Final Report
Project R2004
Development of a Framework for the Sustainable Rivers Audit

Appendix 3
Macroinvertebrates

Review and Development of Aquatic Macroinvertebrate Protocols

Julie Coysh

Richard Norris

Wayne Robinson

This report on the macroinvertebrate theme component of the Sustainable Rivers Audit is presented in two sections. Part A is a review of existing macroinvertebrate programs in each jurisdiction, and presents discussion outcomes from the workshop that was attended by representatives from all jurisdictions in the Basin. Part B develops the indicators and the framework in which the macroinvertebrate theme will be reported. Recommendations from Part B are summarised in the summary at the beginning of this document.

Cooperative Research Centre
for Freshwater Ecology

Table of Contents

Summary

1. Two indicators of condition for the macroinvertebrate theme are proposed, AUSRIVAS O/E taxa and a form of SIGNAL score. O/E SIGNAL and raw SIGNAL as currently calculated have been demonstrated to be insensitive to impacts and consistently overestimate condition. Therefore it is recommended that the AUSRIVAS O/E taxa score is used as the macroinvertebrate indicator for the first year of the Audit. It is proposed a more robust form of SIGNAL is developed in the first year of the Audit by testing regionalised raw SIGNAL scores and calculating O/E SIGNAL using all taxa.

2. It is recommended that existing AUSRIVAS models and associated sampling and processing protocols should be used for assessment in the first year. Existing regional models should be used in preference to statewide models where available and appropriate. Concurrently, existing models should be evaluated using a stepwise process to ascertain whether the existing model is the most appropriate model in each case. Development of regional models for the Basin where appropriate is proposed for the first year of the Audit.

3. The frequency of assessment should maximise the power of the sampling design to detect spatial and long-term temporal trends. Single season models are therefore recommended where taxa numbers are high enough, as sampling density can be increased for the same cost. In Western regions, however, combined season models are recommended to provide an adequate taxon list.

4. Existing Victorian data and models will be used to test the effect of increased taxonomic resolution on taxon richness in lowland zones. The accuracy of assessments can be analysed with existing Victorian models by examining the change in reference sites over time. After testing, genus or species models may be adopted where appropriate.

5. Analysis of both AUSRIVAS O/E taxa and SIGNAL will use comparison to a reference condition. The macroinvertebrate theme should incorporate a measure of departure from reference, and a measure of departure of reference from natural to account for the varying definition of reference condition currently used. Options proposed to measure the departure of reference from natural include using the River Disturbance Index, conceptual models of river function or a narrative description. To measure the departure of a site from reference condition, scoring against reference criteria and measuring the departure of the O/E value from 1 were proposed. These measures would then be turned into alternative health indicators and tested for sensitivity to known disturbances, allowing existing reference sites and models to be used, and providing comparability between different standards of reference.

6. Caution should be used in integrating indicator scores to produce a single score. Preferably, indicators should be reported separately, as they represent different information about the health of a stream. Where a single score is required, reporting of the indicator score that is the further away from reference is recommended. Only the O/E taxa indicator will be reported in the first year. Aggregation will follow the general principles outlined for reporting of theme condition, using the median score for a river valley.

7. Valleys and Valley Process Zones have been proposed as reporting scales. A number of options for sampling design and precision have been proposed, on which a decision is required by the Audit taskforce. The recommended level of change detectable at the river-valley scale for an AUSRIVAS O/E score is 10% and for a SIGNAL score 5%. These are considered appropriate and meaningful levels at which a change should be detectable. A commitment in the order of $714,500 per annum across the Basin would be required to achieve this level of precision at the valley scale. Reporting at the valley scale and theVPZ scale requires considerably more sites, with costs estimated in the order of $1,620,000 per annum.

PART A
Review of existing programs and recommendations

Section A1

Review of existing macroinvertebrate programs in the Basin

The outcomes of this review were discussed at a workshop held at the University of Canberra on 20th March, attended by representatives from each jurisdiction in the Murray-Darling Basin (MDB), the Commission and the CRCFE (see below). Discussion points and changes have been added where necessary in italics. Thus, this document represents the view of all jurisdictions with regard to macroinvertebrate indicators and assessment tools.

State Representatives	**Organisation**	**State**
Eren Turak	EPA	NSW
Natasha Waddell	EPA	NSW
Bruce Chessman	DLWC	NSW
Greg Keen	Environment ACT	ACT
Brian Wilkinson	Environment ACT	ACT
Satish Choy	DNR	QLD
Peter Goonan	EPA	SA
Leon Metzeling	EPA	VIC
Other Representatives		
Sue Grau	MDBC	
Brian Lawrence	MDBC	
John Whittington	CRCFE	
Peter Liston	CRCFE	
Julie Coysh	CRCFE	
Richard Norris	CRCFE	

A1.1 Review of existing programs

The review of existing programs undertaken in Task 1 of the Audit indicated that the following major sampling programs in the MDB incorporate macroinvertebrate indices:

- Index of Stream Condition (ISC)
- Integrated Monitoring of Environmental Flows (IMEF)
- Pressure - Biota - Habitat (PBH)
- MDBC water quality monitoring program
- National River Health Program (NRHP)
- Waterwatch

A number of other programs report macroinvertebrate data (below), but these data are sourced from one of the above programs:

- National State of the Environment Reporting (SOE)
- Water Allocation Management Planning (WAMP) ecological assessment
- Assessment of River Condition (ARC)

Table A1 Macroinvertebrate monitoring programs in the Basin (from Task 1 Report: Review of Existing Programs).

Approach	Monitoring procedure	Where applied in Basin
ISC	AUSRIVAS sampling AUSRIVAS score, probability weighted/raw SIGNAL score	Entire Victorian section of Basin
IMEF	Wetland replenishment study: Method unspecified Conditioning stony beds: Wetting of Terr. Org. matter	Gwydir, Macquarie, Lachlan, Murrumbidgee Murrumbidgee Namoi
PBH	SIGNAL, number of families	Castlereagh, Lachlan, Murrumbidgee
State of Rivers	Not measured	-
MDBC WQ Monitoring program	Artificial substrate	Murray: Above Hume reservoir and below Wellington
National SOE	AUSRIVAS score sourced from NRHP and State programs	Entire Basin
WAMP ecological assessment	AUSRIVAS sampling plus additional analyses on data collected as follows: Taxonomic richness PET taxa SIGNAL index, AUSRIVAS score Functional feeding groups Flow velocity and substrate preference groups	QLD section of Condamine/ Balonne/Culgoa
ARC	AUSRIVAS score sourced from NRHP and State programs	Entire Basin
NSW River Survey	Not measured	-
Wild Rivers	Not measured	-
QLD EPA Guidelines for Waterway Values	Not specified	Not applied yet
NRHP	AUSRIVAS sampling, AUSRIVAS score and probability weighted SIGNAL score	Entire Basin
Waterwatch	AUSRIVAS sampling, AUSRIVAS score, probability weighted/raw SIGNAL score	Sites throughout Basin

The major sampling programs that undertake assessments using macroinvertebrates (ISC, IMEF, PBH, MDBC WQ Monitoring Program, NRHP and Waterwatch) in the MDB are summarised below.

A1.2 Index of Stream Condition (ISC)

The ISC was developed as a holistic measure of river health to be used by natural resource managers and CMAs for benchmarking river health, monitoring rehabilitation efforts, and to set priorities for management action. The ISC provides measures of the health of the aquatic biota and the drivers that may impact on the health of the biota. The ISC incorporates the following 5 indices, that are formed from sub-indices:

- Hydrology
- Physical form
- Streamside Zone
- Water Quality
- Aquatic Life

Macroinvertebrates were selected for the Aquatic Life Index because:

- their ubiquitous and sedentary nature indicates local conditions,
- they integrate the effects of impacts over time, from weeks to years,
- they have a wide range of tolerance to environmental conditions,
- they are well studied, have relatively simple taxonomy and sampling procedures.

Macroinvertebrates were collected according to the sampling protocols of the National River Health Program, which aim to provide a representative sample of the macroinvertebrates at a site. The Aquatic life index is comprised of two sub-indices, the AUSRIVAS O/E score and SIGNAL [O/E] score. The AUSRIVAS O/E score and the SIGNAL score are changed into ratings between 0 and 4, weighted equally and summed to give the Aquatic Life Index. Because of the lower sensitivities that would be expected of macroinvertebrate taxa in more lowland river sections, separate SIGNAL scoring systems have been derived for upland and lowland reaches (Table A2).

Table A2 Ratings for SIGNAL indicator (from Ladson and White 1999)

SIGNAL value (upland reaches)	SIGNAL value (lowland reaches)	Rating
>7	>6	4
6-7	5-6	3
5-6	4-5	2
4-5	3-4	1
<4	<3	0

A1.3 Integrated Monitoring of Environmental Flows (IMEF)

This program was designed to evaluate ecological responses to environmental flows in regulated rivers. IMEF tests a series of hypotheses associated with water management approaches. Hypotheses may include analysis of macroinvertebrates, as well as other biota. IMEF is focussed on understanding the impacts of water management rather than providing a condition assessment of a river. Because different studies in the program are testing different hypotheses, the type of data collected and the analyses vary between rivers.

A1.4 Pressure - Biota - Habitat (PBH)

The PBH program is designed to be a rapid, multi-faceted procedure for assessing ecosystem stress and conservation value in small to medium size streams. PBH has been trialled in 12 NSW sub-catchments. Indicators are grouped under 8 categories; richness, rarity, native abundance, physical structure, water quality, alien biota, sensitivity and hydrological stress. Field trials produced 4 scores; conservation significance, sensitivity, general stress and hydrological stress.

Several components of the biota are measured, including macroinvertebrates. Macroinvertebrates are sampled from 3 habitats using a handnet and live-picked for 30 minutes with the aim of maximising the number of taxa collected. Two macroinvertebrate indicators are incorporated in the components of the biota: richness (number of macroinvertebrate families) and a modification of the SIGNAL score weighted by abundance, the Macroinvertebrate Family Index (MFI). These indices are combined with other biotic indices to produce an average score for components of the biota (Chessman 2001). While these particular assessment tools and indicators were used in the trial of PBH, the program is proposed as a flexible framework in which a range of macroinvertebrate assessment tools and indicators could be used.

A1.5 MDBC water quality monitoring program

The MDBC Water Quality Program is designed to monitor the water quality in the River Murray and several major tributaries by measuring water quality, phytoplankton and macroinvertebrates. Macroinvertebrates are sampled using artificial substrate samplers (ASS) made of plastic mesh, with sampling twice a year. A major limitation of the ASS is that substrates have to be left in the river for at least 30 days for colonisation. During this time the samplers are susceptible to loss and damage because of river level changes and vandalism (MDBC report 1985). A number of indices are reported, including community similarity indices, functional feeding groups (FFGs), species composition and abundance and macroinvertebrate biomass (Bennison *et al.* 1989).

A1.6 National River Health Program

The National River Health Program (NRHP) was designed to assess the ecological status of streams nationwide using macroinvertebrate fauna. As part of the program the First National Assessment of River Health (FNARH) was conducted, with approximately 6000 sites across Australia sampled. Samples are taken with a handnet from either one or two habitats in autumn and/or spring. Macroinvertebrate samples are either lab-sorted or live-picked in the field, (depending on the State/Territory), with the aim of obtaining a representative sample of the invertebrate community at a site to provide an accurate assessment of river health (as opposed to maximising the list of taxa collected). The Australian River Assessment System (AUSRIVAS) is a predictive modelling system that was developed as part of the NRHP to provide an assessment of condition for the nations rivers. AUSRIVAS predicts an expected macroinvertebrate family composition for a site in minimally disturbed condition, based on physical and chemical characteristics, to which the observed macroinvertebrate families are compared. This is reported as a ratio, the AUSRIVAS Observed/Expected (O/E) score. Similarly, an expected and observed SIGNAL score are calculated as part of AUSRIVAS producing an O/E SIGNAL score, providing two indicators of river condition (Simpson and Norris 2000).

A1.7 Waterwatch

Waterwatch is a community-based program for river health assessment. It provides local communities interested in river health with simple, rapid and inexpensive techniques to monitor stream condition including macroinvertebrate protocols. However, not all Waterwatch programs collect information on macroinvertebrates. Macroinvertebrates collected are processed in several ways, depending on the level of expertise and resources available. Sites have been sampled across the Basin, but data and analysis vary in quality and consistency.

A1.8 Other Indicators

While not existing, the development of State Environment Protection Policies (SEPPs) in Victoria were discussed at the workshop, as the indicators to be used are relevant to the Audit. SEPPs will use existing data and new data collected according to the rapid sampling protocols used for AUSRIVAS models. Five macroinvertebrate indicators will be reported for SEPPs:

- *a measure of diversity — number of families*
- *biotic indices — SIGNAL and EPT indices*
- *measures of community structure — numbers of key families and AUSRIVAS predictive models.*

The rationale behind using multiple indicators is that greater confidence can be placed on the outcome if the indicators are in accord; and where a discrepancy occurs, this can be used to indicate a potential impact.

Section A2

Selection of assessment tools and indicators

A2.1 Potential macroinvertebrate indicators

To review potential assessment tools and indicators for macroinvertebrates, it is important to distinguish between the programs that use data derived from macroinvertebrate sampling, the assessment tools used to collect the data and the indicators used to analyse the macroinvertebrate data. To assess macroinvertebrate data there are two main methods of analysis, these are:

- indices and metrics
- predictive models

Indices and metrics refer to standalone measures of river health that are usually interpreted against guidelines or predetermined thresholds. Most indices do not have well defined site specific standards against which they can be compared to determine an effect. Predictive models do provide site-specific assessments by comparing the observed community to a predicted community for a site. This requires a large set of reference sites for comparison. Indices or metrics can also be probability weighted and made site specific using the predictive models. The different assessment tools, indicators and methods of analysis are outlined in Table A3 for each of the programs with a macroinvertebrate component reviewed in Task 1.

Table A3 Assessment tools and indicators used by existing programs with a macroinvertebrate component in the MDB.

Program	Assessment tool	Indicator	Type
ISC	AUSRIVAS sampling	AUSRIVAS O/E score SIGNAL score	Modelled Index/metric
IMEF	Experimental (varying)	Varies depending on experiment	
PBH	Handnet sampling (3 habitats)	Richness Macroinvertebrate Family Index (SIGNAL weighted by abundance)	Index/metric Index/metric
		AUSRIVAS O/E score	Modelled
MDBC WQ monitoring program	Artificial Substrate Sampling (ASS)	Community similarity indices Functional feeding groups Species composition and abundance Biomass	Index/metric Index/metric Index/metric Index/metric
National SOE	AUSRIVAS sampling	AUSRIVAS O/E score	Modelled
WAMP ecological assessment	AUSRIVAS sampling	Taxonomic richness EPT taxa Functional Feeding Groups Flow velocity and substrate preference groups SIGNAL O/E score AUSRIVAS O/E score	Index/metric Index/metric Index/metric Index/metric Modelled Modelled
ARC	AUSRIVAS sampling	AUSRIVAS O/E score	Modelled
NRHP	AUSRIVAS sampling	AUSRIVAS O/E score SIGNAL O/E score	Modelled Modelled
Waterwatch	Varies, some use AUSRIVAS sampling	Varies, some use O/E score Varies, some use SIGNAL O/E score	Modelled Modelled

A2.1.1 Indices and Metrics

Richness Indices

- *Number of EPT taxa*

The number of EPT taxa is the number of macroinvertebrates collected in the insect orders Ephemeroptera (mayflies), Plecoptera (stoneflies), and Trichoptera (caddisflies). Taxa in these orders are considered sensitive, and a decline in the number of EPT taxa can indicate a potential impact.

MDB programs: Only used in Queensland in the WAMP ecological assessment process.

Biotic Indices

Biotic indices indicate the sensitivity of the biota present to impacts. If sensitive taxa are missing, then the ecosystem is likely to be stressed.

- *SIGNAL*

SIGNAL provides an assessment of the sensitivities of the taxa that were collected at a site to various human-caused stressors such as stream salinisation and organic pollution. A SIGNAL score for a site is calculated by summing the SIGNAL sensitivity grades for all taxa present at a site and dividing by the number of taxa at a site to give an average SIGNAL sensitivity score, ranging from 1 to 10.

MDB programs: Used in Victoria for the ISC, Waterwatch and PBH. Probability weighted scores are used as part of AUSRIVAS assessments.

- *Macroinvertebrate Family Index (MFI)*

The MFI is a modification of the SIGNAL index, weighted by abundance. SIGNAL grades for each family are multiplied by the square root of the total abundance of each family and summed for all taxa present at a site. This figure is then divided by the square root of the abundance of all taxa at a site to give an average score between 1 and 10.

MDB programs: The MFI is only used in a subset of NSW sub-catchments by PBH.

Composition Indices

- *Community similarity indices*

Generally provide a measure of the similarity or dissimilarity of communities, based on taxa common or absent from communities or the mathematical distance between the community composition at pairs of samples/sites. Can be used to compare condition between sites or over time. Examples are Sorensen's Index, Czekanowski's Index, Sokal's Measure of Distance, Canberra Metric Dissimilarity Measure, Jaccard Similarity Index.

MDB programs: Only used in the MDBC WQ monitoring program for the River Murray and major tributaries.

- *Species composition and abundance*

The species in the community, their relative proportions and abundance can provide information about the health of the stream. Knowledge of life history and ecology of invertebrates can be used to infer information about the health of a site.

MDB programs: Used in the MDBC WQ monitoring program. Richness also used in NSW for the PBH program and in Queensland for WAMP ecological assessments.

Feeding

- *Functional Feeding Groups (FFGs)*

Categorisation of invertebrates into functional feeding groups should indicate process-level aquatic ecosystem attributes, based on changes in macroinvertebrate food sources and types that occur with distance downstream. Therefore, the macroinvertebrate community present should provide an indication of the food types available or missing and hence some indication of the health of the stream. These groups are generally reported as numbers or percentages (Table A4).

Table A4 Functional Feeding groups

FFG	Description
% shredders	Percent of the macrobenthos that "shred" leaf litter
% collectors/gatherers	Percent of the macrobenthos that "gather"
% filterers	Percent of the macrobenthos that filter FPOM from either the water column or sediment
% grazers and scrapers	Percent of the macrobenthos that scrape or graze upon periphyton
% predators	Percent of the predator functional feeding group. Can be made restrictive to exclude omnivores
% scavengers/generalists	Percent of generalists in feeding strategies

MDB programs: Only used in Queensland for WAMP ecological assessments

Discussion of the use of FFG in Australian streams highlighted a number of potential problems:

- *High variability among species in each family to assign FFG to a family.*
- *Based on Northern Hemisphere FFG assignments.*
- *FFG are highly variable across regions.*
- *Taxa may use a number of different strategies at different stages — this is particularly so in Australian streams as many taxa are generalists.*

If FFG were to be recommended for the Basin, further research would be required and FFG assignments would need to be improved.

Tolerance

- *Flow velocity and substrate preference groups*

These categories give an indication of the preferred environmental conditions of macroinvertebrate taxa in regard to flow conditions and substrates. Substrate preference groups of macroinvertebrates were identified by Queensland DNR and related to flow statistics, resulting in determination of a number of flow preference groups: high, low and no preference. Information about river health may be inferred by the presence or absence of taxa when related to the tolerance groups.

MDB programs: Only used in Queensland for WAMP ecological assessments

Other

- *Biomass*

Biomass is the weight of invertebrates, which may give an indication of density and abundance. Generally in a site in best available condition the abundance of sensitive taxa would be high, but seasonal factors may make abundance and hence biomass an unreliable measure.

MDB programs: Only used in the MDBC WQ monitoring program for the River Murray and major tributaries.

A2.1.2 Predictive Models

- *AUSRIVAS O/E score*

AUSRIVAS is a predictive modelling system, based on the British River InVertebrate Prediction And Classification Scheme (RIVPACS), which predicts an expected taxonomic composition for a site in minimally disturbed condition, based on physical and chemical characteristics, to which the observed community is compared. This is reported as a ratio, the AUSRIVAS Observed/Expected (O/E) score. The probability of occurrence of a macroinvertebrate family at a new test site is calculated in the model by comparison to reference sites in the model. The sum of the probabilities of families predicted at the site gives the number of taxa expected at the site. The observed number of taxa, based on the presence or absence of macroinvertebrate families at the site, is compared to the number expected, providing an observed/expected score. The AUSRIVAS O/E score provides an assessment of river health by comparing the diversity of the community against what would be expected in a minimally disturbed stream with a similar location and physical characteristics, independent of human activities. The AUSRIVAS program also provides a list of expected taxa that are missing, that can be used to infer potential impacts.

MDB programs: Used by all States and Territories in the MDB for NRHP, ARC, National SOE, ISC, WAMP ecological assessments, Waterwatch.

- *SIGNAL O/E score*

The SIGNAL score is also incorporated in the AUSRIVAS program as a second measure of river health, to be interpreted alongside the AUSRIVAS O/E score. In AUSRIVAS, the SIGNAL score is calculated in the same way as it is as an index; however, observed and expected SIGNAL scores are calculated and converted to an observed/expected ratio, giving an assessment in terms of reference condition rather than comparison to predetermined thresholds. Because the SIGNAL score is calculated in the AUSRIVAS framework, the observed and expected scores are site specific. The SIGNAL O/E score provides an assessment of river health by providing information on the sensitivities of taxa collected and also those missing, useful for identifying potential impacts and site condition.

MDB programs: Used by all States and Territories in the Basin for NRHP, ISC, WAMP ecological assessments, Waterwatch.

A2.2 Criteria for selecting assessment tools and indicators

A number of criteria have been proposed for the selection of indicators for each environmental theme in the Audit. Indicators should:

- *build upon existing programs and data as much as possible*
- *be consistent with the conceptual models of river function developed for the Functional Process Zones*
- *be responsive to disturbance*
- *be characterised by measurement and analysis that are rapid (analysis is built into reporting of the indicator)*

- *have standardised methods available and be technically appropriate for State agencies to undertake*
- *have output that can be interpreted relatively unambiguously*
- *have meaning to the wider Basin community.*

The first criterion is also an Audit project objective. Assessment of all the potential indicators against these criteria highlighted two indicators for detailed consideration: the AUSRIVAS O/E score and the SIGNAL O/E score.

A2.3 How the potential indicators meet the criteria for a useful indicator

- *Consistent with the conceptual models of river function developed for the Functional Process Zones*

The conceptual models of river function developed for the Audit represent the physical, chemical and biological entities and processes that occur in each Functional Process Zone (FPZ) in pristine/minimally disturbed condition. The AUSRIVAS score gives an assessment of the diversity of the community against what would be expected in a minimally disturbed (reference) stream with similar location and physical characteristics (independent of human activities) and provides a list of those taxa expected that are missing. The SIGNAL score provides information on the sensitivities of taxa collected and also those missing. Thus, a departure of AUSRIVAS and SIGNAL O/E values from reference condition indicates a likely impact. By examining the list of taxa that were expected but not collected and the sensitivities of these taxa, potential impacts can be identified, providing a focus for further investigation.

It has been suggested in the development of indicators that some environmental themes may be more or less relevant for different Functional Process Zones. With regard to the macroinvertebrate theme, there are no Functional Process Zones in which the use of macroinvertebrates to assess river health is inappropriate. Although macroinvertebrates vary in composition in different zones, as long as the conditions with which they are compared are also varied accordingly, the indicators will remain relevant throughout the zones. AUSRIVAS varies the condition to which a site is compared by using sets of 'reference' sites to which a new 'test' site is assigned a probability of belonging, based on physical and locational catchment variables. A chi squared test of the distance of a new site from each of the reference site groups is used to ensure the new site is similar enough to reference sites comprising the model to be assessed. If the site is too far away from any group the site is allocated "outside the experience of the model" and no assessment is produced.

Some discussion was had at the workshop as to whether structural measures were the most appropriate measures in each zone and whether functional measures may be more appropriate for some zones, for example, lowland zones. It was concluded that a change in function without a change in structure is unlikely, and therefore measures of structure represent functional change anyway.

SIGNAL was originally developed for use in perennially flowing upland streams in Eastern Australia; however, national SIGNAL grades were developed by the NRHP and were applied to a wide range of stream types. The use of SIGNAL in ephemeral streams and large lowland rivers has not been thoroughly assessed. The ISC program developed separate criteria for defining a healthy system in lowland and upland reaches based on SIGNAL scores because of the lower sensitivities of taxa that occur in lowland zones. The use of SIGNAL within the AUSRIVAS program ensures appropriate comparisons, assuming there are an adequate number of lowland reference sites, by generating expected SIGNAL scores that are site specific using a large number of reference sites for comparison. The SIGNAL score used independently of AUSRIVAS does not have this site specificity. However, the sensitivity of SIGNAL may be affected by its use in AUSRIVAS. AUSRIVAS only considers taxa with a 0.5 or greater chance of occurrence, whereas SIGNAL was designed to rate the sensitivity of all taxa present. *The workshop agreed that O/E SIGNAL as it is currently calculated is insensitive and some modified form would have to be developed to be useful for the Audit.*

- *Responsive to disturbance*

AUSRIVAS O/E scores have been used to demonstrate environmental harm in several recent pollution investigation cases in a number of States and Territories in the MDB. For example, Table A5 shows the use of AUSRIVAS in pollution investigations by the NSW EPA from January to June 2000. Impacts detected by AUSRIVAS scores have been assessed for sites along a known trace metal pollution gradient by Sloane and Norris (submitted). This study showed that AUSRIVAS is a sensitive assessment tool and the AUSRIVAS O/E score is a sensitive indicator for detecting mining impacts. AUSRIVAS was also used in the assessment of the impact of an uncontrolled sewage discharge into Perisher Creek from the Perisher Resort in July 2000. AUSRIVAS detected the impacts of the sewage spill above the normal impact of increased resort use at that time.

Table A5 The use of AUSRIVAS for pollution investigations by the NSW EPA January – June 2000 (Source: Turak *et al.* 2000 report).

Issue investigated	**Sites Sampled**
Avgas spill into a creek on the Central Coast	2
Subsidence of rubbish from an illegal tip into a creek in Northwest Sydney	3
Sedimentation of creeks in the Sydney Catchment area caused by construction activities	4
Bentonite release into a river in the Sydney Catchment area during construction activities	2
Overflow of water from tailings dam and mismanagement of runoff from a Bauxite mine in Northern NSW	5

The use of SIGNAL has been validated for the assessment of stream salinisation and organic pollution from sewage treatment plants (Chessman 1995). However, its usefulness for assessing toxic pollution and other types of disturbance is uncertain. Because AUSRIVAS and SIGNAL scores may be sensitive to a different range of impacts (Ladson and White 1999), it is recommended that both indices, AUSRIVAS O/E and SIGNAL O/E, should be used and interpreted together, to ensure the most accurate assessment.

NSW EPA presented an analysis of the robustness of AUSRIVAS and SIGNAL scores as demonstrated by the assessment of NSW sites. Findings showed that in the assessment of reference sites, O/E taxa and O/E SIGNAL produced stable assessments, with a central tendency around one, whereas number of taxa and raw SIGNAL produced very variable *assessments for reference sites (Figure A1, Table A6). Little variation was shown in O/E taxa and SIGNAL scores among regions for reference sites, but significant variation was shown in the number of taxa and raw SIGNAL scores for reference sites among regions (Figure A1, Table A6).*

Both AUSRIVAS O/E and raw SIGNAL scores were regressed against a pH gradient related to a mine site (Figure A2). AUSRIVAS O/E scores showed a good correlation with pH, but raw SIGNAL scores showed no change with changing pH. Thus SIGNAL is insensitive to some mining impacts. NSW DLWC presented scatterplots of raw SIGNAL and O/E taxa against EC, but these were not corrected for confounding factors and thus interpretation of trends was difficult (Figure A3).

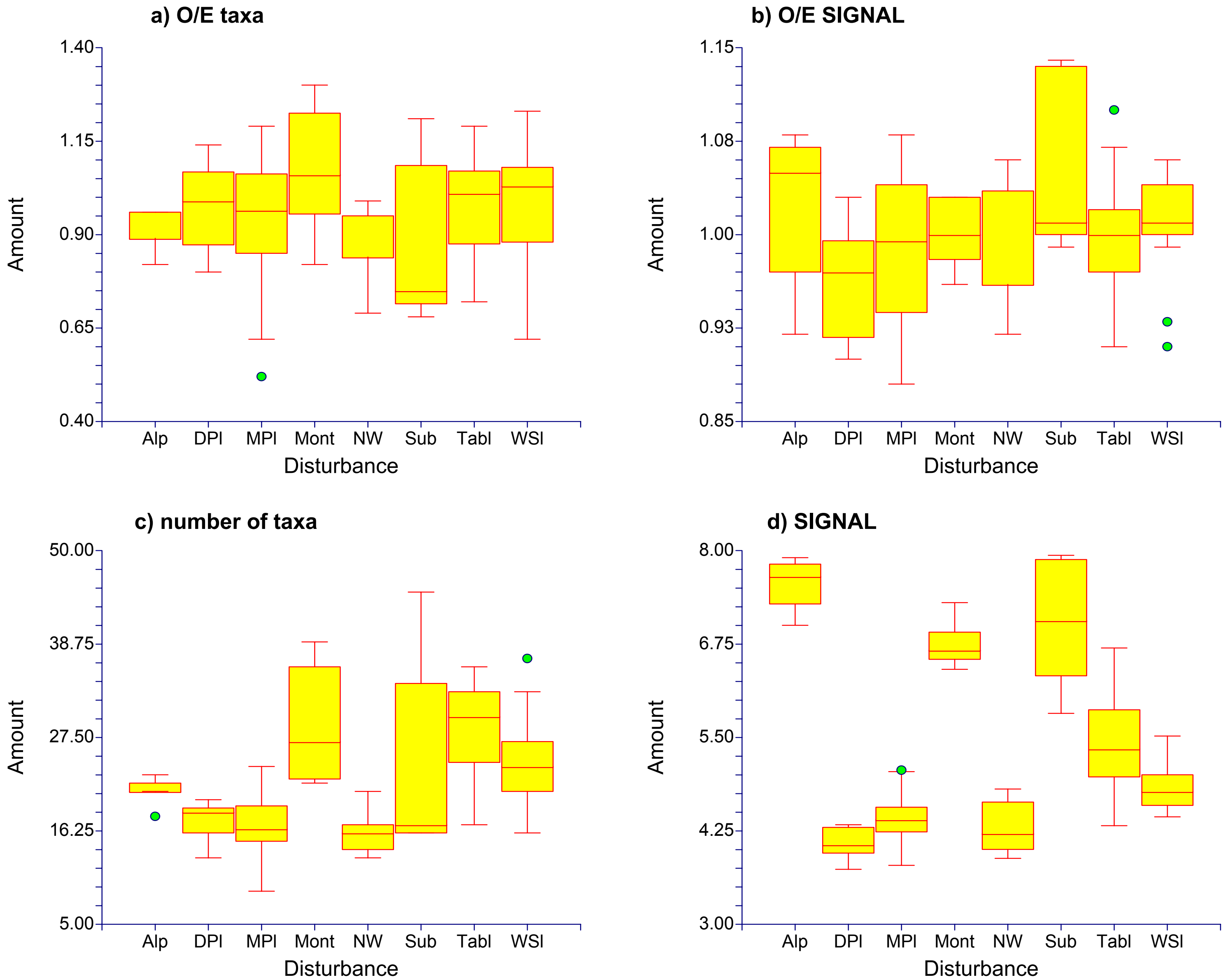


Figure A1 Assessments at reference sites across natural regions (Source: NSW EPA).

Table A6 Pairwise comparisons of assessments at reference sites across natural regions (Source: NSW EPA).

Index	**Probability**	**Significant pairwise differences**
O/E taxa	0.0827	0
O/E SIGNAL	0.0324	1
Number of taxa	<0.0001	9
SIGNAL	<0.0001	21

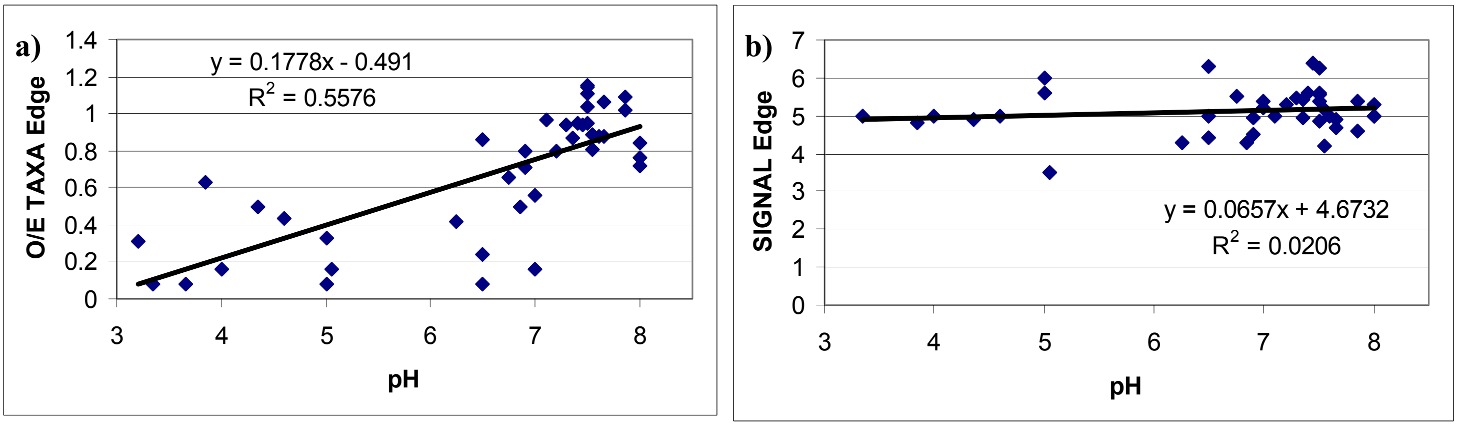


Figure A2 a) pH vs O/E taxa, and **b)** pH vs SIGNAL (Source: NSW EPA).

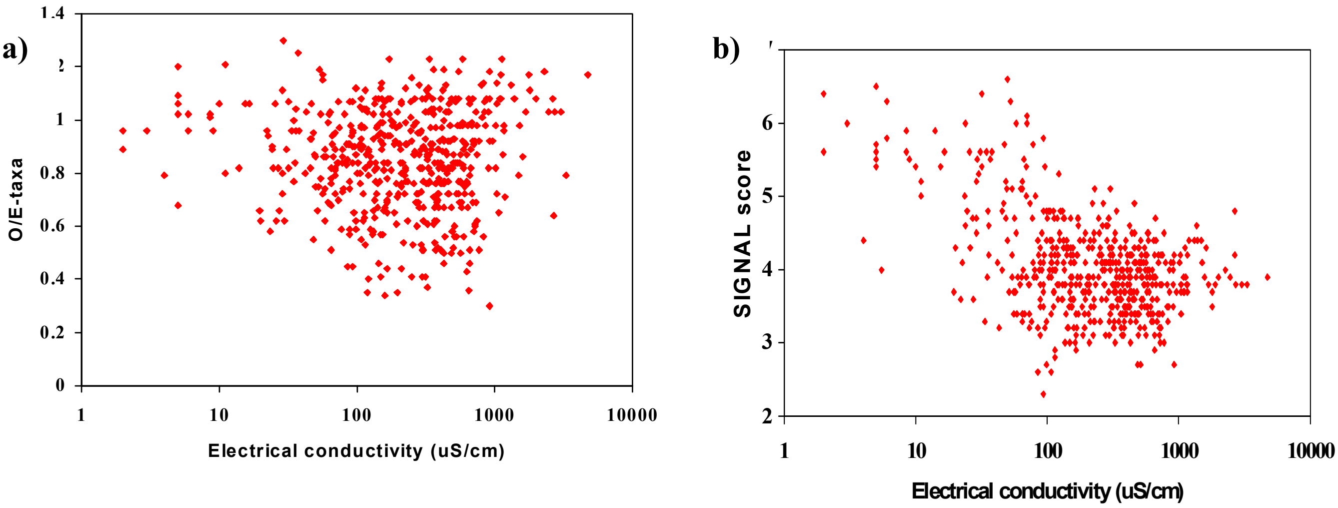


Figure A3 Example scatterplots of **a)** O/E taxa and **b)** O/E SIGNAL against EC for edgewater habitats (Source: NSW DLWC).

- *Measurement and analysis are rapid (analysis is built into reporting of the indicator)*

AUSRIVAS sampling protocols are designed to be rapid, with sampling macroinvertebrates involving a single sweep (edge) or kick (riffle, bed) of 10m of habitat with a handnet. Samples are then sorted in the field by live sort (Victoria, NSW, Queensland) or in the laboratory (ACT/Murrumbidgee catchment upstream of Burrinjuck storage and South Australia). In addition, several habitat and locational variables are measured and recorded on site as predictor variables but extensive habitat assessments may also be collected. For lab sample processing, sites can be sampled within 1-2 hours and the results reported back within a week, depending on the lab. For live-pick sampling, site time is approximately 2–3 hours with an even shorter turn around time of results.

Analysis of macroinvertebrate data is rapid. AUSRIVAS is downloaded from the internet and run locally on the users machine. Users simply input their invertebrate data using the national coding system and the required list of phys/chem data using the AUSRIVAS predictor variable names into the appropriate spreadsheets. The correct model for the region, season and habitat is then selected and run, with outputs displayed immediately. The user may then save these outputs to a local directory. Both AUSRIVAS O/E and SIGNAL O/E scores can be divided into bands of impairment for reporting, from richer than reference to impoverished.

- *Standardised methods are available and are technically appropriate for State agencies to undertake*

A strength of AUSRIVAS is that the techniques for site selection, data collection, laboratory procedures and analysis have been standardised, which means that results are comparable between operators and regions. AUSRIVAS has been used by each jurisdiction in the Basin, and all jurisdictions have staff trained in AUSRIVAS sampling protocols.

- *Output can be interpreted relatively unambiguously*

The AUSRIVAS program has been tailored to suit a variety of users, including resource managers, scientists and community groups. To suit different users, a number of outputs are produced that range in complexity. Outputs include:

- probability of a site's membership to each site group
- probabilities of all taxa at every site
- number of taxa predicted, expected and observed at a site
- observed and expected SIGNAL scores for a site
- O/E taxa and O/E SIGNAL scores for a site
- a list of taxa collected but not used in the model
- a banding for a site (X, A, B, C, D).

Outputs from AUSRIVAS are relatively easy to interpret. Guidance on interpreting AUSRIVAS outputs is provided in the form of a web-based manual located at http://ausrivas.canberra.edu.au.

- *Indicator has meaning to the wider Basin community*

The AUSRIVAS O/E scores provide a meaningful assessment of condition that can be easily understood by the wider Basin community. For example, an AUSRIVAS Observed/Expected (O/E) score of 0.5 would mean that 50% of the macroinvertebrate taxa expected at a particular site, if it were in 'best available' condition, are missing. Similarly, an O/E SIGNAL score of 0.5 would mean that the sensitivity of the macroinvertebrate community on average at a particular site is 50% lower than if it were in 'best available' condition. While sensitive taxa may still be present in the community, the diversity of sensitive taxa has been reduced. In their simplest form, AUSRIVAS O/E scores can be converted to bands, providing a broad assessment of stream condition in comparison to reference that has very obvious meaning about the condition of a stream (Table A7). A similar banding can be produced for SIGNAL O/E scores. The labels and interpretation put on the bands may be varied as appropriate, according to the values of the Basin community.

Table A7 AUSRIVAS bands of impairment

Band	Label	Potential Interpretation
X	**Richer than reference**	More families found than expected. Potential biodiversity "hot-spot" Mild organic enrichment Continuous irrigation flow in a normally intermittent stream
A	**Equivalent to reference**	Expected number of families within the range found at 80% of the reference sites
B	**Below reference**	Fewer families than expected Potential impact either on water and/or habitat quality resulting in a loss of families
C	**Well below reference**	Many fewer families than expected Loss of families from substantial impairment of expected biota caused by water and/or habitat quality
D	**Impoverished**	Few of the expected families remain Severe impairment

A2.4 Summary

The AUSRIVAS sampling and assessment approach builds on existing State and Territory programs, expertise, methods and data. The SIGNAL scoring system is already incorporated into AUSRIVAS and therefore, it is possible to use existing models and data to generate both AUSRIVAS and SIGNAL scores, with no development of new indicators required. AUSRIVAS assessments have already been undertaken by all States/Territories in the MDB and each jurisdiction has staff trained in AUSRIVAS sampling and analysis protocols. Good site coverage of the Basin exists through the NRHP. The NRHP can be used as a database for identifying areas of poor site coverage for the Audit. Thus, it is recommended that the standardised sampling protocols of AUSRIVAS be adopted as the assessment tool of the Audit for the macroinvertebrate theme. As the AUSRIVAS O/E score and the SIGNAL O/E score meet all the criteria proposed for selection of indicators, it is recommended that they be used as indicators for this environmental theme.

The workshop agreed to the use of the AUSRIVAS O/E score as one indicator of condition and some form of SIGNAL as another, but not SIGNAL O/E as it is currently calculated, as it is too insensitive. Either some revision of the calculation of SIGNAL in AUSRIVAS or regionalisation of the raw SIGNAL score would be required. SIGNAL varies naturally over the Basin, therefore different rating systems for different regions would be required if the raw SIGNAL score is used

Section A3

Development of indicators for the Murray Darling Basin

A3.1 How the assessment tool and indicators address each FPZ

Ideally, indicators for the Murray Darling Basin should be appropriate for all the major Functional Process Zones and stream types in the MDB. Consistency in the use of indicators across the Basin will reduce complexity in reporting and ensure comparability across the MDB. Appropriateness of using the proposed indicators in each Functional Process Zone and potential limitations of the indicators are considered in Table A8.

Table 8 How AUSRIVAS addresses each FPZ

Amalgamated Zones	Upland Zones			Mid-slope Zones		Lowland Zones		
Functional Process Zone	**Pool**	**Upland Gorge**	**Armoured**	**Mobile**	**Meander**	**Anabranch**	**Distributary**	**Lowland Gorge**
Reference site coverage of zone by existing models	Good	Good	Good	Fair	Fair	Poor coverage in some areas	Poor coverage in some areas	Poor coverage in some areas
Main macroinvertebrate habitats	Pool/Riffle	Pool/Riffle	Pool/Riffle	Macrophytes, snags, pool, riffle	Macrophytes, snags, pool	Macrophytes, snags, pool	Macrophytes, snags, pool	Macrophytes, pools, wetlands
Sampling area recommended	Main Channel / Edge	Main Channel / Edge	Main Channel / Edge	Edge / Snags	Edge / Snags	Edge / Snags	Edge / Snags	Edge / Snags
Taxa numbers	High	High	High	Med-high	Med-high	Low	Low	Low
Comments	Separate main channel/edge model for these zones			Separate edge/snag model for these zones		More reference sites needed for these zones. Separate edge/snag model for these zones accounting for lower taxa numbers. Combined seasons model.		

A3.2 Jurisdictional Review of existing AUSRIVAS models

AUSRIVAS models have been produced for each of the States/Territories in the Basin for a range of stream types. It has been suggested that AUSRIVAS may not be sensitive in lowland Functional Process Zones, and the sensitivity of AUSRIVAS in lowland streams remains largely untested. In initial AUSRIVAS sampling, lowland areas lacked adequate coverage of reference sites. This led to sites being unable to be assessed from these zones because they fell outside the experience of the models produced. Subsequent sampling rounds recognised this problem and sampled more sites from these zones. Not all States/Territories have incorporated these additional reference sites into their models (see Table A9). Thus, if one or more MDB models were to be built, it could include these sites not in existing models.

The major problems with the assessment of lowland zones in each jurisdiction highlighted by Table A9 are:

- poor existing reference site coverage
- low taxa richness

To address the first point, conditions acceptable for reference can be redefined as the 'best available' in areas with insufficient reference sites, and additional reference sites already sampled but not yet included in models can be added. To address the second point, the list of taxa can be maximised by using combined season models. Another option is to lower the taxonomic resolution to species rather than family level. However, this would require re-identification of all samples already collected to enable use of existing data. While this would entail a significant cost, it would be much less than having to revisit and sample all sites anew.

The question was raised at the workshop; how predictable is the macroinvertebrate fauna in lowland rivers? While insufficient reference sites are considered the main limitation for prediction in lowland streams, it is possible that precision may also be a problem. Climatic influences such as flood/drought cycles are likely to be major drivers of ecosystem function in lowland streams and therefore existing predictor variables may be inappropriate. It was proposed that replication of sampling is required to detect the precision of existing sampling protocols. Other questions raised included: what is the appropriate scale for assessment? Is 10 m of edge habitat enough to provide an indication of river health in each FPZ?

A study on area sampled and replication in Thredbo River (Nichols and Norris submitted) found that for replicated 5, 10 and 20 m rapid collections using a D-net that the number of taxa recovered was not significantly different, using lab-sorting. Total abundance was most variable in the 20 m collections, attributed to possible net clogging and backwash. However, all sampling areas were found to be precise, producing the same site assessment using AUSRIVAS.

A similar study has been done in Victoria by Metzeling and Miller (in press), sampling 4 major habitats (riffle, edge, macrophytes and pool rocks), with sampling replicated for 3, 5, 10 and 20 m using live-picking. An increase in species was detected with sample size, *however this was not significant except for riffles. It was found that 10 m was sufficient sampling area for all the habitats sampled.*

When the reporting scale of the river valley was considered, it was suggested that replication at the scale of within a site may be unnecessary. However, it was concluded at the workshop that some form of replicated sampling at test sites in each zone would be useful to obtain an estimate of the error in existing collection methods in each main habitat.

The possibility of using species level models was considered at the meeting. Results from Victorian species models were presented. Species level models had narrower bands and seem to be more sensitive, i.e. more sites were failed. However, it is equally likely that species level data introduces more noise and the models are just more variable. It was concluded that there was a need to check how reference site assessments changed through time using all models to determine the most accurate models. It was proposed that existing species and genus models could be assessed using Western Victorian data to determine whether changing the taxonomic resolution is effective in increasing taxa richness and improving accuracy of assessments.

Table A9 Jurisdictional review of existing experiences with AUSRIVAS models.

State/ Territory	Reference site coverage of lowland rivers	Model performance and assessment of sites	Model development status of State/Territory	Other issues particular to States
Queensland **Existing state-wide family models:** • Autumn bed • Autumn edge • Autumn pool • Spring bed • Spring edge • no combined season models	-Number of lowland reference sites in existing models is low, extra reference sites have now been sampled since these models were created	-Models appear to be working well for upland and mid river reaches -Uncertain of accuracy of model assessments in lowland reaches. Models assess most sites as poor, which is consistent with expectations of sites. However, these rivers have naturally low numbers of taxa, so may be being assessed as worse than they are. Alternatively, relevant reference sites may be significantly altered and thus provide conservative assessments	-QLD have two out of 7 final versions of models. Funding for the other 5 models has recently been approved and development of these models would be expected in the next 6 months. -Existing versions of the models have high misclassification errors and poor group discrimination. The inclusion of microcrustacea in sampling and models may be responsible for poor predictions as more sensitive taxa are not collected.	-Large lowland rivers in QLD may not be able to be compared to large rivers elsewhere because of differences in water quality, for example, conductivity can be naturally high. -It was suggested that regional models within Queensland would be an improvement on existing State models.
NSW **Existing state-wide family models:** • Autumn edge • Autumn riffle • Spring edge • Spring riffle • Combined season eastern edge • Combined season western edge (lowland model) • Combined season riffle	-Additional reference sites have been added to new models and a separate western region model developed. -The western model is built for lowland rivers (less than 200m altitude) to account for the naturally low numbers of taxa in this area. -Reference sites chosen for the western model were of poorer quality than reference sites in other models.	-Test sites known to be in poor condition from other knowledge of the site are being assessed as poor and reference sites run though the models are coming out as reference, therefore the models appear to be working well -NSW EPA have used AUSRIVAS assessments in a number of pollution cases.	-NSW have all seven statewide models completed. NSW EPA have considerably refined models from earlier versions, with group discrimination and predictor variables improved from earlier versions.	-To obtain sufficient reference sites for models, reference sites were divided into classes of reference: pristine, slightly modified and moderately disturbed -Development of the western lowland model was limited to edge habitat. Taxa lists from two seasons were combined to maximise numbers of taxa

ACT **Existing state-wide family models:** • Autumn edge • Autumn riffle • Spring edge • Spring riffle • Combined season edge • Combined season riffle	-A lack of suitable reference sites for assessment of the Murrumbidgee river has been a problem with existing models. Repeat sampling of existing sites is being used to expand the reference data-set, however, these sites have not been incorporated into the models as yet.	-Test sites assessed by AUSRIVAS along a known gradient of impairment were assessed as impaired by the models -Reference sites assessed by the models are generally assessed as in reference condition, with the exception of reference sites affected by drought conditions	-All six models for the ACT have been developed to final versions.	-In periods of drought, riffle habitat may disappear in ACT, thus edge sampling is relied upon for an assessment.
State/ Territory	**Reference site coverage of lowland rivers**	**Model performance and assessment of sites**	**Model development status of State/Territory**	**Other issues particular to States**
Victoria **Existing state-wide family models:** • Autumn edge • Autumn riffle • Spring edge • Spring riffle • Combined season edge • Combined season riffle **Other models:** • 7 regional family models, edge and riffle • 3 genus models • 2 species models • 2 urban models, family and species	-No models have been specifically built for lowland areas, but 7 regional models have been developed for the State, in addition to the statewide models, which cover lowland areas.	-Both state-wide and regional models appear to be providing similar and accurate assessments -Victoria is one of the few States that has used transformed predictor variables, which seems to be improving predictions.	In addition to developing final versions of regional models, final versions of all state-wide models exist. Preliminary results indicate that regional and state-wide models provide very similar assessments. Where results differ, the regional assessment is taken, as group discrimination appears better. -Victoria have also developed some genus and species models to investigate the effect of increased taxonomic resolution on model predictions but analysis of these models has not yet been undertaken.	

South Australia Existing state-wide family models: • Autumn edge • Autumn riffle • Spring edge • Spring riffle • Combined edge • Combined riffle	-Sites in the lowland constrained zone of the Murray have naturally high salinity and fall outside the experience of existing models because of insufficient reference sites.	-Site assessments for other rivers and zones in South Australia appear to be accurate. -SA EPA have used AUSRIVAS assessments in a number of pollution cases.	-6 existing models for the State are currently used for assessments. -The models are to be rebuilt in March 2001 to include enough naturally salty Murray sites to form a classification group in the models.	-Because lowland sites have naturally lower numbers of taxa, band widths of models may increase and other test sites may be assessed as better than they are. Thus O/E values should be interpreted in preference to bands.

A3.3 Development of models for the MDB versus using existing State/Territory models

Potential modelling options:

1. Use existing State models

A set of predictive models already exists for each State/Territory and could be used for the Audit for the assessment of new or existing test sites. Most States/Territories have a main channel and edge model already developed (Table A9). Thus, it would be possible to use existing models and existing data to generate both AUSRIVAS and SIGNAL scores, with no development of new models required. However, potential problems are insufficient reference sites for the assessment of some lowland zones, as reference sites for existing models are generally defined as 'minimally disturbed', not 'best available' as agreed on for the Audit. Some existing models also have predictor variables that may be inappropriate for lowland streams. Lowland systems are quite different in nature to upland systems, for example in their substrate composition. There are naturally more fines in lowland systems but this can also be an impact. Available habitat in channels and run/riffle/pool ratios is lower in lowland systems and riparian vegetation may be naturally low in lowland systems but scores as disturbed. Although AUSRIVAS models are ideally comparable across the nation, differences in macroinvertebrate sample processing protocols between States and Territories may also have some influence on the comparability of models across the Basin. In the MDB, Queensland, NSW and Victoria all use the live-pick processing protocols, whereas ACT and South Australia use lab-sort processing protocols. However, this should not be a problem provided the methods appropriate to each of the models are followed.

2. Development of a MDB model.

Development of one model for the whole MDB would resolve the differences that currently exist with a range of State/Territory models. However, a potential problem with this option is the large scale the model would have to account for. Modelling experiences to date indicate smaller regional models produce better predictions than large-scale models. At the large scale, locational variables discriminate best between the biota, whereas at smaller scales more specific habitat features become important. Predictor variables would most likely be restricted to locational variables, such as latitude, longitude, altitude, stream order etc. Predictor variables would be restricted to variables common to all State and Territory data-sets (and measured in the same way). Combining macroinvertebrate data from different States and Territories is likely to introduce some bias because of differing sample processing methods, as even data collected using the same sample processing method may have slight collection differences between States and Territories. This issue is relevant to all the modelling options and all the methods considered in this review. One solution to this issue of combining data is to assume that all the methods are basically measuring the same things; thus the result can be turned into a dimensionless score and be comparable.

3. Develop several regional (subcatchment models) within the Basin.

Development of several regional models would reduce the scale of model coverage from the MDB to a subset of Basins. Obviously these splits would make more ecological sense than jurisdictional boundaries on which existing models are based. This option has the same issues regarding combining of physical/chemical and biological data across the Basin as discussed above.

4. Development of models for each Functional Process Zone or Valley Process Zone.

- Because AUSRIVAS matches sites independently for comparison using physical and chemical measures, models for each FPZ are not necessary for choosing reference conditions. However, because different habitats are more appropriate to sample in each Functional Process Zone, different models will be required for each habitat. This may also enable selection of more appropriate predictor variables for upland and lowland sections. Alternatively, models could be created for each Valley Process Zone (upland, mid-slope and lowland). Some regionalisation may be required with this option, as Queensland rivers for example would not necessarily be comparable with Victorian rivers.

The workshop decided unanimously that a single model for the whole Basin would be on too big a scale to provide good predictions. It would also require combining different sampling methods and habitats, possibly requiring resampling of all sites. Regionalisation of models in Victoria and NSW has demonstrated the utility of regional models, with outputs more accurate and robust. It was proposed that regional models be recommended to be developed and tested. A stepwise process to assess the best model option in each case should be used, for example, existing models may be available in some areas that are robust and no new models are required. In other regions new models are needed and considerations include the scale of the region, the taxonomic resolution of the model and the possibility of combining data-sets across State boundaries. It was proposed that one model for the western parts of Victoria, NSW and Queensland may be appropriate, or a further division into FPZs in this area. Initially, combining of data for these States seems possible. All though slightly different protocols are used, all three States use the live-picking method for sample processing. It was proposed that comparability of State models could be assessed by comparing assessments of sites in overlapping regions, for example, assessment of NSW sites in Queensland models and vice-versa.

A3.4 Sample season and habitat

The original scoping study document recommended that edge models be used if existing models are used. However, after consideration of the appropriateness of models for each FPZ (Table A8), it is recommended that main channel or edge models be used for amalgamated upland zones, and edge or snag models for other zones, whether existing or new models are used. Climatic changes may affect the availability of both main channel and edge habitats to macroinvertebrates, with changing wetting and drying regimes. However, these are temporary extremes in climate and the sampling seasons autumn and spring generally avoid the severest of these extremes. Combined model data have been shown to provide the most robust assessment for all zones at the scale of a site, because the taxa list is maximised and seasonal biases are less influential. Using combined data means that new assessments would be delayed until both a spring and autumn sample had been collected from a site. However, existing test site data sampled in two seasons could be assessed immediately. If rapid assessments for new sites were required, existing single season models could be used for this purpose.

Sampling of snags was discussed by the workshop, and it was concluded that the fauna collected on snags are not predictable enough to be modelled. The number of taxa collected is very variable, and taxa numbers found in this habitat are often low. NSW DLWC suggested that existing edgewater models are insensitive to some disturbances. However, NSW EPA presented results from the Western combined edge model that demonstrated the model was working well. Both riffle and edge models for the eastern part of NSW were able to detect impacts of pH related to a mine and have been used to demonstrate a range of impacts in the past.

There was some concern about the recommendation to use combined season models, as the time gap between assessments may mean important changes are missed. As two seasons have to be sampled anyway to produce a combined season assessment, the workshop agreed that single season assessments should also be done where taxa numbers are high enough to enable a robust single season model. It was also recommended that more assessment of single season vs combined models was required. Replication of assessments within a season was proposed as another way of maximising the list of taxa collected.

A3.5 Sampling design, precision and reporting scale options

The workshop was asked to make a recommendation based on best scientific judgement of what would be an acceptable level of precision for indicators. The acceptable amount of change that could confidently be detected was recommended as 0.1 or 10% change in an AUSRIVAS O/E score. As the true mean of a river valley should equal one, this represents a 10% change (i.e. a change in O/E50 of 0.1). To obtain this level of precision, confidence limits of $\pm$ *0.05 around the mean for a river valley would be required. For any form of SIGNAL it was suggested that 5% change in the range of variation would be acceptable, which would require confidence limits of* $\pm$ *0.025 to obtain this level of precision.*

Reporting scale options include:

- reporting at the river-valley scale only
- reporting at the Functional Process Zone scale, which allows the option of aggregating to the river-valley scale as well
- reporting at the scale of Valley Process Zones (upland, mid-slopes, lowland), which allows reporting at the valley scale
- reporting assessments at the river-valley scale, with extra sites in areas of specific interest enabling reporting also at the FPZ scale for these catchments or at sites of specific interest or concern.

While a State or regional model may be used to assess sites, the reporting scale of results can still be at the scale of Valley Process Zones or the river valley, providing the reporting scale is determined prior to the sampling design, to ensure sufficient site coverage and statistical power for robust results.

A3.6 Reference condition

The definition of river health that has been adopted by the Audit emphasises comparison relative to natural habitats within a region: "*the degree to which aquatic ecosystems support and maintain processes and a community of organisms and habitats with a species composition, diversity, and functional organisation relative to that of natural habitats within a region*". Existing AUSRIVAS models use the definition of 'minimally disturbed' for reference condition. This may be a limitation to assessing sites using existing State models with insufficient 'minimally disturbed' lowland sites are available for comparison. Generally, the majority of reference sites in AUSRIVAS models are from smaller stream orders, with fewer reference sites from large lowland rivers. However, test sites from large rivers must be compared to reference sites also from large rivers for a relevant comparison, because lowland large rivers have naturally lower numbers of taxa in comparison to smaller upland streams. The reference condition that has been accepted for the Audit is 'best available' natural habitats (progress report). Thus, development of any new models for the MDB should use this definition. 'Best available' reference for upland sites may be minimally disturbed in many cases. However, 'best available' for lowland streams is likely to be those sites with good management practices for example, considering that for many lowland sites, minimally disturbed conditions are not likely to be encountered.

Therefore, existing reference sites will require redefining to meet the definition of 'best available' and extra reference sites will need to be included in models, whether new or existing models are used. Several approaches are proposed for defining reference condition for macroinvertebrate indicators.

1) Use existing data. Most reference sites for AUSRIVAS predictive models use the definition of minimally disturbed that is equivalent to 'best available'. In areas with insufficient reference sites, for example, lowland zones, define 'best available' reference in conjunction with State agencies (this would require rebuilding of predictive models if additional reference sites were to be added to the models).

2) Another way of defining best available for each river type is to use a reference hierarchy. For example, the NSW EPA divided their reference sites into three classes, A, B and C. The classes A and B indicate near pristine and slightly modified reference sites respectively, with C indicating moderately disturbed sites. Moderately disturbed sites were nominated as the 'best available' reference sites where more appropriate reference sites were not available for that type of river.

3) Use of specific criteria to define 'best available'. The following criteria are an example, they were used to redefine reference sites for the Queensland WAMP process (Table A10). Sites were considered to be in reference condition if they passed all criteria (P), or failed only one criterion (PC - conditional pass) as long as it was not the one relating to upstream dams and weirs. All sites failing this criterion were failed as were those failing a total of two or more criteria (F).

Considerable discussion was had at the workshop and in previous meetings about reference condition. It was concluded that the move away from minimally disturbed as the reference condition was inevitable, and indeed, all States agreed that in existing AUSRIVAS models not all reference sites are minimally disturbed. The management implications of setting reference were considered (downward and upward spirals of *condition) and the importance of recognising we are setting a condition for comparison, we are **not** setting targets. The changing definition of reference over time was also a concern. If an upward spiral model of condition is used where the comparison condition improves over time for example, how will trends be detected over time? It was suggested that all data could be run through new models as they became available and past condition assessments possibly revised. It was accepted that more information would always be revealed over time. The use of Alternative 3 (below) would enable a static point from which to measure change over time.*

NSW DLWC proposed a number of options for dealing with the fact that not all reference sites are near pristene or minimally disturbed.
Alternative 1: Use SIGNAL with a rating system to adjust for natural variation (e.g. rainfall).
Alternative 2: Modify Audit definition of healthy to mean "minimally disturbed" where such rivers exist and "well managed" where they don't.
Alternative 3: Use an assessment method that incorporates both departure from "reference" and departure of "reference" from "natural". To do this;

- *Agree on criteria for "less disturbed" or "better managed" sites*
- *Develop measures of departure of fauna from reference (O/E-taxa and others)*
- *Develop measures of departure of environment from natural (e.g. current vs modelled natural flow regime)*
- *Combine these measures into alternative "health" indicators*
- *Test these alternative indicators for sensitivity to known disturbances*

It was agreed that Alternative 3 would be good solution to the problem of "natural" in the existing definition of river health and a good way to measure how far away we have moved from "natural". It was proposed initially that State models could be used and reference assessed relative to natural using a disturbance index such as RDI or perhaps the Queensland method of scoring each of the criteria used for reference site selection (above). Other methods may also be considered, such as using the conceptual models of river function. This approach would be immediately applicable.

In lowland areas where no appropriate conditions for comparison are available, the notion of a hypothetical natural was considered, possibly using the conceptual models of river function. While it is inevitable that reference is likely to be defined differently for different themes, reference condition for each indicator needs to be the same to ensure comparability. The use of a measure of departure of "reference" from "natural" ensures comparability, even if the "reference" is variable across a region.

Table A10 Criteria used for selecting reference sites in Queensland for WAMP Ecological Assessments.

No.	Reference Condition Selection Criteria
1	No intensive agriculture within 20 km upstream. Intensive agriculture is that which involves irrigation, widespread soil disturbance, use of agrochemicals and pine plantations. Dry-land grazing does not fall into this category.
2	No major extractive industry (existing or historical) within 20 km upstream. This includes mines, quarries and sand/gravel extraction.
3	No major urban area (> 5000 population) within 20 km upstream If the urban area is small and the river large, this criterion can be relaxed.
4	No significant point source waste water discharge within 20 km upstream. Exceptions can again be made for small discharges into large rivers.
5	No dam or major weir within 20 km upstream. Sites within the ponded area of impoundments also fail. Sites failing this criterion automatically fail the overall assessment.
6	Seasonal flow regime not greatly altered. This may be by abstraction or regulation further upstream than 20 km. Includes either an increase or decrease in seasonal flow.
7	Riparian Zone of natural appearance. Riparian vegetation should be intact and dominated by native species.
8	Riparian Zone and banks not excessively eroded beyond natural levels or significantly damaged by stock. Stock damage to the stream bed may be included in this category.
9	Stream Channel not affected by major geomorphological change. Geomorphological change includes bank slumping, shallowing, braiding and unnatural aggradation or degradation.
10	Instream conditions and habitats not altered. This may be altered by excessive algal and macrophyte growth, by sedimentation and siltation, by reduction in habitat diversity by drowning or drying out of habitats (e.g. riffles), by direct access of stock into the river

A3.7 Interpretation of indicators

The AUSRIVAS O/E score and the SIGNAL O/E score have been proposed as the main indicators for the macroinvertebrate theme. However, AUSRIVAS as an assessment tool provides a range of other outputs that could be used for interpretation to improve the understanding of a system. Important ecological information is available by examining the list of taxa collected at a site, and also those expected at a site that weren't collected. Knowledge of the sensitivities of macroinvertebrates to various impacts and their functional roles in an ecosystem is significant and could aid in the interpretation of condition. While O/E scores are proposed as the primary indicators of condition, there is value in having another level of detail underneath these indicators at the interpretation stage, providing understanding of the processes that are occurring in a stream relative to those described in the conceptual models of river function.

A study on the Lithgow mine site by NSW EPA demonstrated the usefulness of other AUSRIVAS model outputs, such as the list of predicted taxa, for interpreting model assessments. They proposed that assignment of taxa to sensitivity categories could be used to detect specific impacts using the known tolerances of specific families/species.

SEPPs in Victoria use a list of key invertebrate families derived for each region, similar to a list of expected families predicted by AUSRIVAS. The list of key families provides an indication of habitat availability as well as water quality. A separate list for each region is necessary to account for differences in altitude, topography, stream size, flow, temperature, etc. The list of key families is based on a range of habitats, stream sizes and stream types within a region; therefore it is unlikely that a site would have all key families present. A percentile of the distribution of reference sites (10^{th}, 30^{th}) is used to set objectives for the indicator.

Key invertebrate families are determined for SEPPs as:

- *typically found in the types of stream in that region*
- *representative of a particular habitat type*
- *representing reasonable to good water quality, tending to disappear as conditions deteriorate*
- *are commonly collected when present (50% or greater chance of occurrence), using the recommended method to sample edges and riffles*

A similar list of key families could be derived for each FPZ or region using the predicted list of taxa for reference sites in a FPZ or region. These taxa lists could be added to the conceptual models of river function and used to provide information about the habitat availability and water quality at a site.

A3.8 Frequency of assessment

The workshop was asked to comment on what they thought would be an appropriate timeframe for assessment. A number of options were proposed but there was little agreement. Options included:

- *A fixed set of sites sampled once every five years*
- *Sampling fewer sites more times (may be more statistically effective)*
- *A mixed sampling design, some sites fixed, some sites only every five years*

It was recommended that the sampling design and frequency be determined by the statistical power required to maximise detection of trends for the reporting scale/s decided on by the taskforce.

Section A4 Summary and Recommendations

Section A1.

- A number of programs exist in the MDB which sample macroinvertebrate data. Of these programs only the NRHP and Waterwatch have a Basin-wide coverage.

Section A2.

- The assessment tool most common to the macroinvertebrate sampling programs is AUSRIVAS. The indicators most common to all the programs are the AUSRIVAS O/E score and the SIGNAL O/E score. This assessment tool and indicators build the most effectively on existing data, which is a Audit project objective, and meet all the other criteria proposed by the CRCFE for selection of indicators. Therefore, it is proposed that AUSRIVAS be used as the macroinvertebrate assessment tool for the Audit and the AUSRIVAS and SIGNAL O/E scores be used as indicators for the macroinvertebrate theme.

There was unanimous agreement on the recommendation to use the AUSRIVAS O/E taxa score as an indicator. Further consideration of the options for SIGNAL was recommended as O/E SIGNAL as it is currently calculated is too insensitive.

Section A3.

- It was proposed that one approach for sampling, analysis and reporting would be to use upland, mid-slope and lowland sections (VPZs) for the macroinvertebrate theme.
- Potential problems identified with existing AUSRIVAS models in each jurisdiction were the potential poor reference site coverage of lowland zones by some models and the lower taxa richness of lowland zones. As reference sites in existing models are defined as minimally disturbed, using the definition of best available and including additional sites sampled but not in the models would substantially increase the reference site coverage of models. Splitting the models into broad zones, using combined season models, and lowering the taxonomic resolution to species level, were proposed to address the lower numbers of taxa in lowland zones.

The merits of using genus/species data will be assessed using Western Victorian data.

- As reference condition for the Audit is defined as 'best available' and existing sites may need to be redefined to meet this definition, a number of approaches were proposed. As well as redefining reference sites, extra reference sites not included in existing models also need to be added. Therefore, redevelopment of existing State models seems necessary if they are to be used for the Audit. Redevelopment of existing models or creation of MDB models does not require collection of new data, therefore these options build effectively on existing data.

It was accepted to use "reference" as currently defined but to use departure from ‘reference’ and departure of ‘reference’ from ‘natural’ as a way of benchmarking condition to ‘natural’.

- One option proposed for development of new models was to create a model for the VPZs in each State, incorporating new reference site data not included in existing models. This would enable appropriate habitat models relevant to zones, while using existing data. Reference sites and predictor variables appropriate to zones could then be used in the models.

Potential Basin models include:

State/Territory	Models
Queensland	Upland, Mid-slope, Lowland
NSW	Upland, Mid-slope, Lowland (NSW already has a western edge combined season models that could be used here)
ACT	Upland
Victoria	Upland, Mid-slope, Lowland
South Australia	Lowland

This would give a total of 11 models for the MDB.

Regionalisation was recommended as the preferred approach to be explored, possibly in combination with existing models, dependent on the ability to combine data across State boundaries. Existing models should be used initially while development issues for regional models are investigated, e.g. taxonomic resolution, sample habitat, season, data compatibility.

- It was proposed that habitat models be developed that are appropriate to the Functional Process Zones and the conceptual models of river function. Thus, for the amalgamated upland zones, main channel or edge was proposed as the sampling habitat. In the mid-slope and lowland zones, the edge was proposed as the sampling habitat.

Snags were not recommended as a suitable sampling habitat for predictive modelling.

- Combined season models are recommended for all lowland zones to obtain the largest possible taxa list and average the effect of seasonal influences, providing the most robust assessment.

It was recommended that combined season models should not represent the only assessment in upland and mid-slope zones where taxa numbers were higher and that single season models should be used in preference to combined season models if only one assessment was done.

The frequency of assessment should be determined by the statistical requirements; the aim being to maximise the power to detect trends.

- The existing NRHP data-set is proposed as a baseline data-set for which gaps in test site coverage may be identified and further sampling recommended.
- Four options for the sampling and reporting scale were proposed:

- reporting at the river-valley scale only
- reporting at the Functional Process Zone scale, which allows the option of aggregating to the river-valley scale as well

- reporting at the scale of Valley Process Zones (upland, mid-slopes, lowland), which also allows the option of aggregating to the river-valley scale
- reporting assessments at the river-valley scale, with extra sites in areas of specific interest enabling reporting also at the FPZ scale for these catchments.

The decision on reporting scale should be determined by the taskforce prior to sampling, informed by the level of precision and number of sampling sites required. The ability to detect a change of 10% was considered the maximum acceptable for an AUSRIVAS O/E score and 5% for a SIGNAL score.

- Additional AUSRIVAS outputs not proposed as indicators were suggested as valuable aids for the interpretation of stream condition and ecosystem functioning, in conjunction with the conceptual models of river function.

The conceptual models of river function should be used in interpretation as well as AUSRIVAS outputs additional to those used as indicators. Key families should be identified for each FPZ or region and added to the conceptual models of river function.

PART B
Development of Methods and Indicators

Section B1
Indicators

To assess condition for the macroinvertebrate theme, measures of structure and sensitivity were chosen. Two indicators have been proposed, the AUSRIVAS O/E taxa score and a version of the SIGNAL score.

O/E SIGNAL as it is calculated currently in AUSRIVAS has been demonstrated to be insensitive in some cases, as has the raw SIGNAL score (see section 2.3 of review and Figures B1, B2 below). However, the SIGNAL index can give valuable information in many cases about the sensitivity of taxa to impacts and is worthwhile including as an indicator. Raw SIGNAL was used in the Index of Stream Condition (ISC), with threshold scores regionalised to resolve differences between SIGNAL scores for upland and lowland sites and different types of rivers. Use of SIGNAL in AUSRIVAS has the advantage of giving a site specific expected SIGNAL score for a site, taking into account other influences such as location and season in the expected score. However, AUSRIVAS only uses those taxa that are predicted with a 50% or greater chance of occurrence in the calculation of a SIGNAL score. It is considered that including all taxa collected may improve the sensitivity of SIGNAL in AUSRIVAS. Thus, development and testing of alternative calculation of O/E SIGNAL scores and regionalisation of scores is recommended for the first year of the Audit. As the SIGNAL O/E value as currently calculated is insensitive and the raw SIGNAL score is not regionalised, use of SIGNAL in either of its existing forms would be inaccurate. For example, Figures B1 and B2 show that SIGNAL O/E as it is currently calculated is insensitive to impairment and consistently over-estimates site condition. Therefore, the first year of reporting should only include the AUSRIVAS indicator while the robustness of the SIGNAL indicator is improved.

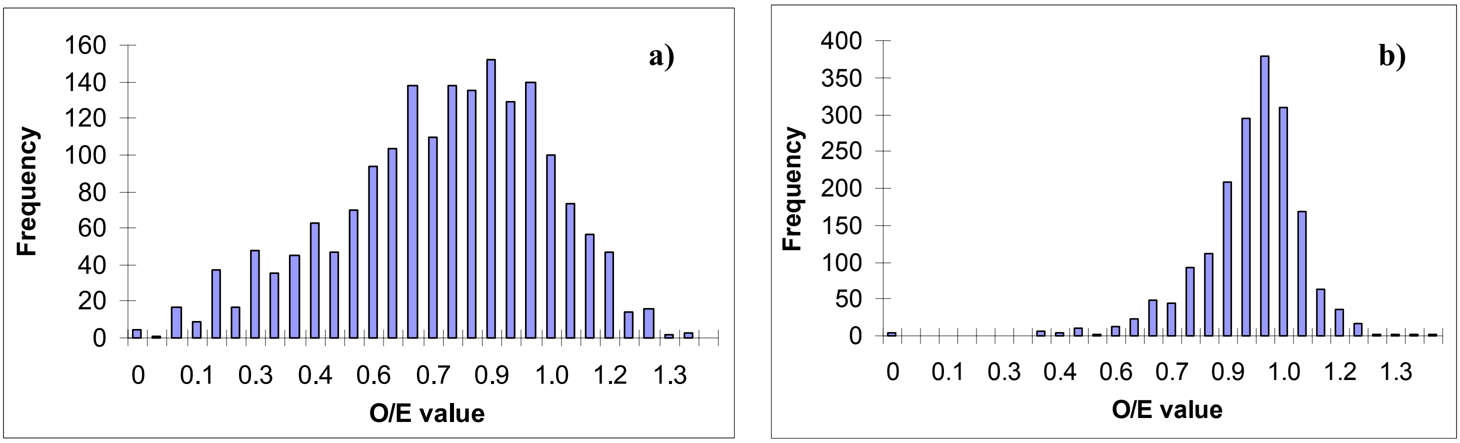


Figure B1 The same NSW AUSRIVAS test sites assessed using a) O/E taxa and b) O/E SIGNAL. The skewed distribution of O/E SIGNAL demonstrates the insensitivity of O/E SIGNAL as it is currently calculated to impacts.

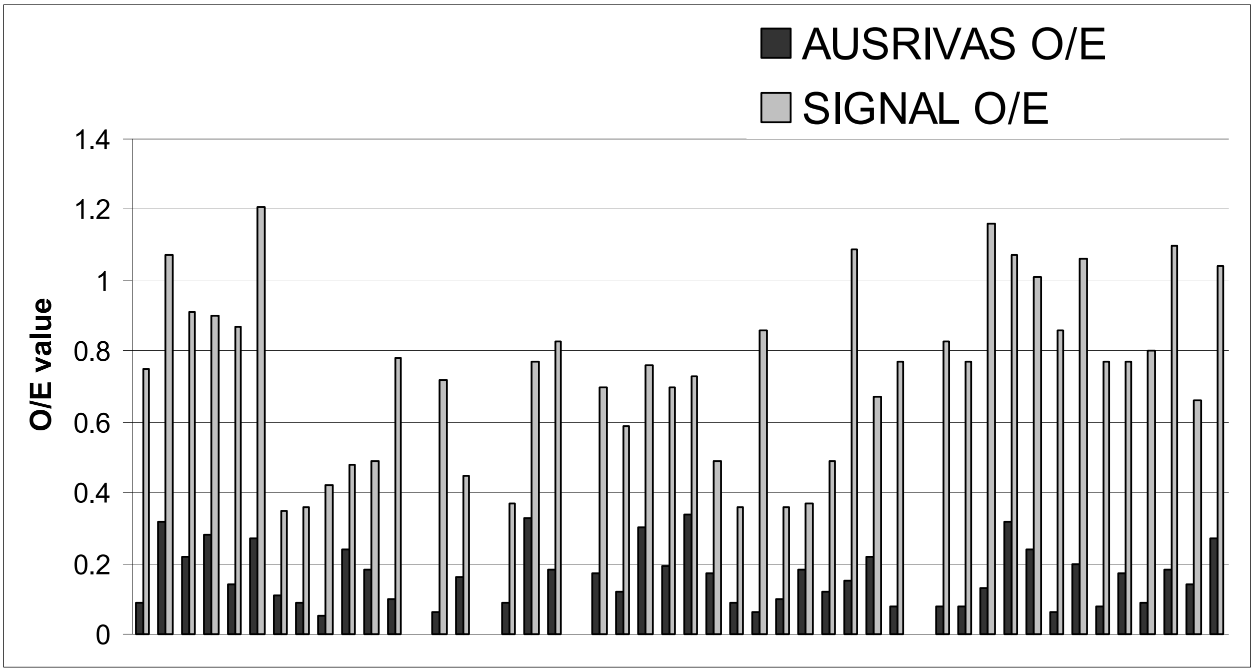


Figure B2 Test sites assessed as Band D by NSW AUSRIVAS models; SIGNAL O/E scores consistently over-estimate condition when compared to AUSRIVAS O/E taxa scores.

Section B2

Methods

B2.1 Sampling Methods

The sampling methods will follow the rapid sampling protocols currently used for AUSRIVAS models. This requires collection of macroinvertebrates from 10 m of a single habitat using a sweep/kick net. Samples are then processed by "live-picking" in the field or "lab-sorting" in a laboratory. Sample collection and processing methods are particular to jurisdictions and in some cases, models. Sample collection, processing and quality control procedures used by each jurisdiction are detailed in manuals available on the AUSRIVAS website (http://ausrivas.canberra.edu.au).

B2.2 Spatial coverage of AUSRIVAS models

The same spatial scale should be used for both AUSRIVAS and SIGNAL indicators, whether AUSRIVAS models or stand-alone indices are used to calculate SIGNAL scores. Scales proposed for consideration for models/indicators were statewide, Basin wide, regional and Valley Process Zones. Regionalisation was recommended by the review as the preferred option for testing and development. However, it is proposed that existing models be used initially for the first year of assessment while development issues for other models are investigated. A stepwise process is suggested for the selection of the most appropriate models for each region to be developed. Existing models should be evaluated and possible options for new models considered. A decision tree for selecting models is proposed in Figure B3. A similar decision tree may be used for determining the most appropriate SIGNAL scoring system across the Basin, e.g. are existing SIGNAL scores available?; are they appropriate for each FPZ?: yes— use existing; no— develop regional signal scores/references appropriate to each model/region.

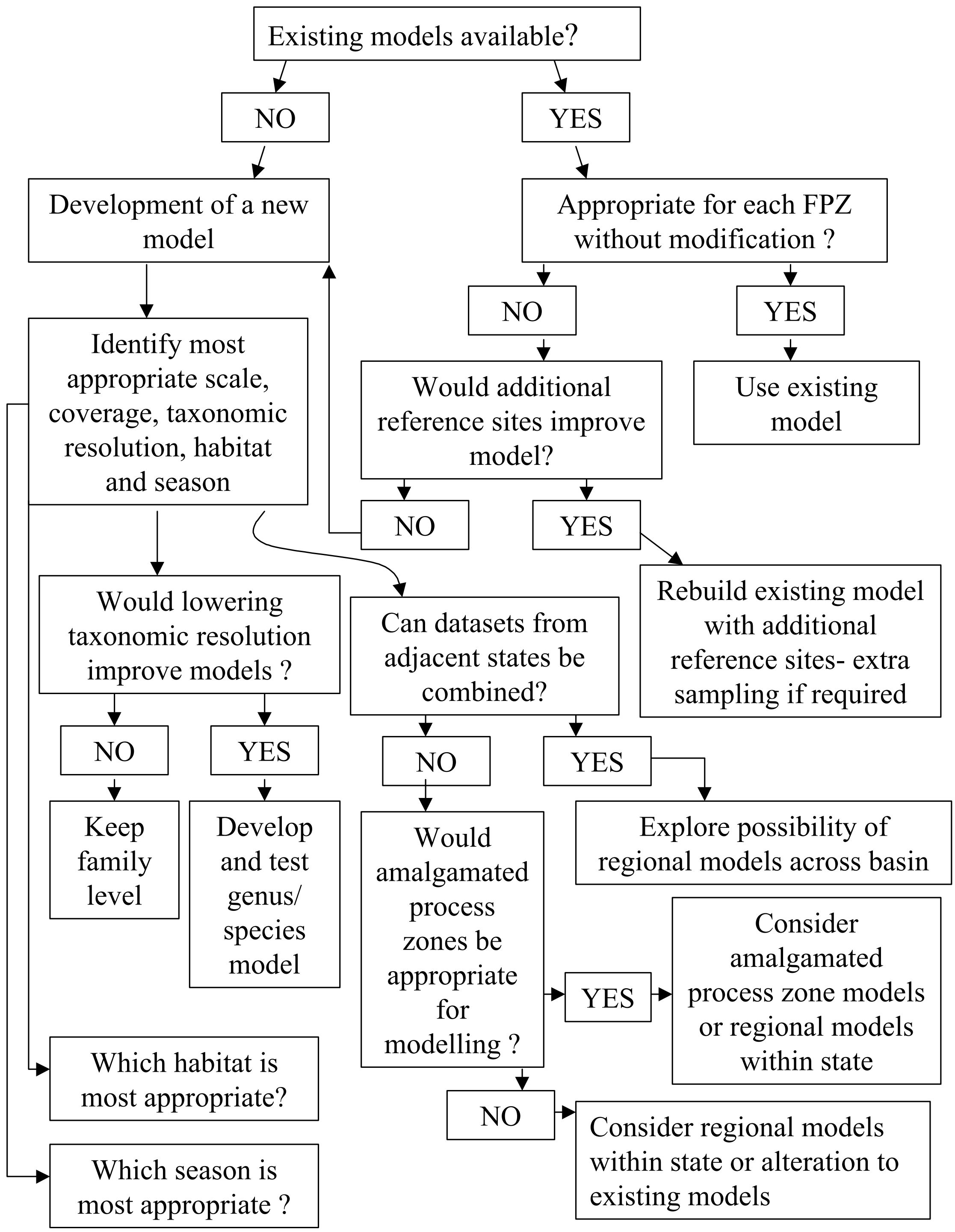


Figure B3 Decision tree for model selection

Table B1 Proposed modelling options for the Murray Darling Basin

Assessment for 1st year of Audit	**Pros**	**Cons**
-Use existing State models (Queensland, NSW, South Australia, ACT)	-can be used immediately	-not appropriate for all FPZs (problems with low taxa nos, ref sites, etc.)
-Use existing regional models for remaining parts of Victoria and western regional model for western NSW	-already developed and tested, improvement on State models, able to be used immediately	
Concurrently in 1st year of Audit		
-Evaluate existing models and explore development of regional model for western Queensland (possibly combine with NSW and Victoria western region models to form a western Basin model)	-increased taxa list and reference sites, more appropriate predictor variables, comparing rivers with low taxa numbers -Victorian and NSW western models already exist	-existing data must be able to be combined if one large western model was adopted
-Develop regional models for remaining parts of NSW and Queensland, South Australia	- improvement on existing State models	-new models required

B2.3 Frequency and season of assessment

The frequency of assessment and the sampling design should reflect the most statistically effective way to maximise the power of the design to detect both spatial and temporal trends. Two main options are available, to sample more intensively and less frequently and to sample less intensively but more frequently.

To sample more intensively and less frequently suggests more interest in quantifying the nature of the long-term trend in the data (Figure B4a). To sample less intensively and more frequently suggests more interest in quantifying the long-term trend and the seasonal fluctuations in the data (Figure B4b). Subsequently, it is likely that the best sampling strategy is dependent on the index being used but ultimately it may be a compromise between two conflicting needs. If quantification of the long-term trend is the priority then increased power is available by reducing variability in the model by sampling only in one season.

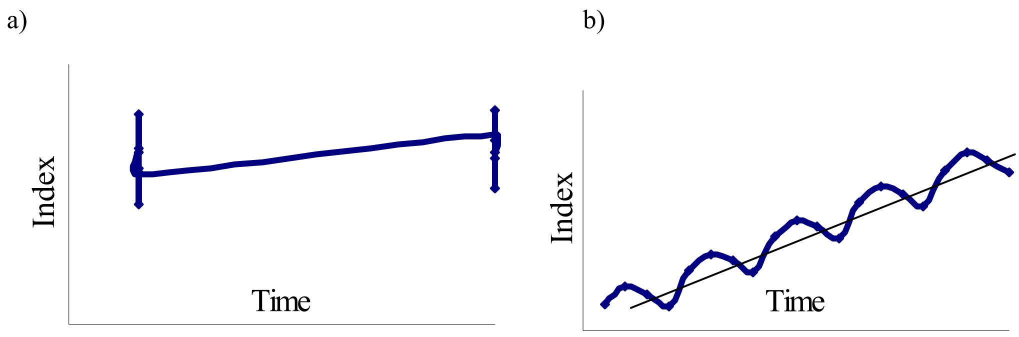


Figure B4 a) Sampling less often but more intensively; gives more accuracy to the estimate of the slope of the trend, **b)** sampling more often but less intensively; also estimates the slope but can estimate or partition out the effects of seasonal fluctuation.

Thus, the consequences of each sampling option are as follows:

1) sampling twice a year (combined season model or two single season assessments)

- more frequent, less intensive; also estimates long term trend but partitions out the effect of seasonal fluctuations
- sampling twice a year required (cost implications)
- the number of taxa collected is maximised, but an averaged assessment of condition from two seasons is provided

2) sampling once a year (single season model)

- less frequent, less intensive, more extensive — enables a greater number of sites to be sampled
- one sample per site per year, more cost effective
- single snapshot measurement of site provided

While seasonal influences have been shown to be an important factor in the analysis of macroinvertebrate data, use of an O/E score takes into account seasonal influences in the expected value. Therefore, if O/E indicators are used it is not necessary to sample more frequently and account for seasonal variation. Thus, of the options proposed, a single season model is likely to be the most effective for detecting long term trends at the valley scale where the numbers of taxa are high enough to provide robust predictions, because sampling density can be increased for the same cost. However, in western regions, combined season models may be required to provide an adequate taxon list. Thus, a cost effective option may be to use single season models in upland areas, possibly sampling the full range of sites in different seasons to even out costs and workload, and using combined season models in lowland areas. However, if individual State agencies see the need for seasonal site assessments they may decide to sample all sites in both seasons and have the opportunity for both individual and combined season assessments. Inevitably, this will reduce the overall number of sites that can be assessed in the available budget. Ultimately the number of sites that can be sampled will be determined by the reporting scale chosen.

B2.4 Sample habitat

It was recommended by the review, that the most appropriate sampling habitat for each FPZ/VPZ should be used. In most cases an appropriate habitat is already being used for assessments, but where this is not the case new models may have to be developed.

B2.5 Taxonomic resolution

It was proposed by the review that species and genus models already developed for Western Victoria should be used to test whether changing the taxonomic resolution is likely to be effective in increasing taxa richness in lowland zones. A main concern is that models are more sensitive, not just more variable, when compared to existing AUSRIVAS models. Accuracy of assessments will be tested by analysing how Western Victorian reference site assessments change over time, assuming that little overall change in assessment would be expected in a reference site that has not been impaired. After testing, models with altered taxonomic resolution should be developed where appropriate.

Section B3

Analysis

Analysis of both indicators, O/E taxa and O/E SIGNAL, will be against a reference condition. Analysis using AUSRIVAS calculates what is expected in terms of reference condition for a particular site and compares this to what is observed (O/E ratio). This indicator generally ranges from 0 to 1, although it is possible to exceed one in some circumstances. Both AUSRIVAS O/E taxa and O/E SIGNAL scores are calculated in this way, giving a score of condition for a site in terms of an appropriate reference. Use of a regionalised raw version of SIGNAL uses a predetermined reference threshold that is not site specific.

B3.1 Reference condition and possible adjustment

It was recommended by the review that assessment methods for the macroinvertebrate theme should incorporate a measure of departure from reference and departure of reference from natural, to account for the varying definition of reference condition currently used. To do this, reference sites can be scored against criteria; for example, the criteria used for selecting reference sites for Queensland in the WAMP process (Table B2), The River Disturbance Index, the conceptual models of river function for each FPZ or even against a narrative description.

Table B2 Selection criteria for reference sites used in the Queensland WAMP process.

No.	Reference Condition Selection Criteria	Score
1	No intensive agriculture within 20 km upstream. Intensive agriculture is that which involves irrigation, widespread soil disturbance, use of agrochemicals and pine plantations. Dry-land grazing does not fall into this category.	3 2 1
2	No major extractive industry (existing or historical) within 20 km upstream. This includes mines, quarries and sand/gravel extraction.	3 2 1
3	No major urban area (> 5000 population) within 20 km upstream If the urban area is small and the river large, this criterion can be relaxed.	3 2 1
4	No significant point source waste water discharge within 20 km upstream. Exceptions can again be made for small discharges into large rivers.	3 2 1
5	No dam or major weir within 20 km upstream. Sites within the ponded area of impoundments also fail. Sites failing this criterion automatically fail the overall assessment.	3 2 1
6	Seasonal flow regime not greatly altered. This may be by abstraction or regulation further upstream than 20 km. Includes either an increase or decrease in seasonal flow.	3 2 1
7	Riparian Zone of natural appearance. Riparian vegetation should be intact and dominated by native species.	3 2 1
8	Riparian Zone and banks not excessively eroded beyond natural levels or significantly damaged by stock. Stock damage to the stream bed may be included in this category.	3 2 1
9	Stream Channel not affected by major geomorphological change. Geomorphological change includes bank slumping, shallowing, braiding and unnatural aggradation or degradation.	3 2 1
10	Instream conditions and habitats not altered. This may be altered by excessive algal and macrophyte growth, by sedimentation and siltation, by reduction in habitat diversity by drowning or drying out of habitats (e.g. riffles), by direct access of stock into the river	3 2 1
	SUM	/30

For example, this scoring method was used for the WAMP process: each criterion is assessed as a ‘pass’ or ‘fail’. A ‘pass’ is automatically assigned a value of ‘3’; a ‘fail’ is then allocated a numerical value according to the level of impact:

1. greatest impact
2. moderate impact
3. no/little impact.

Scores are summed and reference condition assigned to sites where total is >26 (except where Criterion 5 scores a ‘1’ or ‘2’, in which case the site records a ‘fail’).
For the Audit, scores could be used to provide a measure of departure from reference. e.g. if reference is considered a score of 26, a site with a score of 17 would have moved away from reference by 35% (-0.35).
Steps proposed are:

- *Agree on criteria for “less disturbed” or “better managed” sites.*

- *Develop measures of departure of fauna from reference (O/E taxa reference=1) and site from reference (scoring against reference criteria or other).*
- *Develop measures of departure of environment from natural (e.g. existing vs modelled natural flow regime, River Disturbance Index (RDI), conceptual model or narrative description).*
- *Combine these measures into alternative "health" indicators that can be aggregated to the valley scale.*
- *Test these alternative indicators for sensitivity to known disturbances.*

A decision hierarchy (Figure B5) can aid selection of the most appropriate reference sites. This decision tree has already been used informally for reference site selection for existing AUSRIVAS models by most jurisdictions. However, measures of departure from reference and natural have not been used. Therefore, in most cases, more appropriate reference sites are not likely to be available and existing models will be used. Where additional sites are added, the AUSRIVAS models will have to be rebuilt.

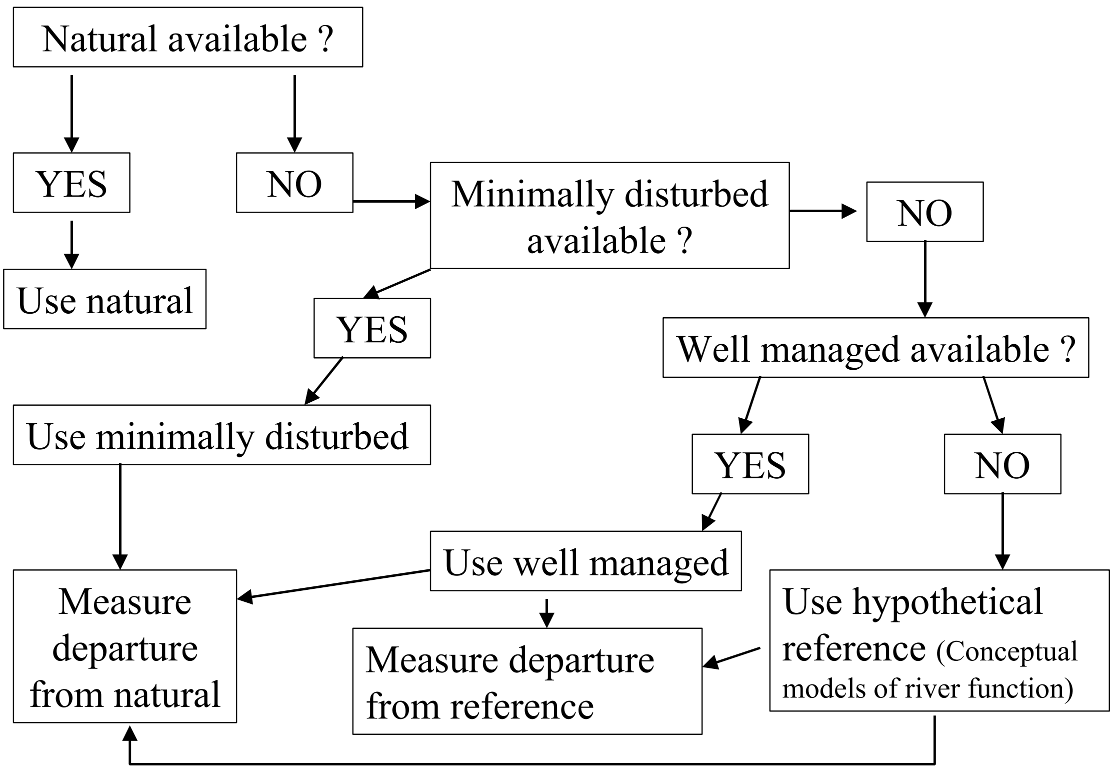


Figure B5 Reference condition decision tree

B3.2 Integration and aggregation of scores to produce an assessment

No integration of indicators will be done in the first year of the Audit to produce a macroinvertebrate assessment, as SIGNAL in its existing forms has been shown to be insensitive in a number of cases. Thus, only the AUSRIVAS score should be reported in the first year while the sensitivity of the SIGNAL score is improved. After the first year, integration of a new version/calculation of SIGNAL indicator with the AUSRIVAS O/E indicator should follow the procedure reported by Barmuta *et al.* (1996) below, to produce a single score if required. However, caution should be used in reporting a single number and where possible indicators should be reported separately, as they represent different information about the health of a stream. The AUSRIVAS O/E taxa score reports on structural features of the community, whereas the SIGNAL score informs on sensitivity of macroinvertebrate taxa. Rather than combine indicators, Barmuta *et al.* (1996) recommend reporting of the indicator score that is the further away from reference, to be conservative.

Barmuta *et al.* (1996) state

> '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 indices could be combined (Institute of Freshwater Ecology 1991), and several options were canvassed in the regional workshops and discussed by the expect panel. 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.'

Aggregation of scores to the river-valley scale/VPZ scale will follow the general principles of aggregation that will be used for each theme, reporting theme condition for a river-valley as a median river-valley score (see Final Report).

Section B4

Sampling design

B4.1 Precision and number of sampling sites

Detection of a 10% change (a change of 0.1 in an O/E score) has been recommended as appropriate for the AUSRIVAS O/E score and 5% change for the SIGNAL score. For a valley assessment, this means that one macroinvertebrate taxon lost out of 10 macroinvertebrate taxa present would be able to be detected. This is considered a significant level of change that should be detectable. However, because the SIGNAL scoring system is less variable, smaller changes of 5% need to be able to be detected in the index (Figures A3, A4). Detection of 10% change in an AUSRIVAS O/E taxa score and 5% in an O/E SIGNAL score on average represents approximately half a band change in each indicator, respectively. Hence, these precision levels are considered the most appropriate and biologically sensitive levels for each indicator. An analysis of the sample size required to achieve this level of precision was performed for the AUSRIVAS O/E score to identify if such levels of precision were realistic. No sample size analysis was performed for the SIGNAL indicator as the version of SIGNAL that will be used has not yet been determined and reporting of this indicator is not required in the first year of the audit.

The number of samples required (sites/reaches assessed) for the Audit depends on the:

- spatial reporting scale of the assessment
- variability of the indicator
- initial condition score of the indicator
- aggregation and reporting statistics used
- desired level of change to detect
- confidence in detecting that change.

The Audit framework attempts to explicitly identify the implications and tradeoffs associated with these sample design issues.

To measure the condition of rivers in the Murray Darling Basin the spatial scale of inference for a measure could be determined and the number of those spatial units that fitted into the largest spatial unit for reporting calculated. For example, if an AUSRIVAS OE50 score is determined to be representative of a 10 km section of river and there are 77358.2 km of river in the MDB, then to sample the MDB precisely would take at least 7736 sites. Obviously this is an unrealistic number of sites and so a sampling regime must be determined that allows inferences to be made at a broad scale (e.g. river-valley), but still allows measurements to be made at the small scale (e.g. site). The sample unit size needs to be known to determine the number of samples required, which is related to the reporting scale and how variable the measure is at the broad scale. For example one could wish to make inference about a measurement at the catchment, river-valley, process zone, reach or site scale and subsequently design a sampling unit size that is different at each scale but the statistical distribution of the measurement (which will affect the number of collections needed) may also differ scale-dependently.

In other words imagine the impairment in river condition is at the Valley Process Zone (VPZ) level. A river valley may have good quality source and transport zones, but the depositional zone may be impaired from say agricultural practices. A good sampling strategy designed for reporting at the VPZ scale obviously should quantify the health in each VPZ and locate the impairment in the depositional zone. A sampling strategy designed for reporting only at the river-valley scale, however, will not necessarily give conclusive evidence as to which VPZ is impaired and in fact may say that on average the river valley is in good condition. Another way of depicting this is to imagine a river valley has a number of impaired reaches randomly distributed across VPZs. If the strategy is to sample sites randomly across the river valley as a whole, or even each VPZ as a whole, the variability of the measure will be a lot greater than if measuring within each reach.

To design an effective sampling strategy, knowledge is required of the distribution of the index at each of the site, reach, process zone and river-valley zone scales. The sample sizes required for each environmental theme are based on an analysis of existing data-sets where they are available. The most comprehensive data-set for this style of analysis is the macroinvertebrate data collected for the First National Assessment of River Health (FNARH). The following discussion refers specifically to this data-set as well as a simulated data-set, (described below), but the principles apply for each environmental theme.

A pseudo data-set of OE50 values for a fictitious river valley was created that consisted of 3 VPZs, 75 reaches (25/VPZ) and 600 sites (8/reach). In each site a normal distribution of 100 possible OE50 scores that could be sampled was generated (mean of 1 and standard deviation of 0.1). For use as the baseline condition and confirmation of the procedure, the following scale dependent sampling strategies were performed on the data-set:

- 20 OE50s randomly from across all 60,000 possible values for the river valley,
- 20 OE50s randomly within the 20,000 possible values within each of the three VPZs
- 20 OE50s randomly within the 800 possible values within each of the 25 reaches within each of the three VPZs
- 20 OE50s randomly within the 100 possible values within each site within each of the 25 reaches within each of the three VPZs.

The sample size of 20 has no bearing whatsoever on the final estimates of sample size for the Audit. It was simply a computationally convenient size and expected to return a reliable estimate of the standard deviation at each sampling scale.

As expected, the samples always returned comparable values for the true mean and standard deviation regardless of sampling scale (Table B3).

The sample sizes required to detect changes in the average OE50 for each scale were then calculated using the iterative formula described in Zar (1984). As the Audit has proposed a stratified random sampling strategy using VPZs (n = 3) for stratification, a further two degrees of freedom were added to each sample size (Table B3).

Naturally the sample sizes calculated on the 'reference' data show that the same sample size is required at each scale of measurement. For example if the precision is required at the river-valley scale then a sample size of n will be enough to detect a change in the mean

OE50 of 0.1. If the same precision is required at the VPZ then a sample of size *n* is required within each VPZ, and so on (Table B3).

It is assumed more likely, however, that impairments could occur in individual sites, individual reaches, across VPZs, or combinations of these scales. Therefore, in a series of subsequent models, sites were manually impaired sites using the following methods:

- ***VPZ Minor:*** Sites in one Valley Process Zone were impaired by subtracting 0.4 of a randomly selected uniformly distributed variate (ranging from 0 to 1) from each OE50 in that VPZ; impaired sites in another valley were created by subtracting 0.2 of a uniformly distributed variate from each OE50; the other Valley Process Zone was left as reference.
- ***VPZ Major:*** Sites in one Valley Process Zone were impaired by subtracting 0.4 of a randomly selected uniformly distributed variate (ranging from 0 to 1) from each OE50 in that VPZ; impaired sites in another valley were created by subtracting 0.8 of a uniformly distributed variate from each OE50; the other Valley Process Zone was left as reference.
- ***Random Reaches Minor:*** 25% of reaches were randomly selected and impaired by subtracting 0.8 of a uniformly distributed variate from each OE50 within an impaired reach; the other reaches were left as reference.
- ***Random Reaches Major:*** 50% of reaches were randomly selected and impaired by subtracting 0.8 of a uniformly distributed variate from each OE50 within an impaired reach; the other reaches were left as reference.
- ***Sites Minor:*** 25% of sites were randomly selected and impaired across the whole catchment by randomly subtracting between 0 and 0.8 (uniformly distributed) from the OE50 values.
- ***Sites Major:*** 50% of sites were randomly selected and impaired across the whole catchment by randomly subtracting between 0 and 0.8 (uniformly distributed) from the OE50 values.
- ***Random:*** All sites were randomly impaired across the whole catchment by randomly subtracting between 0 and 0.8 (uniformly distributed) from the OE50 values.

Finally the actual between-site variation in the OE50 was investigated using existing data from the Ovens, Murrumbidgee and Condamine river valleys. These data allowed estimates of sample sizes required for making inference at the river-valley and Valley Process Zone levels, but not at finer resolutions. The advantage of using the real data is that it allows estimation of the actual sample sizes (i.e. how many sites to sample) in each of the Valley Process Zones, whereas with the artificial data the VPZs would be assumed to be equal for computational simplicity. This procedure allows exact sample size estimates for these three particular river valleys at the river-valley, the VPZ or the river-valley + VPZ scales of reporting. The disadvantage, however, is that it is unknown how much of the observed variability in OE50s is from within-site variation. It is worth noting that the three trial river valleys displayed different types of impairment, with the Ovens showing considerable between-VPZ variation and the other two having relatively even proportions of impaired sites between VPZs. Overall between 0.26 and 0.50 of all sites in each river valley were impaired in the trial data, suggesting that in the final analysis the sampling strategy may have to be determined individually for each river valley. Therefore, the sampling strategy will require review after the first round of sampling, when individual river-valley variability is better understood.

Table B3. Summary of observed variability and sample size requirements for measuring the AUSRIVAS OE50 Index at four spatial scales of measurement from an artificial data-set and from the Audit trial data in three river valleys. For example, if the river valley had about 25% of raches impaired and was randomly positioned among Valley Process Zones (Reach Minor), then to detect a change of 0.15 in the average OE50 whilst sampling and reporting at the river-valley scale with (a) $\alpha = 0.10$ (Type I error = probability of concluding there is a difference when in reality there is not) and (b) Power (1-β) = 0.80 (β =Type II error = probability of concluding there is no difference when in fact there is) there would need to be 14 sites sampled (Boxed). If sampling and reporting were required at the Valley Process Zone level, there would need to be 12 sites sampled per process zone; at the reach level, there would need to be 7 sites per reach, etc. The shaded regions are the best available estimates using the trial data and the traditional values of $\alpha = 0.05$ and Power = 0.80; that is, to report at the river-valley level in the Ovens River would require 21 sites. To report in the Ovens River at the river-valley *and* the Valley Process Zone level would require 45 sites.

					a = 0.10, b= 0.80			a = 0.10, b= 0.90			a = 0.05, b= 0.90			a = 0.05, b= 0.80		
					Size of diference to detect											
	Number of Samples	**Type of Impairment**	**True Mean**	**True Standard deviation**	**0.15**	**0.1**	**0.05**	**0.15**	**0.1**	**0.05**	**0.15**	**0.1**	**0.05**	**0.15**	**0.1**	**0.05**
	60000	Reference	1.000	0.100	6	9	27	6	10	30	7	11	32	7	10	29
		VPZ Minorly	0.900	0.149	9	16	55	10	18	62	11	20	68	10	18	61
		VPZ Majorly	0.800	0.240	18	37	138	20	41	155	22	45	170	20	41	153
		Reach Minor	0.931	0.204	14	28	101	16	31	113	17	34	124	16	30	111
		Reach Major	0.766	0.283	24	50	191	27	56	214	29	61	235	27	55	211
		Sites Minor	0.795	0.278	23	49	184	26	54	207	28	59	226	26	54	204
		Sites Major	0.699	0.283	24	50	191	27	56	214	29	61	235	27	55	211
		Random sites	0.600	0.250	20	40	150	22	45	168	24	49	184	22	44	166
Reporting/ Sampling Scale			Average sample mean	Average sample standard deviation												
River Valley	1	Reference	0.988	0.096	6	9	25	6	10	28	7	10	30	6	10	27
		VPZ Minorly	0.876	0.164	10	19	67	11	21	74	12	23	81	11	21	73
		VPZ Majorly	0.786	0.252	20	40	152	22	45	170	24	49	186	22	45	168
		Reach Minor	0.888	0.202	14	27	98	15	30	110	17	33	121	15	30	109
		Reach Major	0.771	0.319	30	63	242	33	71	272	36	77	298	33	70	268
		Sites Minor	0.746	0.320	30	63	243	33	71	273	36	77	299	33	70	269
		Sites Major	0.835	0.268	22	45	172	24	51	193	27	55	211	24	50	190
		Random sites	0.828	0.261	21	43	162	23	48	182	25	53	199	23	48	180
VPZ	3	Reference	1.003	0.096	6	9	25	6	10	28	7	10	30	6	10	28
		VPZ Minorly	0.885	0.144	9	15	52	10	17	58	10	19	63	10	17	57
		VPZ Majorly	0.781	0.189	13	24	87	14	27	97	15	29	106	14	27	96
		Reach Minor	0.918	0.183	12	23	81	13	25	91	14	28	100	13	25	90
		Reach Major	0.747	0.320	30	63	243	33	71	273	36	77	298	33	70	269
		Sites Minor	0.683	0.292	25	53	202	28	59	227	31	65	249	28	59	224
		Sites Major	0.773	0.292	26	53	203	28	60	228	31	65	250	28	59	225
		Random sites	0.822	0.229	17	34	126	19	38	142	20	42	155	19	38	140
Reach	75	Reference	1.001	0.097	6	9	25	6	10	28	7	10	31	6	10	28
		VPZ Minorly	0.902	0.122	7	12	38	8	13	42	8	14	46	8	13	42
		VPZ Majorly	0.805	0.165	10	19	67	11	21	75	12	23	82	11	21	74
		Reach Minor	0.935	0.123	7	12	38	8	13	43	8	14	47	8	13	42
		Reach Major	0.771	0.187	12	24	86	14	27	96	15	29	105	14	26	95
		Sites Minor	0.717	0.274	23	47	180	25	53	202	28	58	221	25	52	199
		Sites Major	0.804	0.265	22	44	168	24	50	189	26	54	206	24	49	186
		Random sites	0.789	0.222	16	32	119	18	36	133	19	39	146	18	36	132
site	600	Reference	1.001	0.098	6	9	26	6	10	29	7	11	31	6	10	28
		VPZ Minorly	0.902	0.123	7	12	39	8	13	43	8	14	47	8	13	42
		VPZ Majorly	0.802	0.168	11	20	69	12	22	77	13	24	85	12	22	76
		Reach Minor	0.932	0.125	7	13	40	8	14	45	9	15	49	8	14	44
		Reach Major	0.768	0.188	13	24	86	14	27	96	15	29	106	14	26	95
		Sites Minor	0.703	0.215	15	30	111	17	34	125	18	37	136	17	33	123
		Sites Major	0.796	0.177	12	22	77	13	24	86	14	26	94	13	24	85
		Random sites	0.757	0.223	16	32	120	18	36	134	20	39	147	18	36	133
Trial Data River Valley		Ovens	0.933	0.166	11	19	68	12	22	76	12	23	83	11	21	75
		Murrumbidgee	0.779	0.196	13	26	93	15	29	105	16	31	114	15	28	103
		Condamine	0.785	0.194	13	25	91	14	28	102	16	31	112	14	28	101
VPZ		Ovens source	0.985	0.146	7	14	51	8	15	57	8	17	63	8	15	57
		Ovens transportational	1.015	0.105	4	8	27	5	9	31	5	10	34	5	9	30
		Ovens depositional	0.820	0.164	8	17	65	9	19	73	10	21	80	9	19	72
		Murrumbidgee source	0.716	0.218	14	29	112	15	33	126	17	36	138	15	32	125
		Murrumbidgee transportational	0.806	0.090	4	6	20	4	7	23	4	8	25	4	7	23
		Murrumbidgee depositional	0.797	0.198	12	24	93	13	27	105	14	30	115	13	27	103
		Condamine source	0.780	0.197	11	24	92	13	27	104	14	30	113	13	27	102
		Condamine transportational	0.779	0.210	13	27	104	14	30	117	16	33	128	14	30	116
		Condamine depositional	0.793	0.186	10	21	82	12	24	92	13	26	101	11	24	91
VPZ & River Valley		Ovens			21	41	145	24	45	163	25	50	179	24	45	161
		Murrumbidgee			32	61	227	34	69	256	37	76	280	34	68	253
		Condamine			36	74	280	41	83	315	45	91	344	40	83	311

The major finding of the sample size calculations is that considerable variability in sample size can be encountered, particularly within and between river valleys. Looking at the common use values of $\alpha = 0.05$, Power (1-β) = 0.80 and requiring to detect a change in the OE50 of 0.1, the artificial data-set required between 10 and 70 sites to be sampled for reporting at the river-valley scale. Obviously the type and location of impairment immensely influences the variability of the OE50 depending on the scale at which sampling is carried out. Analysis of the trial data found that 21 sites were needed to report at the river-valley scale in the Ovens River, and 28 sites for the other two river valleys. If reporting was also required at each Valley Process Zone the sample sizes were 45, 68 and 83 per river valley (Table B3).

As all the river valleys have different lengths and different proportions of each Functional and Valley Process Zone, a concern regarding the sampling design is that the variability of the index may be influenced by spatial scale. This issue was investigated using the trial data, and results showed the variability of the index was influenced by the type of process zone, regardless of its length.

Summary

The exact sample sizes required to detect changes in the AUSRIVAS OE50 score cannot be precisely calculated in advance of a pilot study because:

- the true within site variability of the OE50 score is unknown
- the samples sizes required at the proposed sampling/reporting scale (river-valley) vary considerably depending on different types and levels of impairment
- the finer the scale of impairment, the more variable the indicator at higher reporting scales
- the samples sizes required at the proposed sampling/reporting scale (river-valley) will certainly be different for each river valley but, based on existing data, 30 sites per river valley would achieve an acceptable level of precision.

All these issues can be addressed after the first round of sampling.

Recommendations

We may possibly speculate that the three trial data-sets are representative of the expected variability in the 29 river valleys to be sampled. This is reasonable because the Ovens is a smaller river valley (1000 km of river) and has a relatively high proportion of Source process zone (Percentage of catchment area that is Source:Transitional:Deposition = 48:21:32), the Condamine is one of the largest river valleys (11000 km) with a relatively large proportion of Depositional process zone (7:34:59) and the Murrumbidgee is in between (6500 km) but has a relatively small proportion of transitional process zone (22:6:72).

Without considering the cost–benefit aspect it is therefore recommend that the ideal sample size for the first round of sampling is 30 sites per river valley. This includes the 28 as determined in the sample size analysis of the trial data and includes 2 extra sites to compensate for rounding in the stratification process and to ensure that a minimum of three sites are positioned in any given VPZ. The MDBC can then be 95% confident of obtaining the true average AUSRIVAS OE50 score for each river valley, knowing that future sampling rounds will also be able to detect changes in river condition at the river-valley scale.

B4.2 Reporting scale

While reporting the mean index score is required at the river-valley scale, models used don't need to be at this scale, because sampling sites form replicates from which results can be amalgamated to provide a valley score.

As well as reporting at the river-valley scale, the options of reporting at the Valley Process Zone scale or the Functional Process Zone scale were also proposed by the review. The sample size analysis indicates that considerably more sites are required to report accurately at these scales. The cost implications of reporting at the valley scale, the VPZ scale and the FPZ scale are estimated in the next section.

Section B5

Approximate costing

B5.1 Reporting at the valley scale

Costs will vary with proximity of sites, condition of river valleys and whether habitat and water quality sampling overlaps with variables required for AUSRIVAS models. Based on the average cost of $750 per site, the approximate cost for sampling and analysis of a river valley for the macroinvertebrate theme is likely to be around $22500 (Table B4). With a fixed cost of $60000 for the 80 additional reference sites needed across the Basin, the approximate cost for sampling the Basin is $714500 for a sampling round.

Table B4 Approximate costing for reporting the macro invertebrate index at the River-valley scale using the Significance level of 0.05 and Power of 0.8.

Macroinvertebrate sampling, processing, identification, analysis	$ 450 per site
Physical/chemical water quality measures required for models	$ 300 per site
Cost per Site	$ 750 per site
Cost per River Valley using recommended 30 sites	$ 22500 per river valley
Added second sample season in lowland zones	$ 7250 per river valley
Basin Wide cost for test site sampling across 22 River Valleys	$654500 for Basin
80 additional Reference sites across entire Basin @$750 each	$ 60000 for Basin
Total Basin Costs	$ 714500

B5.2 Alternative costing and benefits

The major trade off in costing occurs when reporting at finer scales or when detecting small changes in the OE50 score. The differences between using type I error rates of 0.05 or 0.1, or type II error rates of 0.1 or 0.2 are small relative to the difference in reporting scales (Figure B6). As an example, to report with the same levels of type I and type II error and level of detection as above, but at the river-valley and Valley Process Zone level results in a Basin-wide cost of $1.62M whilst the river-valley and Functional Process Zone level reporting is projected to cost $ 3.01M (Figure B6).

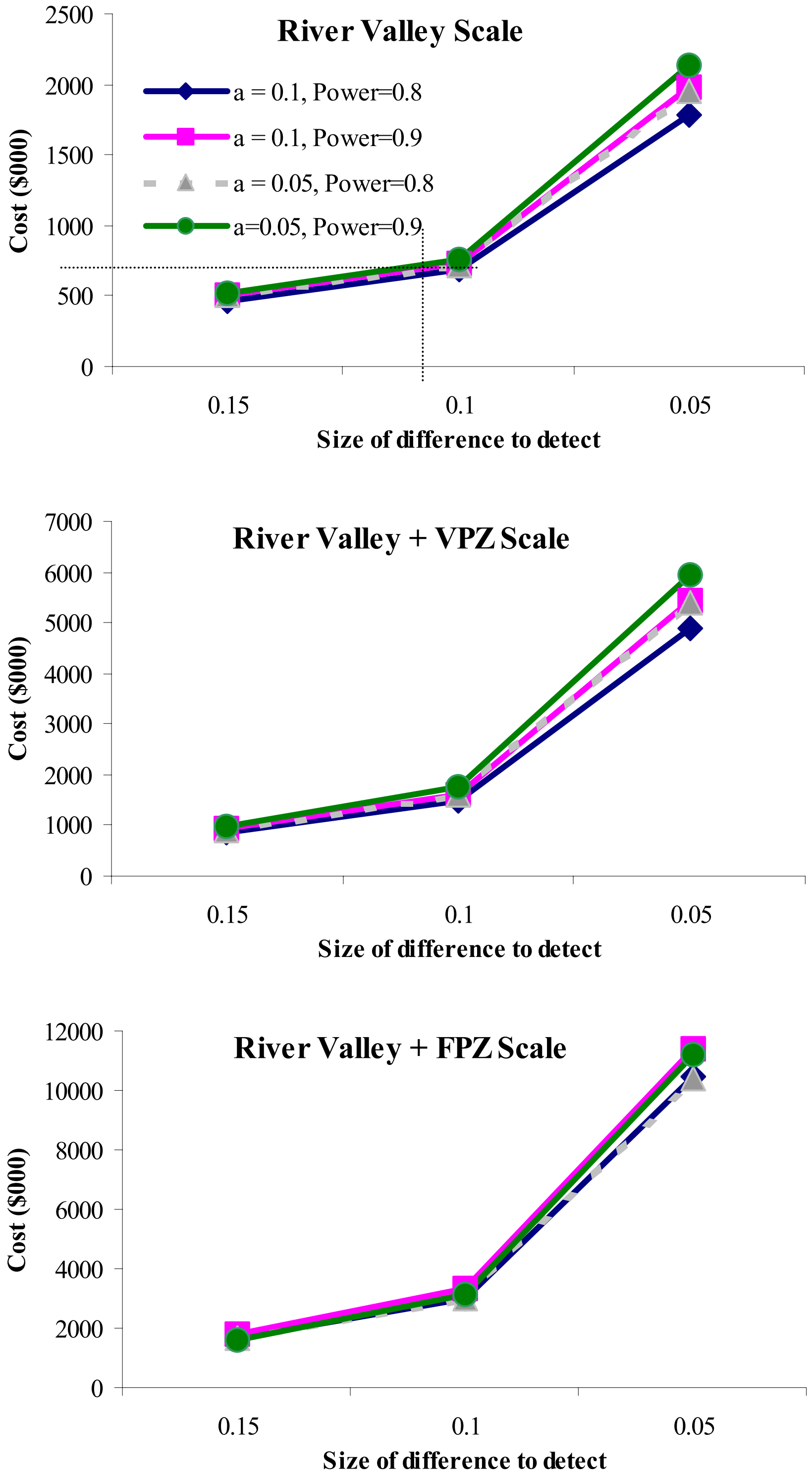


Figure B6: Trade off between total Basin cost and scale of reporting for macroinvertebrate theme for Audit. Note the vertical axes are different for each figure. Dotted lines indicate recommended cost for first round sampling.

References

Barmuta LA, Chessman BC and Hart BT (1996) Interpretation of the Outputs from AUSRIVAS. Second Milestone Report of Project UTA3 to the Land and Water Resources Research and Development Corporation.

Bennison GL, Hillman TJ and Suter PJ (1989) Macroinvertebrates of the River Murray. Survey and Monitoring: 1980–1985. Murray Darling Basin Commission.

Chessman BC (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 BC (2001) Pressure - Biota - Habitat Stream Assessment. Report on field trials in 1999–2000. NSW Department of Land and Water Conservation.

Ladson AR and White LJ (1999) *An Index of Stream Condition. Reference Manual*. 2nd Edn. Department of Natural Resources and Environment, Victoria.

Metzeling L and Miller J (in press). Evaluation of the sample size used for the rapid bioassessment of rivers using macroinvertebrates. *Hydrobiologia.*

Nichols SJ and Norris RH (Submitted) Bioassessment precision using macroinvertebrates: effects of area sampled and replication. *Freshwater Biology.*

Sloane PIW and Norris RH (Submitted) Bioassessment: Are predictive model outputs related to a pollution gradient? *Freshwater Biology.*

Simpson JC and Norris RH (2000) Biological assessment of river quality: development of AUSRIVAS models and outputs. *Assessing the biological quality of fresh waters: RIVPACS and other techniques*. (Eds JF Wright, DW Sutcliffe and MT Furse). Freshwater Biological Association, Ambleside, UK. pp.125–142.

Turak E, Waddell N and Hose G (2000) First National Assessment of River Health; New South Wales Program, Milestone Report 3. NSW Environment Protection Authority, Sydney NSW.

Zar JH (1984) *Biostatistical Analysis* (2nd Edition). Prentice-Hall International.

Final Report
Project R2004
Development of a Framework for the Sustainable Rivers Audit

Appendix 4
Fish

Review and Development of Fish Assessment Protocols

Peter Davies

Cooperative Research Centre
for Freshwater Ecology

Table of Contents

1. Summary

This report recommends a program of work to be conducted within the Sustainable Rivers Audit (Audit) aimed at:

- finalising a standard methodology for fish bioassessment across the MDB;
- conducting the first Basin-wide assessment of river health using fish data.

A specialist workshop, focused on fish-based bioassessment for the Audit, was attended by key personnel from the relevant Murray-Darling Basin agencies (MDBAs) in April 2001. Key issues discussed at the workshop were:

- the absence of and the need for standardised sampling methods across the Murray-Darling Basin (MDB);
- the need for a standard set of variables and derived measures (‘metrics’) which describe fish communities at a range of levels of organisation from the individual to the community level;
- the difficulty in defining reference conditions for fish within the MDB;
- the need for a single analytical method (‘framework’) for making comparisons of metrics against expectations or reference conditions;
- the need for outputs from the assessment which are readily understood and communicable to river managers in the MDB.

All the issues were discussed and agreement reached on a program of activities and surveys to be conducted in the first five-year Audit period in order to implement fish-based bioassessment within the Audit.

All fishery agencies within the MDB use a suite of active and passive fishing gear with survey programs involving varying combinations of electrofishing, nets and traps, with the exception of South Australia which relies only on collection of recreational and commercial fishery data. There was little agreement between MDBA representatives at the workshop on a single sampling methodology, with technical constraints potentially limiting the application of all methods across the diverse range of river types within the MDB. It was agreed that two sampling approaches should be jointly trialled and evaluated in a preliminary phase of the Audit survey — electrofishing (boat and backpack) and passive gear (fyke and gill nets and baited light traps). Formal comparison of catches from the ‘electrofishing only’ and ‘all gear’ (electrofishing plus passive gear) options at the end of the ‘first round’ (2001–2002) of sampling was recommended, with one of the two options to be selected for further sampling rounds within the first five-year term of the Audit. This will allow all data from the first round to be compatible with ensuing sampling rounds.

Sampling for all surveys will be conducted once at each site in the low flow summer–autumn periods. Site lengths will be consistent with the NSW Rivers Survey, but will also be evaluated following the first sampling round. Insufficient data was available to the workshop participants and for this study to allow a detailed evaluation of the number of sites required to be sampled within each river valley for the Audit. A ‘design’ project is therefore recommended to collate all existing and new fish survey data from major MDBA programs, and to conduct power analyses relevant to agreed ‘effect sizes’. This project will recommend final numbers of sites for each river valley which will allow detection of

changes in fish assemblage measures with a known sensitivity. The project must be completed in the 2001/2002 financial year.

A suite of fish and environmental variables was recommended to be measured on each sampling occasion. The fish data will be used to derive values for a total of 29 ‘metrics’ chosen to quantify the state of fish assemblage condition at community, population and individual levels including measures of abundance; biomass; native fish biodiversity; aliens; representation of habitat guilds, trophic, reproductive and migratory guilds; tolerances; abnormalities; and size distribution.

Two analytical frameworks were identified as being potentially suitable for fish-based bioassessment in the Audit — multimetric analysis and multi/univariate predictive modelling. Both frameworks have recently been applied to stream fish assemblages in or adjacent to the Basin. Two methods have been developed within each of the two frameworks:

- *multimetric*: the Index of Biological Integrity, and two fish metrics developed under the NSW DLWC MARA program; and
- *multivariate predictive* — AUSRIVAS/RIVPACS (multivariate), and the regression tree approach (univariate).

None has been fully evaluated and all are still in active development. It was recommended that a project be funded to conduct a comparative assessment of the methods, using Audit fish survey data, whose primary aim will be to develop a final ‘unified’ framework and methodology. The methodology and the form of final outputs will be subject to peer review prior to adoption. The project should also be tasked with analysing the data from the first three Audit fish survey rounds using the final recommended method.

Intrinsic to both analytical frameworks is the concept of reference condition and the need to quantitatively define it in terms of metrics and variable values that are regionally based and representative of an ‘undisturbed’ or ‘least disturbed’ condition. It was recommended that two approaches be used — a ‘best available’ approach using data from the best reference sites/reaches within the MDB following screening for human impacts; and a ‘historical’ approach using expert knowledge and historical sources to define lists of species known to occur in each river valley prior to agricultural development in the MDB. A small review and workshop project is recommended to define the reference condition for fish within the Basin.

Overall, it was concluded that:

- fish bioassessment can readily be developed as an integral part of the Audit;
- much of the background work required to develop a standardised methodology has been done;
- several aspects still require completion and evaluation, under dedicated funding, and preferably with ongoing coordination during the first five-year term;
- this can be done in parallel with the first ‘rounds’ of sampling within the first five year phase of the Audit.

2. Introduction/Context

The purpose of the Sustainable Rivers Audit (Audit) scoping project conducted by the CRC for Freshwater Ecology wasto:

- to identify and/or develop appropriate cost effective and scientifically robust indicators of river health for the proposed Sustainable Rivers Audit, building on existing knowledge and monitoring arrangements;
- to develop a process by which the respective indicators may be measured and interpreted, at recommended temporal and spatial intervals, for each major river within the Basin and reported to Commission;
- to provide indicative costs for undertaking the measurement component of the recommended program; and
- with information provided by the SRA Taskforce and in partnership with State and Territory agencies and the Commission, to undertake a trial assessment that tests the feasibility of the proposed measurement and reporting process.

The Audit will be based on two broad themes with five sub-themes:

Biotic Themes	*Physical Themes*
• Macroinvertebrates • Fish	• Water Quality • Hydrology • Physical Habitat

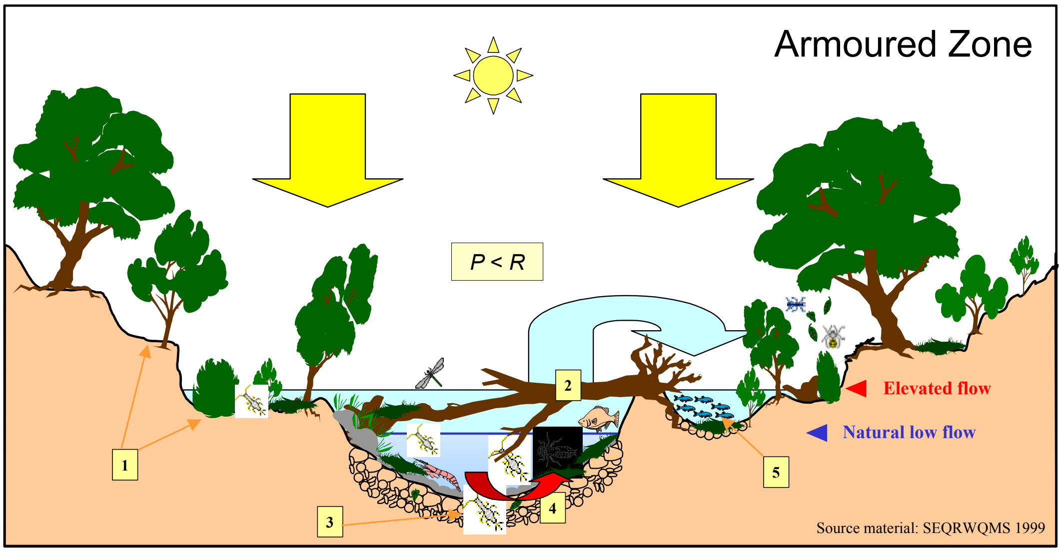
This project focusses on fish, and the need to develop a standardised approach to fish-based assessment of 'river health'. The primary aim for the fish project was to develop fish-community based assessment for the Audit, with the following specific task description:

"Conduct a specialist workshop of State representatives on fish assessments, along with others skilled in survey design, to determine the sampling procedures and approach to sampling and reporting fish community information. This workshop is to advise on sampling effort as well as reporting protocols."

Specialists in riverine fish biology and assessment, including representatives from each of the MDB agencies (MDBAs) attended a 1.5 day workshop in Canberra on 11–12 April 2001. Agreement was sought on the issues of sampling methodology and design, core variables, analytical framework for fish-based bioassessment in the Audit. The list of attendees and agenda for the workshop are at the end of this appendix.

Fish values in the Murray-Darling

Fish are key ecological components of MDB river systems, interacting with and influencing a range of stream ecosystem components and processes (e.g. as in Figure 1).



1 Incised floodplains of different ages formed by channel scouring become inundated at high flow, with submerged terrestrial vegetation and organic litter then available as food and habitat. **2** Fallen timber may create debris dams, trapping organic matter of various sizes, also providing food and habitat for invertebrates, fish and frogs. **3** Armoured cobbles and gravels provide a restricted invertebrate habitat within the substratum and trap detritus, however fish spawning is prevented by the armour layer to low flow areas under logs? **4** Sediment, detritus and nutrients are exported downstream when flow is high enough to move the bed surface and/or scour terraces. **5** Small floodrunners are inundated at high flows, increasing available habitat area.

Figure 1. Conceptual model of armoured functional process zone (FPZ) as used in scoping for the Audit. Fish are envisaged as a top predator with close interactions with in-stream cover substrate and production, as well as habitats available at high flow. Fish are also top-down predators, influencing community and population dynamics of macroinvertebrates and occasionally causing 'trophic cascades', with resulting changes in primary production or algal standing crops.

Fish are also seen as of significant social, recreational and economic value, through recreational and commercial fishing for riverine species. Considerable efforts have been made and continue to be made to maintain and/or rehabilitate river fisheries in the Basin via stocking and habitat management programs. Native fish are seen as having major intrinsic value in themselves, and as indicators of the 'naturalness' of MDB rivers. A number of programs focus on the restoration of native fish populations, again through stocking and habitat rehabilitation. Introduced or 'alien' fish species, especially carp, redfin perch, gambusia and trout, are seen as significant threats to the integrity of MDB river fish populations and ecosystem functioning.

As indicated above, fish communities in the MDB tend to be low in diversity. They are frequently also dominated by alien species, particularly carp in the lowlands and trout in the uplands. The historical changes in MDB fish communities over the last 200 years are believed to be profound (Lake 1982, Gehrke1997, Schiller *et al*. 1997).

Numbers of species recorded in the NSW River Survey (NSWRS, Harris and Gehrke 1997) were low, with 20 freshwater fish species in total across both the Murray and Darling basins (with nine native fish species not observed, but expected). Many native fish species have suffered substantial declines in abundance and in range within the NSW sections of the Basin.

Fish bioassessment

Methods for sampling fish populations are well established. However, there has been little standardisation of sampling methods in previous surveys across the Basin (Faragher and Rodgers 1997).

In addition to a lack of standardisation in sampling methods, there has been little development of a standardised manner of either reporting or analysis of fish population and community data across the Basin. A formal analytical framework was entirely lacking until the advent of the IBI approach as used in the NSW Rivers Survey (NSWRS; Harris 1995, Harris and Silveira 1999). There has been no review or assessment of the strengths and weaknesses of the various analytical frameworks or approaches available for fish bioassessment data in the MDB rivers to date.

A problem specific to bioassessment using fish data in the MDB is the low species diversity at basin, river valley and site scales, as indicated above. An analytical framework which relies strongly on patterns of species occurrence, or on species data alone, is likely to generate outputs which are highly vulnerable to changes in sampling efficiency and which are relatively insensitive to environmental change. It is therefore desirable to use a framework which captures data at a range of 'organisational levels' in order to maximise the information gained from fish survey data (see Adams and Ryon 1994, Adams *et al*. 1996). These can include measures of individual fish health (physiological and/or biochemical measures), growth, condition and energetic status (length–weight and organ–fat–somatic ratios), population size and recruitment, community composition by species and trophic guilds. Together such measures provide an 'integrated' assessment of the condition of fish assemblages, as well as providing key indicators of the effects of a range of human impacts on fish ecology.

Another major problem for any analytical approach to assessing river health using fish data in the MDB is the lack of existing 'reference' or undisturbed fish communities against which to measure changes. Much of the MDB is heavily disturbed at a local scale through physical in-stream and riparian habitat alteration, flow regime change, water quality changes from point and diffuse source pollution. However, major disturbance to key processes in fish populations also operate at reach and valley level scales, particularly through the influence of barriers and storages on fish passage, water quality and thermal regimes. This is particularly relevant to fish species with life histories dependent on migration over substantial distances. In addition, the invasion of substantial portions of the MDB drainage network by alien fish species is seen as a significant impact on native fish assemblages and values. Thus, there are few reaches within the Basin which contain a true undisturbed reference fish assemblage, and none in the predominant lowland regions. Choice and quantification of a reference condition is a major issue for fish-based bioassessment across the Basin.

3. Existing programs/methods

This section briefly discusses the key issues which were examined at the Audit Fish Bioassessment Scoping workshop held in April 2001.

3.1 Sampling methods

A wide range of sampling techniques are used and have been used for riverine fish surveys by the agencies involved in fish monitoring and assessment in the MDB (hereafter the "MDBAs"), with two major distinctions:

- passive sampling (nets, traps, etc.); and
- active sampling (electrofishing, seining, angling, etc.).

Some evaluation of the relative efficacy and cost effectiveness of these approaches has been done. Considerable thought was put into sampling methodologies for the NSW Rivers Survey (Harris and Gehrke 1997), and attempts were made to standardise catch-effort and equipment across a range of river sizes and types. Subsequent comparative investigations have concluded (e.g. Faragher and Rodgers 1997, Pusey *et al*. 1998) that electrofishing, by boat in large river habitats, and backpack in smaller headwater shallow sections, is the most cost-effective approach for sampling a range of fish metrics, possibly combined with seining in smaller snag-free rivers, and can be reasonably standardised between river/reach/habitat types. Electrofishing is not only cost-effective, but also most effective by other yardsticks, such as assessment of abundance, species representation, representation of rare species, representation of various habitat guilds.

Intensive discussion on sampling approaches at the Audit Fish Bioassessment workshop revealed that several agencies were not convinced that these comparative assessments had unequivocally shown that passive gear sampling was not required when sampling across a wide range of river types in the MDB. Concerns were raised about relative efficiency of electrofishing in highly turbid sites and sites with extremes of conductivity, as well as in relation to accessibility of sample sites with electrofishing boats. The discussion concluded that further work was required to confirm the cost-effectiveness and representativeness of electrofishing alone vs combining electrofishing with passive sampling ('all gear') across a range of site types before adopting a single method for the Audit.

Two other issues were relevant to this discussion — the need for high sampling efficiency in river reaches with low species diversity, and the greatly increased resources (in field time, personnel, and hence funding) for sampling program which includes deployment of passive gear.

The efficiency and representativeness of sampling at a site are key issues. For example, the efficiency of electrofishing is dependent on a number of factors (e.g. Davies 1989, Faragher and Rodgers 1997, Pusey *et al.* 1998), most notably fish size and behaviour; conductivity; visibility (water colour, turbidity); habitat type (complexity, depth).

Overall estimates of population size, number of species and size distribution within populations derived from electrofishing are strongly dependent on the number of electrofishing passes through a sample site. Pusey *et al.* (1989) evaluated this for Queensland streams and recommended a minimum of 3–4 passes (accompanied by seine netting where possible). While multiple passes, especially within stop-netted sites, will maximise accuracy and precision of fish-community estimates at the site level, a trade-off is needed between a high commitment of effort at smaller sites and the need for greater site coverage to reduce the variance of overall estimates at a valley scale.

3.2 Sampling design

Size of sites

While sampling efficiency within a site can be controlled by use of multiple passes and standardised effort and 'search strategies', two things determine how well a sample from a fish population represents the situation at reach/VPZ/river-valley scale:

- spatial variability in species/life stage occurrence at the site scale or the reach/valley scale;
- temporal variability (e.g. with season).

The site-scale issue can be addressed by sampling a length of stream adequate to represent the range of habitat types, and hence return a cumulative species/age class/size class number close to an 'asymptotic' value for the reach. Angermeir and Smogor (1995) and Lyons (1992) identified relationships between number of species recovered and length of stream/number of habitat types sampled in larger US rivers. They concluded that sampling needed to cover a site length equivalent to 22–67 and 35 stream widths, respectively, to provide an adequate recovery of species, with the latter figure resulting in sampling of some three complete riffle-pool sequences.

The NSWRS sampled sites varying in length but typically with sampling occurring through a 400 m reach in montane areas, and 1000 m in slopes and lowland river types. These site lengths are comparable with the above studies for slope and upland streams, but somewhat shorter for lowland sites, and some evaluation of the adequacy of these site lengths is needed for lowland situations.

Number and location of sites

Some preliminary analysis will be required to estimate the number of sites required to be sampled within a 'reach' or Valley Process Zone (as defined under the proposed Audit conceptual framework). This will need to address the issues of:

- the number of sites required to produce an assessment that is 'representative' of the reach/VPZ, and
- the 'effect size' which the Audit is required to detect using fish-based bioassessment.

There was insufficient data available for conducting this type of analysis at the VPZ or river valley scale when this report was prepared. Only 20 sites had been sampled in the NSWRS in each of the Murray and Darling basins, and the availability of data for Queensland and Victorian MDB rivers collected with comparable gear was limited. New data is rapidly becoming available from further sampling being conducted in both Queensland and NSW (e.g. in the DLWC IMEF program). The Audit Fish Bioassessment workshop recommended that existing and new data-sets be collated in the first year of the Audit and that a series of analyses (power analyses and accretion curves) be conducted to evaluate the number of sampling sites required in each river valley and VPZ in order to detect specific effect sizes for the agreed variables. This should then be used to refine the sampling program, taking into account the cost implications of the outcomes of the sampling methodology assessment.

I conducted several preliminary power analyses on NSWRS data for three variables: number of species, IBI score and % fish with abnormalities. The results (Table 1) indicate the great variability in the number of sites required to detect ecologically significant changes in fish variables at the river valley/VPZ scale, depending on the metric chosen. For example, the detection of the complete loss (or gain) or a fish species at the river valley scale will require around 20–40 sites per valley (at an alpha level of 0.05, and power of 0.8). This has major implications for the cost of the fish component of the Audit. A review of acceptable effect sizes required to be detected by the Audit fish component must be conducted following completion of a detailed power analysis using a data-set collated from programs across the Basin.

Most sampling programs to date have focused on in-stream fish assemblages, with little emphasis on routine sampling of floodplain/billabong habitats. Existing data-sets on floodplain/billabong fish should be collated and analysed as part of the proposed data/design review, and stratification of sampling site by habitat type considered.

Table 1. Results of preliminary power analyses aimed at determining the minimum number of river sites needed to detect changes in fish variables of specified magnitudes (‘effect sizes’) at the river valley or VPZ scale: alpha = 0.05, beta = 0.8.

Variable	Valley/VPZ	Effect size	No. sites needed
Number of fish species	Darling River, reaches with >12,500 km^2 catchment area (slopes/lowlands)	Loss of 1 species (from site mean of 8.9)	**47**
		Loss of 2 species (from site mean of 8.9)	**13**
	River Murray, reaches with >11,000 km^2 catchment area (slopes/lowlands)	Loss of 1 species (from site mean of 7)	**21**
		Loss of 2 species (from site mean of 7)	**7**
IBI score	Montane (Murray and Darling)	Decline of 6 IBI units – i.e. decline by one band e.g. from poor to below poor	**3**
	Slopes/Unregulated lowlands (Murray and Darling)	As above	**7**
% fish with abnormalities	Darling, montane (most variable values between sites)	Doubling (from 5% to 10%)	**65**
		Trebling (from 5% to 15%)	**18**
	Darling, regulated lowland (least variable values between sites)	Doubling (from 5% to 10%)	**6**
		Trebling (from 5% to 15%)	**3**

3.3 Variables measured

No two MDBA fish survey programs have generated data in the same form. Most current programs provide estimates of the total number of species, the numbers of native and alien species and biomass. Fish length, or size/maturity class is often recorded. Fish weight is not always recorded due to the additional field effort involved.

There is little consistency in derived variables (e.g. fish condition, proportions of trophic groups etc.) reported between MDBA fish survey programs, due in part to differences in aims of the various surveys.

3.4 Analytical Frameworks and Approaches

There are two major approaches to fish-based bioassessment which have been trialled in Australian rivers. They are:

- *multimetric analysis* — IBI and MARA (FSI) approaches
- *multivariate 'predictive' modelling* — AUSRIVAS/RIVPACS, and Regression tree approaches

Neither of these frameworks have been used with large fish data-sets or thoroughly evaluated and tested with independent fish and/or environmental data.

Both approaches use a 'reference' condition as a basis for assessing change (see discussion in Reynoldson *et al.* 1997), although this may not be as obvious or explicit in the metric approaches (e.g. in the case of the FSI).

3.4.1 Multimetric analysis

Basis:

The IBI (Index of Biotic Integrity) approach was developed by Karr (1981, 1987, 1991) and Fausch *et al.* (1986) for rivers in the USA. The IBI uses a series of 'metrics' (univariate descriptors) of aquatic biological communities to provide an overall score for a site. A metric is defined as:

"*a calculated term or enumeration representing some aspect of biological assemblage structure, function or other measurable characteristic that changes in a predictable way with increased human influence.*" (Barbour and Yoder 2000)

A large suite of metrics is available in the IBI approach, with suites for fish and macroinvertebrates derived by Karr (see Fausch *et al.* 1986), and modified by Barbour and others (Barbour *et al.* 1995). A large set of metrics was adopted by the USEPA within its set of Rapid Bioassessment Protocols (Barbour *et al.* 1999) and have been used by a number of other US Federal and state agencies for aquatic bioassessment since the mid 1980s. Metrics are chosen to be:

- ecologically relevant to the assemblage under study and to the program objectives;
- sensitive to stressors;
- responsive in a way that can be discriminated from natural variation.

More than 15 individual metrics are available for fish alone. The choice of individual metrics is largely left to the user, with agencies adopting, and modifying their own particular suite for routine use. Fish metrics include measures of species richness and community composition, trophic structure, abundance and individual fish 'health'. Assessment using measures at a range of levels of organisation is a feature of the IBI approach. The ability to select metrics from a 'shopping list' is regarded by some as adding flexibility to the assessment process, and by others as leading to non-standardisation.

Metrics are derived from standardised sampling of the fish community, combined with internal standardisation against the highest values obtained for each variable for the relevant catchment area. This is done by plotting values against catchment area and fitting lines by eye close to the maximum values (i.e. which lie above 95% of the sites surveyed). This line is referred to as the 'maximum species richness line' or MSRL by Fausch *et al.* (1984), see Figure 2. Once metric values are derived, they are then transformed to standardised scores, and these scores are then added to form a composite score called the IBI. The IBI is then compared to thresholds ('biocriteria'), derived by the user (though often taken directly from the USEPA guidelines). In the US, these are frequently incorporated into state agency regulatory programs.

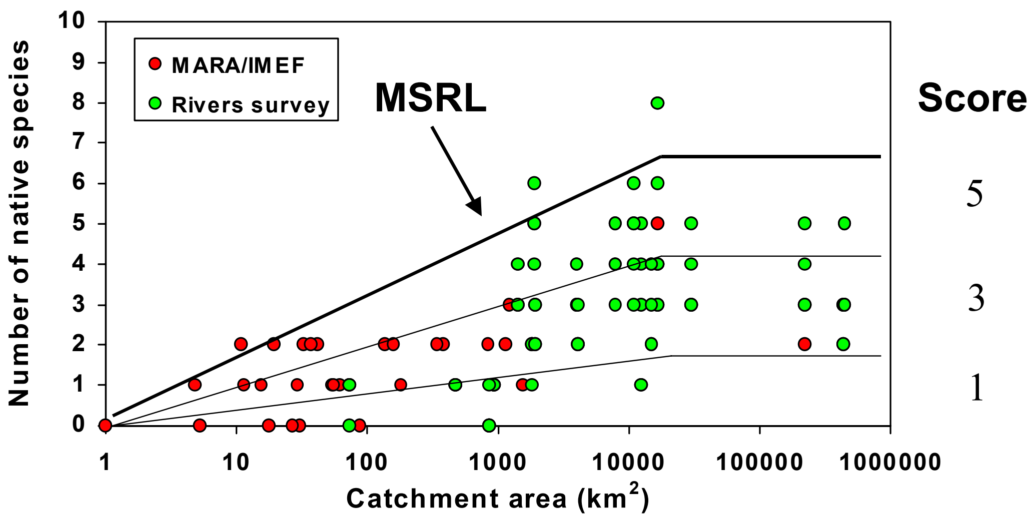


Figure 2. Example of the MSRL technique for defining the upper bounds and scores for the number of native fish species metric at different catchment areas within the Darling Basin (unpub. data from the NSW DLWC MARA, IMEF programs and the NSWRS).

Test sites are selected on a regional basis, and grouped within 'stream classes' (or 'bioregions') based on broad aquatic biological characteristics. While it is sometimes claimed that this classification is based on aquatic reference sites (Barbour and Yoder 2000), in fact this is rarely the case in practice, and many users simply use elevation or some other measure of position in the drainage network as a template for classification and stratification of site selection. A suite of test sites are then selected within each region.

The word predictive is used within some of the IBI literature, but the technique does not comply with strict definitions of a method which results in predictions of a site's condition based on biological–environmental relationships. Thus, while it does not make a 'prediction', the IBI and metric scores are rated in comparison to threshold values derived from data from sites with the highest metric values. These are equivalent to the use of reference data based on the 'best available' sites in the region.

Analysis of the data is generally restricted to comparison of final IBI scores with threshold values. It is often accompanied by examination of spatial trends of individual metrics, and occasionally by correlation with physical/chemical variables. Validation of fish IBI results by such approaches has produced equivocal results (e.g. Shields *et al.* 1995, Frenzel and Swanson 1996, Karr and Chu 1997, Harris and Silveira 1999).

Weaknesses of the IBI approach claimed in the literature include:

- the prevalence of assumptions about the ecological meaning and relevance of metrics, especially when transferred from one region to another;
- the untested validity of standardising metrics for fish using catchment area relationships;
- the semi-arbitrary (i.e. non-statistical) nature of dividing the area under the catchment area curve into three bands of equal area to provide value ranges for standardising metric values;
- the loss of information, potential for summing errors, and the risk of overweighting attribute values (through metric redundancy) when simply summing standardised metrics to form an overall score;
- the uncertain nature of the responses ofsome individual metrics to human disturbance.

Strengths of the IBI approach claimed in the literature include:

- ready use of IBI and metric scores in new situations without the need for complex model building and/or analysis;
- flexibility in choice of metrics and final IBI score thresholds;
- metrics based on ecologically relevant concepts.

Other strengths include:

- the capacity to apply a broad range of knowledge about fish responses to disturbance assessment, in addition to simple species-richness changes;
- the capacity to make comparisons across diverse sites and regions;
- the capacity to refine the approach by measuring the performance of metrics and enhancing the data used in MSRLs.

The second metric-based approach is a univariate metric method using two fish-based metrics, under development by NSW DLWC (Chessman pers. comm.). The first is a fish equivalent of the Australian macroinvertebrate SIGNAL index, called the FSI (fish species index). It assigns 'disturbance tolerance grades' ranging from 1 (tolerant) to 10 (intolerant) to fish species then calculates an abundance-weighted average grade for those species present. There are three versions in development, with one each for assessing water quality, migration and general conditions. The initial grading has been developed based on Harris and Gehrke's (1997) list of intolerant species. New grades are being derived by more standardised procedures (as in Chessman *et al.* 1997).

The second metric being developed by NSW DLWC is the mean rarity score (MRS). Each native species is assigned a rarity grade according to the number of individuals encountered in the NSWRS, ranging from 1 (most common) to 100 (most rare or not recorded). The rarity grades are then averaged for species at a survey site.

Use in Australia:
NSW Fisheries assessed the IBI (Harris 1995, Harris and Silveira 1999) using results of the NSWRS (Harris and Gehrke 1997) at 80 sites in NSW, including 40 sites in the MDB (20 in each of the Murray and Darling drainages, of which 5 were in the montane regions and 5 were in slopes regions in each of the two drainages). A suite of 12 metrics were adopted from Fausch *et al.* (1990), shown in Table 2. One of these metrics (metric 8) was subsequently shown to perform poorly, and Harris and Silveira (1999) recommend the use of only the remaining 11 metrics in future assessments. An IBI sampling and analysis manual is currently being produced (Harris in prep.) for use in south-eastern Australian rivers.

It is unclear to what extent the ‘reference’ condition was adequately identified and sampled in the 40 MDB sites sampled in the NSWRS. There is scope for better defining the relationship between the fish-community attribute values and catchment area in order to provide improved IBI assessments, as illustrated in Figure 2 where new data considerably improves definition of one MSRL.

The use of IBI in the NSWRS was deemed successful (Harris and Silveira 1999) due to:

- a broad match of IBI scores with perceived levels of disturbance and biological impairment;
- a statistically significant correlation between IBI scores and scores of the ‘RDI’ (River Disturbance Index – a measure of physical habitat disturbance as used in the Wild Rivers project, Stein *et al.* 1998), though with low described variance (r^2 = 0.10);
- the ease of use of the approach;
- the consistency in scores between the two annual summer sampling events (interannual correlation in scores with r^2 = 0.564).

Both the FSI and MRS metrics are being developed and used within the NSW DLWC ‘Multi-Attribute Reach Assessment’ (MARA) program, in four NSW catchments (upper Castlereagh, Wollombi, Adelong and on the southern coast). Neither methodology has been fully developed or trialled to date, and data made available for discussion at the Audit Fish workshop was preliminary.

Table 2. Fish assemblage metrics used to calculate the IBI for the NSW Rivers Survey (Harris and Silveira 1999).

Category	Metric	Scores and Criteria		
		5	**3**	**1**
Species richness and composition	1. Total number of native species		Metrics 1–5: Expectations vary with stream size and region	
	2. Number of riffle-dwelling benthic species			
	3. Number of pool-dwelling benthic species			
	4. Number of pelagic pool species			
	5. Number of intolerant* species	> 67%	33 – 67%	< 33%
	6. Percent native fish individuals	> 67%	33 – 67%	< 33%
	7. Percent tentative species			
Trophic composition	8. Proportion of individuals as microphagic omnivores	< 33%	33 – 67%	> 67%
	9. Proportion as microphagic carnivores	> 67%	33 – 67%	< 33%
	10. Proportion as macrophagic carnivores	> 10%	3 – 10%	< 3%
Fish abundance and condition	11. Number of individuals in sample		Metric 11: Expectations vary with stream size and region.	
	12. Proportion of individuals with disease, parasites and abnormalities	0 – 2%	2 – 5%	> 5%

* intolerance to factors including poor water quality and barriers to migration.

3.4.2 Multivariate predictive modelling

Basis:

These two methods (AUSRIVAS/RIVPACS and regression trees) are based on:

- predictive modelling — that is, predictions of assemblage characteristics based on environmental relationships in reference ('least impacted') site data-sets; coupled with
- statistical comparison of observed values at test sites with model (reference) predictions, and
- calculating measures of deviation from the reference site values as indicators of ecological health or intensity of disturbance.

Both methods are highly complementary in data requirements, with the regression tree method focused on univariate descriptions of fish assemblages, and AUSRIVAS/RIVPACS focused on multivariate descriptors (species lists).

RIVPACS (River Invertebrate Prediction and Classification Scheme, Wright 1995) is a multivariate predictive approach, originally developed for macroinvertebrate bioassessment, now incorporated within AUSRIVAS (the Australian River Assessment Scheme, Schofield and Davies 1996, Davies 2000), which can also be used to predict fish assemblage structure (presence/absence of species, life stage). This approach is based on the concept of making bioassessment evaluations, using community compositional data, for a test site relative to an undisturbed community composition 'predicted' by models derived from a data-set collected at a set of reference sites. Development of the predictive model is a separate process from the use of it in assessment, and involves several analytical steps (classification, group designation, discriminant analysis, reference group and taxon probability estimation). The use of the reference condition is explicit within the analytical methodology for RIVPACS, and the appropriate choice of reference sites is a core requirement for its success.

Data from a set of reference sites is classified to define natural biological groupings based on community composition (typically species presence/absence). These groupings are then used to derive the probability of membership of each species in each group.

In addition, 'discriminant functions' are developed (by discriminant function analysis, or DFA) to discriminate the biological groupings using environmental variables which are uninfluenced by human disturbance ('predictor variables', e.g. altitude, distance from source, width). These functions are then used to calculate the probability of membership of a new 'test' site in each reference site group.

The two probabilities are then combined to estimate the probability of a species occurring at the new test site as if the site were in reference ('least impacted') condition. Thus, a list of species can be derived for the test site which are predicted to occur there, each at a given level of probability. The sum of the probabilities is then generated to provide a total number of expected species expected to occur at the test site if it were least impacted.

A count is then made of the number of predicted species actually observed at the test site (using the same sampling methodology). The ratio of the observed number to the expected number (the 'O/E ratio') is then calculated. This ratio spans from 0 (with no expected species present at the test site) to around 1 (with all expected species present). This ratio is the primary output of the RIVPACS approach, and forms the basis of reporting for macroinvertebrate bioassessment under AUSRIVAS. Error bands, incorporating both sampling error and prediction error would provide a basis for objective judgement of the strength of the assessment. Analytical approaches for calculating these are being developed within AUSRIVAS in time for use in the Audit.

The modes of AUSRIVAS/RIVPACS model development and use are quite distinct, and are illustrated in Figures 3 and 4.

The range of O/E values is divided into 'bands' for reporting purposes (Figure 5). A band falling between the 10^{th} and 90^{th} percentile values of O/E for the reference sites is derived as the 'A' or 'equivalent to reference' band. Two bands are then delineated below this band with widths equivalent to the A band — the 'B' or significantly impacted, and 'C' or 'severely impacted' bands. A final 'D' band then falls between the lower bound of the C band and 0.

Criticisms of the RIVPACS/AUSRIVAS approach in Australia have included:

- the need to find suitable reference conditions (sites) for model development;
- the need for intensive stages of model development prior to using the technique;
- problems with incorporating temporal variability within the modelling component;
- problems with low species diversity;
- the complexity of the model development stage which requires expert involvement;
- the need for further 'diagnostic' analysis as an aid in interpretation;
- the need for an intensive data collection from reference sites before assessments can be made.

Other criticisms tend to focus on issues of quality control in sampling and site selection, and pertain more to the history of development of macroinvertebrate-based assessment within AUSRIVAS rather than the RIVPACS approach *per se*.

Key advantages claimed for the RIVPACS/AUSRIVAS approach include:

- the explicit use of a reference framework as a basis for comparative assessment;
- use of regional data in model development and assessment;
- the ease of interpretability of the O/E score as a measure of departure from reference condition;
- the standardisation in sampling and analytical approaches required by the method;
- the standardisation in O/E score outputs and the opportunity to standardise the basis for delineating thresholds (ranges or 'bands' of impairment).

The regression tree approach also uses a data-set from reference sites to derive 'predicted' values of single variables describing the fish community. Regression trees can be developed for a number of univariate descriptors of fish assemblage structure, such as number of native fish species, % microphagic carnivores, total biomass etc., and could also be developed for variables describing life history and fish health characteristics. Sites in the reference data-set are classified using regressions of environmental variables to discriminate values of a single variable (Figure 6). This provides both a classification structure based on value of the environmental variables (e.g. altitude, width etc.), and a statistical distribution for the variable within each reference site classification group.

Test sites are first classified using critical values of the environmental variables into a single reference site group. Values of the fish variable are then compared with the statistical distribution derived from the reference site group data (Figures 6 and 7). Kennard proposes that values outside the 20th and 80th percentile values for the reference site group be identified as in ‘poor condition’, while those sites falling inside the 20–80 percentile range are in ‘good condition’.There is scope for further refining this comparison and providing impairment bands as in RIVPACS.

I. Developing RIVPACS

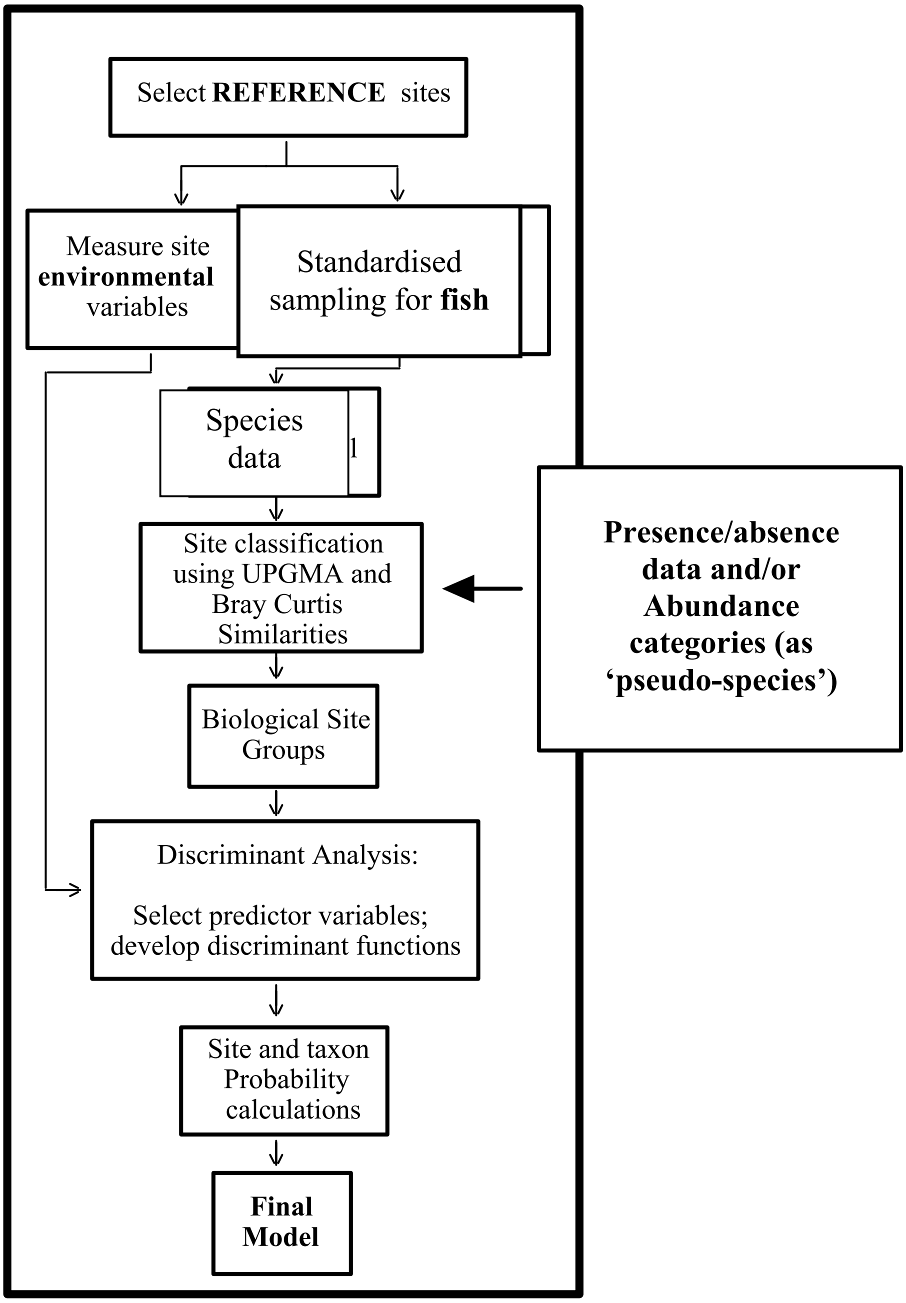


Figure 3. Process of RIVPACS (AUSRIVAS) model development for bioassessment.

II. Using RIVPACS

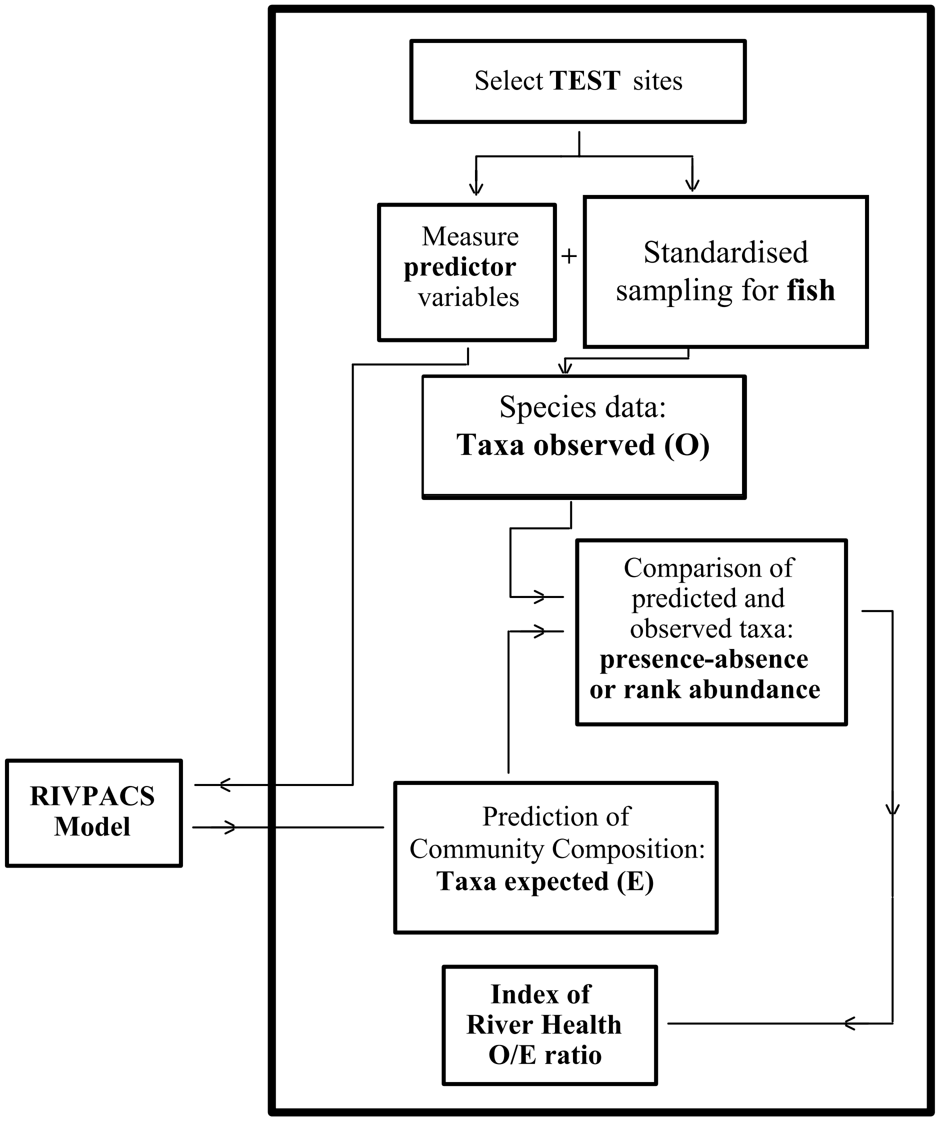


Figure 4. Process of using RIVPACS (AUSRIVAS) models for bioassessment.

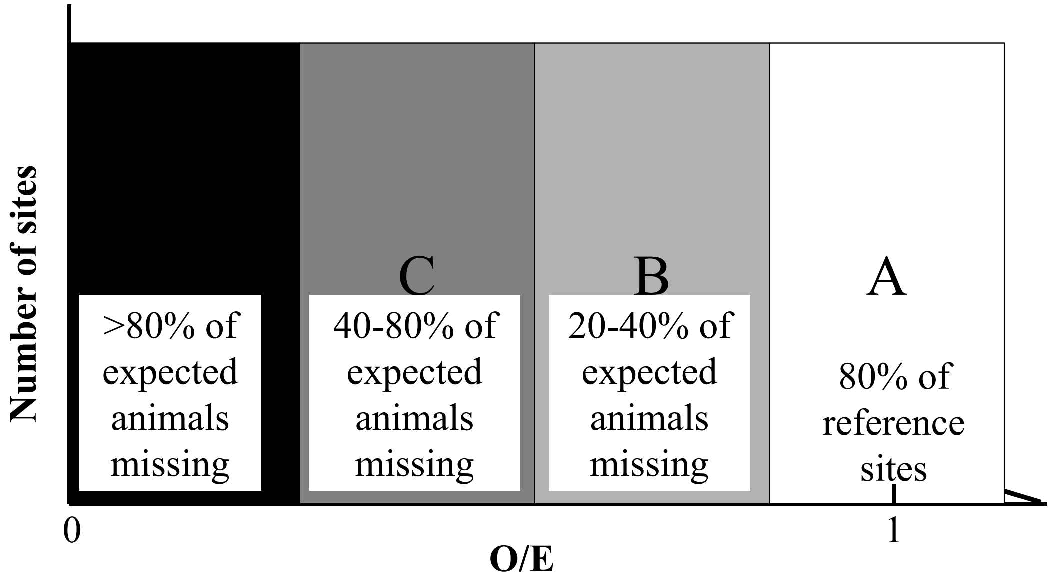


Figure 5. Banding scheme for AUSRIVAS/RIVPACS O/E values.

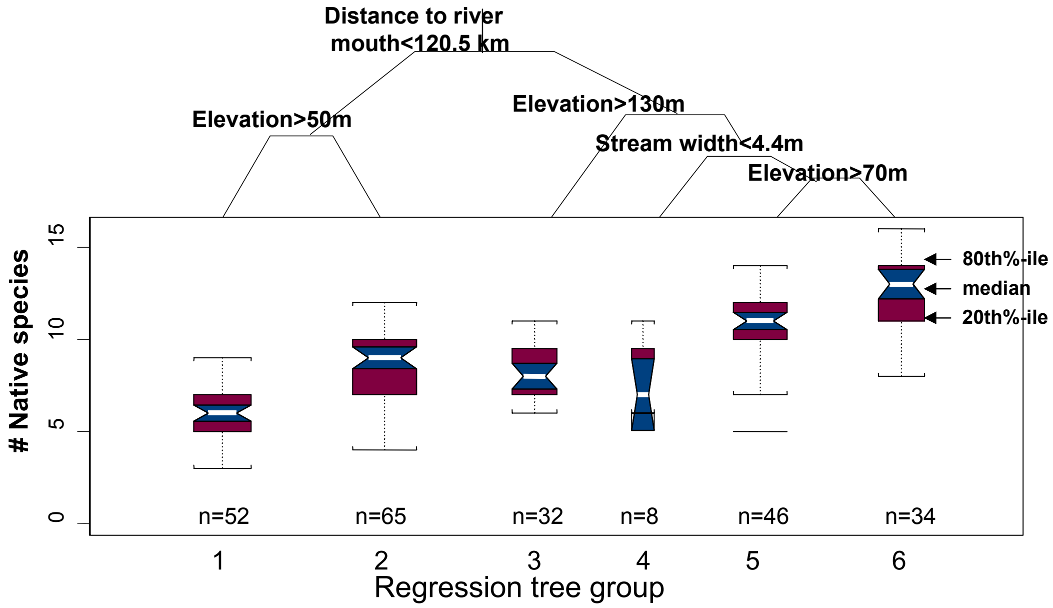


Figure 6. Example of regression tree derived for number of native fish species occurring at reference river sites in SW Queensland (Kennard pers. comm.). Total r^2 = 0.62. Box plots show distribution of number of species in each site group.

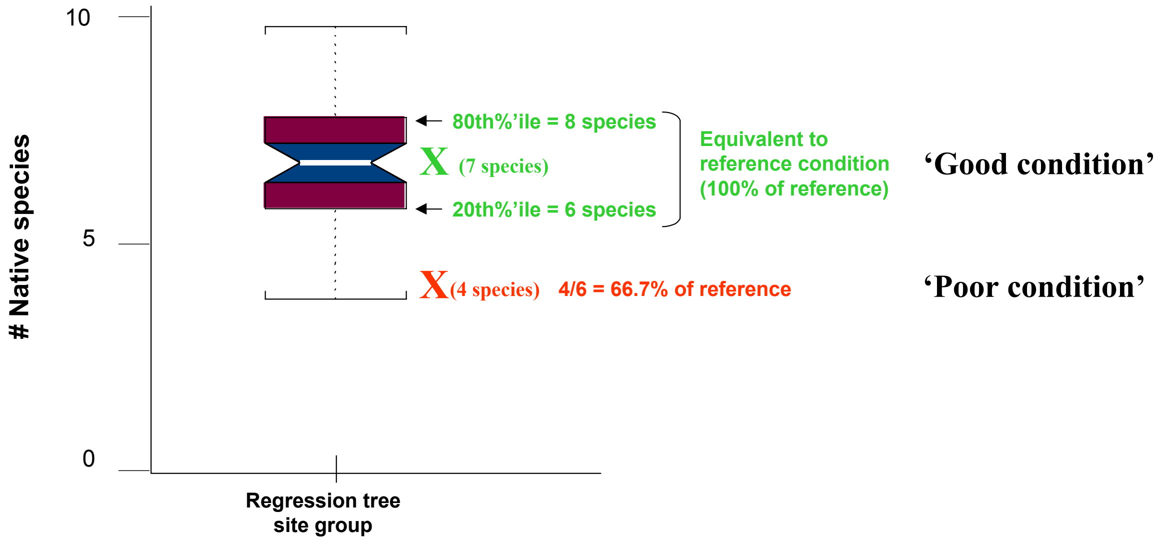


Figure 7. Method for calculating indicator values at new sites in the regression tree approach, as % of reference condition, with a reference or 'good condition' band set at 20th and 80th percentile of reference site values

Uses in Australia:

AUSRIVAS/RIVPACS models have been developed for fish community data in Queensland (Kennard *et al.* unpub. report, Kennard pers. comm.) as well as for New Zealand rivers (Joy and Death 2000). The New Zealand model was highly successful at predicting species occurrence, but was focussed on relatively species rich coastal streams, relatively small catchments, a very strong altitude gradient, and a speciose fauna dominated by migratory fish that are highly responsive to altitude and distance from the sea. The Queensland (Brisbane River) model successfully predicted fish species presence/absence, but the authors noted problems with low species richness. In addition, the ability to incorporate other variables (fish abundance, trophic levels etc.) has not been evaluated with this approach to date, although it is technically feasible.

The regression-tree approach has been used and evaluated as part of the Design and Implementation of Baseline Monitoring project # 3 (South East Queensland Regional Water Quality Management Strategy) by Kennard (unpub. data). Regression tree models were successfully developed using up to eight 'predictor' environmental variables for 6 fish variables, with cross-validated r^2 values ranging from 0.4 to 0.62. The inability to develop valid regression trees and hence make valid predictions, due in part to the inability to discriminate fish variables using the independent environmental variables, is an issue with this approach, and good predictive models could not be developed for four variables including total fish biomass and abundance.

Kennard also developed AUSRIVAS (RIVPACS) style models using the same data as for the regression tree models. The models were successfully validated using independent data, and were based on DFAs with acceptable reclassification errors. O/E values for test sites were related to environmental disturbance indicators by General Linear Modelling. Kennard has concluded that :

- ‘regression tree predictive modelling can explain a large amount of natural variation in many (but not all) univariate fish assemblage indicators’; and that
- ‘multivariate predictive modelling (AUSRIVAS) can explain a large amount of natural variation in fish assemblage composition.’

Evidence to date therefore indicates that both the AUSRIVAS/RIVPACS and the regression tree predictive modelling approaches are suitable as analytical frameworks for fish bioassessment in MDB rivers.

3.4.3 Analytical framework — summary

In summary, there is no single analytical framework that has been universally adopted for fish-based bioassessment in Australian rivers, or within the Basin. None of the methods described above has been used on large spatial data-sets to date, and there has been little external validation (i.e. by use of independent fish and/or environmental data), or formal peer review. The assessments conducted to date for IBI, AUSRIVAS/RIVPACS and the regression tree approach are all encouraging, and no single approach ‘stands out’ to date. Low species diversity, strong temporal variability and the absence of undisturbed reference conditions within the MDB pose problems for all three methods. There is considerable commonality between the approaches, and there are strong grounds for a comparative analysis of a large common data-set to evaluate and refine them. Several criteria for fish-based bioassessment framework within the MDB Audit are desirable. No single method satisfies all of them (Table 3), and some further development is needed in all cases.

Table 3. The ability of the three methods to satisfy key criteria for fish-based bioassessment within the Audit.

Criteria	**Method**		
	IBI	**RIVPACS/ AUSRIVAS**	**Regression tree**
The need to incorporate a reference site/condition as basis for relative assessment of river ‘health’.	**Yes**	**Yes**	**Yes**
The need for measures at several levels of biological organisation including the community, population and individual levels.	**Yes**	**No, but possible**	**Yes**
The desirability of a predictive approach which accounts for natural regional differences in fish assemblage structure and the spatial scale relevant to dynamics of individual species populations (e.g. migratory species).	**No**	**Yes**	**Yes**
The need to incorporate a temporal aspect into the assessment which accounts for key events which influence processes such as recruitment and mortality.	**No, but possible**	**No, but possible**	**No, but possible**
Outputs that are interpretable and can be related to variables describing the intensity of human impacts.	**Yes, but only for individual metrics**	**Yes, but requires development**	**Yes**

4. Proposed Program and Methods for the Audit

4.1 Method development and data collection, in parallel

Because of the need to achieve a better consensus with regard to sampling and analytical methods for fish-based bioassessment, the Audit Fish Bioassessment workshop recommended a two phase program for the first five-year term of the Audit. The first phase will involve a 'first round' of sampling on a subset of sites, but will also include a design review project which will finalise the ongoing sampling strategy (number and layout of sites) and make a final recommendation on sampling methods, based on analysing results from the first round. It will also include a small project to provide interim historical and expert-opinion reference species lists and a spatial reference site list. The second phase will be the commencement of the agreed ongoing program, with a full set of sampling sites sampled prior to the first five-year review over a two-year period. It will also include a project to apply the three possible analytical approaches to the Audit data, review their performance and develop and recommend a final analytical framework and methodology for the Audit Fish component. The project will also conduct the first full analysis of the first Audit fish assessment data in time for the first five-year review.

4.2 Proposed schedule

Thus, the proposed schedule is as follows:

1. Phase 1: 2001–2002

Project 1: Data collation, power analyses and sampling design review
This project, commencing in mid 2001, will collate existing data on fish assemblages within the MDB from all partner agencies. Data of sufficient quality will be incorporated into a series of analyses designed to assess the number of sites and their spatial 'layout' (e.g. stratification by reach type/FPZ) required to provide a representative sampling regime which is able to detect specified differences ('effect sizes') in key fish variables. The project will require a small workshop to agree on effect sizes once the initial power analyses have been conducted, and to agree on the 'trade-off' between program resources, effect size detection (and/or power) and the number of sites in each river valley. This workshop will also review the issue of reference sites (see project below) and how they are to be incorporated as a 'stratum' within the sampling design.

The project will also analyse data from the first sampling round and use it, along with previous data from the NSWRS, to formally assess the cost–benefits of the two sampling options (electrofishing only vs 'all gear' ie electrofishing plus passive gear).

The project will, by September 2002, report on:

- a final agreed spatial sampling design;
- a final agreed sampling methodology.

Project 2: Reference condition project
A small project, commencing in late 2001, will conduct a preliminary assessment of the spatial and historical reference condition for the MDB. The project will collate all available historical data and lists of fish species for the river valleys in the MDB, and

conduct a workshop to discuss and agree on a series of historical reference lists of species for each river valley (or major valley section), with the MDBAs and invited experts.

In addition, the project will develop an initial classification of reference reaches/sub-catchments in the MDB, based on a set of reference site criteria agreed to by the MDBC and MDBAs. This will be used as the basis for developing a preliminary set of reference reaches or sites, finalised at the above workshop, to be used in the design of the first full rounds (rounds 2 and 3) of the Audit fish survey. The agreed list of sites will be provided to the design phase project team to incorporate within the design analysis.

The project will, by mid 2002:

- provide historical and expert-opinion reference lists of fish species for each river valley and major habitat stratum in the MDB;
- provide an accompanying report citing sources and justification and the process and outcomes of the workshop;
- provide a list of initial/interim reference sites/reaches/sub-catchments for use in the Design review project and in the final, stage 2 Analysis project (see below).

Survey: Round 1.
A subset of sites sampled once within all MDB States/Territories in summer–autumn 2001/2 covering a range of site conditions (VPZ, salinity, turbidity etc.) in a stratified design. The intention is not only to conduct sampling that can be incorporated within the first five-year Audit analysis, but also to collect data to allow comparison of the electrofishing vs electrofishing plus set gear method options (see below).

2. Phase 2: 2002–2004

Survey: Rounds 2 and 3.
The first full survey of all sites will be conducted over a two-year sampling period (in two separate sampling rounds), with a single sampling of each site conducted in summer–autumn of 2002/3 and 2003/4. The spatial design and sampling methodology recommended from Phase 1 Design project will be used throughout by all agencies.

Project 3: Data analysis review project
This project, commencing in late 2003, will:
Collate, and screen for quality, all data from sampling rounds 1–3;
Conduct analyses on the combined MDB Audit fish data-set using three methods:

- IBI — multimetric analysis — using the revised IBI methodology (Harris in prep.);
- AUSRIVAS/RIVPACS (and/or the related e-ball) — multivariate predictive modelling — using a range of variables including the metrics from the IBI approach;
- the regression tree approach — univariate predictive modelling — using a range of univariate measures and/or metrics (including those from the IBI approach).

The project should evaluate the three approaches against a set of performance criteria (to be agreed with the MDBC and the MDBAs). The potential for adopting components of each approach within one analytical framework should be actively explored. A final analytical framework and methodology is to be recommended, and the first 3 rounds of the Audit fish data analysed using it.

This project will, therefore, by late 2004, produce:

- a report detailing a final analytical framework and methodology for fish bioassessment within the Audit;
- reports to the MDBAs and MDBC on the results from the first five-year stage of the Audit fish survey data, analysed using the final methodology;
- a report detailing the methods and results of the comparative analysis.

4.3 Sampling methodology

The following sampling methodology will commence in Phase 1, with two components:

1. *Electrofishing*

- Boat electrofishing in large rivers. 15 x 2 minute 'shots', stratified in proportion to the dominant habitats within the site. A minimum of 2 minutes between each shot.
- Backpack electrofishing of 2 x 50 m of pool-edge habitat and 2 x 50 m of riffle-run habitat in smaller systems. 5 x 2 minute passes of edge/snag habitats on each bank, in larger systems in addition to boat-shocking.

2. *Passive gear*

    - Fyke nets in all rivers. 4 nets each with 15 mm stretched mesh. Cod-end out to eliminate platypus mortality.
    - Three multipanel gill nets (slow water only).
    - Baited light traps, all rivers.

One night set for all gear only. Nets to be pulled and re-set through the night to reduce fish mortality.

All fish to be identified and processed in a manner consistent with that used in the NSWRS, and including measuring fork lengths to the nearest mm. Details of all data and of each sampling visit are to be recorded on standard field sheets. All catches are to be recorded separately by gear type.

Following the proposed review of the results of Phase 1 round 1 sampling, i.e. from 2002 onwards, a decision will be made as to whether sampling is reduced to electrofishing only. All subsequent sampling will be conducted in a manner consistent with the outcomes of this review.

A suite of habitat variables will be measured at each site and sampling occasion, consistent with habitat variables required for AUSRIVAS sampling in NSW in addition to on-site measurement of dissolved oxygen, pH, conductivity, turbidity and temperature.

4.4 Timing and frequency of sampling

All sampling is to be done during lower flows in summer–autumn period (December–April inclusive). All sites will be assessed (surveyed) once. Due to the intensity of resources required to conduct fish surveys in the MDB rivers, each full survey of all sites will be conducted over a two-year period, with half the sites being assessed in year 1 and the other half in year 2. It is intended that at least one full survey (each of two, annual sampling rounds), and possibly two (depending on the final number of sites specified for each river valley) could be completed prior to each five-yearly Audit review.

4.5 Variables to be measured

The Audit Fish Bioassessment workshop reviewed the variables that are required to be measured, while being conscious of the practicalities and limitations of fish sampling in MDB rivers. The following variables are to be derived from data recorded for each sampling occasion in the Audit fish component:

- number of species caught;
- abundance of each species in the catch;
- size distribution, as lengths;
- biomass, to be derived from 'standard' length–weight relationships for each species combined with the above length data;
- % of external lesions, abnormalities and parasites;
- variables which influence capture efficiency — EC, temperature, turbidity, fast/slow flow conditions;
- stocking/fishing history of study reach (where applicable).

4.6 Reference condition

The reference concept is particularly problematic for organisms like fish which respond to environmental cues and disturbances at relatively large spatial and temporal scales. Indeed, it is arguable that for some large-scale-migratory fish species, rivers of the scale of the Murray and Darling are required to identify reference conditions. In the absence of lowland river reaches that fit the criteria of a reference condition, assessments must be made relative to the 'best available' conditions. This principle is embedded within the Audit, and within a number of competing/complementary analytical approaches to river health assessment. However it must be recognised that true reference communities are absent within almost all of the Basin, due to the large-scale impacts of barriers, water-quality and habitat changes and interactions with alien species. In addition, considerable care must be taken in identifying issues relating to absence or reduced recruitment of migratory fish upstream and downstream of barriers when selecting candidate reference sites within sub-catchments.

The RFA Fish Bioassessment workshop recommended that two methods for defining the reference condition be used:

- data from 'best available' sites , which are screened against specific criteria (usually indicators of disturbance);
- lists of species believed to have occurred in each river valley/section, derived from expert knowledge and historical records.

The first type of reference 'set' is that used in the current NSW fish IBI and AUSRIVAS/RIVPACS macroinvertebrate bioassessment approaches. However, there was agreement that a historical list is also required to provide a broader conservation context to fish assessment in the Basin. This approach was also taken in the NSWRS (Harris and Gehrke 1997), and while the workshop recognised the problems with bias and fallibility of historical records, it considered that this is still a valid approach to defining a reference condition for fish species presence/absence and diversity, combined with expert-opinion assessment.

4.7 Sampling design & precision

Sampling design and precision will be explored as part of the recommended Design and Analysis projects. The spatial scale at which the indicators are to be reported is consistent with that for the other indicators in the Audit, i.e. the Valley Process Zone and river-valley scales. The analytical methods used for fish will allow aggregation of site assessments to these two scales in a manner consistent with the other indicators/ themes.

4.8 Protocols for assessment

For the purpose of the Audit, river health is considered synonymous with the term ecological integrity. For the purposes of the Audit, river health will be measured as:

> the degree to which aquatic ecosystems support and maintain processes and a community of organisms and habitats relative to the species composition, diversity, and functional organisation of natural habitats within a region.

A suite of derived variables or 'metrics' was recommended by the Audit Fish Bioassessment workshop (Table 4) in order to provide data that could be analysed in a manner consistent with the above definition. These metrics are to be derived from the variables to be recorded for each survey sample (see list above).

4.9 Analysis

The final analytical framework has yet to be decided. It must:

- incorporate a reference site/condition as basis for relative assessment of river 'health';
- use measures at several levels of biological organisation (community, population and individual levels) i.e. all the recommended metrics listed above;
- use a predictive approach which accounts for natural regional differences in fish assemblage structure;
- incorporate a temporal aspect into the assessment which accounts for key large-scale events which influence processes such as recruitment and mortality;
- produce outputs that are interpretable and can be related to variables describing the intensity of human impacts.

4.10 Aggregation of individual indicators to form theme assessment

The aggregation method for fish indicators within VPZs or river valleys will be consistent with the methods used for macroinvertebrates.

Table 4. Derived variables or metrics to be used in analysis of Audit fish bioassessment data.

Concept/Class	Metric
Abundance	Total abundance per unit effort
Biomass	Total biomass per unit effort
Native fish biodiversity	Number of native species Evenness of native species
Aliens	Biomass Abundance Biomass as proportion of all fish Abundance as proportion of all fish
Habitat guilds	Number of species (including aliens) that are: Benthic Pelagic Riffle dwelling Floodplain dwelling
Trophic guilds	Number of species (including aliens) that are: Macrophagic carnivores Microphagic carnivores Omnivores
Reproductive guilds	Number of species (including aliens) that are in reproductive strategy 1,2, 3a or 3b
Migratory guilds	Number of species (including aliens) that migrate at: basin scale Audit river valley scale local (reach) scale
Tolerances	Average scores across all species for: FSI (water quality) FSI (migration) FSI (general) *sensu* Chessman (in prep.)
Abnormalities	Number of individuals (including aliens) that have: visible abnormalities parasites
Size distribution	Number of individuals (list aliens separately) that are adult or subadult.

5. Approximate Costings

The following are indicative costings only for the main elements of phases1 and 2 of the first five years for the Audit fish component. Detailed costings will need to be provided to support budgetary allocations and decisions.

Support for a part-time coordinator is recommended at approx. $12k per annum during the first five years of this component, in order to coordinate MDB agency input, data provision, workshops and ongoing development of the design and analytical components.

Component	Indicative cost*	Responsible
Data collation, power analyses and sampling design review	$50,000	External sub-contractor. MDB agencies to provide data for analysis and attend workshop.
Reference condition project	$25,000	External sub-contractor, MDB agency staff, and experts to attend workshop.
Survey, round 1. Approx. 50 sites	$200,000	Vic, Qld, NSW MDB agencies only.
Survey, rounds 2 and 3. Approx. 200 sites	$400,000	All MDB agencies.
Data analysis review project	$75,000	External sub-contractor.
Program coordination	$60,000	External sub-contractor.

References

Adams SM and Ryon MG (1994) A comparison of health assessment approaches for evaluating the effects of contaminant-related stress on fish populations. *J Aquat. Ecosys. Health* **3**,15–25.

Adams SM, Ham KD, Greeley MS, LeHew RF, Hinton DE and Saylor CF (1996) Downstream gradient in bioindicator responses: point source contaminant effects on fish health. *Can. J. Fish. Aquat. Sci*. **53**, 2177–2187.

Angermeir PL and Smogor RA (1995) Estimating number of species and relative abundance in stream-fish communities: effects of sampling effort and discontinuous spatial distributions. *Can. J. Fish. Aquat. Sci*. **52**, 936–949.

Barbour MT, Stribling JB and Karr JR (1995) Multimetric approach for establishing biocriteria and measuring biological condition. *Biological Assessment and Criteria: Tools for Water Resource Planning and Decision Making.* (Eds Davis WS and Simon TP), pp. 63–77. Lewis, Boca Raton, FL.

Barbour MT and Yoder CO (2000) The multimetric approach to bioassessment as used in the United States of America. *Assessing the Biological Quality of Freshwaters* (Eds Wright JR, Sutcliffe DW and Furse MT), pp. 113–124. Freshwater Biological Association, Cumbria UK.

Barbour MT, Gerritsen J *et al.* (1999) *Rapid Bioassessment Protocols for Use in Wadeable Streams and Rivers: Periphyton, Benthic Macroinvertebrates and Fish.* 2nd Edn. EPAA 841-B99-002. USEPA, Office of Water, Washington DC.

Davies PE (2000) Development of a national river bioassessment system (AUSRIVAS) in Australia. In: "*Assessing the biological quality of freshwaters* (Eds Wright JR, Sutcliffe DW and Furse MT), pp. 113–124. Freshwater Biological Association, Cumbria UK.

Faragher RA and Rodgers M (1997) Performance of sampling-gear types in the New South Wales Rivers Survey. Harris JH and Gehrke PC 1997. *Fish and Rivers in Stress. The NSW Rivers Survey* (Eds Harris JH and Gehrke PC), pp. 251–268. NSW Fisheries Office of Conservation, CRC for Freshwater Ecology, Cronulla NSW.

Fausch KD, Karr JR and Yant PR (1984) Regional application of an index of biotic integrity based on stream-fish communities. *Transactions American Fish. Soc*. **113**, 39–55.

Frenzel SA and Swanson RB (1996) Relations of fish community composition to environmental variables in streams of central Nebraska, USA. *Environmental Management* **20**, 689–705.

Gehrke PC (1997) Species richness and composition of freshwater fish communities in New South Wales rivers. In: *Fish and Rivers in Stress. The NSW Rivers Survey.* (Harris JH and Gehrke PC, Eds). NSW Fisheries Office of Conservation, CRC for Freshwater Ecology, Cronulla NSW, 298 pp.

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

Harris JH and Gehrke PC (1997) *Fish and Rivers in Stress. The NSW Rivers Survey.* NSW Fisheries Office of Conservation, CRC for Freshwater Ecology, Cronulla NSW, 298 pp.

Harris JH and Silveira R (1999) Large-scale assessments of river health using an Index of Biotic Integrity with low-diversity fish communities. *Freshwater Biology* **41**, 235–252.

Joy MK and Death RG 2000. Development and application of a predictive model of riverine fish community assemblages in the Taranaki region of the North Island, New Zealand. *NZ Journal of Marine and Freshwater Research* **34**, 241–252.

Karr JR (1981) Assessment of biotic integrity using fish communities. *Fisheries* **6**, 21–27.

Karr JR (1987) Biological monitoring and environmental assessment: a conceptual framework. *Environmental Management* **11**, 249–256.

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

Karr JR and Chu (1997) Biological Monitoring and Assessment: using Multimetric Indexes Effectively. EPA 235-R97-001. University of Washington, Seattle.

Lake PS (1982) The relationship between freshwater fish distribution, stream drainage area and stream length in some streams of south-east Australia. *Bull. Aust. Soc. Limnol*. **8**, 31–37.

Lyons J (1992) The length of stream to sample with a towed electrofishing unit when fish species richness is estimated. *N. Amer. J. Fish. Management* **12**, 198–203.

Pusey BJ, Kennard MJ, Arthur JM and Arthington AH (1998) Quantitative sampling of stream fish assemblages: single vs multiple-pass electrofishing. *Australian Journal of Ecology* **23**, 365–374.

Reynoldson TB, Norris RH, Resh VH, Day KE and Rosenberg DM (1997) The reference condition: a comparison of multimetric and multivariate approaches to assess water-quality impairment using benthic macroinvertebrates. *Journal of the North American Benthological Society* **16**, 833–852.

Schiller CB, Bruce AM and Gehrke PC (1997) Distribution and abundance of native fish in New South Wales rivers. *Fish and Rivers in Stress. The NSW Rivers Survey.* (Eds Harris JH and Gehrke PC). NSW Fisheries Office of Conservation, CRC for Freshwater Ecology, Cronulla NSW, 298 pp.

Schofield NSJ and Davies PE (1996) Measuring the health of our rivers. *Water* (AWWA Journal) **23**, 39–43.

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

Attendees and Agenda for Audit Fish Bioassessment Workshop

Venue: University of Canberra Room 3B7, University of Canberra.
Dates: 1330 Wednesday April 11 – 1600 Thursday April 12.

Attendees: David Moffatt, Paul Humphries, Mark Kennard, Alison King, John Harris, Mark Lintermans, Tarmo Raadik, Sean Sloan, Peter Gehrke, John Whittington, Richard Norris, Leon Barmuta, Bruce Chessman, Ivor Growns, Julie Coysh, Claire Petekin.

Workshop facilitator: Peter Davies
Wednesday April 11th, 1330–1800
1330 -Gather in CRCFE front office, and walk to venue.
1400 - Welcome and workshop agenda (PD)
1410 - Introduction to the Audit (JW, PD)
1420 - Introduction to Audit Fish bioassessment scoping project (PD)
- Incorporating fish bioassessment into the Audit
- Key questions from the starter document (discussion)
1500 - the reference concept and Basin fish bioassessment.
*1530 - afternoon te*a
1550 - the reference concept continued.
1630 - Sampling methods used by MDB state agencies (discussion)
1700 - Sampling strategy (discussion)
*1800 - Session clos*e

April 12th, 0830–1600
0830 - John Harris – the IBI method
0900 - Richard Norris - RIVPACS style predictive modelling
0930 - Bruce Chessman – the PBH and MARA projects.
1000 - Mark Kennard - SW Queensland DIBM fish component
1030 - morning tea.
1100 - Discussion on core criteria for Audit fish bioassessment, and how each approach fits those criteria.
1300 - Lunch
1340 - Discussion and resolution on best overall approach for the Audit:
- analytical framework and methods;
- sampling.
1500 - afternoon tea
1520 - Final discussion and wrap up.
1600 - Close

Final Report
Project R2004
Development of a Framework for the Sustainable Rivers Audit

Appendix 5
Water Quality

Review and Development of Physico-Chemical Indicators

Ian Lawrence

Cooperative Research Centre
for Freshwater Ecology

Table of Contents

Summary

The selection of physico-chemical indicators is based on the capacity of streams to transform catchment inputs into food forms sustaining higher trophic levels in the stream, and to recycle the in-stream generated detritus.

The indicators reflect the key ecological processes (primary and secondary production & the mineralisation of organic material), and the potential modifiers of these processes (temperature, light or nutrient limitation or stimulation, salinity).

Except in cases of sampling sites established to monitor point source discharges, monitoring sites are predominantly ‘mixed zone’ (riffles, reaches) based.

In addition, given the low frequency of significant flow events, the routine nature of sampling for monitoring purposes means that data is predominantly for low to medium flow conditions. The proposed Audit approach builds on this existing monitoring approach, with data interpreted as reflecting outcomes on in-stream processes.

The adoption of a ‘reference’ based Index (O/E) to assessment of values for the test sites is proposed. In the case of the lowland Valley Process Zones (VPZs), it is generally not possible to identify pristine reference conditions. It is proposed in this case to use process based models to simulate ‘pre-development’ physico-chemical reference conditions.

This appendix elaborates the specific indicators to be measured, the structure of the physico-chemical sustainability index on a VPZ basis, the required number of sites and frequency of sampling, and the estimated annual cost of monitoring across the Basin.

Introduction: Theme context

This appendix addresses the selection of water quality related indicators of stream health, and the development of a physico-chemical index of health.

The Sustainable Rivers Audit has agreed on the development of an assessment framework, as a means of ensuring consistency in approach to the assessment across the Basin, and of guiding the selection of appropriate indicators and integration across themes.

The framework is also seen as an important means of communicating the selection of methods and indicators across a range of stakeholders.

The Audit framework comprises:

- reporting at the river-valley and Valley Process Zone (VPZ) scales;
- the adoption of river valleys, and a three major Functional Process Zones within the valleys, as a basis for stratifying rivers of the Basin into reasonably consistent functional groups for comparison purposes, and for stratification of monitoring sites;
- the development of a narrative (conceptual models) summarising our best understanding of key bio-geochemical processes determining in-stream physical, chemical and biological state responses to catchment inflows;
- the identification of principles guiding the selection of reference systems appropriate to each Functional Process Zone;
- the application of process based models to generate reference conditions where suitable ‘pristine or ‘slightly modified’ reference conditions are unavailable.

The framework identifies three types of process zones (VPZs) in the river valleys:

- upland zones (sediment supply);
- mid-slope zones (sediment transfer);
- lowland zones (sediment deposition or storage).

In approaching the task of selection of appropriate physico-chemical indicators, the primary focus has been on the assessment of river health (outcomes of ecological processes), in association with an assessment of potential physico-chemical modifiers of ecological processes.

Basic steps in undertaking the water quality assessment comprise:

- selection of water quality indicators of key ecological processes or modifiers of processes;
- selection of water quality monitoring (test) sites on a Valley-Process-Zone-basis and catchment-area-basis;
- selection of reference systems for each Valley Process Zone and region (or river-valley) category;
- generation of ‘reference values’ where ‘pristine’ or ‘slightly modified’ reference conditions are not available;

- assessment of available water quality data for designated test and reference sites across Basin, and computation of 'health index' values for each;
- comparison of test' and 'reference' health indices on a river-valley and Functional Process Zone basis.

What is being assessed: Health assessment focus

The Audit project defined River Health as:

i) the degree to which aquatic ecosystems support and maintain processes and a community of organisms and habitats relative to the species composition, diversity, and functional organisation of natural habitats within a region;

ii) the degree to which biological processes incorporate similar amounts of material into the food web as reference systems (productivity); and maintain a food web of similar complexity to that of reference systems (ecological processes) (Gawne 2001).

The second of these definitions provides a rationale and framework for stratifying rivers on the basis of similar functional processes, and similar indicators of structure and ecological processes. This has been adopted as the basis of the development of the physico-chemical assessment of health approach.

The definition also highlights the importance of the stream capacity to transform inputs to the stream into food forms (primary & secondary production) sustaining higher trophic levels in the stream. Productivity of a stream is a function of the organic material and nutrient inputs from its catchment, the in-stream transformation of these materials and recycling rates, temperature and availability of light.

There is also a range of potential modifiers of these processes, either limiting levels of production, or over stimulating production in the case of wastewater discharges high in bio-available nutrients.

In approaching the task of selection of appropriate physico-chemical indicators, the primary focus is on the assessment of river health (outcomes of ecological processes), in association with an assessment of potential physico-chemical modifiers of ecological processes.

Conceptual models

Generally, we observe longitudinal gradients along streams not only in their elevation, but also in their streambed particle size, suspended particle size, size and complexity of organic material, composition of nutrients, and total dissolved solids, to mention a few constituents.

The River continuum concept (RCC) builds on this principle in terms of its description of organic composition and the range of functional feeding groups present in each functional zone. However, the concept needs to be modified in terms of the overlay of riparian and lateral inputs (Riverine productivity model), important during periods of low flow. In addition, within each functional zone, the flow phase plays an important role in determining the dominant functional processes (Flood pulse concept). At the biota level, ‘flow disturbance’ also plays an important role in driving diversity.

In the case of lowland VPZs, the longitudinal processes are further complicated by the changing connectivity structure of the system with changes in flow (water height) rates. The drainage from backwaters or overbank return flows on the falling arm of a flood hydrograph may represent significant lateral inputs of organic material and nutrients to the main channel(s).

For medium to low flow conditions, the physico-chemical water quality at a point in a stream will be a reflection of in-stream bio-geochemical processing of upstream inputs from the catchment (primarily during high rainfall–catchment discharge conditions), together with local inputs from stream corridor and riparian vegetation.

In cases of point source discharges (wastewater or groundwater discharge), stream reaches immediately downstream may exhibit unutilised nutrients or organic material, due to the lag in growth of primary or secondary biomass, lack of suitable substrate, light or nutrient limitations, low temperature or insufficient detention time.

Consequently, the physico-chemical quality at the sampling point reflects the outcome of upstream ecological processes and the potential modifiers.

It is important to note that this approach sets aside the more traditional ‘driver’ approach to water quality monitoring, on the basis that:

- the lack of intensive monitoring of flow events, or surveys of the composition of sediments, prevent any systematic ‘driver’ based assessment of streams;
- the routine monitoring of mid-slope and lowland VPZs represents predominantly medium to low flow conditions, during which in-stream processing of recycled nutrients and organic material is dominant. viz: the physico-chemical water quality reflects outcomes of in-stream ecological processes;
- the stream health focus is more about the state of the system than driver–ecological impairment descriptions.

The proposal to undertake a Pilot Run provides an important means of further developing and testing the approach. The approach is entirely consistent with the conceptual models, risk assessment and reference condition basis of the revised Guidelines for Fresh and Marine Water Quality (ANZECC & ARMCANZ 2000).

Selection of indicators

Two approaches are proposed to the selection of physico-chemical indicators and the development of an index:

- build on the ‘narrative of dominant ecological processes’ developed as part of the overall Audit framework, in identifying key physico-chemical indicators of in-stream bio-geochemical processes;
- build on the methods documented in the Guidelines for Fresh & Marine Water Quality (ANZECC & ARMCANZ 2000) in identifying possible modifiers of ecological processes, and in selecting appropriate indicators of stressors or modifying agents.

Ecological process outcome (physico-chemical) indicators

From a physico-chemical assessment perspective, there are three components:

- primary drivers of in-stream ecological processes: the nutrients, organic material discharged to streams during catchment rainfall events, and in point source and groundwater discharges;
- modifiers of ecological processes: the availability of nutrients & light, temperature (rate of growth), physical mixing, salinity (impact on cell physiology), pH & alkalinity modifiers (impact on chemical equilibrium);
- in-stream ecological (primary & secondary production) response *processes* and associated physico-chemical outcomes (indicators).

As noted above, the primary focus of a 'stream health' assessment is on the physico-chemical water quality 'outcomes' of in-stream ecological processes, and on potential modifiers of these in-stream ecological processes, rather than on an assessment of drivers.

In view of the predominant non-point source nature of stream inputs, point source discharges may be treated as 'modifiers' of in-stream ecological processes.

A further important assumption underpinning the approach is that under the low flow conditions, water quality and biological uptakes at the point of sampling are close to equilibrium in relation to the rates of bio-availability of nutrient and carbon sources.

Drawing on the best available understanding of bio-geochemical processes characteristic of a range of Valley Process Zones, a map of the key bio-geochemical processes as a function of stream flow phase and functional zone has been developed.

Tables 4 A & B describe the dominant bio-geochemical processes specific to each River Valley Zone and flow condition. The process descriptions have been used to identify the key physico-chemical outcomes (indicators) of ecological processes, across the range of Valley Process Zones. They also provide a framework for interpretation of the condition being measured and what the data tells us about stream health.

Summary of indicators of physico-chemical outcomes of ecological processes:

Upland VPZs	nutrient levels & level of mineralisation, level of FPOM & DOM and proportion of total organic material;
Mid-slope VPZs	DO level & diurnal variation, pH level & diurnal variation, nutrient levels & composition, level of FPOM & DOM and proportion of total organic material;
Lowland VPZs	DO level & diurnal variation, pH level & diurnal variation, chlorophyll 'a', nutrient levels & composition, level of FPOM & DOM and proportion of total organic material.

How do indicators meet the criteria for indicator selection:

Consistency with conceptual models of river function:
The indicators have been built directly on the functional processes models described in Appendix 2. It is proposed that the conceptual models be expanded upon in the first pilot year of the Audit to provide more detailed information on the physico-chemical aspects as described in this paper.

Responsiveness to disturbance:
All the outcome indicators are highly responsive to changes in trophic levels and processes, as well as having distinctly different characteristics across different zones.

Measurement and analysis are rapid:
Field probe or sensor (conductivity, DO, pH, NH_4, NOx) based measurements in a number of cases, combustion loss or non-filterable residue (TOC, SS), chemical colorimetric or titrimetric or oxidation (DRP, TKN, DOC, alkalinity) methods.

Standardised methods are available:
Analysis for all indicators is covered by standard methods.

Output can be interpreted relatively unambiguously:
Measurement of major potential modifiers of processes is included to assist the interpretation of the indicator values.

Indicator has meaning to the wider Basin community:
Indicator and O/E values are common across FPZs for VPZs.

A number of the indicators are well established as sensitive measures of net primary production (diurnal DO change, pH change), secondary production & mineralisation (NH_4/NOx, NOx/TN), and the processing of organic material & mineralisation ((FPOM+DOM)/TOC, NOx/TN) in the case of upland Process Zones.
Lawrence *et al.* (2000) demonstrated the sensitivity of the NH_4/NOx ratio as an indication of reducing levels in the case of reservoirs. While similar patterns have been observed in relation to wastewater discharge zones in streams, there has not been extensive application of this indicator to streams at this stage. Further explanation regarding the application of these indicators is required in the Report.

For non-point source based river systems, runoff derived from elevated rainfall events constitutes the major driver of inputs of suspended solids, nutrients and organic material to streams. Research reported by Hart, Grace & Beckett indicates that particulate material rapidly adsorbs nutrients and toxicants, and develops biological coatings of organic material. The particulates with their coating of nutrients, organic material and toxicants, settle to the sediments under less turbulent flow conditions in deeper pools or on the falling arm of the flow event hydrograph.

There has been extensive laboratory and reservoir and lake based demonstration of P release from sediments under low redox conditions. Laboratory based sediment core experiments (Armitage 1995) demonstrated the capacity for a range of carbon sources to reduce sediments, with significant remobilisation of N and P. The research demonstrated the potential for nutrient limitation of the microbial growth, slowing or limiting the sediment reduction and transformation and release of nutrients.

Hart *et al.* 2000 reported that benthic chamber analysis of stream sediment fluxes indicates that monitored P releases do not necessarily increase even when sediments turn anoxic.

Field observations of river sediment release of P are confounded by the heterogeneity of sediments, limited duration (limited redox development) of benthic chamber experiments, lack of redox measurement, limited labile C to drive redox conditions down, and the rapid uptake of a component of released P by bacteria.

The application of sediment diagenesis models, linked to redox conditions, indicates rapid to slow release of P from sediments, depending on the depth of $Fe(OH)_3$ layers and redox conditions (Harper 2001).

There is extensive published material reporting on in-stream N release rates from sediments. De-nitrification occurs at low levels of DO and moderate redox level conditions, and an order higher level of N than P.

Analysis of a range of organic materials indicates algae and some grasses have a labile carbon content some 20 times that of eucalyptus-derived litter. The analysis also highlighted the slow rate of bio-degradation of a range of native vegetation-derived carbon materials, in excess of 100 days in some cases, and the nutrient limitation as a significant factor in determining slow decomposition rate for some materials (Esslemont 2000).

There is substantial similarity between the State of the Environment and Audit water quality indicators (turbidity or suspended solids, salinity, pH, DO, temperature, nutrients). TOC has been added to the Audit indicators, as an important indicator of organic material recycling efficiency, while toxicants have been excluded on the basis of monitoring being beyond the capability of the Audit at this stage.

The selection of indicators was also cognisant of resource constraints faced by monitoring agencies in terms of funds, staff and technical capacity in relation to non-traditional measures, and requiring more complex laboratory analysis. This meant that in a number of cases, the available indicators were sub-optimal.

Potential modifiers of ecological processes indicators

The Guidelines for Fresh & Marine Water Quality (ANZECC & ARMCANZ 2000) identifies the nine major threats (management issues) to aquatic ecosystem functioning and biota, and related indicators of stressors and potential modifiers (Table 1).

Table 1. Summary of threats to aquatic biota and related indicators

Management issue	**Condition indicator**	**Stressor indicator**	**Potential modifiers**
nuisance plant growth	Chlorophyll 'a', change of pH, DO, algal composition	TP, TN, TOC loads (indirect)	Detention time (flow), turbidity, SS (nutrient sorption), pH, temperature, substrate
oxygen depletion	change of DO	TOC or BOD load, NH_4	Mixing (flow), re-aeration (flow), temperature, photosynthesis
elevated suspended solids	turbidity, algal composition, SS concentration	SS load	Flow
salinity changes	EC	Salt load, evaporation losses	Flow
temperature change	change of temperature	temperature of inflows	Flow
pH modification (direct & indirect)	change of pH	acids, bases, photosynthesis, respiration	Alkalinity
changes in optical properties	change of turbidity	SS, nutrient loads (direct), TOC loads (indirect)	TDS, flow
changes in flow regime	seasonal flow regimes	change of seasonal flow duration	
toxicants metals non-metal inorganics	biological effects	Cd, Cu, Pb, Zn NH_4	TDS, DO, SS, DOM, temperature, hardness, pH

Proposed physico-chemical indicators of potential modifiers of ecological processes for the Audit:

—flow, temperature, TDS or EC, turbidity, SS, pH, alkalinity, DO or %saturation, NH_4.

The retention of the monitoring of the modifiers is important:

- to provide a basis for interpretation of observed shifts in the physico-chemical index values;
- to 'normalise' reference and test site measurements in the case of available reference sites; and
- to estimate reference conditions appropriate for the test site conditions in the case of the modelled reference conditions.

Development of Index

There is a range of existing water quality indices, including the Saprobien system of organic pollution measurement, and the trophic system of nutrient enrichment measurement. The indices provide a measure of the cumulative effects of a selected stressor (organic material, nutrients) on the abundance & composition of selected phyla. These indices relate to specific management issues and focus on biological effects.

In view of the major focus of the Sustainable Rivers Audit on the health of the stream, it is important that the physico-chemical index provides a measure of the primary and secondary productivity of the stream, and of potential modifiers to these systems. It is intended that the measures for the test site be compared to the same measures for the reference site, to provide an Observed/Expected ratio.

The physico-chemical indicators of ecological process outcomes, and of potential modifiers to the ecological processes, were identified in the 'Selection of indicators' Section above. The map of the 'key bio-geochemical processes as a function of stream flow phase and functional zone' (Tables 4A, 4B) highlights the central role of primary and secondary production in recycling key elements of life — organic carbon and nutrients.

The productivity of the stream, and the in-stream utilisation of organic material and nutrients discharged from their catchments, are a function of the efficiency of these recycling processes. The measurement of the efficiency of recycling therefore provides a powerful index of stream health, or conversely, departure from expected recycling levels is indicative of either the absence of drivers, or the presence of modifiers potentially impairing the recycling processes. The primary & secondary production is the base of the food web, while the modifiers also relate to the health of the higher trophic levels.

For example, in the case of elevated nutrients, the resultant algal stimulation will result in marked shifts in the DO, pH and NH_4/NOx ratios. Conversely, elevated suspended solids will result in the suppression of biota (smothering, nutrient adsorption & burial, light limitation), and diminution of ranges in DO and pH, and in-stream nutrient residuals.

Similarly, elevated levels of organic material will stimulate secondary production, with depressed levels of DO and pH, and increased levels in the NH_4/NOx ratio and FRP occurring.

In the case of elevated total dissolved salts, shifts will occur in the composition and potentially, in the productivity of biota. Consequently, TDS is being treated as a modifier of ecological processes/health of the stream.

The modifier indicators provide a basis for:

- explanation of observed shifts in indicator values;
- normalisation of monitored and reference sites where 'naturally based' differences occur in flow and temperature;
- calculation of diurnal DO or pH ranges where a single day-time sample only is available.

Impairment due to toxicants will be read in terms of reduced metabolic rates and shifts in fauna composition, e.g. situation of depressed primary production levels in presence of DRP & NOx residuals, and free of light, temperature or retention time constraints.

Given that the wash-off of toxicants will occur during rainfall events (elevated flows), the 'low flow' based monitoring approach would exclude sampling of these discharge conditions. There are also difficulties in relation to the wide range of potential toxicants. It may be that a generic toxicant such as endosulfans could be assessed as representative of a range of crop related pesticides. The possibility of incorporating some background water quality indicators as part of the monitoring program (including toxicants) could be explored as part of the Pilot Project.

The description of key bio-geochemical processes in Tables 4A & B also highlights some important differences between the upland, and the mid-slope and lowland VPZs ecological processes. Building on this approach, the following index measures & structures are proposed:

Upland VPZ

Organic material recycling processes: transformation of CPOM to FPOM & DOM and further uptake by secondary production and higher trophic levels:

i) efficiency measures: (FPOM + DOM)/TOC;

ii) potential modifiers of processes: flow level, temperature, elevated SS.

Nutrient recycling processes: mineralisation of organic forms of nutrients to inorganic forms & further uptake by primary production:

i) efficiency measure: NOx/TN;

ii) potential modifiers of processes: flow level, temperature, elevated SS, elevated nutrients (point source).

Mid slope & Lowland VPZs

Organic material recycling processes: utilisation of FPOM & DOM in secondary production and direct uptake by filter feeders & grazers:

i) efficiency measures: TOC, (FPOM+DOM)/TOC, NH_4/NOx, diurnal range of DO or pH;

ii) potential modifiers of processes: flow level, temperature, elevated SS, elevated organic material (point source).

Nutrient recycling processes: mineralisation of organic forms of nutrients to inorganic forms & further uptake by primary production:

i) efficiency measures: level of primary production (diurnal range of DO or pH), chlorophyll 'a' levels, residual inorganic nutrients (NOx, NH_4, DRP), level of mineralisation (NOx/TN);

ii) potential modifiers of processes: flow level, temperature, elevated SS, elevated nutrients (point source), elevated TDS.

The Index structure is summarised in Figure 1. Outline of Index structure. Note that this approach is significantly different from the current fashion of multi-variate analysis. It provides a much more transparent representation of the in-stream functioning of primary and secondary processes.

Table 2. Calculation of Index

Valley Process Zone	Indicator	Range of levels	Calculation of Index relative to Reference values
Upland	TOC	TOC > Ref TOC < Ref	inverse ratio direct ratio
	(FPOM+DOM)/TOC	low (extensive processing) to high (limited processing)	inverse ratio
	NOx/TN	high (extensive mineralisat) to low (limited mineralisat)	direct ratio
Mid-slope & lowland	diurnal DO	DO > 100%saturation net production DO < 100% saturation net reduction	inverse ratio direct ratio
	pH	pH > Ref net prodtion pH < Ref net reduction	inverse ratio direct ratio
	TOC	TOC > Ref TOC < Ref	inverse ratio direct ratio
	Chlorophyll 'a'	Chlor a > Ref Chlor a < Ref	inverse ratio direct ratio
	(FPOM+DOM)/TOC	low (extensive processing) to high (limited processing)	inverse ratio
	NH_4/NOx	low (well oxidised) to high (severe reducing)	inverse ratio
	DRP	limiting levels — effic growth higher levels — impair growth	inverse ratio
	NOx/TN	high (extensive mineralisation) to low (limited mineralisation)	direct ratio

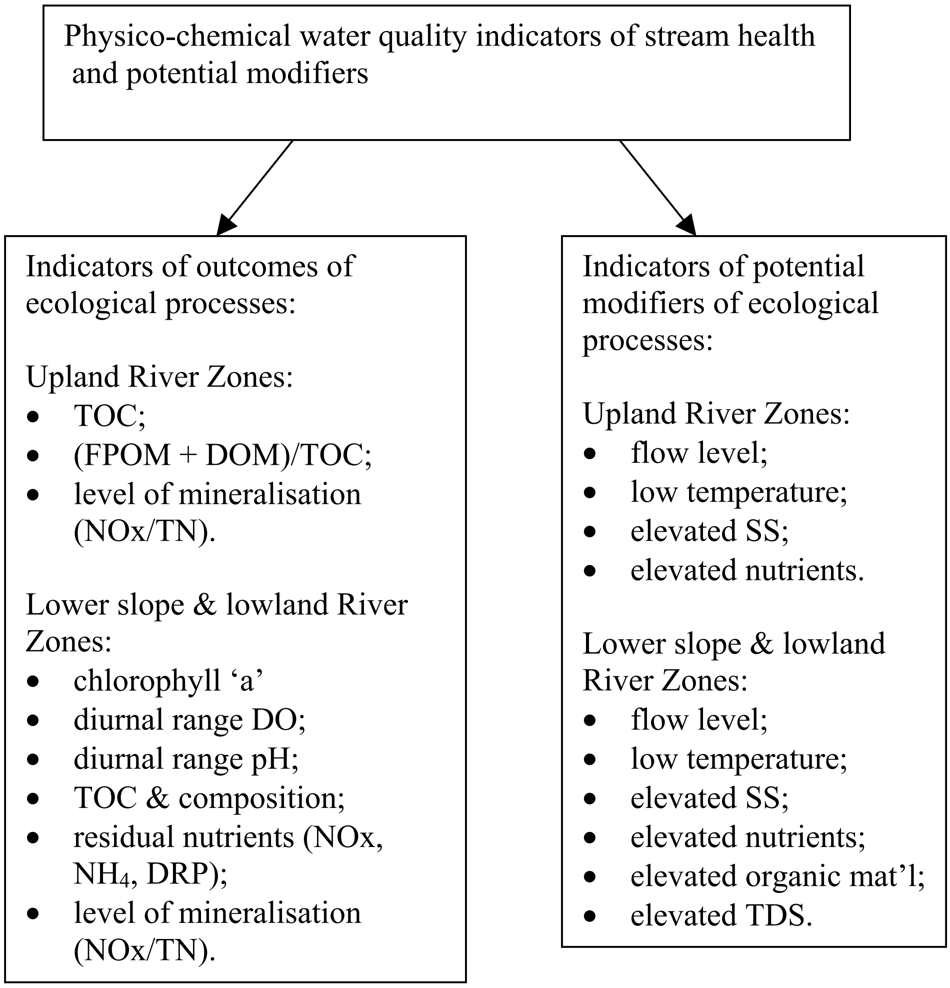
Notes: A number of the indicators are flow and/or temperature dependent, and so require normalisation in relation to differences between monitored site and Reference site 'natural' background factors (based on flow–indicator regression curves for the Reference site).

Composite Water Quality Index calculation:

(sum of the individual indicator indices) ÷ (the number of the indices for the VPZ).

The '20% change values' is a judgement, guided by the ANZECC Water Quality Guidelines 1992 'limits to acceptable change' in relation to the potential for impairment of biota. It is intended as the identification of an 'increment' of change that is likely to be significant in ecological terms, without any overlay of acceptable or unacceptable bands at this stage. The proposed Pilot Project will be invaluable in further testing and developing this aspect of the approach.

Figure 1. Outline of Index structure



Calculation of Water Quality Index:

i) Normalise Reference site ecological indicator values for flows equivalent (percentile) to monitored site flows.

ii) Calculate the O/E Index for each ecological Indicator, based on Table 2.

iii) Calculate the cumulative Water Quality Index for the site, based on the sum of the individual indicator indices, divided by the number of the indices for the Valley Process Zone.

Reference system selection basis

The revised Guidelines (ANZECC & ARMCANZ 2000) require selection of reference sites on the basis of similar bio-geographic & climatic regions, geo-morphological, pedological & topographical characteristics, similar range of habitats and equivalent riparian and aquatic plant communities. In the case of the Audit approach, the framework provides a systematic basis for stratification of river valleys and Functional Process Zones, guiding the selection of reference systems appropriate to the test site.

The selection of reference sites for water quality sampling needs to reflect the Valley, Valley Process Zone and Functional Process Zone of each monitored site Valley Process Zone & Functional Process Zone category, on a valley by valley basis.

In tabulating the list of possible reference sites meeting these criteria, there is a need to exclude non-representative sites such as sites immediately downstream of river junctions or point-source discharges (problem of transverse stratification of flows). A random selection of reference sites is then made from the tabulation of possible sites.

In the case of lowland Process Zones, there are few unmodified streams available for reference purposes. In these cases, it is proposed to generate the best available estimate of reference conditions, drawing on data from modified streams, and estimates provided by the application of an interactive transport, sedimentation, sediment redox & biofilm uptake process based model (daily time step).

The process-based models will be built-up from the conceptual (narrative) models, utilising established physical, chemical & biological (primary & secondary production) relationships. The models will be based on daily time steps and ‘train of channel morphological components’ representative of the Functional Process Zone. The CRCFE has established process-based river models, integrating transport, sedimentation, re-suspension, sediment redox (secondary production), nutrient release, algal uptake (primary production) for mid-slope and lowland rivers. The models are Excel based, with simple input of daily inflows and composition.

A key component of the models will be the temporal changes in physical structure (secondary channel and floodplain connectivity) and processes as a function of flow.

In the case of physico-chemical assessment, indicator values are significantly influenced by flow. Where differences occur between the test site flow and the reference system flow, the test and reference sites are no longer comparable.

Limiting ‘acceptable data’ to baseflow conditions (20 to 80 percentile range), for which a ‘comparable’ reference value is available, does not appear to be a viable solution to this problem, in view of the 20 to 80 percentile flow ratios of 30, 10 & 5 times for upland, mid-slope & lowland streams respectively.

It is proposed that the reference condition should be based on the flow duration condition comparable to that prevailing at the test site at the time of sampling.

This will require the establishment of flow duration curves for both reference and test sites, and the development of a correlation of physico-chemical values with flow in the case of the reference sites.

In this case, a flow–concentration regression curve will need to be developed for each indicator, based on multiple sites and samples for the VPZ. Analysis of the O/E for the test site is then based on the monitored test site value over the regression curve value for the reference VPZ for the equivalent flow condition. Flow condition in this context is the equivalent flow duration probability.

Sampling design

Overview of existing programs & methods

Routine physico-chemical monitoring programs typically are based on sampling from well mixed (riffle or stream reach) zones, and reflect predominantly medium to low flow conditions (flows prevailing for some 97–99% of the time for mid-slope & lowland River Valley Zones).

There is little to no event based monitoring or sediment surveys currently undertaken, against which systematic assessment of drivers of processes could be made: viz. physico-chemical based assessment is limited to assessment of in-stream bio-geochemical process outcomes.

There is an extensive network of water quality monitoring sites for the northern and western slope streams, but limited sites for the western plains streams.

Selection of test sites

There is a choice of two approaches to the selection of test sites:

- a ‘random’ based selection of sites from available monitoring sites for the river valley and Functional Pprocess Zone; or
- a ‘stratified’-based selection, reflecting a standard set of drainage areas.

There is a substantial body of literature describing the inverse relationship between sediment, nutrient, organic loads per km^2 and basin drainage area, reflecting in part the greater proportion of lowland river zones (reduction in transport energy) associated with large lowland areas in the case of the larger basins. While the Functional Process Zones-based stratification approach will go some way towards removing this dependence, it is expected that rivers having extensive lowland reaches, such as the Darling, will show significant differences between the upper and lower reaches of the lowland river functional zone.

It is proposed that a ‘basin drainage area’ based selection of monitoring sites should be adopted as a further ‘stratification overlay’ in these cases.

In the case of the water quality indicators, the site comprises a riffle or well-mixed zone in the case of upland and mid-slope Process Zones, and a channel (pool) reach in the case of lowland Process Zones. Sampling needs to be taken clear of edge effects in each case, in order to reflect channel water quality. In the case of the lowland Process Zones, under low flow conditions, the ‘channel’ water quality will of course reflect local transverse inputs and bio-geochemical processes.

At this point, it is proposed that the only criteria to be used to define sites are the Valley Process Zone and Functional Process Zone categories, and the requirement for sites to be capable of providing a representative (mixed) sample. The framework proposes a random selection of sites from the sites meeting these three criteria, on a valley by valley basis.

Design of number of sites & sampling frequency

The required confidence level in estimates is differences of 20% in the averaged O/E ratio to be detectable at the 10% level.

Number of sites required — confidence levels in respect to differences:

- 6 test sites per River Valley Zone per river valley;
- 3 reference sites per River Valley Zone per river valley.

Timing & frequency of sampling:

- 2 or 3 monthly based samplings across the year (= 4–6 samples/year).

Protocols for assessment

From the CRCFE and other research into in-stream bio-geochemical processes, there is now an appreciation that in-stream physico-chemical data is sometimes a measure of drivers of biological processes, and at other times a measure of the outcome of biological processes. In the unlikely event that a sample reflects a high flow event condition, it is proposed that this data should be excluded on the basis that it does not represent in-stream ecological process outcomes.

Proposal for resolving this issue: Undertake assessment of the water quality data in relation to flow conditions at the time of sampling, and classification of water quality data into drivers or outcomes on the basis of flow levels greater than or less than:

- 15 percentile flows for lowland VPZs;
- 25 percentile flows for mid-slope VPZs;
- 35 percentile flows for upland VPZs.

In the case of the water quality indicators, the site comprises a riffle or well-mixed zone in the case of upland and mid-slope Process Zones, and a channel (pool) reach in the case of lowland Process Zones. Sampling needs to be taken clear of edge effects in each case, in order to reflect channel water quality. In the case of the lowland Process Zones, under low flow conditions, the ‘channel’ water quality will of course reflect local transverse inputs and bio-geochemical processes.

In order to assess the monitored indicator values, in some cases, information is required on a range of associated potential modifiers of bio-geochemical processes. For example, ideally, monitoring is required of the diurnal pattern of in-stream DO and pH, in order to assess primary or secondary production and respiration balance and rates. As a requirement for 24 hours monitoring at each site would be resource intensive, it is proposed to compare a daylight based sample with calculated equilibrium for the prevailing flow, temperature and alkalinity conditions, to assess production & respiration rates.

In the case of the limited depth of the riffle zone based sampling sites for the upland and mid-slope Process Zones, this will not be an issue. In the case of the deeper channel (pool) reaches for the lowland Process Zones, it is proposed that an integrated sampler (tube) be used to integrate variation in indicator values across the depth of the pool.

For each River Valley index, aggregate the VPZs' indices on the basis of the mean value of individual VPZ indices.

Costing

Based on preliminary analysis of water quality variability and the required detectable difference and confidence level, the required number of samples is 6 samples per site (9) per river-valley zone per river valley.

Typically, the water quality index site selection requirements are consistent with the macroinvertebrate site selection requirements. Consequently, water quality sampling could be taken at the same site and time as the macroinvertebrate surveys. Additional water quality samples will be required in order to meet the statistical significance probability criteria.

It is assumed that based on the existing gauging network, and the application of hydraulic models, estimates of flows of sufficient accuracy for the purposes of the Audit can generally be generated without the need for additional gauging stations. Where gauging stations are required, it may be sufficient that staff gauges are installed at sampling sites (officers collecting water quality samples to note level), rather than incurring the high cost of establishing fully automated stations. Costing has included the cost of establishing and calibrating additional staff gauges across 20% of the sites.

Based on the schedule of variables to be analysed, the University of Canberra Ecochemistry Laboratory analysis rates, and an estimate of travel and sampling technicians cost, the annual cost of the physico-chemical sampling and analysis is estimated at $830,000 (see Table 3).

Table 3. Cost of one year's sampling for water quality index.

Physico-chemical sampling cost estimate

Variable	Samples/yr	Analysis cost/sampl	Total cost analysis	No sampling runs	Cost/run	Labour cost	Total cost sampling	Total cost	Avg Valley Total Cost/yr	Basin Total Cost/yr
Upland Valley Process Zone										
SS	54	$8	$432							
TOC	54	25	$1,350							
FPOC	54	25	$1,350							
DOC	54	25	$1,350							
NO3+NO2	54	10	$540							
TN	54	20	$1,080							
TP	54	20	$1,080							
Total			$7,182	6	$250	$400	$3,900	$11,082		
Mid slope Valley Process Zone										
SS	54	$8	$432							
Ec	54	7	$378							
DO	54	7	$378							
pH	54	7	$378							
TOC	54	25	$1,350							
FPOC	54	25	$1,350							
DOC	54	25	$1,350							
TN	54	20	$1,080							
NO3+NO2	54	10	$540							
NH4	54	10	$540							
TP	54	20	$1,080							
SRP	54	10	$540							
Total			$9,396	6	$250	$400	$3,900	$13,296		
Lowland Valley Process Zone										
SS	54	$8	$432							
Ec	54	7	$378							
DO	54	7	$378							
pH	54	7	$378							
TOC	54	25	$1,350							
FPOC	54	25	$1,350							
DOC	54	25	$1,350							
TN	54	20	$1,080							
NO3+NO2	54	10	$540							
NH4	54	10	$540							
TP	54	20	$1,080							
SRP	54	10	$540							
Total			$9,396	6	$250	$400	$3,900	$13,296	$37,674	$678,132

Notes:
Based on 6 test sites & 3 reference sites per Valley Process Zone, and 6 samples/site/yr
Based on the numbers of valley process zones identified for the 22 river valleys
Based on an average travel distance of 500 km/sampling run/Valley Process Zone x 50 c/km
Based on 2 technicians x 1 day x $200/day per sampling run per Valley Process Zone
Analysis costs based on University of Canberra EcoChemistry Laboratory rates 6 July 1999

Table 4A. Map of key bio-geochemical processes as a function of flow phase and functional zone: Upland River Valley Zone

Functional zone	**Flow phase**		
	Flow events	**Post event period**	**Sustained low flow**
Pools Dominant processes:	Major catchment discharge of water, sediment, SS, organic detritus (CPOM). Transport of sediment, SS & organic mat'l. Re-suspension of organic mat'l previously deposited or built-up in riffles post the previous event. Deposition of sediment, SS & organic mat'l in pools.	Mechanical &microbial (secondary prod'n) weathering of organic mat'l. High rates of re-aeration of waters. Uptake of CPOM by biofilm (epilithon). Uptake of released nutrient by epilithon (benthic algae, fungi & bacteria).	Significant riparian vegetation inputs locally. Epilithon uptake of CPOM & FPOM from detrital sources & weathering/decomposition of riparian mat'l.
Potential modifiers of processes:	Flow rates, duration of event, elevated sediment loads	Flow rates, low temperature (suppress biological rates), nutrient point source discharge (elevated primary prod'n/FPOM), consumption of epilithon by periphyton & contribution to detrital pool.	Low temperature (suppress biological rates), nutrient point source discharge (elevated primary prod'n/FPOM), consumption of epilithon by periphyton & contribution to detrital pool.
Riffles Dominant processes:	Re-suspension or sloughing of organic mat'l built-up post the previous event & transport downstream.	Uptake of CPOM & nutrients by epilithon (benthic algae, fungi, bacteria).	Significant riparian vegetation inputs locally. Epilithon uptake of CPOM & FPOM from detrital sources & weathering/decomposition of riparian mat'l.
Potential modifiers of processes:	Flow rates, elevated sediment load, disturbance of cobbles/gravel.	Flow rates, low temperature (suppress biological rates), elevated SS (light limitation), nutrient point source discharge (elevated primary prod'n/FPOM), consumption of epilithon by periphyton & contribution to detrital pool.	Low temperature (suppress biological rates), elevated SS (light limitation), nutrient point source discharge (elevated primary prod'n/FPOM), consumption of epilithon by periphyton & contribution to detrital pool.
Physico-chemical outcomes of bio-geochemical processes (riffle zones)	Sediment & SS levels, organic mat'l levels, nutrient (adsorbed) levels	Mineralisation of organic mat'l. Breakdown of CPOM to FPOM & DOM. SS, residual nutrients & composition, residual FPOM & DOM.	Mineralisation of organic mat'l. Breakdown of CPOM to FPOM & DOM. SS, residual nutrients & composition, residual FPOM & DOM

Table 4B (i). Map of key bio-geochemical processes as a function of flow phase and functional zone: Mid-slope & Lowland River Valley Zones

Functional zone	**Flow phase**		
	Flow events	**Post event period**	Sustained low flow
Pools Dominant processes:	Major catchment discharge of water, SS, adsorbed nutrients, organic mat'l. Transport of sediment & SS. Re-suspension of organic mat'l previously deposited or built-up in reaches & riffles post the previous event. Deposition of sediment, SS, adsorbed nutrients & organic mat'l in large pools.	Decomposition of sedimented organic mat'l (secondary production), with potential release of mineralised nutrients, or adsorption/burial in sediments, or loss to atmosphere. Uptake of released nutrient by benthic biofilm, attached algae, plankton.	Run-down in leakage of nutrients from previous event. Direct recycling of nutrients. Riparian vegetation inputs locally. Biofilm uptake of FPOM & DOM from detrital sources & decomposition of riparian mat'l.
Potential modifiers of processes:	Flow rates, duration of event, elevated sediment loads, elevated SS levels, over-bank flows & returns (dispersion of SS/adsorbed nutrients, elevated CPOM & FPOM returns)	High organic mat'l loads (de-oxygenation, redox), high SS levels (nutrient adsorption/limitation, light limitation), pH modification (chemical equilibrium), elevated TDS (flocculation SS, toxicity), low temperature (suppress biological rates), nutrient point source discharge (elevated primary prod'n), consumption organic prod'n by grazers & collectors.	Time since event, pH (chemical equilibrium), elevated TDS (flocculation SS, toxicity), low temperature (suppress biological rates), nutrient point source discharge (elevated primary prod'n), consumption organic prod'n by grazers & collectors.

Table 4B (ii). Map of key bio-geochemical processes as a function of flow phase and functional zone: Mid-slope & Lowland River Valley Zones

Functional zone	**Flow phase**		
	Flow events	**Post event period**	Sustained low flow
Riffles/reaches Dominant processes:	Re-suspension or sloughing of organic mat'l built-up post the previous event & transport downstream.	Uptake of nutrient & FPOM released from upstream pools by biofilm, benthic algae, attached algae & plankton.	Direct recycling of nutrients. Riparian vegetation inputs locally. Biofilm uptake of FPOM & DOM from detrital sources & decomposition of riparian mat'l.
Potential modifiers of processes:	Flow rates, elevated sediment loads (aggrading/degrading), bank erosion (additional sediment, nutrient & organic mat'l load)	Flow rates, high SS levels (nutrient adsorption/limitation, light limitation), pH (chemical equilibrium), elevated TDS (flocculation SS, toxicity), low temperature (suppress biological rates), nutrient point source discharge (elevated primary prod'n), consumption organic prod'n by grazers & collectors.	Time since event, flow rates, pH (chemical equilibrium), elevated TDS (flocculation SS, toxicity), low temperature (suppress biological rates), nutrient point source discharge (elevated primary prod'n), consumption organic prod'n by grazers & collectors.
Physico-chemical outcomes of bio-geochemical processes (riffle zones or reaches)	Sediment & SS levels, organic mat'l levels, nutrient (adsorbed) levels	SS, DO, pH, residual nutrients & composition, residual FPOM & DOM	SS, DO, pH, residual nutrients & composition, residual FPOM & DOM

References

ANZECC (1992) *Australian Water Quality Guidelines for Fresh and Marine Waters*. National Water Quality Management Strategy paper no. 4. Australian and New Zealand Environment and Conservation Council, Canberra.

ANZECC & ARMCANZ (2000) *Australian and New Zealand Guidelines for Fresh and Marine Water Quality*. National Water Quality Management Strategy paper no. 4 (revised). Australian and New Zealand Environment and Conservation Council & Agriculture and Resource Management Council of Australia and New Zealand, Canberra.

Armitage G (1995) An analysis of the bio-chemical oxidation of carbonaceous materials. Hons Thesis, University of Canberra.

Esslemont G (2000) Plant litter input to riparian systems: Its composition, reactivity & role in sediment nutrient dynamics. CRCFE Project Report.

Fairweather PG and Napier GM (1998) Environmental Indicators for National State of the Environment Reporting: Inland Waters. Environment Australia.

Harper MP (2001) SNAPP: A sorption-based model of phosphorus and nitrogen dynamics in freshwater sediments. *Ecological Modelling*

Hart B, Bourgues S, Grace M, Harper M & Cole B (2000) Nutrient release from river sediments: Phase II — Validation and application of a sediment-release model. *Land & Water Australia*.

Lawrence I, Bormans M, Oliver R, Ransom G, Sherman B, Ford P and Wasson R (2000) Physical and nutrient factors controlling algal succession and biomass in Burrinjuck Reservoir, Land & Water Resources Australia.