indices for diagnostic purposes rather than for formal reporting.
  - there was some doubt expressed about the ability of *O/E* SIGNAL to discriminate effects when only family level data are available. It was felt that more R&D was needed to identify the sensitivity of various families to particular pollutants
  - a number of groups were concerned at a lack of information on responses when only family level data was available (Note: some states are resolving to the species level). Some groups felt that there could be a different level of discrimination in different regions
  - a number of groups commented on the need to be able to assess the quality of the references sites chosen

3. Banding

The groups addressed the following questions: How many bands are appropriate? Where and how should bands be set? Should we have different bands for different uses? Should bands be nationally consistent? Should bands have descriptive labels?

- Summary of comments.
  - there was consensus that bands should be technically defensible and agree with scientific reality (ie. band indicating poor quality should refer to situations that are easily recognisable)
  - most groups opted for five bands – it was felt that this would allow for some incentive for improvement to move to next band; three bands might be too wide to allow this even though considerable improvement might have occurred
  - there was no consensus on how the bands should be set, although many groups felt the statistical method was the most appropriate. There was generally agreement that only one method should be used nationally and that this should be 'transparent' so that all could understand the basis for its adoption

- the consensus view was that the banding system recommended should be the same nationally – there needs to be compatibility between States and Territories
- one group cautioned that the rapid biological AUSRIVAS outputs are rather simplistic, but are being used to assess quite complex issues
- there was general agreement that bands should be labelled, but there was no consensus whether these should be simple letter labels (eg. X, A, B, C) or more descriptive labels (eg. excellent, modified, degraded, severely degraded) – it was argued by a number of groups that people will inevitably give the bands descriptive labels whether or not they are so labelled
- one group suggested that the banding scheme might need some form of 'community market survey' to assess whether it is acceptable

4. Borderline sites

The groups addressed the following questions: how should we deal with borderline sites (eg. treat conservatively, apply 'benefit of doubt' or more investigation)? How to balance 'false alarm' and 'false security'?

- Summary
  - most groups argued for a conservative approach that would see borderline sites being downgraded to the lower band. It was felt that this approach was the best for balancing 'false alarm' (ie. Type I error) against 'false security' (ie. Type II error)
  - some groups argued that borderline sites should be reassessed (possibly by re-sampling the site), taking into account any other relevant information on the site condition

5. How to treat future changes and modifications

The groups addressed the question: How are we to cope with improvements and changes to the indices that will occur with time?

- Summary of comments:
  - there was overwhelming agreement on the importance that the AUSRIVAS program be maintained as a national program with regular review
  - most groups felt that changes in taxonomy are inevitable, and it would be best if collected material were kept so that it can be reanalysed at a higher level of taxonomic resolution if needed in future
  - one group felt it may be useful in the future to include more general information on macroinvertebrate biology to assist with the interpretation of the indices (Note: the CRC for Freshwater Ecology have recently produced simple guide on macroinvertebrates – (Hawking & Smith 1997)
  - it was a general feeling of all groups that the capacity to respond to future changes and challenges will be linked to continuity of funding for this program
  - a number of groups raised the need to protect reference sites from anthropogenic changes – perhaps legislation is needed to protect these sites
  - a number of groups identified a desirability that the diagnostic capacity of the indices be improved over time

6. Outputs/Displays

The groups addressed the questions: Should these be user related? What options are available and what would people prefer (eg. simple text, geographically referenced, other)? What access should there be to the data (WWW, restrictions)?

- Summary of the comments:
  - a range of output products will be needed depending upon the user (eg. trends for So*E* reporting, simple *O/E* for community groups, colour coding of rivers for catchment management authorities). Extension of the reporting unit to river lengths rather than 'spot sites' was regarded by some as desirable for environmental reporting and catchment planning activities.
  - there was general agreement that the data should be easily available to interested groups (including data in electronic and hard copy forms) – data should be stored so it is nationally compatible
  - there was strong support for the data to be geographically referenced (eventually with some form of GIS support); because different agencies use different software and different data management procedures, the minimum requirement was that AUSRIVAS output plain text files with identifier codes that could then be imported simply into the host agency's procedures.
  - the groups generally felt that we should work towards internet/WWW access to the interpreted data and to GIS display systems – but don't delay development of simple outputs waiting for the 'ultimate' GIS. Limited access by community groups to the internet is an impediment especially in rural and remote areas, so simple, printed outputs should be readily available.
  - a number of groups identified the need for flexibility because the types of output products will evolve as people start to use the data
  - one group suggested that there a project should be established to develop a range of output products (eg. trends, simple pictures of the indices, etc)
  - a number of cost-sharing principle may have to be developed regarding use of the data (eg. what charge to consultants wishing to use data collected with public funds?)
  - One suggestion was to develop an expert system approach for incorporating taxon-specific, verbal information for diagnostic use by agency biologists. This could be adapted to local or regional conditions and could provide more informative support for agency biologists than using indices alone.

The main outputs from the workshop are conveniently summarised in Table 3. These are arranged under five headings – use of information, the best index or indices to use, what banding to use, what to do with borderline sites & desirable outputs for the data – for each of the five main uses identified by the workshop groups. A list of all the participants in the workshops is given in Appendix B, and the background materials supplied to the organisers in Appendix C.

4.1.4 summary

Overall, four broad issues were canvassed: the number of indices used for reporting, the number of bands or categories into which indices were divided, and the procedure to be followed for test sites that fell close to the borderline between two bands. Table 3 summarises the consensus views that emerged from these workshops. For most of the uses of AUSRIVAS, both *O/E* FAMILIES and *O/E* SIGNAL were deemed useful summary indices and that five bands would give sufficient discrimination for summary reporting. The only exception to this was for environmental reporting, where three bands might suffice. The justifications advanced for this were that for such

Table 3 Summary of the general consensus of comments to emerge from the regional workshops

USE OF INFORMATION	INDEX/INDICES	BANDING	BORDERLINE SITES	DESIRABLE OUTPUTS
Environmental reporting (eg. SoE, state of rivers)	*O/E* FAMILY	Three bands (eg. A, B, C)	Use conservative approach	Spatial differences/ Geographically reference display Extend reporting unit to lengths of river rather than spot sites Trends over time
Regulatory activities (eg. compliance, environmental impact assessment)	*O/E* FAMILY *O/E* SIGNAL	Five bands (eg. X, A, B, C, D)[1]	Allow benefit of doubt (ie. put borderline sites into higher band)	Quote band level for particular site Summary information on family composition at site with supporting narrative from agency biologist(s)
Catchment planning & management (including water allocations & environmental flows)	*O/E* FAMILY *O/E* SIGNAL[2]	Five bands (eg. X, A, B, C, D)	Conservative approach (ie. put borderline sites into lower band)	Geographically referenced display Extend reporting unit to lengths of river rather than spot sites
Community monitoring	*O/E* FAMILY *O/E* SIGNAL	Five bands (eg. X, A, B, C, D) Sub-division of bands or report index values	Conservative approach (ie. put borderline sites into lower band)	Geographically referenced display Summary information on family composition at site with supporting narrative from agency biologist(s) Index values to report trends
Education & training	*O/E* FAMILY *O/E* SIGNAL	Five bands (eg. X, A, B, C, D) Index values Supporting raw information		Range of displays

1 No consensus on whether band to be given simple letter labels or more descriptive labels.
2 Useful if further diagnostic indices for sewage, heavy metals, etc. were developed.

broad-scale uses, three categories were adequate (corresponding to 'as good as reference', 'moderately impacted' and 'severely impacted'), and that outputs from AUSRIVAS were likely to be only one of several indicators of overall river health in future SoE-style exercises. There was also a sense that if SIGNAL values were likely to change in the foreseeable future, use of *O/E* SIGNAL could be problematic if comparability of outputs was to be assured for documenting trends.

For sites which fell close to borderline values of bands, most uses would be best served by adopting a conservative approach whereby sites were allocated to the lower quality band and thus more likely to be judged as impacted. The possible exception to this rule was for some forms of regulatory use, where allowing benefit of doubt may lead to a greater chance of improved compliance in the future.

For AUSRIVAS to be useful for all these purposes, a variety of outputs with varying levels of detail are necessary. At the very least the final band that a site is allocated to needs to be output together will information that allows that site to be mapped in a geographically-referenced display by the host agency's data management and software. A plain text report for each site documenting index values and a list of the families expected and those observed provides invaluable diagnostic information for agency biologists which allows them to provide supporting narrative to help interpret the outputs for different audiences. At present the reporting unit for AUSRIVAS is a site on a stream or river reach. For environmental reporting and catchment planning and management, many felt that the reporting unit should be kilometres of river falling within a band rather than number of sites falling within a band (*cf.* National Rivers Authority 1991; National Rivers Authority 1994). Further research would need to be undertaken to permit this change of reporting unit.

4.2 **review** of proposed outputs

4.2.1 **procedure**

A telephone link-up was held on Thursday 6th March 1997 from 1:00 – 3:30 PM Eastern Standard Summer Time to discuss the draft proposed outputs of AUSRIVAS. The panel involved in this review included one representative nominated by the project leader in each State or Territory and A/Prof. R. Norris and Mr J. Simpson of the related project, Statistical support for the Monitoring River Health Initiative. The participants are listed below.

NAME	AFFILIATION
Jane Suggit	Northern Territory
Peter Goonan	South Australia
Peter Liston	Australian Capital Territory
Leon Metzeling	Victoria
David Oldmeadow	Tasmania
Eren Turak	New South Wales
Satish Choy	Queensland
Michael Smith	Western Australia
Leon Barmuta	University of Tasmania
Barry Hart	Monash University
Bruce Chessman	N.S.W. Department of Land and Water Conservation
Justen Simpson	University of Canberra
Richard Norris	University of Canberra

A draft set of proposed outputs were presented by Leon Barmuta at the first coordination meeting of FNARH on 4–5th February 1997 and copies of the proposals were circulated to all participants to discuss with their colleagues. In response to several e-mailed comments, a final set of proposals was either e-mailed or faxed to each of the participants approximately two weeks before the phone conference.

The outcomes of the discussions on the key issues identified in the proposals are summarised under each heading below.

4.2.2 **band** widths and cut-offs: statistical v. non-statistical criteria

The issue of whether to use statistical or non-statistical criteria to determine band boundaries was discussed extensively, and the arguments presented for the two positions have been summarised in Section 2 of this report. Overall, there was consensus to use statistical criteria to set the width of the reference band (band A), although whether to use percentile values v. a multiple of the standard deviation was a decision left to the project team to determine in the light of further analyses carried out in conjunction with A/Prof. Norris's team.

There was agreement that the width of band A be used to set the width of the lower bands, and that sites that were richer than reference also be identified in the banding scheme.

4.2.3 **combining** seasonal and habitat assessments

The strategy of preferring to use data from two seasons in combined-seasons models was reiterated, although the necessity to use single-season data and models in some circumstances was noted. Accordingly, using the seasonally most appropriate model was the approach adopted. There was some discussion about whether ultimately one of the habitats would be redundant and, therefore, whether AUSRIVAS should just focus on the 'most sensitive' habitat. This was countered by view that different disturbances would have differing impacts in the two habitats, and that in some regions or years one habitat may not be available for sampling. Accordingly, retaining both habitats for assessments wherever possible was judged to be the best approach, although single habitat rules needed to be included in the outputs for those circumstances when only one habitat was available or when it was only appropriate to sample the one habitat.

4.2.4 **naming** of the bands

There was little support for avoiding verbal descriptors for the bands because this complicates interpretation for a lay audience. However, there was little agreement about the precise form of names for the bands, although there was strong support for names that identified band A as 'as good as reference' so that it was not mistaken to mean 'pristine'; avoiding implicit value judgements in band names was also seen as important. It was also emphasised that the band names should be consistent with the way the band boundaries and widths were determined.

4.2.5 **reporting** assessments when the reference conditions are known to be degraded

This was acknowledged to be most problematic for So*E* reporting purposes, and requires further work to evolve some means of identifying and reporting situations where the reference conditions are already impacted. Development of other, independent assessments of 'river condition' was seen as one potential solution, or using alternative methods to predictive modelling to set biological targets was seen as another.

4.3 **integration** with related projects

Since the submission of the first milestone report from this project, the following integration activities have been undertaken with related MRHI/NRHP projects.

Statistical support for the Monitoring River Health Initiative, project leader: A/Prof. R. Norris. We have liaised closely with A/Prof. R. Norris and his project team, Justen Simpson and Paul Blackman. Integration activities have included supply of SIGNAL scores for incorporation in the AUSRIVAS software (November 1996), supplying information on band widths and values (February–April 1997), and supporting documentation for inclusion in printed and on-line materials for AUSRIVAS software (March–June 1997). We have involved A/Prof. Norris and Mr Simpson in the expert panel which commented on draft index and banding proposals and have enjoyed a productive and close collaboration on many of the issues surrounding the AUSRIVAS software and development of the outputs. They have also conducted some analyses for us as part of developing the outputs, and Mr Blackman is currently integrating the final single-site reporting output with the AUSRIVAS software.

Investigation of alternative approaches to linking habitat variables with site classification in a RIVPACS model, project leaders Dr D.R.B. Stockwell and Dr D.P. Faith. We provided feedback and comments on the final report from this project, and discussed potential means of developing and further testing DDRAM and e-ball. Our recommendations in Section 3 are a direct result of these discussions.

Development and Implementation QA/QC protocols fora sampling and sorting components of the MRHI agency bioassessment programme, project leaders: Dr C. Humphrey and Dr A. Storey. Drs Humphrey and Storey kept us closely informed about the results of their investigations into the real rates and types of identification errors in agency data. This information helped formulate the principles and rules for combining index values and for drafting the banding scheme. In return, we commented extensively on their criteria for assessing errors and the impact these would have on the integrity of models.

Relating River Macroinvertebrate Communities to Specific Changes in Water Quality, project leaders: Dr B.C. Chessman and Mr P. McEvoy. This project investigated alternative SIGNAL measures, and these were made available to us for testing for diagnostic purposes. Information from this project provided valuable background for discussions in the regional workshops in September 1996.

In addition, one or more of this project team have participated in the annual MRHI workshop, steering committee meetings of *Statistical support for the Monitoring River Health Initiative,* meetings of the Technical Advisory Committee of the MRHI, and the first coordination meeting of the FNARH. These were effective fora for us to communicate our research progress, solicit feedback on our proposals and obtain information relevant to the progress of our research project.

5. **listing** of lead agencies

The following table lists the contacts within the lead agencies of each State or Territory for the distribution of the final version of the information in this document about the interpretation of outputs from AUSRIVAS.

CONTACT(S)	AGENCY
Peter Thompson Satish Choy	Queensland Department of Natural Resources
Eren Turak	NSW Environment Protection Authority
Peter Liston	Environment ACT
Leon Metzeling	Environment Protection Authority, Victoria
David Fuller	Department of Primary Industry and Fisheries, Tasmania
Peter Goonan	Environment Protection Authority, SA
Stuart Halse Michael Smith	Conservation and Land Management, WA
Jane Suggit	Department of Lands, Planning and Environment, NT

6. references

Anonymous (1994) River Bioassessment Manual. Version 1.0. National River Processes and Management Program: Monitoring River Health Initiative. Published jointly by Department of Sport and Territories, Land & Water Resources Research & Development Corporation, Commonwealth Environment Protection Agency, Canberra, ACT, Australia.

Armitage, P.D., Moss, D., Wright, J.F. & Furse, M.T. (1983) The performance of a new biological water quality score system based on macroinvertebrates over a wide range of unpolluted running-water sites. *Water Research* 17, 333–347.

Arthington, A.H., Conrick, D.L., Connell, D.W. & Outridge, P.M. (1982) *The ecology of a polluted urban creek.* Australian Water Resources Council Technical Paper No. No. 68. Australian Water Resources Council, Canberra, ACT, Australia.

Barbour, M.T., Gerritsen, J., Griffith, G.E., Frydenborg, R., McCarron, E., White, J.S. & Bastian, M.L. (1995) A framework for biological criteria for Florida streams using benthic macroinvertebrates. *Journal of the North American Benthological Society* 15, 185–211.

Barbour, M.T., Plafkin, J.L., Bradley, B.P., Graves, C.G. & Wisseman, R.W. (1992) Evaluation of EPA's rapid bioassessment benthic metrics: Metric redundancy and variability among reference stream sites. *Environmental toxicology and chemistry* 11, 437–449.

Barmuta, L.A., Chessman, B.C. & Hart, B.T. (1996) *Interpreting the output of the National River Bioassessment Protocol to support management decisions.* Milestone 1 Report for Project UTA3 to the Land and Water Resources Research and Development Corporation.

Barton, D.R. & Metcalfe-Smith, J.L. (1992) A comparison of sampling techniques and summary indices for assessment of water quality in the yamaska river, quebec, based on benthic macroinvertebrates. *Environmental Monitoring and Assessment* 21, 225–244.

Cairns, J., Jr & Pratt, J.R. (1993) A history of biological monitoring using benthic macroinvertebrates. *Freshwater Biomonitoring and Benthic Macroinvertebrates.* (Eds D.M. Rosenberg & V.H. Resh), pp. 10–27. Chapman and Hall, New York, NY, USA.

Camargo, J.A. (1993a) Dynamic stability in hydropsychid guilds along a regulated stream – the role of competitive interactions versus environmental perturbations. *Regulated Rivers – Research & Management* 8, 29–40.

Camargo, J.A. (1993b) Macrobenthic surveys as a valuable tool for assessing freshwater quality in the Iberian Peninsula. *Environmental Monitoring and Assessment* 24, 71–90.

Campbell, I.C. (1978) A biological investigation of an organically polluted urban stream in Victoria. *Australian Journal of Marine and Freshwater Research* 29, 275–291.

Castella, E., Bickerton, M., Armitage, P.D. & Petts, G.E. (1995) The effects of water abstractions on invertebrate communities in U.K. streams. *Hydrobiologia* 308, 167–182.

Chapman, P.M., Farrell, M.A. & Brinkhurst, R.O. (1982) Relative tolerances of selected aquatic oligochaetes to individual pollutants and environmental factors. *Aquatic Toxicology* 2, 47–67.

Chessman, B. (1995) Rapid assessment of rivers using macroinvertebrates: A procedure based on habitat-specific sampling, family level identification, and a biotic index. *Australian Journal of Ecology* 20, 122–129.

Chessman, B.C., Growns, J.E. & Kotlash, A.R. (1997) Objective derivation of macroinvertebrate family sensitivity grade numbers for the SIGNAL biotic index: application to the Hunter River system, New South Wales. *Marine and Freshwater Research* 48, 159–172.

Chessman, B.C. & Robinson, D.P. (1987) Some effects of the 1982–83 drought on water quality and macroinvertebrate fauna in the lower La Trobe River, Victoria. *Australian Journal of Marine and Freshwater Research* 38, 289–299.

Chutter, M. (1994) The rapid biological assessment of stream and river water quality by means of the macroinvertebrate community in South Africa. In M.C. Uys (ed.), *Classification of rivers and environmental health indicators. Proceedings of a joint South African/ Australian workshop, February 7–14, Cape Town, South Africa.* Water Research Commission Report No. TT 63/94, Pretoria, South Africa.

Clarke, R.T., Furse, M.T. & Wright, J.F. (1994) *Testing and further development of RIVPACS. Phase II: Aspects of Robustness.* A report to the National Rivers Authority No. R&D 243/7/Y. National Rivers Authority, Bristol, U.K.

Clarke, R.T., Furse, M.T., Wright, J.F. & Moss, D. (1996) Derivation of a biological quality index for river sites: comparison of the observed with the expected fauna. *Journal of Applied Statistics* 23, 311–332.

Cosser, P.R. (1988) Macroinvertebrate community structure and chemistry of an organically polluted creek in south-eastern Queensland. *Australian Journal of Marine and Freshwater Research* 39, 671–683.

Department of Natural Resources and Environment, ID&A Pty Ltd & Hydrology, C.f.E.A. (1997) *An Index of Stream Condition: User's Manual*. No. 3 volumes: User's Manual; Reference Manual; Trial Applications. Waterway and Floodplain Unit, Department of Natural Resources and Environment, Melbourne, Victoria, Australia.

Diamond, J.M., Barbour, M.T. & Stribling, J.B. (1996) Characterizing and comparing bioassessment methods and their results: a perspective. *Journal of the North American Benthological Society* 15, 713–727.

Ewell, W.S., Gorsuch, J.W., Kringle, R.O., Robillard, K.A. & Spiegel, R.C. (1986) Simultaneous evaluation of the acute effects of chemicals on seven aquatic species. *Environmental Toxicology and Chemistry* 5, 831–840.

Furse, M.T., Clarke, R.T., Winder, J.M., Symes, K.L., Blackburn, J.H., Grieve, N.J. & Gunn, R.J.M. (1995) *Biological assessment methods: controlling the quality of biological data. Package 1: The variability of data used for assessing the biological condition of rivers.* National Rivers Authority, Bristol, U.K.

Furse, M.T., Wright, J.F., Armitage, P.D. & Moss, D. (1981) An appraisal of pond-net samples for biological monitoring of lotic macro-invertebrates. *Water Research* 15, 679–689.

Growns, J.E., Chessman, B.C., Jackson, J.E. & Ross, D.G. (1997) Rapid assessment of rivers using macroinvertebrates: cost and efficiency of six methods of sample processing. *Journal of the North American Benthological Society* 16, 682–693.

Growns, J.E., Chessman, B.C., McEvoy, P.K. & Wright, I.A. (1995) Rapid assessment of rivers using macroinvertebrates: case studies in the Nepean River and Blue Mountains, NSW. *Australian Journal of Ecology* 20, 130–141.

Hannaford, M.J. & Resh, V.H. (1995) Variability in macroinvertebrate rapid-bioassessment surveys and habitat assessments in a northern California stream. *Journal of the North American Benthological Society* 14, 430–439.

Hawking, J.H. & Smith, F.J. (1997) *Colour Guide to Invertebrates of Australian Inland Waters.* CRC for Freshwater Ecology, Murray-Darling Freshwater Research Centre, Albury, N.S.W.

Hilsenhoff, W.L. (1987) An improved biotic index of organic stream pollution. *Great Lakes Entomol.* 20, 31–39.

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

Hogg, I.D. & Norris, R.H. (1991) Effects of runoff from land clearing and urban development on the distribution and abundance of macroinvertebrates in pool areas of a river. *Australian Journal of Marine and Freshwater Research* 42, 507–518.

Institute of Freshwater Ecology (1991) *Testing and further development of RIVPACS.* Interim Report, R&D Project 243. National Rivers Authority, Bristol, U.K.

Jones, J.R., Tracy, B.H., Sebaugh, J.L., Hazelwood, D.H. & Smart, M.M. (1981) Biotic index tested for ability to assess water quality of Missouri Ozark streams. *Transactions of the American Fisheries Society* 110, 627–637.

Kerans, B.L. & Karr, J.R. (1994) A benthic index of biotic integrity (B-IBI) for rivers of the Tennessee Valley. *Ecological Applications* 4, 768–785.

Mackey, A.P. (1988) The biota of the River Dee (Central Queensland, Australia) in relation to the effects of acid mine drainage. *Proceedings of the Royal Soceity of Queensland* 99, 9–19.

Marchant, R., Graesser, A., Metzeling, L., Mitchell, P., Norris, R. & Suter, P. (1984) Life Histories of Some Benthic Insects from the La Trobe River, Victoria. *Australian Journal of Marine and Freshwater Research* 35, 793–806.

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

National Rivers Authority (1991) *The Quality of Rivers, Canals and Estuaries in England and Wales.* National Rivers Authority, Bristol, U.K.

National Rivers Authority (1994) *The quality of rivers and canals in England and Wales (1990 to 1992) as assessed bye a new general water quality assessment scheme.* Water Quality Series No. 19. National Rivers Authority & HMSO, London, U.K.,

Nicholas, W.L. & Thomas, M. (1978) *Biological release and recycling of toxic metals from lake and river sediments.* Australian Water Resources Council Technical Paper No. 33. Australian Government Publishing Service, Canberra. ,

Norris, R.H. (1986) Mine Waste Pollution of the Molonglo River, New South Wales and the Australian Capital Territory: Effectiveness of Remedial Works at Captains Flat Mining Area. *Australian Journal of Marine and Freshwater Research* 37, 147–157.

Norris, R.H. (1994) Rapid biological assessment, natural variability and selecting reference sites. In M.C. UYS (ed.) *Classification of Rivers and Environmental Health Indicators. Proceedings of a joint South African/Australian workshop, February 7–14, Cape Town, South Africa.* pp 129–166 Water Research Commission Report No. TT 63/94, Pretoria South Africa.

Norris, R.H., Lake, P.S. & Swain, R. (1982) Ecological effects of mine effluents on the South Esk River, north-eastern Tasmania 111. Benthic macroinvertebrates. *Australian Journal of Marine and Freshwater Research* 33, 789–809.

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

Pearson, R.G. & Penridge, L.K. (1987) The effects of pollution by organic sugar mill effluent on the macro-invertebrates of a stream in tropical Queensland, Australia. *Journal of Environmental Management* 24, 205–215.

Pinder, L.C.V. & Farr, I.S. (1987b) Biological surveillance of water quality – 3. The influence of organic enrichment on the macroinvertebrate fauna of small chalk streams. *Archiv für Hydrobiologie* 109, 619–637.

Plafkin, J.L., Barbour, M.T., Kimberly, D.P., Gross, S.K. & Hughes, R.M. (1989) *Rapid Bioassessment Protocols for Use in Streams.* United States Environment Protection Agency, Washington, DC, USA.

Resh, V.H. & Jackson, J.K. (1993) Rapid assesment approaches to biomonitoring using benthic macroinvertebrates. *Freshwater Biomonitoring and Benthic Macroinvertebrates.* (Eds D.M. Rosenberg & V.H. Resh), pp. 195–233. Chapman and Hall, New York, NY, USA.

Schofield, N.J. & Davies, P.E. (1996) Measuring the health of our rivers *Water* 23 (May/June), 39–43.

Slooff, W. (1983) Benthic macroinvertebrates and water quality assessment: some toxicological considerations. *Aquatic Toxicology* 4, 73–82.

Stark, J.D. (1985) A macroinvertebrate community index of water quality for stony streams. *Water and Soil Misc. Publ* 87, 1–53.

Stark, J.D. (1993) Performance of the Macroinvertebrate Community Index: effects of sampling method, sample replication, water depth, current velocity, and substratum on index values. *New Zealand Journal of Marine and Freshwater Research* 27, 463–478.

Stein, J.L., Stein, J.A., Nix, H.A. & Hutchinson, M.F. (in prep.) *The identification of Wild Rivers in Australia: database development and methodology*. CRES Report Centre for Resource and Environmental Studies, Australian National University., Canberra, ACT, Australia.

Stockwell, D.R.B. & Faith, D.P. (1996) *Investigation of alternative approaches to linking habitat variables with site classification in a RIVPACS model – Final Report.* Land & Water Resources Reseach & Development Corporation, Canberra, ACT.

US Environmental Protection Agency (1996) *Revision to rapid bioassessment protocols for use in streams and rivers: periphyton, benthic invertebrates, and fish.* Draft revision Office of Wetlands, Oceans, and Watersheds, United States Environmental Protection Agency, http://www.epa.gov/owowwtr1/monitoring/AWPD/RBP/bioasses.html.

Victorian Environmental Protection Authority (1995) *Protecting water quality in the Yarra catchment.* Publication No. 471. Victorian Envirnomental Protection Authority, Melbourne, Victoria.

Watson, J.A.L., Arthington, A.H. & Conrick, D.L. (1982) Effect of sewage effluent on dragonflies (Odonata) of Bulimba Creek, Brisbane. *Australian Journal of Marine and Freshwater Research* 33, 517–528.

Wright, J.F., Armitage, P.D. & Furse, M.T. (1989) Prediction of invertebrate communities using stream measurements. *Regulated Rivers: Research & Management* 4, 147–155.

Wright, J.F., Armitage, P.D., Furse, M.T. & Moss, D. (1988) A new approach to the biological surveillance of river quality using macroinvertebrates. *Verhandlungen Internationale Vereinigung für Theoretische und Angwante Limnologie* 23, 1548–1552.

Wright, J.F., Furse, M.T. & Armitage, P.D. (1994) Use of macroinvertebrate communities to detect environmental stress in running waters. *Water quality and stress indicators in marine and freshwate ecosystems: linking levels of organisation – individuals, populations, communities.* (Ed D.W. Sutcliffe), pp. 15–34. Freshwater Biological Association, Ambleside, U.K.

Wright, J.F., Furse, M.T., Clarke, R.T. & Moss, D. (1991) *Testing and further development of RIVPACS*. IFE Interim Report to the National Rivers Authority No. R&D 243/1/Y. National Rivers Authority, Bristol, U.K.

Wright, J.F., Furse, M.T., Clarke, R.T., Moss, D., Gunn, R.J.M., Blackburn, J.H., Symes, K.L., Winder, J.M., Grieve, N.J. & Bass, J.A.B. (1995) *Testing and Further Development of RIVPACS.* National Rivers Authority, Bristol, U.K.

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

Yoder, C.O. & Rankin, E.T. (1995) Biological response signatures and the area of degradation value: new tools for interpreting multimetric data. *Biological assessment and criteria. Tools for water resource planning and decision making.* (Eds W.S. Davis & T.P. Simon), pp. 263–286. Lewis, Boca Raton, FL, USA.

appendix A

list of participants in regional workshops

Perth

Organiser: Dr Stuart Halse	*MRHI Group, CALM*
Rob Donahue	Water and Rivers Commission
Jeff Kite	Water and Rivers Commission
Dr Luke Pen	Water and Rivers Commission
Tom Rose	Water and Rivers Commission
John Ruprecht	Water and Rivers Commission
Christine Easton	Agriculture WA
Jerry Parlevliet	Agriculture WA
Andrew Higham	Department of Environmental Protection
Winston Kay	MRHI group, CALM
Adrian Pinder	MRHI group, CALM
Mick Smith	MRHI group, CALM
Jim Lane	Department of Conservation and Land Management
Dr Don Edward	University of Western Australia
Dr Jenny Davis	Murdoch University
Dr Pierre Horwitz	Edith Cowan University
Kim Richardson	Edith Cowan University

Adelaide

Organiser: Ian Kirkegaard	*EPA (SA)*
Jenny Neill	SA Museum
Nicola Begg	KESAB
Jonathan Noble	KESAB
Kathryn Bellette	EPA (SA)
John Cugley	EPA (SA)
Peter Goonan	EPA (SA)
Bob McLennan	EPA (SA)
Ben Moretti	EPA (SA)
Geoff Fisher	Torrens & Patawalonga Catchment Water Board
Jim Burston	DENR
Andrew Emmett	DENR (Water Resources)
Michael Good	DENR (Water Resources)
Stuart Pillman	DENR
John Rolls	DENR
Greg Rowberry	DENR
Bob Myers	Upper River Torrens Landcare Group
Chris Madden	SA Water

Peter Schultz	SA Water
Tim Thompson	SA Water
Amy Blaylock	University of Adelaide
Sarah Farrelly	University of Adelaide
Dr Keith Walker	University of Adelaide
Prof Bill Williams	University of Adelaide
Jill Brooks	Natural Resources Council
Jackie Venning	State of the Environment Dept
Rebecca Fisher	Fishing Industry Council
Ian Grant	Waterwatch
Stuart Pillman	Wildlife
Gerry Butler	Mt Compass
Leon Broster	Murray Darling Association
Jane Suggit	NT Power and Water Authority
Michael Lawton	NT Power and Water Authority

Brisbane

Organiser: Peter Thompson	*Department of Natural Resources*
Don Alexander	Department of Natural Resources
Don Begbey	Department of Natural Resources (Catchment Management)
Sam Brown	Department of Natural Resources (Catchment Groups)
Gary Burgess	Department of Natural Resources
Satish Choy	Department of Natural Resources
Jim Dale	Department of Natural Resources (Operations Manager)
David Free	Department of Natural Resources
Heather Hunter	Department of Natural Resources
Rena Lloyd	Department of Natural Resources
Glen Moller	Department of Natural Resources (State of the Rivers Project)
George Raymond	Department of Natural Resources (CRC Sugar Production)
Chris Robinson	Department of Natural Resources (Planning)
Jim Roblatin	Department of Natural Resources
Dawn Talber	Department of Natural Resources
Brian Bycroft	Brisbane City Council
Paul Macwitz	Brisbane City Council
Prof Angela Arthington	Griffith University
Christy Coffey	Department of the Environment
Andrew Moss	Department of the Environment
Denise Tranier	Department of the Environment
Glennis Spence	Candanni
Bob Simpson	Department of Primary Industries (Freshwater Fisheries)
Rob Laylor	Total Catchment Management Group (South Johnson)
Roger Shaw	DSS
Peter Justin	No affiliation given

Sydney

Organiser: Russell Cowell	*EPA*
Tony Church	EPA (Inland Catchments)
Derek Elms	EPA
Helen Hoffmann	EPA
Trevor Jones	EPA (South Coast)
Dr Klaus Koop	EPA
Christian Mercer	EPA
Dr Euen Turak	EPA
Andrew Brooks	Department of Land and Water Conservation
Richard Denben	Department of Land and Water Conservation
Michael Mohoney	Department of Land and Water Conservation
Dr Peter Cornish	NSW State Forests
Robert Towler	NSW State Forests
Dr Alan Jones	Australian Museum
David Robinson	Australian Museum
Peter Serup	Australian Museum
Alan Dodds	Sydney Water
Peter Schneider	Sydney Water
Peter Wilson	NSW National Parks
Dr Peter Gherke	NSW Fisheries/ CRC for Freshwater Ecology
Peter Code	Catchment Management Committee
Peter Liston	ACT Department of Environment and Land Protection

Melbourne

Organiser: Lisa Dixon	*EPA*
Dr Peter Breen	Melbourne Water/ CRC for Freshwater Ecology
Dr Chris Walsh	CRC for Freshwater Ecology
Marryn Kelly	Department of Natural Resources and Environment
Paul Wilson	Department of Natural Resources and Environment
Chris Bell	EPA
Lisa Dixon	EPA (Freshwater Science)
Keston Osmars	EPA (Dandenong)
Narelle Martin	EPA (North West)
David McKenzie	EPA (Gippsland)
Leon Metzeling	EPA
Leonie Walker	EPA (North West Region)
Chris Chesterfield	Melbourne Water
Vince Pettigrove	Melbourne Water
Dr Kumar	Water Ecoscience
Tim Doeg	MFRI
Ross Wilson	Waterwatch (Geelong)
Ian Everist	CLPC
Stan Pickering	Mallee CLP Board

appendix B

source material from regional workshops

The following was the background material used for each of the regional workshops. This material was distributed to state organisers prior to the workshops.

National River Health Program

Interpreting the output of the National River Bioassessment Protocol to support management decisions

Background paper for the State and Territory workshops, August 1996

Leon Barmuta, University of Tasmania

Barry Hart, CRC for Freshwater Ecology and Monash University

Bruce Chessman, Australian Water Technologies

what is the National River Health Program?

The National River Health Program (NRHP) was established by the Commonwealth Government in 1992.

Its goals are:

- To monitor and assess the health of Australian rivers.
- To enhance the management of river flows, water allocation and water use to ensure the sustainability of river and floodplain ecosystems.
- To encourage active management to improve the health of Australia's rivers, based on a sound understanding of ecological and hydrological processes.
- To evaluate the effectiveness of river management actions on a national scale.

The NRHP is administered by:

- Land and Water Resources Research & Development Corporation
- Environment Australia
- Urban Water Research Association of Australia.

what is the National River Bioassessment Protocol?

The National River Bioassessment Protocol (NRBP) is the set of scientific methods used in the State and Territory component of the Monitoring River Health Initiative (MRHI). The MRHI is part of the NRHP.

The protocol is based on the successful British RIVPACS model (River InVertebrate Prediction and Classification Scheme). It has the following main features:

- **Studying** many **reference** (least disturbed) and **test** (possibly polluted or otherwise unhealthy) river **sites.**
- **Sampling macroinvertebrates** from precisely defined habitats (eg. riffles, macrophyte beds) using rapid, semi-quantitative methods at each site.
- **Identifying the macroinvertebrates** collected (to family level rather than species in most States).
- Measuring **natural environmental features** of the sites such as altitude and stream width.
- **Classifying** the reference sites into groups such that sites in each group have a similar type of macroinvertebrate community.
- **Building a predictive model** that can determine the macroinvertebrate families likely to be present at an unpolluted site with particular natural environmental features.
- **Using the model** to output the macroinvertebrates expected at a test site (if it is not impaired).
- **Comparing** the **actual or observed** macroinvertebrate community at the test site with the expected natural community.
- **Calculating an index** to measure the difference between the actual and predicted communities.
- Using the index to **assign the site to an environmental quality category**, eg. unimpaired, slightly impaired, grossly impaired.
- Using macroinvertebrate and environmental information to make **inferences about causative factors** responsible for reduced river health, and so point the way to **remedial management action** (decision making).

why use bioassessment?

- **Australia's rivers are degraded and deteriorating.** The increasing occurrence of bank erosion and collapse, sediment deposition and infilling, algal and cyanobacterial blooms, loss of native fish and invertebrate populations and invasion of carp are signs of decline.
- **Past programs have failed.** Based solely on physical, chemical and microbiological assessment of water quality, these programs often have not protected rivers and their biological resources. The physical, chemical and microbial constituents of wastewater discharges are only part of the causes of damage to our rivers. The past emphasis on regulation of these constituents has resulted in other factors such as land degradation, diffuse pollution and the effects of flow changes being overlooked.
- **Without bioassessment, management action is hard to decide on and prioritise.** We lack the holistic understanding of the systems we aim to protect. As a result we cannot determine the measures needed to protect them, or direct the available funds to the most cost-effective measures. This sometimes results in large amount of money being spent for no appreciable environmental benefit.

why use macroinvertebrates?

- Macroinvertebrates are **widely and successfully used** in Britain, continental Europe and North America for biological assessment of rivers.
- The **technology and expertise** for macroinvertebrate-based assessment are **well established** and available throughout Australia.

- Macroinvertebrate sampling and analysis are **technically feasible and affordable** for covering a large number of sites with limited resources.
- We know from past case studies that Australian river **macroinvertebrates are sensitive** to the cumulative impacts of a wide range of disturbances. These include flow alteration, urban and agricultural development, salinisation, siltation, sewage and other organic wastes and heavy metals.

Macroinvertebrates are not sensitive to all human impacts, however. For example, macroinvertebrates do not respond to human pathogens or barriers that affect fish migration. Macroinvertebrate bioassessment should therefore be implemented as part of an integrated monitoring program including physical, chemical, microbial and other forms of bioassessment appropriate to local conditions.

Physical and chemical studies are an essential adjunct to macroinvertebtrate monitoring in order to identify causative factors, and are part of the MRHI.

objectives of this project and the workshops

This project addresses the last two components of the bioassessment protocol listed above. That is, it aims to apply indices to combined test site survey data and outputs from the predictive models in order to assign environmental quality ratings and guide management action.

The formal aims are:

- To compare and evaluate numerical techniques for summarising and interpreting the output of the national river bioassessment protocol.
- To develop appropriate delivery systems that maximise the interpretability of the national river bioassessment protocol.

The objectives of these workshops are to obtain the views and preferences of potential users (both managers and scientists) on:

- Suggested indices
- Methods and formats for environmental quality ratings (EQRs)
- Desirable formats and media for presentation of EQRs.

the raw output format

The MRHI predictive models are being developed by the various State and Territory agencies with support from a team at the CRC for Freshwater Ecology led by Dr Richard Norris.

The raw output from the model and the survey data from a test site will typically be in a form similar to the following example (the number of families has been greatly reduced for purposes of illustration).

Barramutta River – riffle sample – March 1997

FAMILY	MODELLED PROBABILITY OF OCCURRENCE (Y)	ACTUAL OCCURRENCE IN SAMPLE (Z)
Aeshnidae	0.4	–
Baetidae	**0.8**	1
Chironomidae	**0.9**	1
Dugesiidae	–	1
Ephydridae	–	1
Glossosomatidae	**0.6**	–
Hydridae	0.1	–
Isostictidae	**0.6**	1
Janiridae	0.3	–
Kokiriidae	**0.8**	–

In this example, the model output lists eight families that have some chance of occurring in a riffle in the Barramutta River in autumn if the river is in its natural state (expected probability of occurrence greater than zero). Only five of these eight have a probability greater than 0.5, that is the model predicts they are more likely to occur than not to.

In the actual sample five species were observed (those with a '1' for 'present'). Some of these had a high probability of occurrence in the model prediction (they were expected) and some had a low or zero probability (they were not expected).

This sort of output is not likely to be very useful to managers, especially when dozens of sites are involved. It is obvious that the observed families were not quite what was expected, but how big and significant is the departure from expectation?

Some sort of index needs to be applied to the data to summarise the difference between what was observed and what was expected for a sample of this type from a natural river.

the ideal index?

What features do we need in the ideal index? We believe that the following ones are important:

Sensitivity: it should discriminate between sites with various levels of environmental quality.

Reliability: it should accurately reflect the environmental condition of a site.

Robustness: it should not be unduly influenced by natural fluctuations in the biota.

Generality: it should not rely on faunal features that may be inapplicable to particular parts of Australia.

Interpretability: it should indicate not only the degree but also the type of impact at a site.

Objectivity: it should not depend on subjective interpretation by the user.

These criteria have been used in a preliminary assessment of alternative indices leading to the indices proposed.

a simple index: *O/E* FAMILY OCCURRENCE

A simple index can be constructed by comparing the families that were actually observed with the families expected from the model prediction. This is called observed over expected (*O/E*) family occurrence. It is illustrated below using the same example as before.

Barramutta River – riffle sample – March 1997

FAMILY	MODELLED PROBABILITY OF OCCURRENCE (Y)	ACTUAL OCCURRENCE IN SAMPLE (Z)
Aeshnidae	0.4	–
Baetidae	**0.8**	1
Chironomidae	**0.9**	1
Dugesiidae	–	1
Ephydridae	–	1
Glossosomatidae	**0.6**	–
Hydridae	0.1	–
Isostictidae	**0.6**	1
Janiridae	0.3	–
Kokiriidae	**0.8**	–
Total	***4.5***	***5***
Number of families expected (P > 0.5)	*5*	
Sum of probabilities of expected families	*3.7*	
Number observed of those expected		*3*

O/E **FAMILY OCCURRENCE = 3/3.7 = 0.81**

There are various possible ways of calculating the ratio of observed to expected families, but for this example we have divided the number of expected families that were actually observed by the sum of probabilities of the expected families. It seems that the river is not too bad as our observed is 81% of our expected.

However, this index has some drawbacks:

- It does not take account of the types of families that were expected but were missing. Why were these particular families not there? What does their absence tell us about the types of impact that may be occurring at the site?
- It does not consider the families that we observed that were not expected. What does their presence tell us about impacts?

an alternative index: *O/E* SIGNS

The SIGNS index stands for Stream Invertebrate Grade Number Sum.

A grade number between 1 and 10 is assigned to each family of macroinvertebrates. A high number means that the family is sensitive to a particular pollutant such as sewage. A low number means that the family is tolerant. SIGNS is the sum of the grade numbers of all the families expected or observed.

Using the same example as before, we can compare the observed SIGNS score to the expected SIGNS score (*O/E* SIGNS).

Barramutta River – riffle sample – March 1997

FAMILY	SENSITIVITY NUMBER (X)	PROBABILITY OF OCCURRENCE (Y)	CALCULATE EXPECTED SIGNS INDEX	ACTUAL OCCURRENCE IN SAMPLE (Z)	CALCULATE OBSERVED SIGNS INDEX
			x •y	x •	z
Aeshnidae	6	0.4	2.4	–	–
Baetidae	5	**0.8**	4.0	**1**	5.0
Chironomidae	1	**0.9**	0.9	**1**	1.0
Dugesiidae	3	–	–	**1**	3.0
Ephydridae	2	–	–	**1**	2.0
Glossosomatidae	8	**0.6**	4.8	–	–
Hydridae	4	0.1	0.4	–	–
Isostictidae	7	**0.6**	4.2	**1**	7.0
Janiridae	5	0.3	1.5	–	–
Kokiriidae	10	**0.8**	8.0	–	–
Total		***26.2***	***18.0***		

***O/E* SIGNS = 18.0/26.2 = 0.69**

The expected SIGNS is calculated by multiplying the model probability of occurrence of each family by its grade number, and summing the products. The observed SIGNS is calculated by summing the grades for each family present. Dividing the observed SIGNS by the expected SIGNS gives the *O/E* SIGNS.

In this example the *O/E* SIGNS is only 0.69, so it seems that the river may actually be polluted by sewage, or affected by some similar impact.

However, one disadvantage of SIGNS is that because it is a sum, it depends on sampling effort. If the river was sampled by someone who was particularly good at finding rare families they would calculate a higher observed SIGNS value than someone who found only the abundant families.

a standardised index – *O/E* SIGNAL

SIGNAL stands for Stream Invertebrate Grade Number Average Level.

SIGNAL differs from SIGNS in that it is the average of the grade numbers of the families present rather than their sum. So it is not dependent on sampling effort.

We can again use the same example to compare the observed SIGNAL score with the expected SIGNAL score (*O/E* SIGNAL).

Barramutta River – riffle sample – March 1997

FAMILY	SENSITIVITY NUMBER (X)	PROBABILITY OF OCCURRENCE (Y)	CALCULATE EXPECTED SIGNS INDEX	ACTUAL OCCURRENCE IN SAMPLE (Z)	CALCULATE OBSERVED SIGNS INDEX
			x •y	x •	z
Aeshnidae	6	0.4	2.4	–	–
Baetidae	5	**0.8**	4.0	**1**	5.0
Chironomidae	1	**0.9**	0.9	**1**	1.0
Dugesiidae	3	–	–	**1**	3.0
Ephydridae	2	–	–	**1**	2.0
Glossosomatidae	8	**0.6**	4.8	–	–
Hydridae	4	0.1	0.4	–	–
Isostictidae	7	**0.6**	4.2	**1**	7.0
Janiridae	5	0.3	1.5	–	–
Kokiriidae	10	**0.8**	8.0	–	–
Total		***4.5***	***26.2***	***5***	***18.0***
Average			*5.8*		*3.6*

O/E SIGNAL = 3.6/5.8 = 0.62

The expected SIGNAL is calculated by multiplying the model probability of occurrence of each family by its grade number, summing the products and dividing by the total of the probabilities. The observed SIGNAL is calculated by summing the grades for each family present, and dividing by the number of families. Dividing the observed SIGNAL by the expected SIGNAL gives the *O/E* SIGNAL.

In this example the *O/E* SIGNAL is only 0.62, so again it seems that the river may actually be polluted by sewage, or affected by some similar impact.

a combination of indices?

It is possible that a combination of the three indices could be used to provide some diagnosis of the type of impact at a test site. The following table shows the hypothetical responses of sensitive and tolerant families and each index to different types of impact.

Hypothetical effects of different impacts on number of sensitive and tolerant families (P = No. predicted) and index values

CONDITION OF TEST SITE	SENSITIVE FAMILIES	TOLERANT FAMILIES	O/E FAMILY OCCURRENCE	O/E SIGNS	O/E SIGNAL
Average of reference	P	P	1	1	1
Better than average	> P	< P	1	> 1	> 1
Mild nutrient/organic enrichment	> P	> P	> 1	> 1	1
Strong nutrient/organic enrichment Altered flow; sediment	< P	> P	1	< 1	< 1
Toxicants	< P	< P	< 1	< 1	< 1

At reference sites both sensitive and tolerant families are present. At a test site that is similar to an average reference site, the sensitive and tolerant families present will be much the same as predicted. Therefore all index values will be approximately 1.

At a test site that is actually in better condition than the average reference site, there will be more sensitive families and fewer tolerant families. This should not affect the *O/E* occurrence index because it is not based on family sensitivity. However, the other indices will assume values greater than 1 because they increase with greater representation of sensitive families relative to tolerant families.

In cases of mild enrichment of rivers by nutrients and organic matter, such as occur with mild sewage and agricultural impact, both sensitive and tolerant families become more prevalent. This is because they benefit from increased food sources. In such cases *O/E* family occurrence and *O/E* SIGNS will be greater than 1. However, *O/E* SIGNAL should not change because it is an average, not a sum.

In cases of stronger nutrient and organic enrichment, such as below sewage treatment works that do not discharge toxicants, or in more heavily altered agricultural areas, sensitive families start to be replaced by tolerant ones. *O/E* family occurrence should not be greatly affected in such cases, but may decrease slightly if the tolerant replacement families are few, or are ones that are uncommon at reference sites. The other indices will decline sharply. Similar responses are likely in streams affected by altered flow regimes and sediment deposition.

In instances of toxic pollution, such as rivers receiving sewage effluents such as ammonia or chlorine, or mining wastes containing heavy metals, families will be lost with little replacement. In such cases all indices should decline.

These hypotheses are yet to be tested with actual NRHP data.

banding of index values

Index values by themselves have little meaning unless they can be associated with 'bands' indicating various degrees of impairment of river health. The following table gives an example of a possible banding scheme.

Example of a possible 'banding' scheme for *O/E* SIGNAL

O/E SIGNAL SCORE	BAND	INFERRED CONDITION OF THE MACROINVERTEBRATE COMMUNITY
> 1.05	A1	Similar to reference (above average)
0.85–0.95	A2	Similar to reference (below average)
0.65–0.75	B	Mildly impaired
0.45–0.55	C	Moderately impaired
0.25–0.35	D	Severely impaired
< 0.15	E	Grossly impaired

Note that in this scheme, there are gaps between the bands. For example, an *O/E* SIGNAL score of 0.60 falls between bands B and C. Sites with such borderline scores would require further assessment before they were categorised, for example resampling or examination of physical and chemical data. This would reduce the risk of misclassification.

However, occasional misclassification is unavoidable. Index scores, like any measurement, are subject to measurement error and statistical variability.

There are various statistical means available for setting band limits. For example, the lower limit of band A could be set at one standard deviation below the mean score for reference sites. This would result in approximately 5% of reference sites being banded below band A.

The statistical criteria used will determine the rates of Type 1 and Type II errors.

Type I errors are 'false alarms', when we conclude that a site is impaired when it really is not. Putting a reference site below band A is a Type I error.

Type II errors are 'false security', when we conclude that a site is not impaired when it really is. Putting a polluted site in band A would be a Type II error.

Both Type I and Type II errors will be unwanted by managers. The first could lead to management action being taken unnecessarily, and the second would result in a problem not being addressed.

Unfortunately, it is a statistical fact of life that for a given sampling effort, as we move the band limits to reduce Type I errors we increase Type II errors, and *vice versa*.

We need feedback from managers on which type of error is most unacceptable.

displaying the findings

Once bands are assigned, and inferences are made about the types of impacts causing each site to be banded as it is, we are some way to answering the key management questions:

- Is this site healthy or not?
- If it is unhealthy, how unhealthy?
- If it is unhealthy, why is it unhealthy?
- What can we do to fix it?

To provide the answers to these questions we need to decide on ways of displaying the bands and other information. There are various possibilities such as using spatially referenced information (eg. GIS technology).

Options will be canvassed at the workshops.

appendix C
results of analyses for banding criteria

introduction

The final decisions about the number of bands used in reporting and the values used to designate band cut-offs needed to consider both the reporting requirements identified by potential end-users in the regional workshops and the constraints imposed by the variability amongst reference sites in the final models.

methods

The analyses were restricted to those data sets for which final combined seasons models were available within the timelines imposed by this project and the necessity to deliver final recommendations in a timely manner to A/Prof. Norris's group so that band boundaries could be incorporated into the final software package for AUSRIVAS Accordingly combined seasons models for the ACT, Tasmania and WA were used. These data sets encompassed a wide variety of environments and model types. The ACT was from a cool to warm temperate area, while the WA data included information from temperate, Mediterranean climates, semi-arid regions and the wet-dry tropics. Both sets of models showed well-defined group structure, with the ACT data expected to show the least variation amongst index values for reference sites and WA the greatest. The data from Tasmania was from a cool, wet temperate climate but the group structure in the classification phase of the analysis was less well defined, suggesting that the reference sites formed more of a continuum rather than the discrete groups evident in the other data sets. Including these data in the analysis provided information about whether such continuous faunal patterns prevented useful outputs from AUSRIVAS using the present RIVPACS-like procedures.

Data from the ACT and Tasmania also included a variety of test sites with known impacts so that the performance of the chosen indices could be verified. No data from test sites were available for WA. As criteria for the number and width of bands were developed, these were checked against values emerging for models as they were developed for the remaining States and Territories so that the rules evolved would be consistent nationally.

There were three steps to these analyses. The first analysis was to determine whether *O/E* SIGNS was strongly correlated with *O/E* FAMILIES as suggested by Clarke *et al.* (1996) and therefore redundant. This was achieved by calculating the Pearson product-moment correlation coefficient between these two indices for the reference sites for each model. Scatterplots were also inspected to detect any non-linear relationships. The distributions of the indices for reference sites in each model were inspected using histograms and normal probability plots. If the distributions were close to normal, then using a multiple of the standard deviation of the distribution for the reference sites would be a direct and easy method of setting the band width of the reference band, A. If the distributions showed non-normal behaviour, using percentile values would be a more robust approach. Once provisional band boundaries and widths had been set, the third set of analyses involved examining the results from the test sites in the ACT and Tasmanian data to test different scenarios for band widths and boundaries and different rules and procedures for combining habitat assessments. These analyses were initially performed 'blind' (ie. without knowledge of the nature of the human disturbances at the test sites), and the results and interpretations checked with the providers of the data. Because of the small number of test sites available, formal statistical procedures were not used in this third set of analyses.

redundancy of *O/E* SIGNS

O/E SIGNS was strongly correlated for all models (smallest Pearson's correlation coefficient, r = 0.894; see Table C1) and there were no curvilinear relations. For all references sites combined across all the models the relationship remained strongly linear (r = 0.916; Figure C1)

Table C1 Pearson product-moment correlation coefficients for each combined seasons model

STATE OR TERRITORY	HABITAT	NO. OF REFERENCE SITES	CORRELATION, *r*
WA	Macrophyte	73	0.907
	Channel	126	0.894
Tasmania	Edge	82	0.949
	Riffle	82	0.983
ACT	Edge	41	0.948
	Riffle	35	0.919

Figure C1 Scatterplot of *O/E* SIGNS v. *O/E* FAMILIES for all reference sites from the combined seasons models for WA, Tasmania and ACT

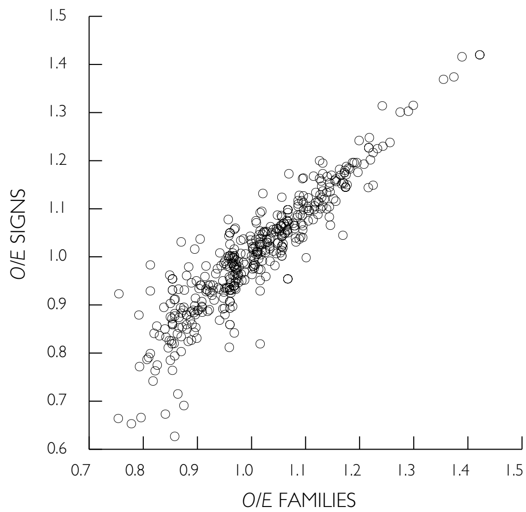
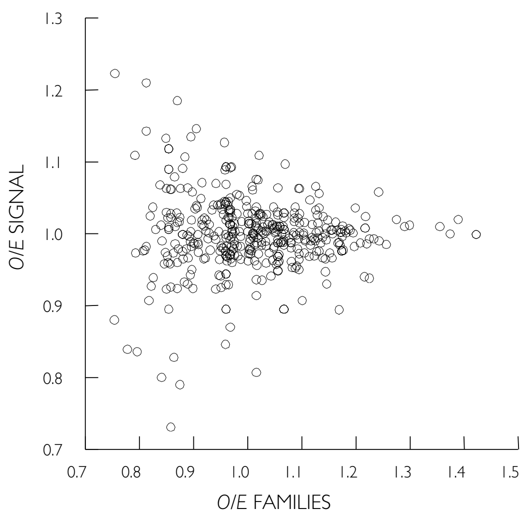


Figure C2 Scatterplot of *O/E* SIGNAL v. *O/E* FAMILIES for all reference sites combined seasons models for WA, Tasmania and ACT.



From these results, it is clear that *O/E* SIGNS is redundant, since it simply weights the component families in *O/E* FAMILIES by the SIGNAL value resulting in a slightly wider range than *O/E* FAMILIES. This is similar to the result of Clarke *et al.* (1996) for the equivalent indices of RIVPACS in Great Britain. Consequently there was no value in retaining this index, and no further analyses were performed using it. By contrast, *O/E* SIGNAL promises to provide different information particularly at lower values of *O/E* FAMILIES (Figure D2).

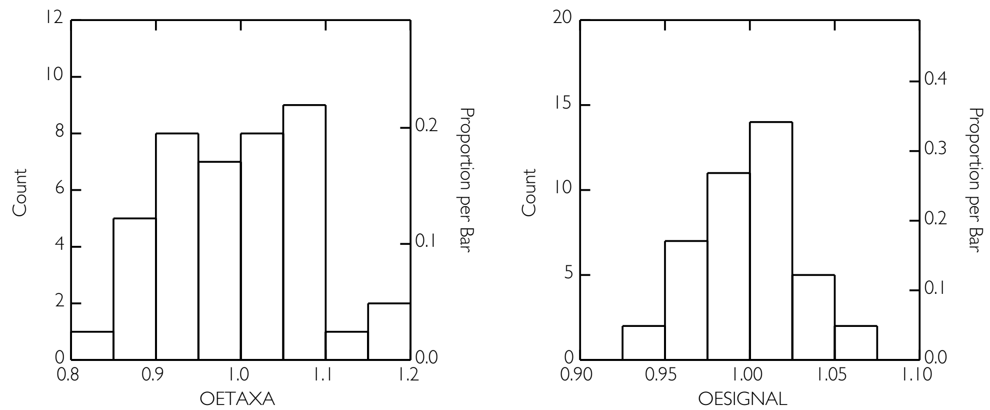
departures from normality

Although some of the distributions of the indices *O/E* FAMILIES and *O/E* SIGNAL were approximately normal, some showed either left- or right-skewness (Figure D3). Accordingly using a multiple of the standard deviation may result in either too high or low a value for the division between band A and band B. Use of designated percentile values is therefore the more robust option. The only potential disadvantage of this procedure is that for models with a small number of reference sites (say <50), percentiles may fall into large gaps in the actual data values.

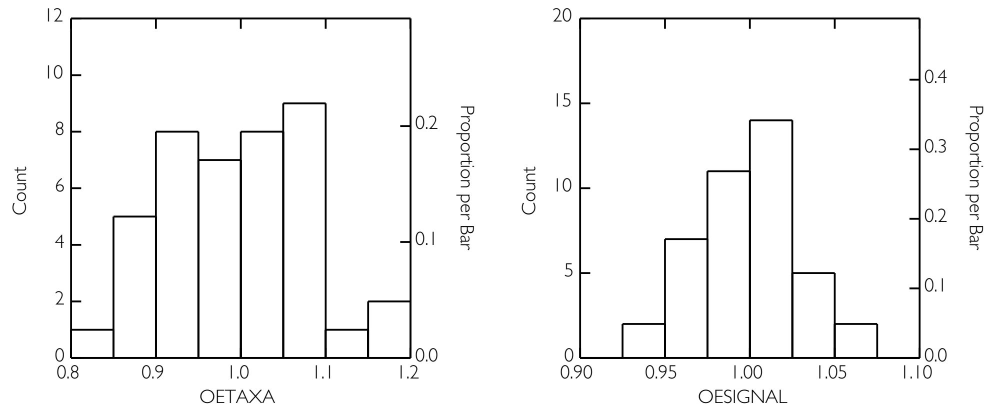
On the following pages histograms are presented for each model for the two indices together with normal probability plots. On the normal probability plots, points that deviate from the diagonal straight line indicated deviations from normality. OETAXA refers to *O/E* FAMILIES and OESIGNAL refers to *O/E* SIGNAL.

Figure C3

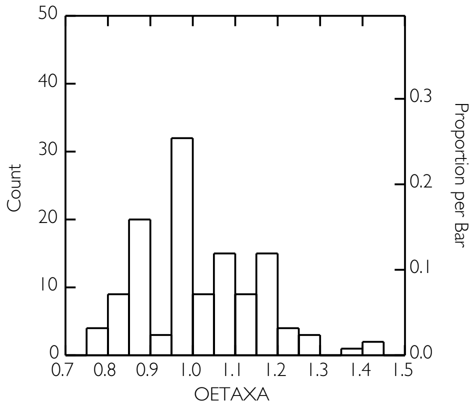
WA macrophyte

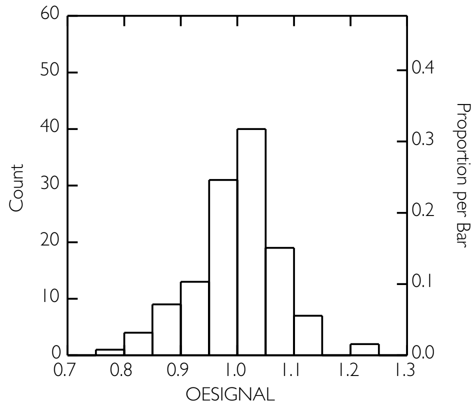


WA macrophyte normal probability plot.

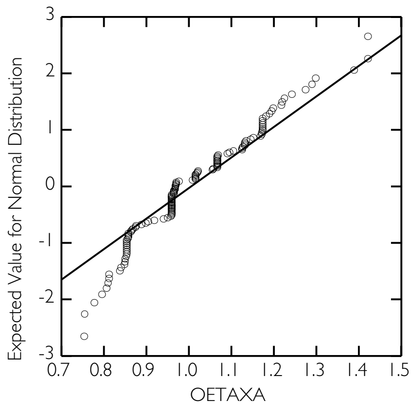


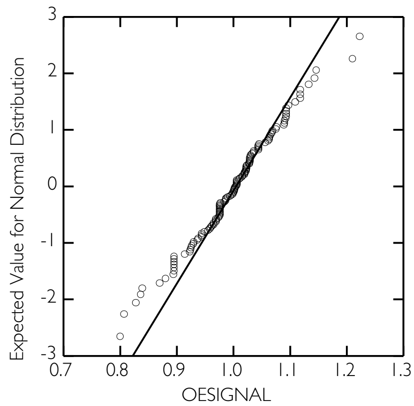
WA channel



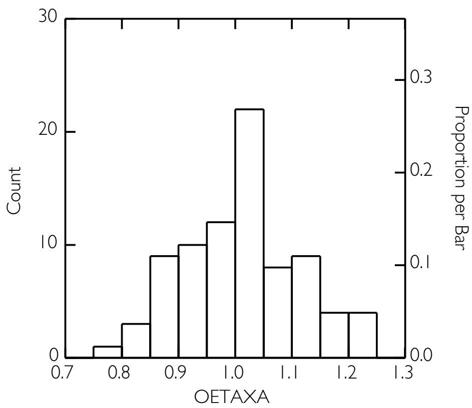


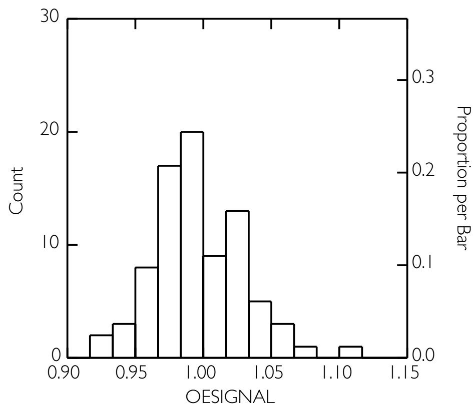
WA channel normal probability plot



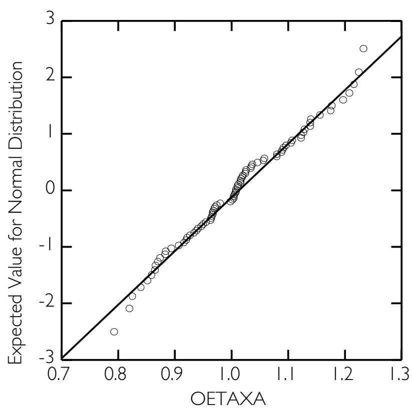


Tasmania edge





Tasmania edge. normal probability plot.



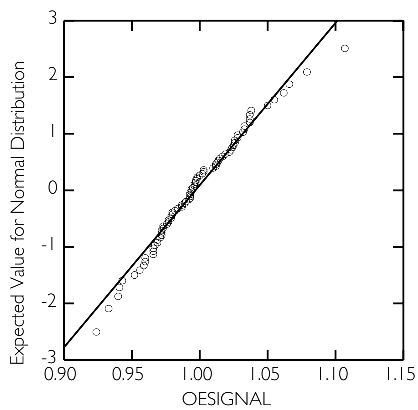
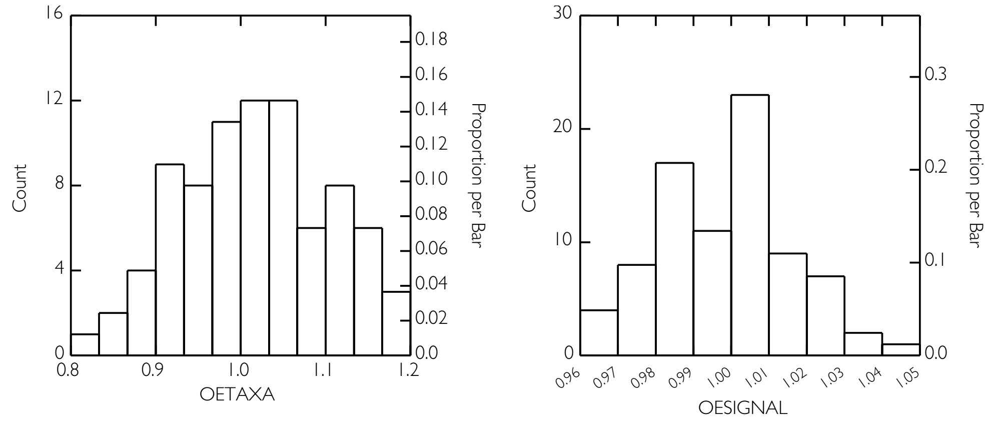
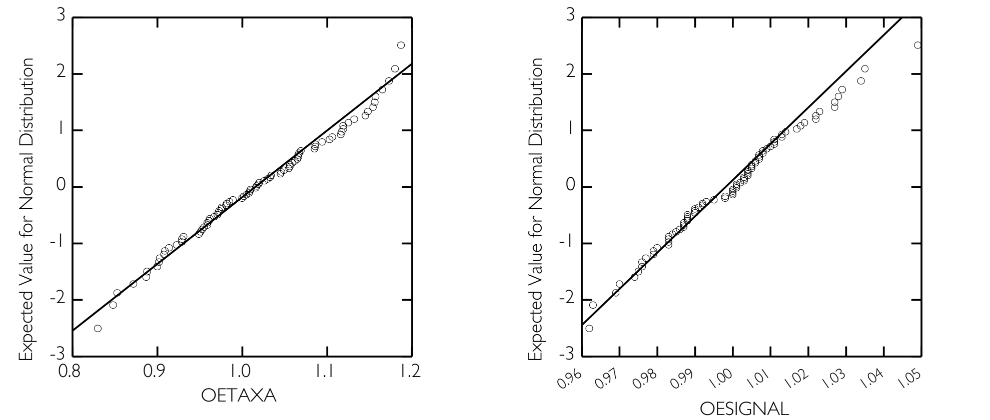


Figure C3

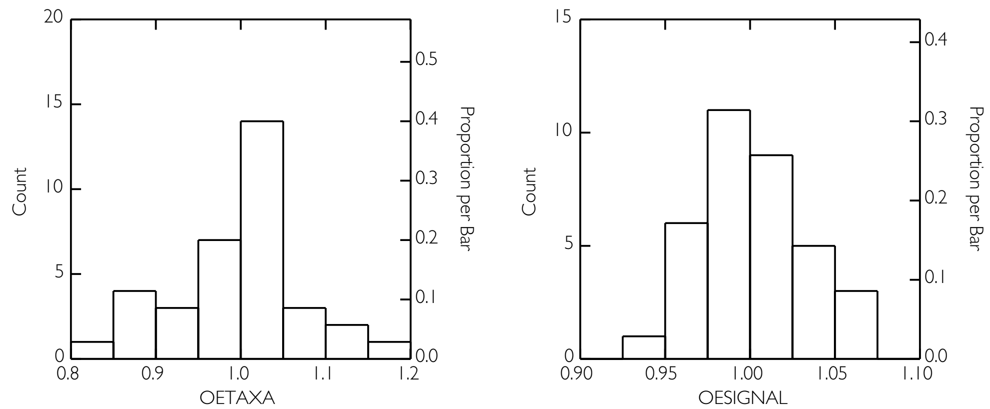
Tasmania riffle



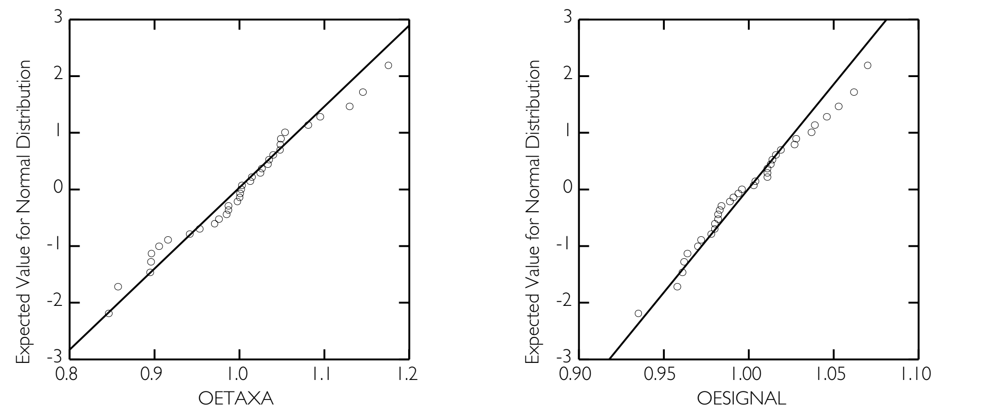
Tasmania riffle normal probability plot



ACT riffle



ACT riffle normal probability plot



options for selecting the band width

Table D2 compares the different options for defining band boundaries across the different models. The WA data sets were most variable and the ACT least variable. This probably reflects the environmental heterogeneity of the former and relative homogeneity of the latter. Accordingly the width of the lowest quality band is smallest for WA and widest for ACT. An alternative of dividing the index values below band A into equal intervals was considered but rejected because the widths of bands B and C would be unequal for different models, whereas national consistency of bandwidths was the consensus preference of participants in the regional workshops.

Table C2 Band boundaries for *O/E* FAMILIES using different criteria for the six models.

	WA MACROPHYTE			WA CHANNEL			ACT EDGE			ACT RIFFLE			TAS EDGE			TAS RIFFLE		
	4 × Standard deviation	5%-ile–95%-ile	10%-ile–90%-ile	4 × Standard deviation	5%-ile–95%-ile	10%-ile–90%-ile	4 × Standard deviation	5%-ile–95%-ile	10%-ile–90%-ile	4 × Standard deviation	5%-ile–95%-ile	10%-ile–90%-ile	4 × Standard deviation	5%-ile–95%-ile	10%-ile–90%-ile	4 × Standard deviation	5%-ile–95%-ile	10%-ile–90%-ile
Width of Band A	0.510	0.395	0.313	0.573	0.433	0.332	0.327	0.241	0.199	0.307	0.273	0.199	0.412	0.357	0.274	0.349	0.285	0.2423
No of bands possible	1.5	2.1	2.7	1.2	1.9	2.6	2.6	3.6	4.5	2.8	3.1	4.5	1.9	2.4	3.2	2.4	3.1	3.7
Upper bound of Band A	1.255	1.217	1.167	1.286	1.243	1.182	1.163	1.104	1.093	1.153	1.130	1.095	1.206	1.197	1.140	1.174	1.157	1.145
Lower bound of Band A	0.745	0.822	0.854	0.714	0.810	0.850	0.837	0.863	0.894	0.847	0.857	0.896	0.794	0.840	0.866	0.826	0.872	0.902
Lower bound of Band B+	0.490	0.625	0.698	0.427	0.594	0.684	0.673	0.743	0.795	0.693	0.721	0.797	0.588	0.662	0.729	0.651	0.730	0.781
Lower bound of Band B–	0.235	0.427	0.541	0.141	0.377	0.518	0.510	0.622	0.695	0.540	0.584	0.697	0.381	0.483	0.592	0.477	0.587	0.659
Lower bound of Band C+	–0.020	0.230	0.385	–0.145	0.161	0.352	0.347	0.502	0.596	0.386	0.448	0.598	0.175	0.305	0.455	0.303	0.445	0.538
Lower bound of Band C–	–0.275	0.032	0.228	–0.432	–0.056	0.186	0.183	0.381	0.496	0.233	0.311	0.498	–0.031	0.126	0.318	0.128	0.302	0.416

Width of Band A refers to the width of the similar to reference band for each index for *O/E* FAMILIES (2 × standard deviation (column headed 4 × Standard deviation), using the 5th and 95th percentiles and the 10th and 90th percentiles. The values for *O/E* FAMILIES for the lower bound of each of bands B and C are given. These two bands are subdivided in half to form, for example band B+ and B-. This subdivision is not recommended for use in formal reporting. The preferred option is not greyed-out.

performance of indices with known test sites

Generally there was fair agreement between band allocation based on the two indices, with over 60% of the test sites from the Tasmanian and ACT models being allocated to identical bands on the two indices within each habitat type. For the mis-matches neither *O/E* FAMILIES nor *O/E* SIGNAL banded consistently higher or lower except for the ACT riffle model for which 6 test sites were banded higher on *O/E* SIGNAL compared with *O/E* SIGNAL. (However, 4 of these 6 sites were banded identically by both indices for the edge habitat.)

Although these results are encouraging, discussions with those who collected the data for the mis-matched test sites revealed further information that either confirmed the lower band allocation or suggested that the site would need reassessment if the recommended reporting procedure was followed because of unusual circumstances encountered during sampling or site selection.

references

Clarke, R.T., Furse, M.T., Wright, J.F. & Moss, D. (1996) Derivation of a biological quality index for river sites: comparison of the observed with the expected fauna. *Journal of Applied Statistics* 23, 311–332.

appendix D
SIGNAL grades

The following spreadsheet details the SIGNAL grades as they stand at present. '95' denotes the SIGNAL 95 grades; '97G' are generalised grades that are implemented in the AUSRIVAS software; and '95M' are the provisional grades for use when metals contamination. 95M grades are not yet available for all taxa, nor are they used in the AUSRIVAS software.

	Version	
Family name	95	97G
Aeshnidae	6	7
Ameletopsidae	10	8
Amphipterygidae	8	8
Amphisopidae		10
Ancylidae	6	5
Antipodoeciidae	10	10
Aphroteniinae		9
Arrenuridae		7
Athericidae	7	7
Atriplectididae	10	7
Aturidae		6
Atyidae	6	5
Austroperlidae	10	10
Baetidae	5	6
Belostomatidae	5	4
Bithyniidae		8
Blephariceridae	10	10
Branchipodidae		7
Brentidae		3
Caenidae	7	5
Calamoceratidae	8	8
Calocidae	8	9
Carabidae		4
Ceinidae	5	5
Ceratopogonidae	6	5
Chaoboridae		3
Chironomidae	1	1
Chironominae		1
Chlorolestidae		7
Chrysomelidae		6

	Version	
Family name	95	97G
Cirolanidae		5
Clavidae		2
Coenagrionidae	7	2
Coloburiscidae	10	10
Conoesucidae	8	9
Corbiculidae	6	4
Corduliidae	7	5
Corixidae	5	3
Corophiidae		3
Corydalidae	4	7
Culicidae	2	3
Curculionidae		6
Diamesinae		8
Dixidae	8	8
Dolichopodidae		5
Dugesiidae	3	3
Dytiscidae	5	3
Ecnomidae	4	5
Elmidae	7	8
Empididae	4	5
Enchytraeidae		6
Entomobryidae		1
Ephemerellidae		8
Ephydridae	2	2
Erpobdellidae	3	3
Eusiridae	8	8
Eustheniidae	10	10
Eylaidae		5
Gammaridae	6	6
Gelastocoridae	6	7

continued

Family name	Version 95	Version 97G
Gerridae	4	4
Glossiphoniidae	3	3
Glossosomatidae	8	8
Gomphidae	7	6
Gordiidae	7	7
Grapsidae		1
Gripopterygidae	7	9
Gyrinidae	5	5
Haliplidae	5	5
Haplotaxidae	5	5
Hebridae	6	6
Helicophidae	10	9
Helicopsychidae	10	10
Heteroceridae		2
Hydrachnidae		4
Hydraenidae	7	5
Hydridae	4	4
Hydrobiidae	5	4
Hydrobiosidae	7	9
Hydrochidae	7	5
Hydrodromidae		6
Hydrometridae	5	6
Hydrophilidae	5	4
Hydropsychidae	5	6
Hydroptilidae	6	5
Hydryphantidae		6
Hygrobatidae		7
Hygrobiidae	5	5
Hymenosomatidae		3
Hypogastruridae		1
Hypsimetapodidae		10
Hyriidae	6	5
Isostictidae	7	6
Isotomidae		1
Janiridae	5	5
Kokiriidae		9
Leptoceridae	7	8
Leptophlebiidae	10	10
Lestidae	7	6

Family name	Version 95	Version 97G
Lestoideidae		5
Libellulidae	8	5
Limnephilidae	8	9
Limnesiidae		6
Limnichidae		4
Limnocharidae		6
Lumbriculidae	1	1
Lymnaeidae	3	3
Megapodagrionidae	7	7
Mesamphisopidae		7
Mesoveliidae	4	2
Microsporidae		6
Mideopsidae		8
Momoniidae		9
Muscidae	3	4
Naididae	1	2
Nannochoristidae	10	8
Naucoridae	5	4
Neoniphargidae		10
Nepidae	5	5
Neurorthidae	8	8
Noteridae		5
Notodromadidae		3
Notonectidae	4	4
Notonemouridae	8	9
Ochteridae		4
Odontoceridae	8	8
Oeconesidae		10
Oniscidae		6
Oniscigastridae	10	10
Onychiuridae		1
Ornithobdellidae		1
Orthocladiinae		2
Osmylidae	8	8
Oxidae		6
Palaemonidae		5
Paracalliopidae		8
Paramelitidae		7
Parastacidae	7	6

Family name	Version 95	97G
Perthidae		10
Philopotamidae	10	9
Philorheithridae	8	9
Phreatoicidae		9
Phreatoicopsidae		6
Phreodrilidae	5	5
Physidae	3	2
Pionidae		5
Planorbidae	3	3
Plectrotarsidae		10
Pleidae		5
Podonominae		9
Polycentropodidae	8	8
Prosopistomatidae		10
Protoneuridae	7	6
Psephenidae	5	6
Psychodidae	2	5
Ptilodactylidae	10	9
Pyralidae	6	6
Richardsonianidae		5
Saldidae		4
Sciomyzidae		3
Scirtidae	8	8
Sialidae	4	4
Simuliidae	5	5
Siphlonuridae		7
Sisyridae		5
Sminthuridae		1
Sphaeriidae	6	5
Sphaeromatidae		4
Spionidae		1
Spongillidae		6
Staphylinidae	5	5
Stenopsychidae		10
Stratiomyidae	2	2
Sundatelphusidae		1
Synlestidae	7	7
Synthemidae	7	7
Syrphidae		5

Family name	Version 95	97G
Tabanidae	5	4
Talitridae		6
Tanypodinae		3
Tasimiidae	7	8
Temnocephalidae		6
Tetrastemmatidae		1
Thaumaleidae	7	7
Thiaridae	7	3
Tipulidae	5	6
Torrenticolidae		6
Tubificidae	1	2
Unionicolidae		7
Veliidae	4	5
Viviparidae		7