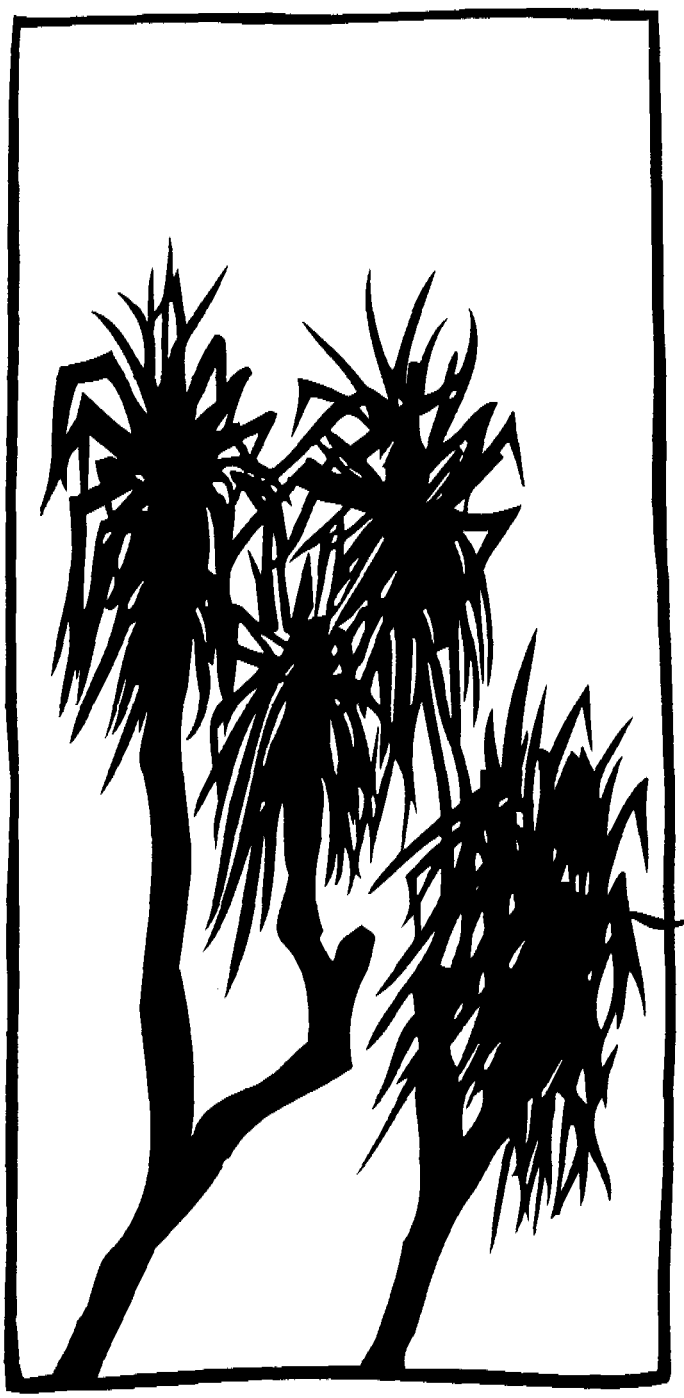
internal report **247**



LWRRDC Milestone Report 2: Development and implementation of QA/QC protocols for sampling and sorting components of the MRHI agency bioassessment program.

Chris Humphrey
Andrew Storey
Lisa Thurtell

supervising scientist

LWRRDC Milestone Report

LWRRDC project reference no.: ARR2

Project title: Development and implementation of QA/QC protocols for sampling and sorting components of the MRHI agency bioassessment program

Principal investigators: Dr Chris Humphrey, ERISS & Dr A Storey, Univ WA

Project duration: 1 October 1995 to 30 June 1997

Due date for milestone report: 31 May 1996

Project objectives

- Provide a literature review (and summary of relevant opinion) on Quality Assurance/ Quality Control (QA/QC) methodology and criteria for sampling and sorting of macroinvertebrate samples.
- Identify and recommend components of agency sampling and sorting protocols for internal and external QA/QC audit.
- Recommend interim QA/QC methods and acceptance/ rejection criteria for internal and external audit.
- Implement internal and external QA/QC programs.
- Evaluate and adopt suitable laborat
- Undertake R&D to refine acceptan
- Recommend protocols and criteri 'QC of ongoing and future National/ State monitoring program
- Assess and report on agency QA/Q ttee.

Report IR247

Alteration to original objectives: N

Milestones and achievement criteria

Milestone 1, 31 May 1996

a) Complete literature review and consultations with personnel involved with MRHI.

b) Have: recommended components of agency sampling/sorting protocols for internal/external auditing; designed, initiated and coordinated internal audits; selected interim QA/QC acceptance criteria; and commenced implementation of external QA/QC for each State and Territory MRHI lead agency.

c) Evaluate and select appropriate subsampling methods for ongoing QA/QC.

d) Complete R&D to refine acceptance criteria for QA/QC performance.

e) Report results of commissioned R&D and external auditing to NRHP annual workshop.

f) Provide progressive feedback to agencies on their performance for the 3rd MRHI sampling

Achievement criteria

a) Literature review, summary of consultations and recommendations on internal and external QA/QC protocols;

b) Progress results of subsampling evaluation (R&D);

c) Progress results of R&D for refining QA/QC acceptance criteria;

d) Progressive QA/QC performance of State/ Territory agencies for their 3rd sampling.

Achievement of milestone criteria 1

Literature review, summary of consultations and recommendations on internal and external QA/QC protocols

Attachment 1 contains a report entitled: 'Quality assurance/ quality control in rapid bioassessment projects with preliminary guidelines for implementation in the Australian Monitoring River Health Initiative' (Storey & Humphrey, 1996a). In this report, a number of objectives, in common with those in milestone criteria 1, are met, namely:

- A summary of QA/QC programs applied to overseas rapid bioassessment protocols and a review of methods/approaches used in these programs.
- Identification of aspects of the MRHI protocol that should be subjected to QA/QC.
- Recommendation of basic approaches, designs and analyses to be applied by agencies when implementing QA/QC programs as part of the MRHI.
- Selection of interim criteria (data quality objectives) for acceptance/rejection of QA/QC conditions as applied in overseas studies, and a summary of remedial action recommended in the event of non-compliance.
- A summary of QA/QC programs currently being undertaken as part of the Australian Monitoring River Health Initiative.

Note that we regard this report as interim and preliminary only. Thus, whilst extensive consultations were carried out with MRHI agencies in compiling this report, it remains now to distribute its contents to agencies and, if possible, to relevant experts in Australia and overseas seeking their comment for improvement. Naturally, QA/QC for the MRHI will evolve with time, particularly as empirical data from related R&D projects and external QA/QC come to hand. Hence, it is our desire to have a continuing role in devising and documenting QA/QC for the MRHI and for the current document to form the basis of formalised protocols to this end. In the short-term (to project completion in June 1997), we will update the document and submit revised versions with ensuing milestone reports.

Achievement of milestone criteria 2

Progress results of subsampling evaluation (R&D)

Attachment 2 contains a report entitled: 'Assessment of the efficiency of four types of device for subsampling of aquatic macroinvertebrate samples: preliminary results' (Storey & Humphrey, 1996b). This report presents a comparative statistical evaluation of results to date from processing one of three 'mega-samples', of known macroinvertebrate composition, through four different subsampling devices. The work aims to investigate sources of error in subsampling devices used by State and Territory (S/T) agencies that preserve samples in the field for subsequent laboratory sorting. The design of the study is a statistically rigorous one with results from the single mega-sample processed to date showing high precision and accuracy of each of the devices in characterising community composition and structure. Some of these results have been used to set interim criteria for acceptance/rejection of QA/QC conducted on subsampling and sorting

components of the MRHI protocol. Processing of, and reporting of results from, the remaining 2 mega-samples will be completed by December 1997.

Achievement of milestone criteria 3

Progress results of R&D for refining QA/QC acceptance criteria

For most components of the MRHI protocol, interim criteria upon which to assess the performance of S/T agencies are based on some empirical data, namely results from associated R&D evaluating the efficiency of laboratory subsampling devices. These results, in particular, have been used to set criteria to apply to field and laboratory sorting procedures. Nevertheless, it is possible that the interim QA/QC criteria are overly-conservative and unnecessarily stringent. Adoption of such criteria in the Australian MRHI could result in a high failure rate that is both disillusioning to S/T agencies and, more importantly, unnecessary insofar as improving the quality of, or altering conclusions drawn from, model output.

Attachment 3 contains a report entitled: 'Proposed analytical approach to assess the effect of different QA/QC criteria for sampling efficiency on the robustness of MRHI models' (Storey & Humphrey, 1996c). The report describes R&D designed to better define QA/QC criteria for key aspects of the MRHI protocol. Acceptance criteria will be refined for these aspects by way of (i) classification and model outcome (observed/expected) results after error simulations performed on existing agency data sets, and (ii) classification and o/e results following incorporation of "corrected" agency data (as derived from external QA/QC). Thus, through existing agency data sets and models, thresholds will be sought at which the error rate begins to compromise the robustness, accuracy and predictive capability of models. Results can then be used to ensure that the error rate in the data acquisition phase remains within identified limits.

Contents of the report in Attachment 3 were distributed to members of the MRHI Technical Advisory Committee for comment. Hence, the report has undergone extensive revision and now describes methodology that has been agreed upon by peers. This R&D will be conducted and reported by December 1996.

Achievement of milestone criteria 4

Progressive QA/QC performance of State/ Territory agencies for their 3rd sampling

Apart from R&D, the current project has an implementation phase of external QA/QC auditing of agency sorting procedures. The report in Attachment 4: 'External QA/QC of MRHI agency subsampling and sorting procedures: Results to 30 May 1996' (Thurtell & Humphrey, 1996) describes progress in this phase of the project.

It was the original intention in the project proposal that samples used for both sorting and identification QA/QC be shared to reduce costs. It became apparent, however, that to meld the QA/QC in this way would compromise the integrity of the identification QA/QC (i.e. agencies would have very advanced warning of samples required for identification QA/QC) and a decision was made to decouple both QA/QC projects. Whilst this slightly reduces the number of samples that the sorting QA/QC worker has been able to process (i.e. 3% of samples as opposed to 5%), the decision to separate the work otherwise did not affect the LWRRDC milestone reporting and hence it was not necessary to inform LWRRDC of this decision at the time.

The external QA/QC study aims to cross-check community structure reported by agencies (from live-sorting or laboratory subsampling) against that representative of the whole sample. At the onset, it was recognised that most of the error arising in agency sorting procedures would arise in live sorting - as opposed to laboratory subsampling and sorting of preserved samples. Consequently, priority was given to processing live-sorted samples so that initial results reported in Attachment 4 pertain only to these samples; results arising from QA/QC of laboratory-sorted subsamples will be reported at a later date.

The conclusions drawn from initial QA/QC results of agency live-sorted samples are discouraging insofar as the viability of this sorting procedure and quality of results for MRHI modelling are concerned. The results were presented to members of the Steering Committee for the Norris modelling project on July 1, 1996. Members agreed that the significance of the seemingly poor results (high failure rates against interim criteria) to the accuracy of developing agency models would not be known until associated R&D to refine acceptance criteria had been conducted - as described in Attachment 3. Even so, liberal 'worst-case' criteria were used in the assessment of relative/rank abundance data where high failure rates were observed. This would suggest that, with the possible exception of riffle habitat, the potential has been lost for development of models based on relative/rank abundance for those agencies using live-sorting procedures. This is of some concern given that such models could potentially be more sensitive to detection of human disturbance in streams. Even for ongoing and future development of presence/absence models, we have advised against further use of the existing live-sorting protocol by agencies without substantial modification (see Attachment 4).

At this stage, state agencies have not been advised of the results of external QA/QC. Seemingly poor results in live-sorting are endemic amongst agencies and until the significance of these results is assessed through associated R&D and through comparison with QA/QC conducted for those agencies sorting laboratory subsamples, there seems little point in advising agency staff of results. In any case, accompanying such advise would need to be the decision on if, how and when sorting protocols were to be revised. The advice of the national NRHP coordinator will be sought on this matter.

Additional comments on results of external QA/QC and implications are made in the 'Summary' section below.

Variations required to future milestones

Much of the R&D for the project was not completed for the current Milestone report as had been anticipated originally. (The LWRRDC coordinator for the NRHP program was made aware of this and accepted a revised reporting schedule accordingly.) As a consequence, additional reporting of these aspects of the project will be required in future Milestone reports (December 1996 and June 1997). Milestones and achievement criteria will need to be altered to reflect these changes. In addition, external QA/QC for agency samples derived from rounds 3 and 4 will not be completed by December 1996 as appointment of the professional officer to conduct this work was not made until February 1996 (as opposed to December 1995 as originally proposed). The advice of the LWRRDC will be sought in altering future milestones.

Financial issues

In March 1996, the NRHP committee approved a request from the ERISS for additional funds to complete R&D on assessing the efficiency of subsampling devices used by MRHI agencies (described in Attachment 2). An additional $6,000 will be provided to ERISS to conduct this work.

Human resource issues

Human resource issues are linked to 'Financial issues' raised above and in the 'Summary'.

Communication achievements

So far, all communication has centred on NRHP technical advisory and steering committee meetings. Progress results will be reported at the NRHP workshop to be held in October 1996.

Listing of attachments

Attachment 1

Storey, AW & Humphrey, CL (1996a). Quality assurance/ quality control in rapid bioassessment projects with preliminary guidelines for implementation in the Australian Monitoring River Health Initiative.

Attachment 2

Storey, AW & Humphrey, CL (1996b). Assessment of the efficiency of four types of device for subsampling of aquatic macroinvertebrate samples: preliminary results.

Attachment 3

Storey, AW & Humphrey, CL (1996c). Proposed analytical approach to assess the effect of different QA/QC criteria for sampling efficiency on the robustness of MRHI models.

Attachment 4

Thurtell, L & Humphrey, C (1996). External QA/QC of MRHI agency subsampling and sorting procedures: Results to 30 May 1996.

Other comments

See 'Summary'.

Summary

A comprehensive report on 'Quality assurance/ quality control in rapid bioassessment projects with preliminary guidelines for implementation in the Australian MRHI' has been completed and submitted as an attachment. Further revisions of this document will be provided to LWRRDC in ensuing Milestone reports. Preliminary results on an 'Assessment of the efficiency of four types of device for subsampling of aquatic macroinvertebrate samples' has also been submitted as an attachment to this report. This subsampling R&D is progressing satisfactorily.

Of most potential concern to the MRHI is the apparent poor representativeness of invertebrate community structure arising in agency live-sorted samples. This is the preliminary assessment of work arising from external QA/QC of agency sorting procedures (Attachment 4). It is not possible at this stage to assess the significance of these results as they might affect the accuracy of developing agency models until associated R&D to refine QA/QC acceptance criteria has been undertaken (described in Attachment 3). This R&D includes (i) classification and model outcome (o/e) results after error simulations performed on existing agency data sets, and (ii) classification and o/e results following incorporation of "corrected" agency data that is derived from external QA/QC.

Most members of the MRHI Technical Advisory Committee have been made aware of preliminary results arising from external QA/QC and the R&D required to assess the significance of the results. It has been generally agreed by TAC members that the R&D approach that would best and most definitively address this issue is that arising from (ii) above, i.e. assessing the effect upon existing agency classifications and model outcomes when 'corrected' data (derived from external QA/QC) are seeded in agency data sets and modelling conducted using these altered data sets. In order to expedite this R&D, it will be necessary to provide additional funds towards the conduct of external QA/QC; as reported above, existing staff resources are available only to process relatively few of the available agency residues and these resources must also cover QA/QC of preserved subsamples. For this evaluation, there is the possibility of processing live-sort residues from all 50 samples gathered by one WA sub-agency (Murdoch University) in round 1, as well as the additional residues that are available from other agencies. The advice of the NRHP coordinator

will be sought on this matter and a number of scenarios provided to him on additional samples available, and funds required, for this work.

10 Milestones

Date **Description of Stage**

30 May 1996 Milestone 1

(a) Complete literature review and consultations with personnel involved with MRHI.

(b) Have: recommended components of agency sampling/sorting protocols for internal/external auditing, designed, initiated and coordinated internal audits, selected interim QA/QC acceptance criteria, and commenced implementation of external QA/QC.

(c) Evaluate and select appropriate subsampling method for ongoing QA/QC.

(d) Complete R&D to refine acceptance criteria for QA/QC performance.

(e) Report results of commissioned R&D and external auditing to NRHP annual workshop.

(f) Provide progressive feedback to agencies on their performance for the 3rd MRHI sampling.

ACHIEVEMENT CRITERIA:

Submission of milestone report to LWRRDC containing:

(i) Literature review, summary of consultations and recommendations on internal and external QA/QC protocols;

(ii) Results of subsampling evaluation (R&D);

(iii) Results of R&D for refining QA/QC acceptance criteria;

(iv) Progressive QA/QC performance of State/ Territory agencies for their 3rd sampling.

31 December 1996 Milestone 2

(a) Complete external QA/QC of agencies' 3rd and 4th samplings.

(b) Provide feedback to agencies on their performance over the 3rd and 4th samplings.

ACHIEVEMENT CRITERIA:

Submission of milestone report to LWRRDC detailing QA/QC performance of State/ Territory agencies.

30 June 1997 Milestone 3

Complete all reporting to agencies and NRHP committee on external QA/QC and associated R&D.

ACHIEVEMENT CRITERIA:

Submission of milestone report to LWRRDC detailing the proficiency of each agency throughout the project and providing future directions for QA/QC of biological monitoring programs in Australian streams using macroinvertebrates.

ATTACHMENT 1

Costs for Phase A, Project Development

- Collaborative consultant: 27 days @ $500/ day	$13,500
- Costs for consultant to visit ERISS for 1 week	2,000
- Technical officer (subsampling R&D): 40 days @ $163/ day	6,500
- Consultant (R&D for acceptance criteria): 28 days @ $600/ day	16,900
- Attendance of consultant at annual NRHP workshop	2,000
- Loan/ hire/ construction of subsampling equipment, incidentals	2,000

Subtotal = $42,900

Costs for Phase B, Project Implementation

- Professional/ science officer (for QA/QC audit of 2 remaining agency samplings): one year @ $40,000/ year	$40,000
- Attendance of PO at annual NRHP workshop	2,000

Subtotal = $42,000

Grand total = $84,900

ATTACHMENT 1

Quality Assurance/Quality Control in Rapid Bioassessment Projects with Preliminary Guidelines for Implementation in the Australian Monitoring River Health Initiative

A.W. Storey[1] & C.L. Humphrey[2]

Report prepared for the Land and Water Resources Research and Development Corporation

[1] Department of Zoology, The University of Western Australia, Nedlands, W.A. 6907.
[2] Environmental Research Institute of the Supervising Scientist, Private Mail Bag 2, Jabiru, N.T. 0886

TABLE OF CONTENTS

INTRODUCTION

Growing concern for the ecological state of Australia's wetlands, particularly following toxic algal blooms in the Murray-Darling river systems, combined with the inclusion of biological indicators in water quality guidelines (ANZECC, 1992) has seen a rapid increase, in recent years, in the extent and use of biological assessment of water quality in Australia. A major stimulus for this increase has been the provision of funds by the Australian Government for the development of the 'National River Health Programme' (NRHP). Part of this programme is the 'Monitoring River Health Initiative' (MRHI), involving government agencies from all Australian States and Territories in a national program to develop a standardised and coordinated rapid bioassessment approach to biological monitoring of water quality in Australian streams and rivers.

The success of such an extensive program will depend on the development, acceptance and implementation of standard protocols encompassing all aspects of data acquisition (Davies, 1994). Even though standard protocols have been adopted across all States and Territories (S/Ts), there may be differences in the way these protocols are interpreted and implemented by different personnel/organisations. Such variations may affect the integrity of the data gathered. Therefore, there is a need for continual review of the quality of the data being collected. The accepted approach for reviewing data quality is to implement an ongoing Quality Assurance/Quality Control (QA/QC) program.

QA/QC is recognised as an essential component of any large project involving many different parties. Its purpose is to ensure that methods for data collection are standardised, that data are of a consistent and high quality and that this quality is maintained throughout the project. Plafkin *et al.* (1989) defined "quality control" as the routine application of procedures for obtaining prescribed standards of performance in the monitoring and measurement process, and "quality assurance" as the incorporation of the quality control functions in a totally integrated program designed to ensure the reliability of monitoring and measurement data. They continued by stating that "quality assurance and control (QA/QC) should be a continuous process implemented throughout the entire bioassessment program. All aspects of the study, including field collection, habitat assessment, laboratory processing, and data analysis are subject to QA/QC procedures". Plafkin *et al.* (*op. cit.*) considered that quality assurance was accomplished through the establishment of thorough investigator training, protocol guidelines, comprehensive documentation and management of field and laboratory data, verification of data reproducibility, and instrument calibration.

General aspects of rapid bioassessment protocols that require QA/QC, as identified by Plafkin *et al.* (*op. cit.*), include: training of all personnel in a consistent manner so that assessments are conducted in a standardised manner; collection of replicate samples at selected sites to check on the accuracy of the collection effort; resampling of selected sites by different teams and comparison and assessment of the reproducibility of results; assessment of the accuracy of taxonomic identifications; standardisation of field data sheets and abbreviations used on data sheets to reduce errors; filing of field and laboratory data sheets and final reports; appropriate training of personnel in any technique for the subjective assessment of habitat with periodic cross-checks conducted among investigators to promote consistency; and, regular calibration of instruments used for measuring physico-chemical parameters with known standards and recording of instrument type and serial number against all field measurements.

Given the broad scale and objectives of the MRHI, and the similarity of the rapid bioassessment protocols to those used overseas, the development and implementation of QA/QC programs were seen as logical and necessary components of the Australian program. However, before a QA/QC program could be implemented it was necessary to:

a) identify specific aspects of the data acquisition phase (protocols) that require auditing,
b) develop approaches and methods to use in QA/QC programs,
c) select performance criteria on which to base decisions as to the acceptance or rejection of data quality, and where possible,
d) give advice upon appropriate remedial action to take in the event of rejections/sub-quality data.

In addressing these issues this report will:

- Summarise QA/QC programs applied to overseas rapid bioassessment protocols and review methods/approaches utilised in these programs.
- Identify aspects of the MRHI protocols that should be subjected to QA/QC.
- Recommend basic approaches, designs and analyses to be applied by agencies when implementing QA/QC programs as part of the MRHI.
- Select interim criteria (data quality objectives) for acceptance/rejection of QA/QC conditions as applied in overseas studies, and summarise remedial action recommended in the event of non-compliance.
- Summarise QA/QC programs currently being undertaken as part of the Australian Monitoring River Health Initiative.

APPROACHES FOR QA/QC ADOPTED IN OVERSEAS PROGRAMS

There are two major overseas freshwater biological monitoring programs in place overseas which have been subjected to QA/QC: firstly, the British National Water Quality Survey, using the RIVPACS models and associated sampling methodology (Wright, 1995), for which QA/QC has been described by van Dijk (1994), and secondly, the US Geological Survey's (USGS) National Water-Quality Assessment (NAWQA) Program, using rapid bioassessment protocols (Plafkin *et al.*, 1989; Cuffney *et al.* 1993a (cited Cuffney *et al.* 1993b)), for which QA/QC guidelines have been detailed by Cuffney *et al.* (1993b). Because the assessment protocols used in each study differ, it is necessary to provide a general description of the protocols before reviewing the QA/QC procedures utilised by each study.

The British National River Quality Survey and QA/QC procedures

Description of program and protocols

The National River Quality Survey (NRQS) is a quinquennial survey of the chemical and biological quality of rivers throughout Great Britain. In 1990, it was decided to incorporated into the NRQS biological assessment of the quality of rivers, utilising the **R**iver **I**nvertebrate **P**rediction and **C**lassification **S**ystem (RIVPACS) (Wright, 1995).

Development of RIVPACS commenced in 1977 at the River Laboratory of the Institute of Freshwater Ecology (IFE) (previously the Freshwater Biological Association (FBA)). Since then, the system has evolved through a number of stages incorporating expansions and modifications to statistical/analytical methods. Originally, the system was based upon the sampling of an initial set of 268 sites of high biological quality. The aims of the study were to a.) develop a biological classification of unpolluted running-water sites in Great Britain based on macroinvertebrate fauna, and b.) assess whether the type of macroinvertebrate community at a site could be predicted using physical and chemical features (Wright *et al.*, 1984). The approach required a simple but flexible sampling procedure that could be applied to a wide range of streams over the full range of habitat types. For this purpose, a kick and sweep technique (Furse *et al.*, 1981) with a standard FBA pond net (900 μm mesh, 230 x 255 mm frame, 275 mm bag depth and 1.5 m handle) was used to sample all major habitats at a site in proportion to their occurrence. Sampling was carried-out for 3 minutes, with the intention of gaining as comprehensive a species list as possible, and was repeated in spring, summer and autumn to obtain representatives of all elements of the fauna. At deep riverine sites, a light-weight version of the medium naturalist dredge was used, and this was complemented with pond netting along the margins (Wright, 1995). Sites were located at 5, 10, 20, 30 and 40 km from the source of the river and then at 20 km intervals thereafter along the remaining undisturbed reaches of the river. All biological sampling was conducted by water industry biologists under detailed guidelines set by the IFE.

Once collected, samples were preserved in the field in formaldehyde and sent to the River Laboratory for sorting and identifications. On each visit, the biologists completed a comprehensive form detailing environmental features of the river, riparian zone and adjacent catchment. Chemical data for the site (or nearby locations) were acquired from water industry chemists (Wright, 1995).

All sorting and identifications were performed by IFE staff based at the River Laboratory. Each sample was examined for approximately 2 h by placing aliquots in a white tray and sorting through the material by eye. Large numbers of specimens were removed to ensure that most, if not all, species present in the sample were available for identification (Wright *et al.*, 1984). "A consistently high level of taxonomic resolution was achieved by the use of highly trained permanent staff" (Wright, 1995). Where adequate keys were available, identifications were made to species level, and in cases where adequate keys were unavailable, or taxonomy was problematic, standard protocols were adopted to ensure amongst-site comparability (see Furse *et al.*, 1981; Wright *et al.*, 1984).

Following development and testing of the initial RIVPACS, more sites were added to increase geographical coverage (n=370) and a new procedure was developed and incorporated for prediction of the probability of taxon occurrence. In the mid-1980s, biological indices based on the Biological Monitoring Working Party (BMWP) system were incorporated into the prediction system (*viz.* allocation of scores in the range 1-10 to selected families, according to their known tolerance to organic enrichment), and a microcomputer version of the package (RIVPACS I) was developed and released for testing by water industry biologists. Following RIVPACS I, the data set was again enlarged to include deep lowland rivers and some smaller streams, and a new classification and prediction system, based on 438 sites in 25 classification groups, was developed (basis of RIVPACS II). Between 1989 and 1990 an operational version of

RIVPACS II was developed for use on PC and mainframe computer in time for the package to be incorporated in the 1990 NRQS, conducted throughout the UK (Wright, 1995). The results of the 1990 NRQS have since been summarised and the RIVPACS II has been comprehensively tested by the IFE, using an independent data set.

Subsequently to RIVPACS II, the data set has been further increased (~ 700) to provide a more comprehensive geographical coverage, including reference sites in Northern Ireland and additional small streams throughout Britain. The procedures for enhancing the robustness of the system have been further evaluated, alternative methods of classification and prediction have been examined and these changes assimilated into RIVPACS III for use in the 1995 NRQS (Wright, 1995).

While predictions based on species-level identifications were preferred, the water industry required a technique that allowed rapid site appraisal at low cost. Such an approach was subsequently incorporated in RIVPACS I by downgrading predictions from species to BMWP family level and then producing predictions of BMWP score, number of taxa expected and the average score per taxon (ASPT, determined by dividing the total BMWP score at a site by the number of scoring taxa). These indices proved particularly useful for rapid site assessment because the ratio of the observed value for a site divided by the expected, as predicted by RIVPACS, provided an indication of whether the site was under stress. An unstressed site would have an O/E ratio close to unity and the ratio would decrease as environmental stress adversely affected community composition and taxa richness. A summary of the development of the RIVPACS approach and the details of the statistical methods for classification and prediction are reviewed by Wright (1995).

QA/QC Procedures

Initial QA/QC procedures for biological assessment using RIVPACS were largely 'in-house' checks conducted by IFE staff at the River Laboratory, with no published prescriptions. Relevant publications allude to some of these controls. For instance, sampling was conducted by water industry biologists 'following detailed guidelines set by the IFE' (Wright, 1995) and, Wright *et al.* (1984) comment that during the sample sorting phase large numbers of specimens were removed to ensure that most, if not all species present in the sample were available for identification. Similarly, Wright (1995) notes that a consistently high level of taxonomic resolution was achieved by the use of 'highly trained permanent staff', and two substantial checking procedures (undescribed in the literature) were invoked to ensure both accuracy of identifications and consistency in level of identification (Wright *et al.*, 1984). These types of checks reflect the developmental nature of RIVPACS during the late 1970s and early 1980s, whereby errors were minimised through the use of a small group of highly qualified and dedicated staff. Once the package was made available for the independent use by water industry biologists, it was seen necessary to standardise methodologies in order to reduce errors. This was achieved through training of water industry staff.

Staff of the IFE assisted the National Rivers Authority (NRA, now part of The Environment Agency, a non-departmental public body established by the U.K. Environment Act, 1995) in making a video for presentation to water industry biologists. The training video was designed to help standardise protocols for biological sampling and for recording physical features of habitat for use in RIVPACS predictions. The video was shown to staff before RIVPACS II was incorporated in the 1990 NRQS (Wright, 1995).

However, there still had been no formal development of QA/QC for any stage of the protocols.

In 1990, Kinley & Ellis (1991, cited van Dijk, 1994) reported results of an IFE audit of macroinvertebrate sampling, which examined the frequency of missed taxa. This audit was repeated in 1991 and 1992 (van Dijk, 1994), but it was not until 1994 that a procedural manual describing a proposed Analytical Quality Control (AQC) system to analyse freshwater macroinvertebrate samples to BMWP level was published (van Dijk, 1994). The procedure was designed for application in all NRA biological laboratories, with the manual proposing that quality be assessed by AQC inspectors, who would be fully qualified biologists from either within the laboratory or the region. The inspectors would perform a second processing and analysis of a sample and record any taxa missed or incorrectly identified by the original analyst. Hence, the quality of the primary analysis of the sample was defined by the number of missed taxa identified in the secondary analysis. The aim of the program was to ensure that the average number of missed taxa over all analysts within a laboratory remained at or above an acceptable limit, with unacceptable quality being detected promptly to allow remedial action.

The proposed AQC program takes a statistical approach to assessing error rates to allow for the inherent variability in the effectiveness with which samples are processed. The actual number of missed taxa found in successive samples is known to vary about an average value, referred to as the **process average**. Kinley & Ellis (*op. cit.*) reported that the variability of missed taxa recorded in the IFE audit in 1990 approximated to a Poisson distribution. This was reconfirmed in subsequent audits in 1991 and 1992 (van Dijk, 1994), and was adopted as an underlying assumption for the AQC program (i.e. the Poisson distribution adequately described the variability of the observed numbers of missed taxa when the process average is constant). van Dijk's (1994) approach and terminology for the AQC program have been adapted from quality control applied to industrial processes. The AQC program requires definition of an **Acceptable Quality Level (AQL)**, which is the worst (poorest quality) process average that is still acceptable, and a **Rejectable Quality Level (RQL)**, which is the process average at or below which results are regarded as unacceptable and require remedial action. If the process average is as good or better than the AQL, then no alarms will be triggered, however, if the process average falls to a level that is at or below the RQL, then an alarm should sound, followed by the necessary remedial action. Given the inherent variability in sample processing, there will be occasional false alarms when the average quality is better than the AQL, and conversely, there will be some occasional delays in sounding the alarm when average quality is worse than the RQL. To prevent the repetitive sounding of false alarms or the absence of an alarm when the process average is worse than the RQL, the AQC takes into account the **Average Run Length (ARL)**, which is the average number of samples processed between alarms for a given process average. Ideally, the ARL will be long at process averages better than the AQL, and conversely, short when the process average is below the RQL.

To achieve the objective of acceptable processing quality, with appropriately-timed alarms, the AQC program utilises the cumulative sum or 'Cusum' technique combined with a Decision Interval Approach. Cusum is defined as the cumulative sum of successive differences of observed results from a fixed reference value, and basically means that as each result becomes available, a given quantity (known as the **Reference Value (R)**) is subtracted from the result and the differences accumulated. R may be any value but is

often an assumed average or target value lying between the AQL and RQL and intended to provide good discrimination between process averages at these levels (van Dijk, 1994). These values may be subjectively assigned or empirically derived from an established database with known error rates (*sensu* van Dijk (1994) using results of IFE audits of water authority sample processing). If observed results are running close to R, then successive differences will be positive or negative in roughly equal numbers and the sum of the differences will not stray far from zero. However, if the observed results are repeatedly higher than R then the cusum will increase over time. Conversely, a decrease will occur if the observed results are repeatedly below R. Cusums may be used to detect changes in the average level of a measure, determining the point of onset of those changes, as well as obtaining a reliable estimate of the current average value of the measure (van Dijk, 1994). A decision process, called the **Decision Interval (D)** must be combined with this technique to determine at what point performance is unacceptable. The value for D is chosen to achieve desirable performance in terms of ARL at the two targets, the AQL and the RQL (van Dijk, 1994).

van Dijk (1994) summarises application of the AQC using the cusum technique with the Decision Interval Approach in six steps:

1. R is selected, checking of samples commences and the cusum is set to zero
2. When a value greater than R is recorded, the cusum is set to the difference between R and the observed value.
3. With successive samples, the cusum continues to accumulate for as long as the observed values in a sampling run continue to exceed R.
4. If the observed values are less than R and the cusum returns to zero then the process stops and the operator returns to Step 1.
5. If the cusum continues to accumulate and exceeds the value of D then an alarm sounds.
6. After an alarm signals, remedial action is taken, and the process recommences at Step 1 to assess the effectiveness of the remedial action.

van Dijk (1994) notes that the values for R and D will depend on the purpose of the investigation and the quality targets required (i.e. the AQL and associated ARL); both values will be set by management, based on scientific and management grounds. A collection of AQC schemes, with three nominal ARLs of 50, 100 and 200 samples, AQL values ranging from 1 to 5 missed taxa and associated values of R and D are presented and 'it is envisaged that the collection of schemes should cover most of the NRA's requirements' (van Dijk, 1994).

In the application of this approach to sorting and identifying freshwater macroinvertebrates, van Dijk (1994) highlighted the need to select an AQC controller, AQC inspectors, suitable forms for recording data and a random sampling procedure. The controller would conduct the scheme, selecting the test samples by way of a random selection procedure, providing the sample to the inspector (sample suitably identified to keep the original laboratory processor anonymous), updating the Cusum record, informing management of any transgressions and controlling paperwork. The inspector would be fully trained, qualified and experienced in detecting and identifying all taxa likely to be encountered by the target laboratory. The inspector would be provided with a list of taxa and a vial containing identified specimens together with the original sample from which the specimens were derived. Adequate time will then be allocated to finding and identifying any missed taxa. Using a standard form, the inspector reports results of the secondary analysis to the controller, who updates the Cusum record for the relevant

laboratory. Details of the method required to calculate 95% c.i. about the process average are also presented (van Dijk, 1994).

In the event that an alarm signals, van Dijk (1994) recommends various possible courses of remedial action, although the final decision will rest with management. Reference to the types of taxa being missed (i.e. difficult families mis-identified by all analysts), the state of the sample (i.e. large quantities of detritus) and the analyst that conducted the primary analysis (i.e. inexperienced in all or some aspects of taxonomy) will assist in making a decision. Options on how to treat data collected during the deterioration in quality are also discussed, namely: 1) scrap some or all the results, 2) rework some or all of the samples, 3) accept the results with the proviso that their quality is low. Finally, van Dijk (1994) recommended that the AQC inspectors should be audited occasionally by an external body, such as the IFE, to assess their quality.

The USGS NAWQA Program and QA/QC procedures

Description of program and protocols

The USGS NAWQA program is a perennial program designed to produce a comprehensive, multi-faceted assessment of the quality of the nation's flowing-water resources. The program targets 60 study units (coupled ground- and surface-water systems) located in the USA, including Alaska and Hawaii. Investigations within each unit use consistent national guidelines for selecting sampling sites and collecting physical, chemical and biological data. This consistency allows an integrated assessment of the status and trends in the nation's water quality (Cuffney *et al.* 1993b).

The biological component of the NAWQA program comprises mainly ecological surveys of stream habitat and community characterisations of benthic invertebrates, as well as fish and algae. As part of the program, nationally consistent guidelines have been developed to ensure that the study units collect comparable data. These guidelines call for the processing of samples by contract laboratories that are responsible for all phases of sample processing, such as identifying and quantifying benthic invertebrates. National consistency and standardisation in the processing of samples, particularly for those constituents processed entirely by contract laboratories was deemed essential. Therefore, nationally-consistent guidelines and criteria, including QA/QC procedures, were developed for the processing of benthic invertebrate samples collected as part of the USGS's NAWQA program (Cuffney *et al.* 1993b).

The US NAWQA program involves personnel from three groups: i) the study unit which collects, pre-processes and supplies the biological samples to the contract laboratory; ii) the contract laboratory which receives and processes the samples and returns data to the study unit for analysis; and iii) the Biological Quality Assurance Unit (BQAU) of the USGS which evaluates QA/QC issues of the program.

A study unit will typically collect three samples at a site - two semi-quantitative samples, one each from "richest-targeted habitats" (RTH) and "depositional-targeted habitats" (DTH), and a qualitative multi-habitat sample (QMH) being a composite sample from as many habitats as can be practicably sampled in the reach. Each sample is pre-processed in the field. Pre-processing involves three activities: a) visual inspections (at various

opportune times) to remove and preserve large specimens that are rare or may interfere with subsequent processing (**large-rare component**), b) elutriation to separate sand and gravel (**elutriate component**) from invertebrates and organic material (**main-body component**), and, c) subsampling (sample splitting) to reduce the volume of each main-body sample to $\leq$ 750 mL (**main-body and split components**) (Figure 1). The large-rare and main-body components are sent to the contract laboratory for processing (identification and enumeration) and the elutriate and split components are sent to the BQAU for QA/QC purposes. However, when a sample volume is small and field subsampling is not necessary (i.e. main-body sample volume less than 750 mL) there is no split sample available for QA/QC.

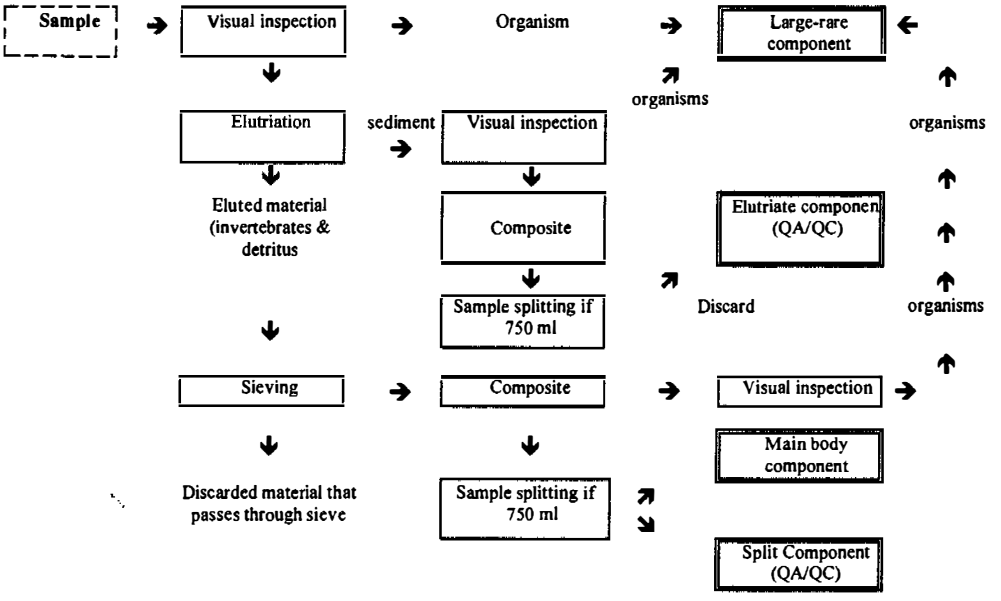


Figure 1. Field sample processing activities in the USGS's NAWQA Program (after Cuffney *et al.*, 1993b).

Field elutriation and sample splitting provide the contract laboratory with a main-body sample that is relatively consistent in volume and condition. Processing of these samples involves separation of invertebrates from organic and inorganic matter, and identification and enumeration of the organisms. Samples that contain large volumes of fine detritus or large numbers of small organisms usually invoke subsampling. The approved protocol is the division of the sample into quarters. This may be carried-out using a variety of approved methods (sample splitters, cone subsampler or a gridded tray) and the contract laboratory conducts its own QA/QC check for subsampling, as described below. Following removal of invertebrates from the detritus in the main-body component (or subsample fraction(s)), the sample remnant is returned to the BQAU for further QA/QC checking.

QA/QC Procedures

The USGS's NAWQA Program is underpinned by comprehensive QA/QC, addressing all major components of the protocols. Initially, 10 percent of elutriate and split-sample components are randomly selected and subjected to QA/QC checks. Thus:

Elutriate samples are processed within the BQAU to determine the number and identity of invertebrates in the gravel and sand. The QA/QC criteria for each sample elutriate are:

- the number of additional taxa and abundance of organisms in the elutriate must be $\leq$ 10% of the estimates for the whole sample (combined large-rare and main-body components).

If a sample fails this check, then the BQAU works with the relevant study unit to correct the elutriation problems through additional training and/or modifications of the elutriation procedure.

Split sample components are processed either by the BQAU or a contract laboratory (in the latter case the sample is presented as if it is an original main-body sample). The QA/QC criteria are:

- the number of taxa must lie within 10 percent of the estimate of the main-body component, and
- the index of community similarity (complement of the Bray-Curtis dissimilarity index) between the split and the main-body component must be at least 90 percent.

Subsampling QA/QC checks conducted by the contract laboratory are intended to ensure that the estimates of the number of taxa and proportion of each taxon obtained through subsampling meet minimum QA/QC requirements. After the main-body component is split into quarters, two quarters are selected at random, sorted, identified and enumerated. Criteria used to evaluate subsampling are:

- the number of taxa that the subsamples have in common must be at least 90 percent of the combined number of taxa, and
- the similarity between the community structure of the two subsamples must be at least 90 percent.

If the subsamples do not meet the criteria then an additional quarter is sorted and compared to the half already sorted. If the criteria still are not attained then the entire sample is processed. Hence, this method can reduce the volume of sample processed by 25 to 50 percent depending on the number of fractions that need to be sorted.

Sample remnants (either main-body or subsample fractions) are returned to the BQAU for QA/QC checking. Two criteria are used to evaluate the effectiveness of sorting:

- the number of new taxa missed during the sorting process must be $\leq$ 10 percent of the number of taxa on the original data sheet for that sample/subsample, and
- the total number of organisms found in the sample remnant must be $\leq$ 10 percent of the total number removed from the sample component.

Semi-quantitative samples are evaluated using both criteria, but qualitative samples use only the first criterion.

Identification and quantification QA/QC checks are intended to establish that invertebrates have been correctly identified and enumerated. For a specific sample selected for QA/QC checking, all invertebrate fractions pertaining to that sample that have been sorted and identified are re-identified and counted by a different taxonomist to the one that originally conducted the identifications. Three criteria are applied:

- at least 90 percent of the specimens must be correctly identified,
- the number of new taxa added to the taxa list must be < 10 percent of the original number of taxa, and
- the Bray-Curtis community similarity index comparing the original and the corrected data must be at least 90 percent.

The first two criteria are applicable to semi-quantitative and qualitative samples, whereas the last criterion is applicable only to semi-quantitative samples. Identification and enumeration QA/QC checks are performed on the large-rare and main-body components separately and then on the whole sample after the components have been combined. QA/QC checks on the large-rare and main-body components are based on the portions of the sample processed in the laboratory and do not include corrections for field or laboratory subsampling. The subsequent QA/QC checks on the whole sample are intended to represent the cumulative error arising on constituent components contributing to the whole sample and, as a consequence, whole-sample estimates are corrected for field and laboratory subsampling. Cuffney *et al.* (1993b) state, "it is not anticipated that a whole-sample QA/QC check would fail unless either or both of the QA/QC checks on the constituent subcomponents fail".

QA/QC checks on identifications, enumeration and sorting effectiveness are based on a block of ten or fewer samples, with 10 percent of samples being checked (i.e. one sample selected at random). If the selected sample fails a QA/QC criterion then the whole block of ten samples is re-processed and re-checked until the criteria are met. When re-checking is required, a different sample from the original is selected at random from the block of ten.

Sample volume is subjected to a QA/QC check to determine the volume of material remaining in subsamples compared to the original volume of the main-body sample component. The criterion is:

- the volume of material in unprocessed subsamples (i.e. 3rd & 4th quarters) combined with remnants (organic and inorganic material) in processed subsamples (1st & 2nd quarters) must sum to within 15 percent of the original main-body component volume.

If this criterion is not met then the laboratory must account for the difference, which may be due to the removal of large debris, such as sticks during processing. The check is intended to ensure that subsamples are not lost during laboratory processing, that any lost material is accounted for, and that all sample remnants are returned to the BQAU for review.

The QA/QC checks are summarised on a QA/QC summary check list. This provides an overview of the checks on identification and quantification, sorting effectiveness, and sample volume. The types of errors encountered during sample processing (i.e. processing errors such as the accidental spilling of part of a sample) and results of the QA/QC checks are entered whilst any remedial action taken is also recorded.

The USGS's NAWQA Program also calls for QA/QC on specimen reference or voucher collections maintained by study units. The collections are seen as permanent accounts of conditions within the study unit that can be compared with future collections from this and other study units. Identifications of all specimens must be confirmed by an independent taxonomist and this information is recorded on a database, along with the reference to the taxonomic keys used to identify the specimens.

QA/QC PROGRAMS BEING CONDUCTED UNDER THE MRHI

For the MRHI, QA/QC has been categorised into "external" and "internal" - the former being carried out by external agencies (i.e. Environmental Research Institute of the Supervising Scientist (ERISS) and the Murray-Darling Freshwater Research Institute (MDFRC)) and "internal" being the responsibility of the individual State and Territory agencies (hereafter S/T agencies). This section summarises the external and internal QA/QC programs being undertaken as part of the MRHI.

Separate internal and external QA/QC, even for the same aspect of the protocol being audited, is regarded as a critical component of rapid bioassessment programs (or indeed, any large project involving many different parties). Such internal and external checks of critical aspects of the monitoring protocol are also followed in the UK and the USA with rationale as follows: The independent external QA/QC invariably audits only a very small percentage of agency samples (in the UK, 1%); it is in place as a 'safety-net' should internal QA/QC be failing. External QA/QC is never instigated as a means of monitoring the progress of individual agency workers. Internal QA/QC, on the other hand, is best placed to target agency needs. In particular, good internal QA/QC programs can target individual performance, detect problems at an early stage and thereby allow intervention and training to proceed before the quality of data gathered in the program is seriously compromised.

Aspects of (internal) QA/QC that agencies should focus efforts lie mainly in operator efficiency (i.e. any aspect of the MRHI protocol for which significant sources of error can arise as a result of the involvement of different workers). Other QA/QC that is clearly agency-based is that relating to techniques that are customised for, or specific to, a particular agency. Collectively, such aspects of the MRHI protocol that should be targeted include, operator efficiency in sampling, sorting efficiency (live-sorting or laboratory subsampling and subsequent sorting), identifications as well as other 'oddities' such as, for example, sampling with rakes vs kicking, or live-sorting with the naked eye vs use of jeweller's 'vision visors'

External QA/QC

Two critical components of the MRHI protocol have been targeted for external QA/QC. Thus, the performance of all S/T agencies is currently being assessed with regard to taxonomic identifications and efficiency of procedures used to sort samples. Prioritisation

of these aspects of the bioassessment protocol for external QA/QC is consistent with practise adopted in similar programs overseas (described above).

Taxonomic QA/QC. This work is being coordinated by the MDFRC, in collaboration with ERISS and specialist taxonomists throughout the country. The project has four main objectives:

1. to convene taxonomic workshops so as to facilitate training of agency staff in family-level identifications;
2. to cross-check a random selection of S/T samples (5%) to verify accuracy of identifications;
3. to inform agencies of results of QA/QC checks with advise, where necessary, on ways to improve sample identifications; and,
4. to report results of QA/QC to the NRHP committee.

Efficiency of sample sorting procedures. The second external QA/QC project will examine the efficiency of sample sorting procedures used by the S/T agencies. Two sorting protocols are currently employed by agencies, either field sorting of live samples or laboratory sorting of preserved samples. The external auditors are cross-checking community structure reported by agencies (from live-sorting or laboratory subsampling) against that representative of the whole sample. For live-sorting procedures, two taxa abundance lists are combined to provide the whole-sample estimate of community structure, namely: (i) that derived from agency residues after further processing by the external auditor (subsampling, sorting, specimen identification, enumeration and standardisation to sample unity) and (ii) that reported for the agency-sorted component (see full descriptions below). External auditors have selected and are processing an agency residue each from 5% of sites. Residues have been selected randomly after stratifying by sampling protocol (live sorted or preserved), catchment (or bioregion) and habitat.

Laboratory sorting of preserved samples normally involves subsampling of the finer fraction of each sample, for which the S/T agencies utilise a number of different subsampling devices. Therefore, a preliminary aspect of the external QA/QC will be the comparison of the performance of the different subsampling devices.

It is recognised that some taxa will be missed by the agencies in the initial sorting of the samples, particularly taxa with a low level of occurrence. In addition, there is the potential in the live-sorting procedure in particular, to obtain biased estimates of community structure that might alter to a significant degree the rank order of abundance of the different taxa. Whilst the external QA/QC described above is designed to check agency sorting efficiency, it is inappropriate to simply transfer to the MRHI the various criteria used in similar projects overseas governing the level of accuracy that is acceptable/required (see below, 'Recommended interim QA/QC criteria to be adopted for the MRHI'). Thus, part of the current external QA/QC project will be a review of the different criteria, selection of interim criteria, testing of the performance of these criteria on known data sets and refinement of the criteria before applying these new values to samples collected by the States and Territories in MRHI. Similar to the taxonomic QA/QC, agencies will be informed of the results of the QA/QC checks with advice, where necessary, on ways to improve efficiency of sample sorting. The performance of the sub-sampling devices and the results of the QA/QC and associated R&D will be reported to the NRHP committee.

Internal QA/QC

From a review of QA/QC protocols and approaches used in other countries/projects, the efficiency of different operators in collecting samples (hereafter 'sampling efficiency') was identified as a MRHI protocol which required QA/QC. Given the considerable expense and difficulties that would be entailed in externally auditing this aspect of the protocol - requiring field location of external auditors - it seemed more appropriate and justifiable that this be addressed at the State/Territory (internal) level.

In the section above ('QA/QC programs being conducted under the MRHI'), aspects of the MRHI protocol requiring internal (agency) QA/QC were listed. As at the time of preparing this report, S/T agencies had progressed in various ways towards adopting QA/QC programs for MRHI. To this end, some attempt has been made in this study to document and, where possible, coordinate the different QA/QC projects. A summary of all internal QA/QC projects being undertaken by each State and Territory is presented below ('Summary of State/Territory Internal QA/QC Programs').

RECOMMENDED INTERIM QA/QC CRITERIA TO BE ADOPTED FOR THE MRHI

Following review of the protocols and their QA/QC assessment in the British and US rapid bioassessment programs, together with a thorough understanding of the Australian MRHI protocols, it was possible to identify components of the MRHI program that required QA/QC auditing. These components are described above. Once identified, it was then necessary to set interim criteria upon which to assess the performance of S/T agencies.

Given the similarity of the Australian MRHI protocols to the British and US bioassessment programs, it was intended to select interim QA/QC criteria based on those applied to the relevant protocols of the above programs. The British criteria are complicated to apply, requiring an established database from which the process average, Acceptable Quality Level, Rejectable Quality Level, Reference Value and Decision Interval values are derived. Conversely, the US approach uses conservative but nonetheless arbitrary criteria, based on community dissimilarities, numbers of taxa and abundance of animals which are applied to each aspect of the sampling and processing protocols. These arbitrary values were selected by the USGS in the absence of relevant empirical data.

It is worth noting that the QA/QC criteria selected for the USGS's NAWQA program are acknowledged by Cuffney *et al.* (1993b) to be preliminary, interim criteria to be evaluated on the basis of empirical results. For instance, Cuffney *et al.* (1993b) note that criteria for checking elutriate components will be revised and applied to the overall estimate of data variability if there is a high rejection rate which cannot be improved by modifying the procedure. Similarly, criteria for split sample QA/QC "may not be achievable by field processing; consequently, the results of the split sample program will be used to empirically derive expectations for sample splits" (Cuffney *et al.*, 1993b). QA/QC criteria for other components of the protocols are also open to review and possible modification (Cuffney *et al.*, 1993b).

The US approach to QA/QC methodology and assessment criteria has been selected as the preferred model to follow for the Australian MRHI. The rationale for this decision is based on 3 factors: (i) like the US program, there has been up to the present only limited empirical data gathered in the Australian MRHI for selection of assessment criteria; (ii) preliminary, interim criteria are conservatively based; and (iii) the US criteria take into consideration the assessment of community structure (*viz* dissimilarity measures) as well as standard compositional indices. (While present modelling of agency data in the MRHI is concerned only with family presence/absence, this might be expanded in future modelling to incorporate measures of family relative or rank abundance.)

A disadvantage in selecting, *a priori*, conservative interim QA/QC criteria is the possibility that the measures are overly- and unnecessarily stringent. Thus, adoption of such criteria in the Australian MRHI could result in a high failure rate that is both disillusioning to S/T agencies and, more importantly, unnecessary insofar as improving the quality of, or altering conclusions drawn from, model output. For most components of the MRHI protocol, interim criteria upon which to assess the performance of S/T agencies are based on some empirical data and hence are an improvement on the arbitrary values selected in the USGS's program. Preliminary results from associated R&D evaluating the efficiency of laboratory subsampling devices have been used to set upper benchmark ('best possible') criteria to apply to sampling and field and laboratory sorting procedures. These values replace, for many components of sample processing, unrealistical and overly-conservative values set in the USGS guidelines.

Nevertheless, because the QA/QC criteria recommended for MRHI are based on limited empirical results, it is stressed that these are interim only and will be reviewed following additional R&D, including modelling on existing data sets/models (ACT, WA.), and on results of internal and external QA/QC programs. It will not be the intention to set criteria which are more stringent than can be achieved by best possible practice.

Interim QA/QC criteria for components of the MRHI protocols and associated methodology

Sampling efficiency:

Efficiency of all MRHI staff involved in sampling a habitat at a site should be assessed against that of an experienced agency operator, on a routine basis. This would involve operators, including an experienced agency operator, each collecting one replicate sample from the same habitat at a site and comparing the samples. Potential confounding errors can arise in the assessment because of two factors: (i) within-habitat variation in community structure (the result of habitat heterogeneity); and (ii) intra- and inter-personnel biases in sorting (especially live-sorting) and identification. To minimise potential problems arising from habitat variation, habitat for QA/QC assessment should be selected on the basis of sufficient uniformity and coverage to enable all operators to collect an adequate sample. To prevent confounding errors arising during sorting and identifications, all residues should be preserved and later sorted, and identifications cross-checked (details of these procedures are outlined below).

For data analysis pertaining to the first three of the QA/QC criteria listed below, proportional abundances of taxa present in the samples to be compared (field live-sorted samples or laboratory subsampled portions) are scaled such that the total invertebrate count of the samples is the same as that of the smallest of the samples. (In scaling down, abundances are rounded to the nearest whole animal; thus a value ≥ 0.5 is equivalent to at least one animal, a value < 0.5 is equivalent to zero animals.)

QA/QC criteria to use in the assessment of operator efficiency should include:

- the number of taxa encountered in any one of the samples must lie within 10 percent of the number of taxa recorded in the other;
- the number of taxa common to both samples must be ≥ 90 percent of the total number of taxa in both samples combined;
- the community similarity index (Bray-Curtis) comparing the two samples must be at least 80 percent for presence/absence data, and 75 percent for relative abundance data; and (if desired),
- for quantitative data, the total number of organisms found in one sample must be within 90 percent of the total number recorded in the other sample.

Field sorting of live samples:

For internal and external QA/QC, a small percentage of the residues remaining after field sorting is retained and preserved for later laboratory processing. These samples are randomly selected on the basis of bioregion, catchment, habitat and/or operator in such a manner that the sorting operator is only made aware of which residues are required for QA/QC processing *after* field sorting is completed. For external QA/QC, this selection procedure has been carried out by way of sealed envelopes sent from the external auditor to the MRHI agency operators, the contents of which indicate whether or not the samples are required for processing (5% of samples). For the most part, a sealed envelope has been associated with, and has accompanied, every sample sorted in the field since August 1995. Sample identity (location and habitat) is labelled on the outside of the envelope and after completion of field sorting, the operator opens the envelope to determine whether preservation of the sample residue is required.

For assessment of field live-sorting efficiency, a subsample of the residue is taken and its contents sorted and identified. The size of the subsample should be such that its total invertebrate count is equivalent at least to that live-sorted in the field. The aim in assessing field live-sorting efficiency is to compare the live-sorted component with an equivalent-sized component representative of the *whole* sample. (This is an analogous procedure to that used in the assessment of laboratory subsampling and sorting of preserved samples - see below.) 'Rare' taxa are discounted from the comparisons. To standardise data accordingly, the following steps are carried out:

(i) A 'subsample' of the agency live-sorted sample, equivalent in percentage terms to that taken of the agency residue, is drawn from the live-sort taxa abundance list by direct proportion. (Abundances are rounded to the nearest whole animal; thus a

value ≥ 0.5 is equivalent to at least one animal, a value < 0.5 is equivalent to zero animals.)

(ii) The community structure of the subsampled live-sort and residue components are then combined to derive a best proportional estimate of the whole sample (WS_n, where n = size of subsample in percentage terms). Proportional abundances of taxa present in WS_n are further scaled down so the total invertebrate count of WS_n is the same as that of the live-sorted component. (Again, abundances are rounded to the nearest whole animal.)

(iii) For QA/QC analyses, only those taxa are included whose combined abundance in WS_n and live-sort component contribute $> 1.0\%$ of the total abundance of WS_n and live-sort component combined.

QA/QC criteria to use in the assessment of field live-sorting efficiency should include:

- the number of taxa encountered in the live-sort component must lie within 20 percent of the number of taxa recorded in the best proportional estimate of the whole sample (WS_n);
- the community similarity index (Bray-Curtis) comparing the live sort component and best proportional estimate of the whole sample (WS_n) must be at least 20 percent for presence-absence data;
- the community similarity index (Bray-Curtis) comparing the live sort component and best proportional estimate of the whole sample (WS_n) must be at least 50 percent for relative abundance data; and
- a Spearman Rank Correlation comparing the community structure of the live-sort component and best proportional estimate of the whole sample (WS_n) must be significant at $P < 0.05$.

Laboratory sorting of preserved samples:

Fine fractions of agency-preserved samples are usually subsampled (1/4, 1/8 or smaller) and the subsample sorted. There are two aspects of the laboratory subsampling and sorting procedures that require QA/QC checks, (i) efficiency of agency sorting procedures and (ii) efficiency of subsampling procedures. For (i), internal (agency) checks are conducted of animals remaining in the residue of the original subsample following sorting. In practice, it may be desirable to combine QA/QC for (i) and (ii) as described in the following section.

For internal and external QA/QC, a small percentage of the residues remaining after sorting of the subsample is retained for further laboratory processing (3% for external audits). These samples are randomly selected on the basis of bioregion, catchment, habitat and/or operator. From each of the residues retained for further processing, internal and external auditors select a further subsample *of the same size* as that originally subsampled by the agency. (If the additional subsample is not removed at the same time as the initial agency subsample (i.e. in the case of external auditing), then the audit subsample will be a larger proportion than that taken originally to allow for the subsample already removed.) The subsample is sorted and all taxa identified and enumerated.

QA/QC criteria to use in the assessment of laboratory subsampling and sorting of preserved samples should include:

- the number of taxa encountered in the second subsample must lie within 20 percent of the number of taxa recorded during the initial agency sorting;
- the community similarity index (Bray-Curtis) comparing the original subsample and the second subsample must be at least 80 percent for relative abundance data and 85 percent for presence/absence data; and,
- for quantitative data, the total number of organisms found in the second subsample must be within 90 percent of the total number removed from the original subsample.

Specimen identifications:

For internal and external QA/QC, specimens from a small percentage of the samples identified by agencies are re-identified by competent practitioners to verify accuracy of identifications and enumerations (5% of samples for external audits). These samples are randomly selected on the basis of bioregion, catchment, habitat and identifier.

QA/QC criteria to use in the assessment of accuracy of specimen identifications include:

- at least 90 percent of the specimens must be correctly identified;
- the number of new taxa added to the taxa list must be ≤ 10 percent of the original number of taxa; and,
- the community similarity index (Bray-Curtis) comparing the original and the re-identified sample must be at least 90 percent based on abundance data.

Until acceptance/rejection criteria are better defined (from ongoing and related R&D), the best advice that can be forwarded to agencies from the results of external QA/QC is that they concentrate on results at the 'poorer' end of the scale, evaluate possible causes of the (excessive?) error and set in place procedures to redress the problem (i.e. review protocols, train staff etc). Feedback will be provided to agencies on a regular basis on progress made with refining the acceptance/rejection criteria.

SUMMARY OF STATE/TERRITORY INTERNAL QA/QC AND/OR RELATED PROGRAMS

At the time of reporting, State and Territory agencies had conducted, were conducting, or had indicated their intentions to conduct, a number of internal QA/QC and/or related R&D programs.

Australian Capital Territory:

The ACT program is coordinated by the Department of Environment, Land & Planning, ACT Planning Authority, with sample collection, processing and identifications sub-contracted to the CRC for Freshwater Ecology, University of Canberra. The CRC/University of Canberra has had a continuous involvement in developing biological monitoring techniques and has undertaken extensive QA/QC and related R&D programs on aspects of sampling, mostly pre-dating the MRHI. Much of this work was undertaken

as student projects under supervision of qualified staff, and the results are available in unpublished reports/theses.

Sampling Efficiency

Sampling for the MRHI is undertaken principally by CRC students trained in aquatic ecology and under general supervision. A study of the effect of different sampling operators on data quality has been previously undertaken; results indicate that changes in operator make little difference to the resulting classification of invertebrate data. Based on this result, the ACT does not see any need to implement a program of checks on operator efficiency for MRHI. If new, untrained staff are recruited to undertake sampling, then it is understood that a QC program would be initiated to check sample collection following initial training.

Additional QA/QC and related R&D

The ACT also conducts an internal QA/QC program for taxonomic identifications, separate from the external program being coordinated by MDFRC. Pre-identified samples are randomly selected and identifications cross-checked by experienced biologists. The number of samples selected depends on the size of the original data set, but ranges from 5% to 20%, with a higher proportion of samples from small data sets being selected for cross-checking than large sets.

In addition, the ACT (University of Canberra) has examined the efficiency/reproducibility of their box subsampler, used by other agencies (e.g. South Australia, but in a highly modified version). Even though all samples collected by the ACT are routinely preserved and returned to the laboratory for processing, the agency compared the efficiency of the subsampler for laboratory (preserved) processing to equivalent-sized 'subsamples' derived from using standard live-sorting procedures in the field. Modelling, using classification and ordination techniques on presence/absence and abundance data demonstrated that laboratory processing of preserved subsamples containing 100 animals provided similar results to those containing 200 animals, but the smaller subsample required less effort to process and identify. Samples of 200 animals obtained from live-sorting in the field were the least suitable sample type, producing variable data at greater expense than laboratory sorted subsamples. A predictive model constructed using data from laboratory-processed subsamples of 100 animals provided an efficient and effective tool for the detection and assessment of biological impairment in aquatic ecosystems (Simpson, 1995).

New South Wales:

The New South Wales EPA is the lead agency for MRHI and they, in association with three sub-agencies - Charles Sturt University (CSU), Department of Land and Water Conservation (DLWC) and University of New England (UNE) - collect and process all samples.

Sampling Efficiency

The NSW lead agency, together with one of its sub-agencies, DLWC, is implementing pairwise comparisons of experienced (three) and inexperienced operators (four) in round three of sampling. It was considered that there was insufficient area of habitat within any reach at a site to allow more than two independent replicate samples of any habitat to be

collected. Therefore, pairwise comparisons would be made across a total of 30 sites. For each pair of operators sampling a specific habitat, a total of six sites would be sampled, with each operator collecting one replicate sample. The experienced biologist would collect the first replicate at three of the sites and the inexperienced person would collect the first replicate at the remaining three sites. In addition, as a test of within-site/within-operator variability, an experienced biologist from the EPA will collect three replicates from the dominant edge habitat at an additional three sites.

In the fourth sampling round, NSW will broaden the above program to compare sampling efficiency of the most experienced biologist in each of the four agencies/sub-agencies (i.e. EPA, DLWC, CSU & UNE). This assumes that the experienced biologists have trained staff within their sub-agency to their own level of competence and therefore staff within each sub-agency will sample with the same efficiency.

Additional QA/QC and related R&D

Currently, NSW is not undertaking any additional QA/QC programs. However, if the external QA/QC taxonomic checks undertaken by the MDFRC detect a high error rate in identifications, then the State will implement an independent taxonomic checking procedure.

Northern Territory:

The Northern Territory agency (Water Resources Division of the Northern Territory Power & Water Authority) has been understaffed for much of the MRHI program and, currently (February 1996), are behind schedule with sample processing. Allowing the agency to undertake intensive QA/QC programs would only generate additional samples which they would be unlikely to process. Therefore, the agency has directed available resources to investigate QA/QC issues directly influencing the Northern Territory data collection, namely, sampling efficiency and sampling methodology.

Sampling Efficiency

The NT will implement a program in the third sampling round to test the efficiency of different operators in using the sweep net to collect macroinvertebrates from sand bed habitat. Because of the threat from crocodiles, the sampling methodology utilised by the Northern Territory involves two people in the stream, one person disturbing the habitat with a heavy garden rake ("rake operator") whilst a second person uses the pond net to collect the sample ("net operator"). The majority of sites are lentic/slow flowing, particularly at the end of the dry season and therefore the net operator actively sweeps the area as it is disturbed by the rake operator. It was envisaged that the greatest source of error would be associated with the person sweeping with the net. Therefore, in the proposed test of sampling efficiency, it was decided to standardise the disturbance component of sampling by using the same rake operator, and assess different net operators. The efficiency of three experienced operators and a total novice, with no prior experience of macroinvertebrate collection methods, will be compared. Each operator will collect three replicate samples from each of two sites on the Daly River with sampling restricted to sand bed habitat. Whilst this habitat is the second most dominant habitat type in the NT, it has the advantage of being more homogeneous than the dominant edge habitat and hence is a more accurate test of real operator differences. The samples will be processed and results reported subsequently. Riffle reaches, which are more

heterogeneous, are sampled by the normal kick method and therefore are not subjected to raking.

The above program examines the efficiency of different sweepers and assumes that personnel rake with the same intensity. An additional operator efficiency study will be implemented to compare raking efficiency. With sampling restricted to the more difficult-to-sample and more heterogeneous edge habitat, the net operator will remain constant and the rake operator changed. Two rake operators, suspected as having differences in raking intensity, will each disturb three replicate 10 m reaches and the net operator will collect standard sweeps from each operator. This exercise will be repeated at a second site, within 2 kms in the same stream. Comparisons will be made between rake operators and the data reported subsequently.

Additional QA/QC and related R&D

As noted above, sampling by the NT has been standardised to "rake and sweep samples", with one person disturbing the habitat with a rake whilst a second person sweeps with the pond net. An additional R&D program is to be implemented in the third sampling round to assess the efficiency/representativeness of raking. At two dominant habitats (edge habitat and sand bed habitat) at which the rake/sweep method is normally utilised, three replicate kick/sweep samples will be taken in addition to three replicate rake/sweep samples, using the standard kick/sweep methodology described in the National River Bioassessment Manual. At each site (habitat), the rake/sweep and kick/sweep samples will be interspersed to avoid spatial biases in invertebrate community structure.

Another ongoing R&D program is a cooperative study between the NT agency and the Environmental Research Institute of the Supervising Scientist (ERISS). Both organisations sample riffle habitat at the same three sites on the upper South Alligator River, the former organisation following MRHI protocols and the latter using a Surber sampler as part of a long term, established monitoring program. Data collected by both methods will be compared to assess the sampling efficiency of the MRHI protocols compared to the quantitative Surber sampling.

In addition, the NT agency plan to implement internal checks on sorting efficiency of all personnel involved in sample processing, by checking sample residues for missed taxa, and an internal check on identifications in cooperation with ERISS.

Queensland:

The lead agency for the MRHI in Queensland is the Resource Management Group of the Queensland Department of Primary Industries. The majority of sampling is conducted by hydrographers from the Group, under direction from experienced biologists.

Sampling Efficiency

The majority of sampling in the Queensland program is conducted by hydrographers who initially had no experience in collecting biological samples from streams. These staff underwent a one-week training program at the Australian Centre for Tropical Freshwater Research, James Cook University, prior to the first sampling round. Early in the MRHI program, it was seen as a priority to assess the efficiency of different operators in sample collection and live-sorting. The design of the overall QA/QC program facilitated inter-operator comparisons conducted within and amongst habitats at different sites. For a

particular habitat at a given site, percentage similarity of invertebrate communities were calculated amongst samples collected and sorted by different personnel. Appropriate statistical tests on number of taxa, their relative abundance and taxa presence/absence data collected by different operators were performed. The samples have been processed and preliminary results are available. The Queensland internal QA/QC program has been interactive, with follow-up training and advice offered where seen necessary. The findings of the Queensland QA/QC program will be made available in late 1996.

Live-sorting Efficiency

Inter-operator differences in live-sorting were assessed as part of the overall QA/QC program conducted in round two of sampling. An experienced biologist collected replicate samples from a particular habitat at a given site. Experienced and inexperienced personnel then live-sorted different replicates with the data compared to assess the efficiency of live-sorting. This was replicated across habitats to assess inter-habitat differences. Data from this exercise will be analysed and the results reported as per sampling efficiency (described above).

Additional QA/QC and related R&D

Queensland also conducts an internal QA/QC program for taxonomic identifications, complementary to the external program being coordinated by MDFRC. Pre-identified samples have been randomly selected and identifications cross-checked by experienced biologists.

South Australia:

The South Australian Environment Protection Authority is the lead MRHI agency in South Australia, with the majority of sampling conducted by experienced biologists.

Sampling Efficiency

The South Australian team consists of two experienced biologists who conduct all sampling, with one operator sampling lotic sites and the other operator, the lentic sites. Because of this consistency, the SA agency does not see operator efficiency as an issue for QA/QC. It has been agreed that, in the future, if the team should change as a result of a new or inexperienced person becoming involved, then the agency will need to adopt a QA/QC program to address operator efficiency.

Additional QA/QC and related R&D

South Australia and Victoria are collaborating on a QA/QC project to compare the efficacy of the two sorting methods (laboratory sorting of preserved samples vs live-sorting) at jointly-sampled sites. This project commenced in late 1994, but the number of sites sampled (n=2) was restricted because of drought conditions. It was the intention that additional sites be sampled in spring 1995 following good rains. The project is based on replicate samples being taken of the same habitats at the shared sites, with both teams sampling on the same day. Taxa lists then are compared between laboratory sorted samples (South Australia) and field live-sorted samples (Victoria). Preliminary results suggest that laboratory sorting collects more taxa. Previously the residues of live-sorted samples were not retained. However, it is planned to retain all residues from shared sites sampled in future and these will be re-sorted to identify taxa missed in the live-sorting.

In addition, the SA agency has undertaken a QA/QC study to assess the performance of the box subsampler utilised in its MRHI program. Again, this has been carried out in collaboration with Victoria, with both teams processing and identifying separate subsamples (10%) from a common sample and calculating percentage similarity between the subsamples. Preliminary results are encouraging, and indicate a Czeckanowski Similarity Coefficient of ~ 0.86. The results from this project will be subsequently reported.

Tasmania:

The Tasmanian lead agency is the Tasmanian Department of Primary Industry & Fisheries. Two field teams each of two personnel collect the samples, with one member sampling and live-sorting and the other measuring and taking site and habitat environmental parameters. These roles are held constant throughout each sampling round.

Sampling Efficiency

The lead agency has initiated an inter-operator comparison at three sites, upper Meander (Central Coast region), West Queen (West Coast region) and upper Sth Esk (North East region) rivers, in which the two personnel, one from each team, will each take four replicate samples from each habitat sampled. (Samples from the West Queen River were also collected concurrently with consultant biologists using the MRHI protocol to sample stream invertebrates for the Mount Lyell Remediation, Research and Demonstration Program.) To facilitate sampling, sites have been selected with sufficient habitat from which the extra replicates may be taken. All samples will be live-sorted in the field, but the residues will be retained for further laboratory processing so that a complete taxa listing may be compiled for each replicate. Sampling efficiency of each operator at the three sites will be assessed by within and between operator comparisons.

Additional QA/QC and related R&D

Tasmania have implemented a series of additional internal QA/QC programs:

- In-house cross-checking of taxonomic identifications has highlighted several 'problem' groups that require reworking. As these groups are noted, staff are trained/updated on their taxonomy, past samples rechecked and, where appropriate, the database updated.
- On each sampling round, the two field teams assemble at three sites and cooperatively sample the sites to compare and ensure that standard approaches are used by the field teams.
- A cross-check of description and measurement of habitat and physical parameters at two sites (i.e. those sites used in assessing sampling efficiency) is conducted, whereby the sites descriptors are independently assessed by different personnel with the results compared. Any major differences in site descriptions are investigated and methods/approaches re-standardised to avoid any repetition of such discrepancies.

Victoria:

The Victorian lead agency, the Victorian Environmental Protection Authority has been conducting broad-scale biological monitoring in Victoria for many years prior to the MRHI, using standardised methodologies similar, but not identical, to those specified in the National River Bioassessment Manual. To maintain consistency and compatibility

with existing databases, the EPA has continued using its established monitoring protocols. Sampling in Victoria is also conducted by the sub-agency, Water Ecoscience (formerly State Water Laboratory), also using the pre-MRHI protocols.

Sampling Efficiency

The Victorian lead agency does not see the need to implement an internal QA/QC program to investigate operator (sampling) efficiency. The agency employs experienced biologists to conduct all sampling and a series of workshops and sampling expeditions has been used to fully standardise sampling protocols. It is understood that if new personnel, particularly inexperienced personnel, join the teams then in-house QA/QC will be implemented after an initial training program.

With LWRRDC (MRHI) funding, the Victorian EPA, in collaboration with AWT Ensight (Sydney), are conducting an evaluation of rapid bioassessment for RIVPACS modelling with the aims of a) assessing inter-operator sampling and field live-sorting efficiency (in particular, experienced versus novice operators), and b) comparing the information derived from rapid bioassessment samples with that obtained from traditional quantitative sampling techniques.

The first half of the study involved two experienced biologists and two novices sampling concurrently riffles, edges and pool rocks at the same locations. Each person sampled then live sorted and identified his/her sample to family level. A deficiency in the design was the lack of preservation of residues for additional laboratory processing, leading to potential inability to determine at what stage in sample collection and further sample processing operator errors had arisen. Fortunately, preliminary data for sites sampled in Victoria indicated that, although the number of animals removed from samples differed between individuals, there was no marked difference in the number of taxa encountered. Moreover, multivariate analyses indicated a high similarity in community structure amongst the data of the four operators sampling the same habitat of a given site, whilst in all cases the data correctly discriminated between unimpacted and impacted sites. Preliminary assessments of the identifications conducted by the novices have shown them to be reliable with about an 8% error rate in correctly naming the taxa but generally negligible error in enumerating taxa (Metzeling, 1995). The numbers of animals misidentified amounted to about 2% of the total, these being mainly the less common or "rare" taxa. This project is on-going and will be written-up and reported subsequently. As noted above, future studies with a similar design would benefit from the preservation and subsequent processing of sample residues to enable a full interpretation of the data. This is particularly relevant to interpreting any discrepancies arising amongst operators.

Additional QA/QC and related R&D

Victorian agencies conduct two internal taxonomic QA/QC programs. The first, conducted by the EPA, involves random selection of 10% of samples for re-identification by the senior taxonomists. Results to date show no errors at the family level with approximately 11% error rate at the species level. This QA/QC program is on-going and is independent of the external taxonomic checks coordinated by the MDFRC.

The second taxonomic QA/QC study is a cooperative program between the two agencies, the Victorian EPA and Water Ecoscience. Each team samples the same habitats at one shared site per MRHI sampling round, collecting samples as per routine procedures. Back in the laboratory, the two people in each agency team independently identify their

respective sample, then swap the samples with the other agency team for re-identification. As a result, each sample is identified four times. These data will be used to maintain consistency in identifications and to target problem groups for in-house or other training. Results from the cooperative study also facilitate some comparative assessment of sampling and/or live-sorting efficiency by each group.

Victoria and South Australia are collaborating on a QA/QC project to compare the efficacy of the two sorting methods at jointly-sampled sites. In addition, Victoria are assisting South Australia in a QA/QC program to assess the performance of the box subsampler utilised in the South Australian laboratory. Descriptions of these projects are provided in the South Australian summary above.

Western Australia:

Sampling in Western Australia is conducted by the lead agency, the Western Australian Department of Conservation & Land Management (WADCALM) and three sub-agencies: University of Western Australia, Murdoch University and Edith Cowan University. Each institution samples a different bioregion and all sampling is conducted by experienced biologists.

Sampling Efficiency

WA agencies have convened several 'workshops' in which sampling efficiency has been addressed; they have "gone to some lengths to ensure sampling is thorough and consistent among all agencies"(M. Smith, WADCALM, pers. comm.). They are not planning to carry out any additional QA/QC program to cover this issue.

Additional QA/QC and related R&D

WA agencies are relying on the external QA/QC project on sorting efficiency to identify any inconsistencies arising from this aspect of the MRHI protocol. An initial intention to examine the effectiveness of live-sorting has been abandoned because of the constraints of time and money and the desire not to duplicate the external QA/QC project examining sorting efficiency.

CONCLUSIONS FROM S/T QA/QC PROGRAMS

Following discussions with lead agencies from each State and Territory, it has been possible to compile a summary of internal QA/QC programs already conducted, currently being conducted and/or planned for the near future (Table 1).

This report demonstrates reasonable but disparate coverage of most of the issues identified for internal QA/QC (Table 1). Whilst agencies complain of the lack of resources made available for internal QA/QC, there are still misunderstandings amongst some agencies as to the need for *independent* and *routine* internal and external QA/QC. In particular, there may be some lack of appreciation at this stage amongst agencies of potential high error that can arise in live-sorting of samples; there are no agencies at present committed to routine QA/QC of this aspect of the MRHI protocol.

Maximum value from these studies described above will be attained by an external review of the results and conclusions arising from the individual internal QA/QC programs. This is because the studies tend to have a regional limitation due to State and

Territory boundaries. A review will identify common trends and develop common recommendations which may be applied at a national level. With this end-point in mind, it is desirable that all States and Territories agree to such a review and, importantly, on common deadlines to facilitate this.

Table 1. Summary of internal QA/QC and R&D projects undertaken/proposed by each S/T agency

State		Project Description
ACT	(a)	Previously undertaken study to assess the effect of different operators on sampling efficiency
	(b)	Internal taxonomic QA/QC.
	(c)	Assessing performance/reproducibility of a box subsampler.
NSW	(a)	Amongst-operator comparison of sampling efficiency by experienced and novice personnel.
	(b)	Internal taxonomic QA/QC will be implemented if external QA/QC indicates shortfalls.
NT	(a)	Amongst-operator comparison of 'sweeping' efficiency for the rake/sweep sampling method.
	(b)	Amongst-operator comparison of 'raking' efficiency for the rake/sweep sampling method.
	(c)	Comparison of "rake/sweep" sampling to standard "kick/sweep" sampling.
	(d)	Comparison of data produced by MRHI protocols to quantitative samples by Surber.
	(e)	Internal taxonomic QA/QC.
	(f)	Internal check on operator sorting efficiency.
QLD	(a)	Amongst-operator comparisons in sampling and live-sorting efficiency.
	(b)	Internal taxonomic QA/QC.
SA	(a)	Comparison of field live-sorting to laboratory sorting in partnership with the Victorian agency.
	(b)	Assessing performance/reproducibility of a box subsampler.
TAS	(a)	Amongst-operator comparison of sampling efficiency
	(b)	Internal check on operator live-sorting efficiency.
	(c)	Internal taxonomic QA/QC re-identifying 'problem' groups from all samples.
	(d)	Amongst-operator check on site descriptions (habitat & physical parameter description).
	(e)	Regular inter-agency checks to ensure standardised application of MRHI sampling protocols
VIC	(a)	Amongst-operator comparisons of sampling and live-sorting efficiency.
	(b)	Comparison of data collected by MRHI protocols to traditional quantitative sampling methods.
	(c)	Internal taxonomic QA/QC re-identifying 10% of randomly-selected samples.
	(d)	Inter-State check on operator sampling and sorting efficiency, and taxonomic consistency.
	(e)	Comparison of live-sorting to laboratory sorting in partnership with the South Australian agency.
	(d)	Assessing performance/reproducibility of the box subsampler.
WA	(a)	An initial inter-agency check to ensure standardised application of MRHI sampling protocols

ACKNOWLEDGMENTS

The authors thank all lead agency coordinators and other personnel involved in the MRHI for their time, opinions and cooperation in implementing internal QA/QC and in discussing the aims and designs of their programs to facilitate the preparation of this review.

REFERENCES

ANZECC (1992) National water quality management strategy. Australian water quality guidelines for fresh and marine waters. Australian and New Zealand Environment and Conservation Council.

Cuffney, T.F., Gurtz, M.E. & Meador, M.R. (1993a) Methods for collecting benthic invertebrate samples as part of the National Water-Quality Assessment Program. U.S. Geological Survey, Report No. 93-406, 66p.

Cuffney, T.F., Gurtz, M.E. & Meador, M.R. (1993b) Guidelines for the processing and quality assurance of benthic invertebrate samples collected as part of the National Water-Quality Assessment Program. U.S. Geological Survey, Report No. 93-407, 80 p.

Davies, P. E. (1994) Monitoring River Health Initiative river bioassessment manual. National river processes and management program, Freshwater Systems, Tasmania.

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

Kinley R.D. & Ellis, J.C. (1991) The application of statistical quality control methods to macroinvertebrate data. NRA Report. CO 2905.

Metzeling, L. (1995) National River Health Program - Monitoring River Health Initiative: Victoria. **Aust. Soc. Limn. Newsletter** 33 (2), p.33

Norris, R.H. & Norris, K.R. (1995) The need for biological assessment of water quality: Australian perspective. **Australian Journal of Ecology** 20:1-6.

Plafkin, J.L., Barbour, M.T., Porter, K.D., Gross, S.K. & Hughes, R.M. (1989) Rapid bioassessment protocols for use in streams and rivers: benthic macroinvertebrates and fish. U.S. Environmental Protection Agency, Washington, D.C. Report No. EPA/440/4-89/001.

Simpson, J. (1995) The effect of subsampling method on the efficiency, accuracy and precision of a predictive model for the assessment of aquatic ecosystem health. Hons. Thesis, University of Canberra.

van Dijk, P. (1994) Analytical quality control for macroinvetebrate enumeration. R&D Note 331 to the National Rivers Authority, Bristol, U.K.

Wright, J. F. (1995) Development and use of a system for predicting the macroinvertebrate fauna in flowing waters. **Australian Journal of Ecology** 20:181-197.

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

ATTACHMENT 2

Assessment of the efficiency of four types of device for subsampling of aquatic macroinvertebrate samples: preliminary results

by

Andrew Storey [1] & Chris Humphrey [2]

1 Department of Zoology, The University of Western Australia, Nedlands, WA 6907

2 ERISS, Locked Bag 2, Jabiru, NT 0886

Assessment of the efficiency of four types of device for subsampling of aquatic macroinvertebrate samples: preliminary results

Introduction

This report presents a comparative statistical evaluation of results to date from processing one of three mega-samples of aquatic macroinvertebrates with associated organic and inorganic material through four different subsampling devices. This work is funded by LWRRDC as part of R&D for MRHI QA/QC assessment, and aims to investigate sources of error in subsampling devices used by State and Territory (S/T) agencies that preserve samples in the field for subsequent laboratory sorting.

From the literature it appears that any similar work has not been carried out with such statistical validity, with no studies using truly independent replicate subsamples. Further, previous studies do not appear to have undertaken multivariate evaluations of results (i.e. use of community dissimilarity measures to produce ordinations and classifications) which is of direct relevance to RIVPACS-type modelling currently being developed as part of the MRHI.

Ultimately, two additional mega-samples, differing mainly in residue and macroinvertebrate composition and structure will be processed to assess the accuracy and precision of each subsampling device. The resulting data will be evaluated, along with information on time taken to subsample with each device to select the optimal subsampling device for future QA/QC assessments for sampling under the MRHI. The results of all analyses will be reported to the NRHP Committee in late 1996.

Methods

One 'mega-sample' collected from riffle habitat of the upper South Alligator River (NT), with very low levels of detritus, was split into four 'main' samples using a geological splitting device (described below). Each main sample was then dedicated to one of the four subsampling devices and passed through the device to produce subsamples comprising 1/8th of the main sample. The size of the mega-sample (i.e. total abundance of organism) was determined such that an 1/8th subsample from the main sample would comprise approximately 250 animals. Following each subsampling run, the subsample was sorted, macroinvertebrates identified to family level and enumerated, and then macroinvertebrates and organic/inorganic material remixed with the main sample and the process repeated a total of five times to produce five independent replicate subsamples of each main sample dedicated to each device. Following the subsampling exercise, the whole of each main sample was sorted to determine number and abundance of each taxa to allow comparisons of the five replicate subsamples to the original, known sample.

Three of the four subsampling devices selected are those commonly used by S/T agencies for MRHI:

1. ACT box subsampler. This device is modified from Marchant (1989) and is utilised by the Australian Capital Territory agency. It has an inner surface area of 1225 cm^2 and a depth of 13.2 cm, and is divided by thin metal partitions into 100 cells, each cell with an area of 11.56 cm^2. The sample, with additional water if necessary, is poured into the device to the top of the cell partitions (2.8 cm depth), and then the device is tipped from side to side (45°) until the main sample is 'evenly spread' across the cells. Cells are selected at random and the contents (animals and organic/inorganic matter) evacuated using a vacuum pump. Normally sufficient cells are selected to provide a subsample of approximately 200 animals. However, to allow

comparison to the two sample splitters, which split to 1/8s (12.5% of sample), 12 cells were selected at random as an approximation to 1/8th (12 vials = 12% of the surface area, requiring a scaling factor of taxon abundance values of x 8.333).

2. SA box sampler. This device is utilised by the South Australian agency and consists of a sealed, plastic rectangular box, with an inner surface area of 784 cm^2, packed tightly with 100 circular plastic vials. Total inner surface area is 598 cm^2, while total interstitial area is 186 cm^2. The device is filled with ethanol (or water, for a sample preserved in formalin) to a level just covering the vials (5.4 cm depth), and the sample is poured into the sampler. A tightly-fitting plastic grill (light-diffuser, with partitions 1.5 mm thickness, and 1.56 cm^2 spacings) is placed over the vials to hold them in place, the box is sealed and then agitated vigorously from side to side (~90°) to disperse the sample. The partitions of the plastic grill are vertically-aligned (to 1.25 cm depth) so that during agitation detritus and organisms are free to pass in and out of the vials and interstitial spaces. Sixteen vials were selected at random to represent a 12.5% subsample and the contents removed (animals and organic/inorganic matter) (one vial = 0.763% of the total surface area of the device, 16 vials = 12.21%, requiring a scaling factor of taxon abundance values of x 8.19).

3. Jug splitter. This device is utilised by the Western Australian lead agency and consists of a cylindrical jug, to the inside wall and base of which has been attached a plastic vertical 'V'-partition running the full inside height of the jug, with the tops of the two arms of the V-partition fixed to the inner wall of the cylinder. The sample is poured into the jug with sufficient water to bring the sample volume up to a pre-determined height in the jug. This height is determined such that, with the jug tilted with the internal V-partition aligned on the lowest side of the jug, the tip of the partition located at the base of the cylinder just becomes exposed at the point at which the sample would pour from each side of the partition at the mouth of the jug. With the sample in the jug, it is stirred and the jug then tipped so that the sample contents settle to either side of the internal division. Each half of the sample is then poured out to either side of the partition into separate receptor containers. This procedure is then repeated with each half until 1/8s are derived. One 1/8th is selected at random and sorted (1/8th = 12.5% of the main sample, requiring a scaling factor of taxon abundance values of x 8.0).

4. Geological splitter (hereafter, 'geo-splitter'). This device currently is not utilised by any S/T agency, but was selected as an additional comparison, particularly with the performance of the jug splitter. The device is a metal structure consisting of a horizontal multi-partitioned grill (total area 30.6 x 15 cm) with 2 basal drawers, one on each side of the device. There are a total of 36 spaces across the top of the grill (18 per split); each space is 6.8 mm wide, between vertically-aligned partitions of thickness 1.7 mm. A sample is poured evenly across the grill with alternate spaces/partitions draining to a different basal drawer, thereby giving rise to 2 even halves. This process is then repeated with each half until 1/8s are derived. One 1/8th is then selected at random and sorted (1/8th = 12.5% of the main sample, requiring a scaling factor of taxon abundance values of x 8.0).

Each replicate subsample was sorted to remove all macroinvertebrates. Specimens were identified to family level and enumerated, and abundance values scaled-up to be comparable to the main sample abundances (see above).

Bray-Curtis dissimilarity measures between each pairwise combination of main sample were calculated to determine their similarity/comparability. This was performed on presence/absence and untransformed abundance data sets using PATN software (Belbin, 1993). In addition, the proportion of the total number of taxa in each main sample with an abundance of less than 8 individuals was determined. This indicates the proportion of taxa which could not occur in all 1/8th subsample portions.

The total abundance of macroinvertebrates as determined by each device was compared to the known total abundance in the main samples (after subsamples were scaled-up by the appropriate factor to allow valid comparison). The proportion of taxa in the main sample that were not recorded in each subsample was also determined.

Bray-Curtis dissimilarities between each replicate subsample and its respective main sample were calculated for each subsampling device using presence/absence and untransformed abundance data sets, with subsamples scaled-up to main sample abundances. UPGMA classification and SSH ordination of the replicates from all devices, together with the main samples, were performed using default options in PATN (Belbin, 1993). These analyses allowed comparisons to be made within and amongst devices.

Results

Main samples derived from the mega-sample demonstrated very high mean pairwise similarities, both for presence/absence (mean pairwise comparisons = 97.7%; Table 1) and abundance (mean pairwise comparisons =93.2%, Table 2) data sets, indicating a relatively accurate split of the mega-sample in quarters.

Table 1. Bray-Curtis dissimilarities for pairwise comparisons between each main sample derived from the initial mega-sample (presence/absence data).

	ACT multi-cell	SA multi-cell	Jug splitter
SA multi-cell	0.000		
Jug splitter	0.023	0.023	
Geo-splitter	0.023	0.023	0.048

Table 2. Bray-Curtis dissimilarities for pairwise comparisons between each main sample derived from the initial mega-sample (abundance data).

	ACT multi-cell	SA multi-cell	Jug splitter
SA multi-cell	0.0743		
Jug splitter	0.0625	0.0653	
Geo-splitter	0.0645	0.0809	0.0613

Approximately 26% of taxa in the main samples had a total abundance of less than 8 individuals and could not have occurred in all subsample portions when the main sample was split into eighths (Table 3). Assuming an even distribution of individuals amongst subsamples, the probability of this 26% of taxa occurring in subsamples will depend on their actual abundances in the main sample. A taxon with an abundance of seven individuals will have a 0.875 probability of occurring in a subsample, a taxon with four individuals, a probability of 0.5, and a taxon with one individual, a probability of 0.125 of occurring in one subsample. Conversely, a taxon with an abundance of seven individuals will have a 0.9999 probability of occurring at least once across all five replicate subsamples, a taxon with four individuals, a probability of 0.969, and a taxon with one individual, a probability of 0.487 (Table 4).

Table 3. Proportion of taxa in each main sample with a low-occurrence (i.e. taxa with an abundance of <8 individuals and which could not occur in all eight fractions of a subsampling run).

Subsampling device	Number of taxa in main sample	No. taxa with abundance <8 individuals	Proportion of total no. taxa with a low occurrence
ACT multi-cell	22	6	0.273
SA multi-cell	22	6	0.273
Jug splitter	21	5	0.238
Geo-splitter	21	5	0.238

Table 4. Probabilities of a taxon occurring at least once across five replicate subsamples at varying abundances levels for taxa with less than eight individuals in the main sample.

No of individuals of a taxon in the main sample	Probability of the taxon not occurring in one subsample	Probability of the taxon not occurring in five replicate subsamples	Probability of a taxon occurring at least once in 5 replicate subsamples
7	0.1250	0.0001	0.9999
6	0.2500	0.0010	0.9990
5	0.3750	0.0074	0.9926
4	0.5000	0.0313	0.9688
3	0.6250	0.0954	0.9046
2	0.7500	0.2373	0.7627
1	0.8750	0.5129	0.4871

The ACT multi-cell device accurately estimated total abundance in the main sample, with relatively low variance indicating high precision (Figure 1). The SA multi-cell device also had high precision, but overestimated abundance suggesting either a scaling error (i.e. inaccurate estimation of vial surface area to interstitial area) or a possible preference for sample residues to settle in vials as opposed to interstices. The two splitting devices had comparable accuracy to the ACT multi-cell device, with main sample abundance falling within the 95% confidence intervals of the subsample estimates. However, there was relatively high variance on these estimates, indicating low precision, with both splitting devices tending to overestimate total abundance (Figure 1).

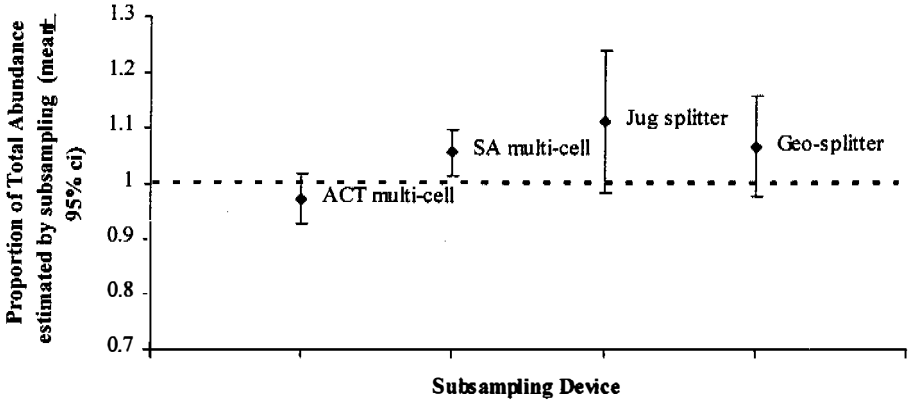


Figure 1. Estimates of total abundance of animals in the main sample by each subsampling device.

Individual replicate subsamples, from each device, failed to sample approximately 20% of the fauna in the main samples (Figure 2). This was not surprising, considering that 26% of the fauna could not have occurred in all subsample divisions (Table 3). The actual probabilities of individual taxa not occurring in any subsample will depend on their abundance in the main sample (Table 4), with the abundance of each taxon varying between each main sample. Overall, however, the SA multi-cell device was the only device not to record all taxa over the five replicates, missing one taxon (Leptoceridae; two individuals occurred in the main sample). There was no obvious difference in either the precision or accuracy with which each device detected taxa present in the respective main samples (Figure 2).

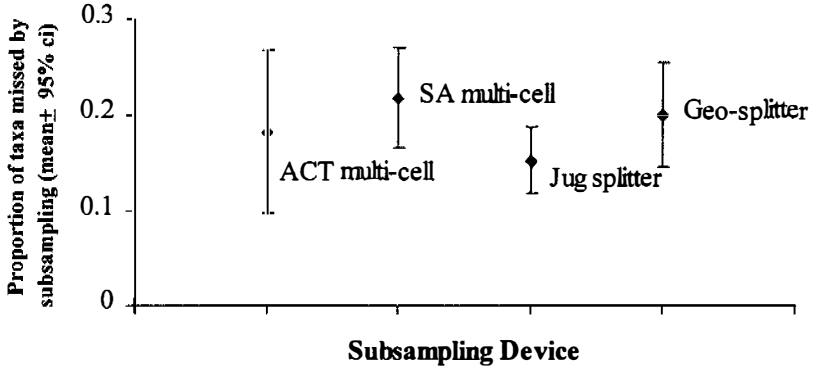


Figure 2. Proportion of taxa in main samples that were not recorded in each subsample (mean ± 95 % ci).

There was no significant difference amongst devices in levels of dissimilarity between replicate subsamples and respective main samples on presence/absence data (Figure 3). Mean dissimilarities of 0.103, 0.123, 0.083, 0.112 were recorded for the ACT multi-cell, SA multi-cell, Jug splitter and Geo-splitter respectively. Again, these small inter-device differences were

to be expected, particularly as the taxa being missed were the same across all four main samples (i.e. Gomphidae were under-represented in all four main samples). Classification and ordination (Figures 4 & 5 respectively) showed no systematic separation of replicate subsamples either from the main samples or between devices. Mean pairwise dissimilarities (± 95% ci) amongst subsamples within each device, on presence/absence data were 0.119 (± 0.034), 0.099 (± 0.015), 0.079 (± 0.025) and 0.107 (± 0.021) for the ACT multi-cell, SA multi-cell, Jug splitter and Geo-splitter respectively.

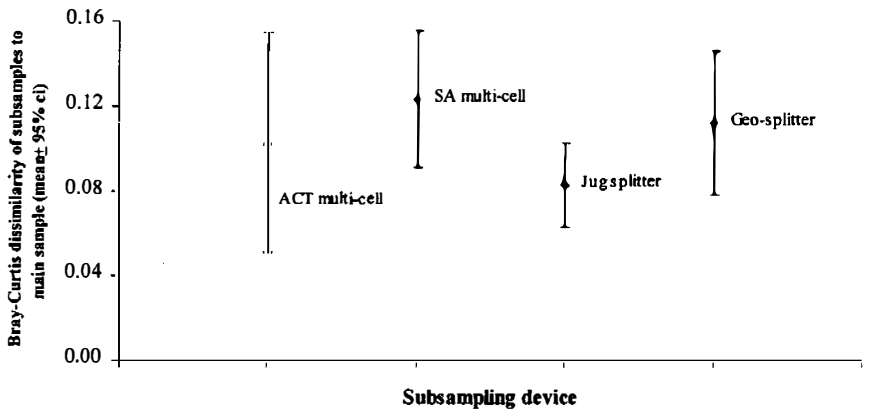


Figure 3. Mean Bray-Curtis dissimilarities of replicate subsamples (n=5) to main sample for presence/absence data (x ± 95% c.i.)

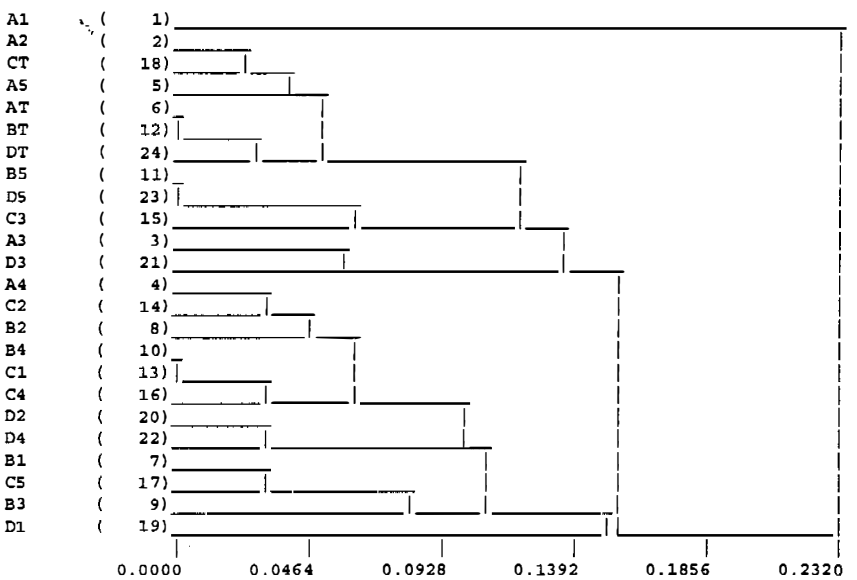


Figure 4. UPGMA classification of replicate subsamples with main samples for each subsampling device (presence/absence data; A = ACT multi-cell device, B = SA multi-cell device, C = Jug splitter, D = geo-splitter; 1 - 5 = replicate subsamples; T = main sample).

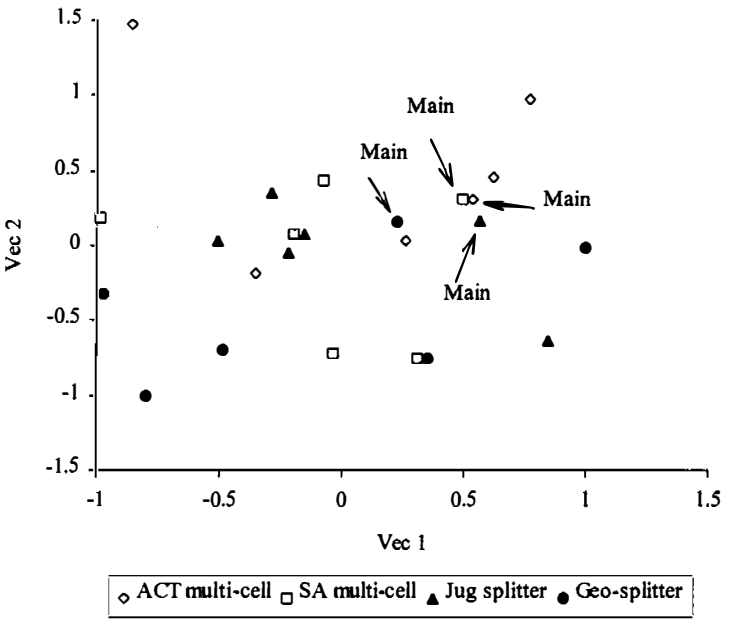


Figure 5. Vector 1 by vector 2 SSH ordination of replicate subsamples with main samples for each subsampling device (presence/absence data) (optimum solution was achieved with three dimensions with a stress of 0.141).

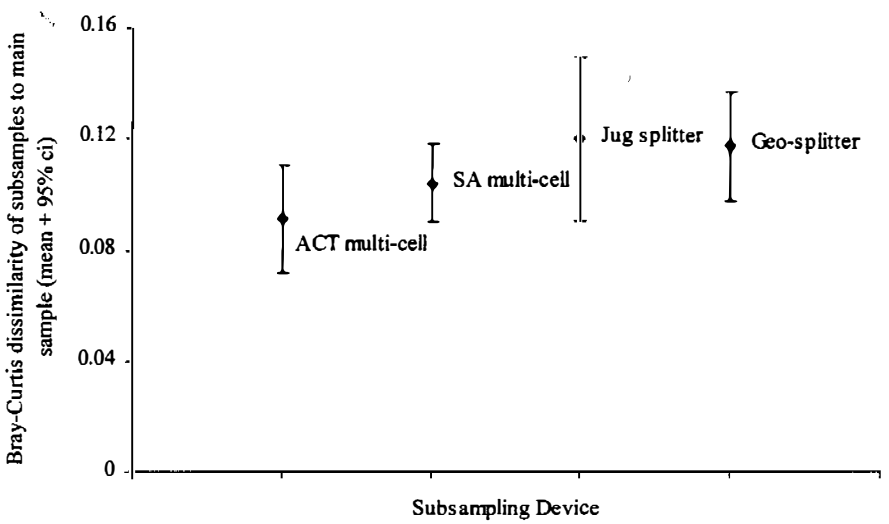


Figure 6. Mean Bray-Curtis dissimilarities of replicate subsamples (n=5) to main sample for abundance data (mean ± 95% c.i.)

Mean dissimilarities of replicate subsamples to main samples, using abundance data (Figure 6), demonstrated no significant differences amongst devices in levels of dissimilarity. Mean dissimilarities of 0.091, 0.104, 0.120, 0.117 were recorded for the ACT multi-cell, SA multi-

cell, Jug splitter and Geo-splitter respectively. Classification and ordination (Figures 7 & 8) showed some systematic separation of replicate subsamples. Compared with the presence/absence ordination, there was higher precision within subsampling devices, particularly the ACT multi-cell device, with tighter clustering of replicates within each device. However, the main samples tended to separate from the subsamples, forming a separate group within the UPGMA classification. This reflected the high similarity amongst main samples, with a progressive decline in similarity between subsamples and the respective main sample, and amongst subsamples from different main samples. Mean pairwise dissimilarities (± 95% ci) amongst subsamples within each device, on abundance data were 0.124 (± 0.017), 0.098 (± 0.008), 0.162 (± 0.016) and 0.151 (± 0.020) for the ACT multi-cell, SA multi-cell, Jug splitter and Geo-splitter respectively.

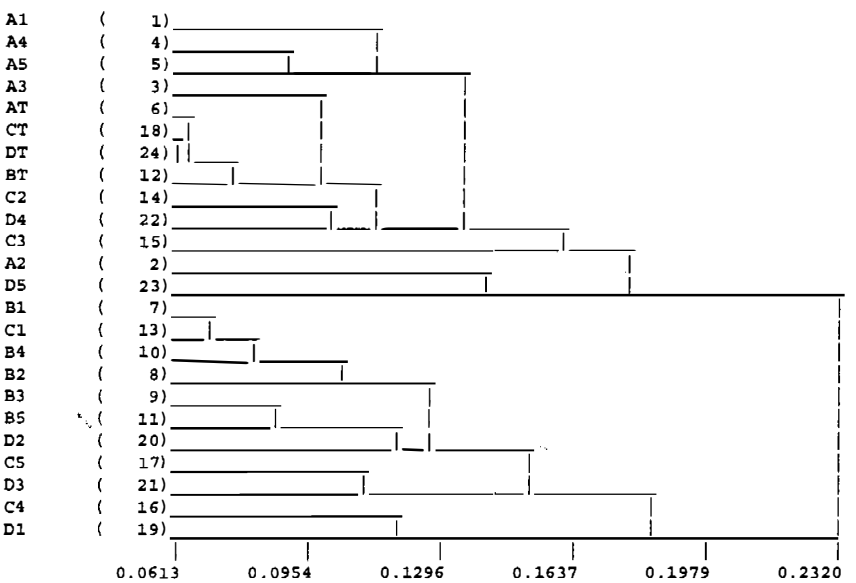


Figure 7. UPGMA classification of replicate subsamples with main samples for each subsampling device (abundance data; A = ACT multi-cell device, B = SA multi-cell device, C = Jug splitter, D = geo-splitter; 1 - 5 = replicate subsamples; T = main sample).

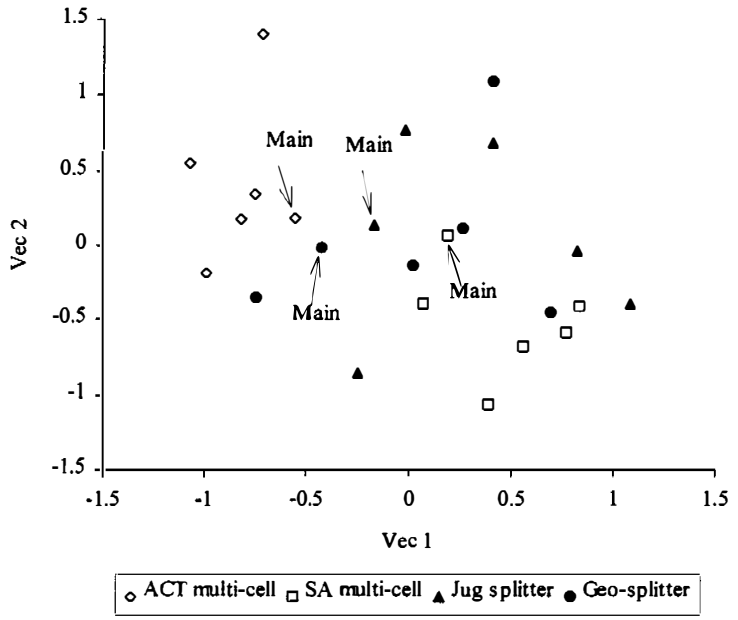


Figure 8. SSH ordination of replicate subsamples with main samples for each subsampling device (abundance data) (optimum solution was achieved with three dimensions with a stress of 0.139).

Discussion

The option of using one mega-sample split into four main samples was a compromise between statistical rigour (*viz.* collection of independent replicate subsamples from each device) and logistical constraints. The ideal design would have involved passing a single sample through each subsampler five times. However, initial trials with such a design resulted in substantial deterioration of the sample, to such an extent, that many taxa were no longer identifiable before the sample had passed through all subsampling devices. Therefore, the above design was used as a compromise, as a means of obtaining independent replicate subsamples for each device. Bray-Curtis dissimilarity comparisons of the four main samples demonstrated a very high level of similarity. Therefore, performance of each subsampling device should not be influenced by a main sample being very different in composition from the other main samples.

Scaling of the abundance values for each taxon represented in subsamples derived from the SA multi-cell device proved problematic. This related to the non-uniformity of circular vials set in a square box with uneven interstitial areas, and into which animals may fall. The area of vials will be re-estimated to calculate the number required to obtain a subsample equivalent to ~ 12.5% of the main sample and analyses repeated for the final report. This may lead to minor alterations in calculations based on abundance data.

Results from the processing of one mega-sample indicate no outstanding differences amongst devices at this stage. However, the riffle sample from the South Alligator River had very low levels of detritus. It is possible that the subsampling devices may perform differently for a sample with large quantities of detritus and/or a different community composition (i.e. more or less rare taxa or 'heavy' taxa (i.e. Gastropoda)). These aspects will be assessed through the selection of appropriate mega-samples for remaining testing proposed under this project.

QA/QC criteria proposed by Cuffney *et al.* (1993) for the US rapid bioassessment protocols adopt a maximum dissimilarity of 10% between components (e.g. between any two equivalent-sized subsamples drawn from a main sample). Although these criteria were subject to review based on empirical data, the current project indicates that these criteria would be difficult to meet, and any criteria set for aspects of the Australian MRHI rapid bioassessment protocols need to consider results from the present analyses.

Time taken to process each sample through each subsampling device will be part of the evaluation process to select the optimum subsampling device for QA/QC testing. Where two devices provide a similar efficiency in terms of main sample reproducibility, but one device is faster to use and therefore provides an economic saving, then this needs to be considered.

Two additional mega-samples, selected to represent different types of sample structure (e.g. quantity of detritus and community composition) will be processed and analysed using the above protocols and all results will be reported to the NRHP committee. These additional analyses will complement work completed to date to identify the most efficient subsampling device for use in future QA/QC assessments for the MRHI. The results will also indicate the level of error being introduced to the agency data sets through laboratory subsampling of preserved samples (reference and test samples) collected as part of the MRHI.

Acknowledgements

The authors gratefully acknowledge the assistance of Robin Galbreath and Stacey Braund of ERISS who undertook all of the subsampling, sorting, identifications and enumerations.

References

Belbin, L. (1993) PATN: Pattern analysis package. CSIRO, Division of Wildlife & Ecology, Canberra, ACT.

Cuffney, T.F., Gurtz, M.E. & Meador, M.R. (1993) Guidelines for the processing and quality assurance of benthic invertebrate samples collected as part of the National Water-Quality Assessment Program. U.S. Geological Survey, Report No. 93-407, 80 p.

Marchant, R. (1989) A subsampler for samples of benthic invertebrates. **Bulletin of the Australian Society for Limnology** 12: 49-52.

ATTACHMENT 3

Proposed analytical approach to assess the effect of different QA/QC criteria for sampling efficiency on the robustness of MRHI models

by

Andrew Storey [1] & Chris Humphrey [2]

[1] Department of Zoology, The University of Western Australia, Nedlands, WA 6907

[2] ERISS, Locked Bag 2, Jabiru, NT 0886

Proposed analytical approach to assess the effect of different QA/QC criteria for sampling efficiency on the robustness of MRHI models

Table Of Contents

Background

This note outlines proposed analytical methods to be used in an R&D project designed to test the effect of different QA/QC criteria on the robustness of models developed from data produced by the MRHI. This R&D is funded by LWRRDC as part of the NRHP.

The MRHI program ultimately will result in the development of a series of models for different habitats sampled in different seasons in each region (bioregion) of Australia. Currently, the basis for each model will be a UPGMA classification of family-level presence/absence macroinvertebrate data, with discrimination between site groupings (based on UPGMA & TWINSPAN) using environmental data (*sensu* RIVPACS, Wright, 1995), and subsequent prediction of taxa occurrences (Wright, 1995). However, an alternative model ('GOANNA') based on ordination of family-level macroinvertebrate data (presence/absence and abundance) is also being developed as part of R&D for MRHI (Stockwell and Faith).

The accuracy with which a model can a) relate environmental data to macroinvertebrate community assemblages (*viz.* classification groupings), and b) predict community composition from environmental data, will depend very much on the accuracy with which community composition at each site has been characterised during the data acquisition phase (this includes field and laboratory procedures used to collect data from reference and test sites).

QA/QC has identified various potential sources of error associated with the data acquisition phase (i.e. operator efficiency during sample collection, sorting efficiency of live-sorted samples, precision and accuracy of laboratory sub-sampling devices,

accuracy of laboratory subsample sorting, taxonomic identifications). The level of errors in these various stages is not yet known, but will be revealed following ongoing QA/QC programs designed to look specifically at these issues. Remedial action, such as additional training and modification of approaches/techniques/methods will be used to reduce error rates that appear to be unacceptably high. However, the problem is knowing the level at which an error rate becomes unacceptable. It is unreasonable to expect no errors during the data acquisition phase. Even if this could be achieved through extensive, time consuming and expensive protocols, it is unlikely to result in a model that is more accurate than one produced at less cost, but with a known and acceptable error rate. Conversely, a data set with a high error rate will result in poor relationships between macroinvertebrate assemblages and environmental parameters and is unlikely to provide good predictive capability. The challenge, therefore, is to identify the thresholds at which the error rate begins to compromise the robustness, accuracy and predictive capability of the model and to ensure that the error rate in the data acquisition phase remains within the identified limits. An error rate lower than necessary indicates that procedures could be modified and QA/QC criteria relaxed. This also may result in a cost saving.

Approach

It is proposed to test the effects of different QA/QC criteria on agency models developed from two existing MRHI data sets; a taxa-poor fauna from south-western WA and a relatively taxa-rich fauna from the ACT. RIVPACS-type models have already been applied to both data sets by independent workers (R&D consultant in collaboration with the respective MRHI agency). Development of the alternative ordination-based model ('GOANNA') is not as far advanced as that for the RIVPACS-based models. Consequently, we have only considered at this stage the effects of different QA/QC criteria on RIVPACS-based models. An approach in which the effects of different QA/QC criteria are also tested on the GOANNA model will be pursued.

Errors at different stages of the data acquisition process, of course, are cumulative, resulting in an "**overall error rate**". The error rates at each individual stage (i.e. live sorting, laboratory subsampling and taxonomy) will be identified from the current external QA/QC programs from which an overall error rate may be determined and related back to the present R&D. This R&D will test different hypothetical overall error rates on model development and accuracy of outputs. Once an acceptable overall error rate has been identified, the QA/QC criteria can be stipulated to ensure data acquisition falls within the required criteria.

The following is the proposed iterative procedure that will be followed to test the effects of different error rates on the above-mentioned models. These models are based on data which contain an unknown level of sampling and processing error. However, for the purposes of this study it must be assumed that the data are error free and may be used as a bench-mark against which the effects of 'simulated errors' may be assessed.

For each of the two data sets, it is intended to reclassify randomly-selected test sites, that have been "altered", into the UPGMA classification of reference sites upon which the models are based, and to determine rates of site misclassification. The "altered" conditions will be data manipulations that simulate errors in the data acquisition phase. Two main errors are envisaged:

- loss of taxa (especially rare/low occurrence taxa), and
- alterations to community structure (through under/over estimates of abundances of common and rare taxa respectively).

It is envisaged that MRHI will progress to develop models based on rank or relative abundance of taxa (i.e. both RIVPACS-type models and GOANNA, the ordination-based model referred to above). These models may be more precise than those developed on presence/absence data, providing greater site/group resolution in the classification and greater sensitivity to detecting impact. Therefore, recommendations on QA/QC criteria based on presence/absence data sets alone may be misleading - particularly if they recommend to agencies a lowering of the rigour, when in fact, rank abundance data are more sensitive to error rates and require more precise data. Thus, models derived from both data types need to be tested.

Presence/absence data

As a hypothetical example of the proposed approach, a model developed from a data set of 100 reference sites will be used, from which 10 sites are selected at random to act as test sites. At each test site, taxa are sequentially deleted (i.e. 1, 2, 3, 4, ... etc taxa) and the effect on the model (classification structure, site groupings and observed/expected taxa ratios) iteratively assessed.

There are inherent problems in the above approach. In the sequence of removing the 10 test sites from the data set, altering the data at these sites, then reclassifying the whole data set (90 remaining reference sites + 10 test sites (all altered)), it is likely that the resulting classification will be very different from the original. Hence, there is no way to know the reason why the 10 test sites have misclassified. Therefore, the approach needs to be able to add a site to test for misclassification without altering the original classification.

It is proposed that the hypothetical model based on the 100 reference sites would be left in its original form but 10 test sites would be selected at random and using one test site at a time, the whole data set is reclassified using each test site as an additional sample (i.e. 100 reference sites (unchanged) + 1 test site (changed: taxa sequentially deleted)). Two criteria will be used to determine if a test site has reclassified correctly:

- a) whether the test site classifies with its original site (i.e. site 'A1-unchanged' and 'A1-changed' classify as a pair), and,
- b) whether the test site classifies into the same UPGMA/TWINSPAN site grouping as the original site (i.e. site 'A1-changed' does not pair with 'A1-unchanged' but falls within the same classification group).

This exercise is repeated 10 times, once for each of the 10 test sites, and the percentage rate of misclassification recorded (i.e. not pairing and not classifying into the same group). The process is then conducted for various hypothetical overall error rates (i.e. 1, 2, 3, 4, etc taxa missed). This will result in reclassification of the 10 test sites for each level of deletion, providing percentage error rates for 1, 2, 3, 4 etc missed taxa and will determine at what level the structure of the classification is likely to deteriorate. In addition, Bray-Curtis dissimilarity between the test site and the original unchanged site will be determined at each level of deletion, providing an indication of the level of dissimilarity at which misclassification occurs.

The taxa to be deleted will be selected in order of increasing relative abundance (although the classifications/models will be on presence/absence data), with the least abundant taxa at the test site being deleted first and so on. This taxa-deletion procedure is based is on the assumption that low abundance (*viz.* rare) taxa are most likely to be missed in the sample processing phase (picking/sorting/subsampling), although this will not always be the case (i.e. large and obvious taxa such as Decapoda and Odonata). An alternative approach is to base deletions on the results of external QA/QC checks. For example, very small cryptic taxa, that are not necessarily rare, but are regularly overlooked in live-sorting may be targeted for deletion. This possibility will depend on the identification of such groups in the external QA/QC program.

An additional stage in the testing process will involve calculating the O/E ratios for each of the 10 test sites, at each level of taxonomic exclusion (i.e. 1, 2, 3 etc taxa deleted), by running the test sites through the original model (100 reference sites). It is possible that a test site will classify to the correct UPGMA group, but the O/E ratio may be reduced to such an extent that the test site appears as "impacted" (e.g. O/E ratio < 0.8). This would indicate a failure, i.e. a reference site is inadequately sampled/processed so that it appears as impacted.

A final stage in the analysis will be to use the two models to test additional samples that have had processing errrors corrected by external QA/QC. Paired samples, comprising uncorrected data (which include agency sampling/processing errors) and corrected data (which have had processing errors rectified during QA/QC) will be tested through the appropriate model and UPGMA group affinities and O/E ratios determined for each pair. This will indicate the effects of actual errors in data collection on model predictions. Data that have been subjected to QA/QC were collected in the third and fourth MRHI sampling rounds. They represent a limited selection from the same reference sites on which the current models are based, which were sampled in the first MRHI sampling round. The first round data will not have been subjected to QA/QC assessment. As a result, the models will incorporate inherent sampling/processing errors. Therefore, the uncorrected, rather than the QA/QC-corrected data may be more similar to data from the same reference sites sampled in the first sampling round. The implications of this will be assessed during the analyses.

The analyses described above will be conducted on model "templates" developed by Richard Norris (i.e. flexible UPGMA using Bray-Curtis dissimilarity and removal of 'infrequently-occurring' taxa) to maintain maximum relevance.

Rank/relative abundance data

Initially, the same procedure as described above will be invoked (same data set and same 10 test sites), except that all data will be rank or relative abundance (i.e. UPGMA classification and the 10 test sites). However, an added source of error is the alteration of the relative abundance order, either through under picking or over picking of certain taxa. There are several possible approaches by which relative abundance of taxa within each test site may be systematically altered.

1. Total reverse of the rank order, whereby, for a community of 20 taxa, taxa number 1 becomes 20, 2 become 19, 3 becomes 18 etc.
2. Total randomisation of the relative abundance order.
3. Intermingling of ranks, whereby the order for 20 taxa would be changed from 1, 2, 3, 4, 5, ...20, to 1, 20, 2, 19, 3, 18, 4, 17 etc.
4. Elevation of one taxon, for sake of argument the middle ranked taxon in each sample, to the top rank (e.g. taxon 10 becomes 1).

An additional approach is to base any alterations in rank order on the spectrum of errors commonly encountered in external QA/QC evaluation of the MRHI data. The actual method by which this is achieved will depend upon empirical data which will become available from the QA/QC program.

At this stage, it is not possible to progress further than testing reclassification to a UPGMA classification derived from rank abundance data because there is no suitable mechanism for developing predictive capability (*viz.* O/E ratios) from abundance data. It is suspected that if classification/model groupings are developed on the basis of abundance data (rank or raw) that the outputs which only report presence/absence will result in over- or under-predictions on O/E ratios.

The end result of the above analyses will be the identification of QA/QC acceptance criteria for overall error rates in MRHI sample processing. The criteria will be the dissimilarity values derived between altered and unaltered data and the thresholds at which misclassification has an unacceptable effect on model structure and outputs, for presence/absence and relative abundance data sets. These threshold values will be used to modify interim criteria set by Storey & Humphrey (in prep.).

Basis of recommendations on QA/QC criteria

Results of manipulations made to model input data

The results of the above analyses will be used to recommend QA/QC criteria for number of taxa missed in the whole sample processing stage as well as to provide a 'first-pass' estimate of the critical error about which models are sensitive to alterations in community structure (rank abundances).

At present, the cumulative error represented in the QA/QC criteria can only be accounted for in operator errors arising at sample sorting and identification procedures of the MRHI protocol (because these are the only aspects of the protocol currently being audited). Results of external QA/QC to date indicate, generally, very high

accuracy of agency-identified specimens. It is the expectation, therefore, that a very large percentage of the error arising in the whole sample processing will be associated with the sorting of samples, particularly, live-sorting. Thus, thresholds of unacceptable deterioration in data quality based on presence/absence data may be related back to agency sorting procedures in the internal and external QA/QC procedures that determine the number of taxa missed by agencies during live-sorting. Bray-Curtis dissimilarities between live-picked and corrected data, considered together with the number of taxa missed during sample processing, may be related to threshold values derived from the above R&D to determine if agencies are achieving the desired data quality.

Thresholds of unacceptable deterioration in data quality for relative or rank abundance data also will be represented by a dissimilarity criterion determined from the above R&D. This value will be compared with a dissimilarity derived from QA/QC procedures for agency sorting to assess data quality by comparing community structure of the agency-sorted sample with that representing the best estimate of the entire sample (agency-sorted + residue component).

Additional basis for setting QA/QC criteria

It is anticipated that the current MRHI models will be tolerant to quite a degree of error introduced at the input stage. With future MRHI model refinement, however, through greater site density, temporal replication and improved model sensitivity, models may be developed that are more responsive to the quality of input data. Hence, it is proposed to base future QA/QC acceptance criteria on the most conservative of 2 sets of results. Firstly, the thresholds derived from the above R&D, and secondly, the actual sorting errors (with a measure of their variability) derived from present agency sorting - as measured in the external QA/QC programs. If thresholds identified from the above R&D are more rigorous than levels currently being achieved by S/T agencies, then quality of data acquisition will have to be improved. It will not be the intention to recommend QA/QC acceptance criteria that might lead agencies to 'relax' their present procedures for sorting.

Summary

External QA/QC checks of sample processing by S/T agencies will evaluate the efficiency of field live sorting and laboratory sub-sampling and subsequent sorting. Using the QA/QC thresholds derived from the above R&D, it will be possible to assess the effects of the observed error rates associated with each protocol on model group membership and O/E outputs. QA/QC acceptance criteria will be recommended on the basis of threshold values derived from this R&D. Assuming S/T agencies achieve the QA/QC criteria, errors in model outputs resulting from reduced data quality will be minimised.

References

Storey, A.W & Humphrey, C.L. (in prep.) Quality Assurance/Quality Control in Rapid Bioassessment Projects with preliminary guidelines for implementation in the Australian Monitoring River Health Initiative. Milestone Report to LWRRDC.

Wright, J. F. (1995) Development and use of a system for predicting the macroinvertebrate fauna in flowing waters. **Australian Journal of Ecology** 20:181-197.

ATTACHMENT 4

External QA/QC of MRHI agency subsampling and sorting procedures
(Results to 30 May 1996)

by

Lisa Thurtell & Chris Humphrey
ERISS, Locked Bag 2, Jabiru, NT 0886

External QA/QC of MRHI agency subsampling and sorting procedures

(Results to 30 May 1996)

Introduction

For the Monitoring River Health Initiative (MRHI), quality assurance and control (QA/QC) for components of the bioassessment protocol have been categorised into 'external' and 'internal' - the former being carried out by external agencies and 'internal' being the responsibility of the individual State and Territory agencies (hereafter S/T agencies). The rationale for separate external and internal QA/QC is provided in Storey and Humphrey (1996a). A critical component of the MRHI protocol requiring external QA/QC is the efficiency of S/T agencies in procedures used to sort samples. The ERISS has been commissioned by the National River Health Program committee to conduct such an external audit.

Two sorting protocols are currently employed by agencies, either field sorting of live samples or laboratory sorting of preserved samples. The current external QA/QC study aims to cross-check community structure reported by agencies (from live-sorting or laboratory subsampling) against that representative of the whole sample. At the onset, it was recognised that most of the error arising in agency sorting procedures would arise in live sorting - as opposed to laboratory subsampling and sorting of preserved samples. Consequently, priority was given to processing live-sorted samples and initial results reported below pertain only to these samples; results arising from QA/QC of laboratory-sorted subsamples will be reported at a later date.

Methods for QA/QC of agency live-sorting procedures

Selection and procurement of residues for processing

Agencies that incorporate live-sorting in the MRHI protocol are VIC, NSW, QLD, WA and TAS. For external QA/QC, a small percentage of the residues remaining after field sorting of samples from rounds 3 and 4 was retained by agencies and preserved for later laboratory processing. These samples were randomly selected on the basis of bioregion, catchment, habitat and/or operator in such a manner that the sorting operator was only made aware of which residues were required for QA/QC processing *after* field sorting was completed. For the current QA/QC, this selection procedure was carried out by way of sealed envelopes sent from the external auditor (ERISS) to the MRHI agency operators, the contents of which indicated whether or not the samples were required for processing (5% of samples). For the most part, a sealed envelope has been associated with, and has accompanied, every sample sorted in the field since August 1995. Sample identity (location and habitat) was labelled on the outside of the envelope and after completion of field sorting, the operator opened the envelope to determine whether preservation of the sample residue was required.

Appendix 1 lists the sites and habitats from which agency residues were requested, the percentage requested from the total number of habitats and current status of the requested residue. The live-sort component of some residues has also been received and those samples processed/identified in entirety (live-sort and residue) are indicated (Appendix 1). Agencies which are sorting preserved subsamples were given a list of residues to retain from rounds 3 and 4 (Appendix 1). These residues are being forwarded to ERISS following removal and processing of required subsamples by agencies.

Processing of residues and analysis of data

For assessment of field live-sorting efficiency, a subsample of the residue was taken using a modified Marchant subsampler (see Storey and Humphrey, 1996b) and its contents sorted and identified. The size of the subsample was such that its total invertebrate count was equivalent at least to that live-sorted in the field. The aim in assessing field live-sorting efficiency is to compare the live-sorted component with an equivalent-sized component representative of the *whole* sample. (This is an analogous procedure to that used in the assessment of laboratory subsampling and sorting of preserved samples - see Storey and Humphrey (1996a).) For the present report, analyses of data with and without 'rare' taxa were conducted and results compared.

To standardise data accordingly, the following steps were carried out (after Storey and Humphrey, 1996a); for ease of understanding, reference is made to the relevant calculations in the worked example shown in Table 1:

(i) From the agency live-sort taxa abundance list (step (1), Table 1), a 'subsample' equivalent in percentage terms to that taken of the agency residue (step (2)), is drawn by direct proportion (step (3)). (Abundances are rounded to the nearest whole animal; thus a value ≥ 0.5 is equivalent to at least one animal, a value < 0.5 is equivalent to zero animals.)

(ii) The community structure of the subsampled live-sort and residue components are then combined to derive a best proportional estimate of the whole sample (step (4): WS_n, where n = size of subsample in percentage terms). Proportional abundances of taxa present in WS_n are further scaled down so the total invertebrate count of WS_n is the same as that of the live-sorted component (step (5)). (Again, abundances are rounded to the nearest whole animal.)

(iii) For QA/QC analyses involving removal of rare taxa: only those taxa are included whose combined abundance in WS_n and live-sort component contribute $> 1.0\%$ of the total abundance of WS_n and live-sort component combined (steps in Table 1 involving 'Common taxa only').

Interim QA/QC criteria to use in the assessment of field live-sorting efficiency follow from Storey and Humphrey (1996a). In general, preliminary results from associated R&D evaluating the efficiency of laboratory subsampling devices (Storey and Humphrey, 1996b) have been used to set criteria to apply to field and laboratory sorting procedures using presence/absence data. However, for relative abundance data, liberal criteria were set on the basis that if the thresholds could not be met, it would not be possible to use the data in models based upon rank abundance. Thus:

- the number of taxa encountered in the live-sort component must lie within 20 percent of the number of taxa recorded in the best proportional estimate of the whole sample (WS_n);
- the community similarity index (Bray-Curtis) comparing the live sort component and best proportional estimate of the whole sample (WS_n) must be at least 50 percent for relative abundance data and 80 percent for presence-absence data;
- a Spearman Rank Correlation comparing the community structure of the live-sort component and best proportional estimate of the whole sample (WS_n) must be significant at $P < 0.05$.

Results

Sorted and identified residues and corresponding live-sort components from 36 agency samples, representing 4 habitats, have been used in a preliminary assessment of the representativeness of field live-sorting. Results of QA/QC analyses are presented in Tables 2 and 3 and Figure 1. (Results of analyses conducted for the worked example shown in Table 1

can be found in Table 2, Macrophyte sample 4 from QLD.) Note that analyses based upon relative abundance data that are presented in Tables 2 and 3 employ a *dis*similarity index (the inverse of the similarity criteria described in the Methods above). Results may be discussed according to presence/absence and relative abundance data as follows:

Presence/absence data

Two taxa 'presence/absence' criteria have been developed for QA/QC assessment of live-sorting procedures: (i) taxa in the live-sort component must lie within 20 percent of the number of taxa recorded in the best proportional estimate of the whole sample ('number of taxa' criterion); and (ii) the Bray-Curtis similarity index comparing the same 2 components must be at least 80 percent ('shared taxa' criterion).

The 'number of taxa' criterion is analogous to a RIVPACS observed/expected (o/e) taxa ratio. Not only is the 20% acceptance level achievable in laboratory subsampling and sorting but an o/e ratio outside of this range, *sensu* RIVPACS, would probably and intuitively be regarded as unacceptable - particularly for the lower threshold. When all taxa are included for each sample, 36% of samples failed to meet this criterion whilst with inclusion of only common taxa, there was a 25% failure rate (Table 2, Fig. 1). For 'common taxa', best results followed the order riffle> edge> macrophyte> pool (Table 3, Fig. 1).

With the 'shared taxa' criterion (similarity index), 92% of samples failed to meet the acceptance criterion when all taxa were included in QA/QC analyses. However, with inclusion of common taxa only, there was a substantial improvement in results such that only 22% of samples failed to meet the criterion (Table 2). Best results with and without rare taxa were riffle and edge habitats (Table 3).

Relative abundance (community structure) data

Two taxa 'relative abundance' criteria have been developed for QA/QC assessment of live-sorting procedures: (i) the Bray-Curtis similarity index comparing the live-sort component and best proportional estimate of the whole sample must be at least 50 percent; and (ii) a Spearman Rank Correlation comparing the community structure of the same 2 components must be significant at $P < 0.05$.

With analyses conducted using the Bray-Curtis similarity index, 47% of samples failed to meet the acceptance citerion whilst with inclusion of only common taxa, there was only a marginal improvement (44%) (Table 2). Using Spearman rank correlation, 50% and 97% of samples failed the criterion when rare taxa were included and excluded respectively (Table 2). (Significance levels for this correlation are dependent upon sample size and hence removal of rare taxa resulted in a substantial increase in the failure rate of QA/QC samples. Even so, for 75% of samples, the correlation value (R) was reduced after exclusion of rare taxa, Table 2). A majority of samples from riffle habitat only, met both of the relative abundance criteria when all taxa were included in the analyses (Table 3).

Discussion and conclusions

QA/QC acceptance criteria

- Presence/absence data: It is not possible at this stage to assess the significance of relatively high failure rates for the 'number of taxa' criterion applied to agency samples (reported above). Acceptance criteria are currently being refined for this critical end-point by way of (i) classification and model outcome (o/e) results after error simulations performed on existing agency data sets, and (ii) classification and o/e results following incorporation of "corrected" agency data (from external QA/QC) (see Storey and Humphrey, 1996c). Further, QA/QC results from live-sorting would need to be compared

with those from agencies that have sorted preserved samples; results from the latter would presumably set an 'upper-level' benchmark for acceptance criteria applied elsewhere.

The acceptance criteria applied to agency data above, however, were based upon some empirical logic and are regarded as not unreasonable. There is some initial concern, therefore, in high failure rates arising after application of preliminary/interim criteria, i.e. 20-60% of samples (according to habitat) lying outside the 20% live-sort/whole-sample estimate criterion.

For the 'shared taxa' criterion (similarity index), results are only encouraging for data which have had rare taxa removed (when only 22% of samples failed to meet the criterion). However, as discussed below, removal of rare taxa in this manner, prior to modelling, might not be warranted nor justified.

- Relative abundance data: Recovery of relative or rank abundance data after live-sorting is particularly poor. Despite liberal, 'worst-case', acceptance criteria, 33-60% of samples (depending upon habitat) failed the dissimilarity criterion. Recovery of relative/ rank abundance data followed the order: riffle > edge > macrophyte > pool; it is unlikely that data from habitat other than riffle could be used in models based upon rank abundance.
- Rare taxa: Exclusion of rare taxa did not substantially improve results with the exception of the 'shared taxa' criterion applied to presence/absence data. In any case, culling of rare taxa is currently being conducted for agency models on the basis of 'site occurrence'. Additional culling on the basis of abundance in samples, therefore, might not be warranted.

Nature of errors

- It was apparent that QA/QC results from 'experienced' agencies (e.g. VIC) were no better than those from 'inexperienced' agencies (e.g. QLD) (Table 2). Further consultation needs to be carried out with agencies to determine whether experience of individual field-sorting staff can account for variation in the results reported above. However, it is unlikely that this factor could account for significant variation given that the quality of results appeared to be mostly dependent upon the type of habitat sampled.
- Live-sorting over- and under-estimated taxa richness in equal proportion (Table 2, Fig. 1). Hence, there appears to be no consistency in errors arising in the live-sorting procedure.
- Small taxa including chironomids, other diptera and oligochaetes, and many cryptic families (especially hydroptilids), were consistently under-represented in all live-sorted samples. Conversely, large and conspicuous 'rare' taxa, as well as mobile taxa (e.g. mites), were generally over-represented in samples. Results were particularly poor for habitats with high amounts of detritus associated with samples, i.e. edge, macrophyte and pool. Small and cryptic taxa simply become 'lost' amongst detritus.
- Some live-sorted samples were represented by fewer than 100 animals (Table 2), despite high abundances of invertebrates found in corresponding residues. The number of taxa recorded in a sample will be, to a large extent, proportional to the number of invertebrates removed from the sample and hence very small or, for that matter, very large samples may seriously bias estimates of taxa number.

  (For reference sites where animals in samples are numerous but where the live-sort component has few animals, different QA/QC approaches may be needed. In this case, a future requirement in QA/QC analyses might be the comparison of small (< 100 animals) agency live-sort components with whole-sample estimates of a pre-set minimum size. Thus, instead of the present method of scaling down the whole-sample estimate to the sample size of the live-sort component, this latter component instead would need to be scaled up to the same minimum sample size as the whole-sample estimate. In doing so a

penalty would be incurred by agencies because taxa in the live-sort component would be fewer than those in the whole-sample estimate. The same QA/QC approach might also be considered for extraordinarily large live-sort samples where a pre-set maximum sample size is applied.)

Concern for MRHI models

- The greatest concern for the MRHI in the results reported above is potential for irretrievable loss of quality control and "contamination" of current models with low quality data. Agencies live-sorting samples in the field have not been preserving residues and hence it will be impossible to correct past, ongoing and future errors detected in QA/QC programs.
- It is unlikely that agencies using live-sorting procedures will be able to develop rank abundance models except, perhaps, for riffle habitat.
- The early impression gained from conducting the external QA/QC program is that the live-sorting procedure used by MRHI agencies is 'undisciplined' and subject to wide-ranging biases of individuals. Such biases appear to be endemic amongst all agencies, experienced and inexperienced, and the impression reached is that extensive training of agency staff in live-sorting procedures would not be particularly fruitful without significant modifications to the sorting technique. There also appear to be attitudinal problems of agency staff in field live-sorting that are evident in the data - especially low sample size (e.g. 31 animals removed from a VIC riffle sample, Table 2). Otherwise, there appears to be a lack of suitable checks that might include, for example, a deliberate focus for some of the time allocation on small and cryptic taxa.

Possible remedial action for future MRHI sampling and sorting

Even on the basis of the early results reported above and without having undertaken R&D to assess the significance of seemingly 'poor' results as they affect MRHI models (Storey and Humphrey, 1996c), it would be difficult to justify continued live-sorting by MRHI agencies in the manner conducted to date. Even preservation of all future live-sort residues would not be a viable remedial option - live-sorting biases appear so widespread that most of the residues would need to be processed to correct errors.

A preliminary exercise that should be conducted prior to further MRHI sampling is a cost-benefit analysis to compare the resource effort in sorting of preserved vs live samples e.g. ACT/SA vs other states. Of course, a component of the analysis would need to consider quality of data retrieved. If the difference was minimal, presumably further use of the live-sorting method would be unwarranted, particularly for detritus-rich habitat.

If it were concluded that live-sorting was the most cost-effective method of sorting MRHI samples, we would recommend that R&D be carried out quickly to develop less subjective procedures for live-sorting. Such procedures might include, for example:

- Removal of invertebrates from randomly-selected cells of a "gridded" tray, or from randomly-placed, raised perspex rings placed on the floor of the tray. These methods impose a discipline upon sorting operators by ensuring a concentrated and objective focus on the sample. An additional advantage of these methods would include recovery of absolute abundance data.
- Use of jeweller's vision visors to detect small and crytic taxa.
- Use of larger mesh sizes in the sampling protocol (to avoid small taxa).
- Specification of a minimum and maximum sample size.

References

Storey, AW & Humphrey, CL (1996a). Quality assurance/ quality control in rapid bioassessment projects with preliminary guidelines for implementation in the Australian Monitoring River Health Initiative. Milestone Report to LWRRDC, May 1996.

Storey, AW & Humphrey, CL (1996b). Assessment of the efficiency of four types of device for subsampling of aquatic macroinvertebrate samples: preliminary results. Milestone Report to LWRRDC, May 1996.

Storey, AW & Humphrey, CL (1996c). Proposed analytical approach to assess the effect of different QA/QC criteria for sampling efficiency on the robustness of MRHI models. Milestone Report to LWRRDC, May 1996.

Table 1 Procedures used to compare live-sorted component with an equivalent-sized component representative of the whole sample (data from an agency macrophyte sample).

SAMPLE FRACTION	TAXON																							
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Tot
All taxa																									
(1) Live-sort	***1***	***7***	***8***	***2***	***10***	***3***	***15***	***4***	***1***	***3***	***17***	***3***	***2***	***1***	***1***										***78***
(2) Residue$_{10}$		5	27	34	6	8	20		5	4		1				84	3	7	6	2	1	1	2	2	218
(3) Live-sort$_{10}$		1	1		1		2				2														7
(4) WS$_{10}$ = (2)+(3)		6	28	34	7	8	22		5	4	2	1				84	3	7	6	2	1	1	2	2	225
(5) WS$_{10}$ * 78/225		***2***	***10***	***12***	***2***	***3***	***8***		***2***	***1***	***1***					***29***	***1***	***2***	***2***	***1***			***1***	***1***	***78***
Common taxa only																									
(1) + (5)	1	9	18	14	12	6	23	4	3	4	18	3	2	1	1	29	1	2	2	1			1	1	156
Taxa >1% [(1)+(5)]		√	√	√	√	√	√	√	√	√	√	√	√			√		√	√						
(1) Common		***7***	***8***	***2***	***10***	***3***	***15***	***4***	***1***	***3***	***17***	***3***	***2***												***75***
(5) Common		***2***	***10***	***12***	***2***	***3***	***8***		***2***	***1***	***1***					***29***		***2***	***2***						***74***

Table 2 Comparison of agency live-sorted component and best proportional estimate of the whole sample, with and without rare taxa removed, using Bray-Curtis dissimilarity index (on abundance and presence/absence (p/a) data), Spearman Rank Correlation and number of taxa.

State/ Habitat	N (live-sorted)	Including Rare Taxa				Rare Taxa Removed			
		Dissim. (p/a)	% Live-sort taxa	Dissim. (abund.)	Spearman (R)	Dissim. (p/a)	% Live-sort taxa	Dissim. (abund.)	Spearman (R)
QLD									
Edge 1	141	0.3125	120	0.4786	0.443*	0.1146	112	0.4511	0.324
Edge 2	782	0.2388	108	0.3678	0.637***	0.0357	92	0.3553	0.617*
Edge 3	150	0.4559	117	0.3872	0.283	0.2121	136	0.3569	0.471
Edge 4	146	0.1857	62	0.4034	0.539**	0.0313	93	0.3813	0.523
Edge 5	504	0.3487	82	0.5507	0.462*	0.0000	100	0.5335	0.446
Edge 6	97	0.6548	93	0.5440	-0.177	0.5417	83	0.5269	-0.136
Pool 1	114	0.3450	211	0.6228	0.372	0.2479	144	0.6055	0.162
Pool 2	235	0.2727	100	0.6026	0.478	0.1429	100	0.5930	0.548
Pool 3	145	0.2471	113	0.4483	0.704***	0.0000	100	0.4234	0.706*
Pool 4	154	0.4167	133	0.5635	0.447	0.0714	116	0.5442	0.327
Riffle 1	208	0.1176	130	0.2692	0.917***	0.0000	100	0.2556	0.734*
Riffle 2	204	0.3795	90	0.3105	0.565**	0.0385	92	0.2598	0.588
Riffle 3	414	0.3509	94	0.3867	0.385	0.1111	100	0.3663	0.517
Riffle 4	218	0.2968	95	0.5392	0.259	0.1071	78	0.5294	0.242
Macro 1	289	0.2471	113	0.2159	0.667**	0.0556	112	0.2000	0.605
Macro 2	221	0.3985	110	0.2739	0.206	0.1736	112	0.2668	0.325
Macro 3	106	0.2708	133	0.6698	0.483	0.2262	117	0.6651	-0.006
Macro 4	78	0.4187	93	0.6410	0.182	0.2500	100	0.6242	-0.127
Macro 5	64	0.4773	109	0.6535	0.121	0.3939	81	0.6423	-0.205
NSW									
Pool 1	121	0.5068	55	0.4215	0.218	0.1667	66	0.3739	0.681
Edge 1	65	0.4170	68	0.3835	0.167	0.1538	69	0.3387	0.209
Edge 2	142	0.2816	95	0.7352	0.428*	0.0333	93	0.7246	0.226
TAS									
Riffle 1	250	0.3324	92	0.3187	0.638**	0.0556	88	0.3035	0.552
Riffle 2	282	0.1900	80	0.4619	0.646***	0.0357	92	0.4475	0.378
Riffle 3	287	0.2424	81	0.4107	0.616**	0.0357	92	0.3975	0.300
Riffle 4	102	0.2929	81	0.3431	0.556*	0.1736	88	0.3267	0.295
Edge 1	145	0.3542	93	0.5017	0.455	0.1667	66	0.4872	0.500
Edge 2	109	0.3095	107	0.6600	0.484*	0.1556	111	0.6683	0.032
Edge 3	171	0.2874	68	0.5814	0.407	0.1071	78	0.5740	0.145
VIC									
Riffle 1	257	0.3500	160	0.6770	0.541**	0.0417	109	0.6681	-0.085
Riffle 2	31	0.3194	75	0.7049	0.206	0.3194	75	0.7049	-0.399
Riffle 3	123	0.2381	77	0.5403	0.608**	0.0000	100	0.5083	0.199
Riffle 4	50	0.2381	128	0.3265	0.669*	0.2381	128	0.3333	0.584
Edge 1	162	0.4118	100	0.5741	0.296	0.1736	88	0.5658	-0.080
Edge 2	134	0.2929	81	0.4889	0.523**	0.1099	92	0.4724	0.305
Edge 3	94	0.2721	106	0.4583	0.391	0.1987	108	0.4382	0.195

* P<0.05
** P<0.01
*** P<0.001

Table 3 Summary of agency results (all taxa/ common taxa only) according to QA/QC criteria.

Habitat	Sample size	Dissimilarity distribution for presence/ absence data (%)				Percent live-sort taxa: cumul. proportion (%) of samples lying within			Dissimilarity distribution for abundance data (%)			Spearman P < 0.05 (%)
		< 0.2	0.2-0.3	0.3-0.4	> 0.4	10%	20%	30%	< 0.4	0.4-0.5	> 0.5	
Riffle	12	16/ 84	42/ 8	42/ 8	-/ -	25/ 58	50/ 75	92/ 100	50/ 58	17/ 9	33/ 33	75/ 8
Edge	14	7/ 86	35/ 7	29/ -	29/ 7	45/ 36	73/ 64	73/ 73	21/ 36	29/ 28	50/ 46	54/ 9
Macro-phyte	5	-/ 40	40/ 40	20/ 20	40/ -	60/ 20	80/ 100	80/ 100	40/ 40	0/ 0	60/ 60	20/ 0
Pool	5	-/ 80	40/ 20	20/ -	40/ -	20/ 40	40/ 60	40/ 60	0 /20	40 /20	60/ 60	20/ 20

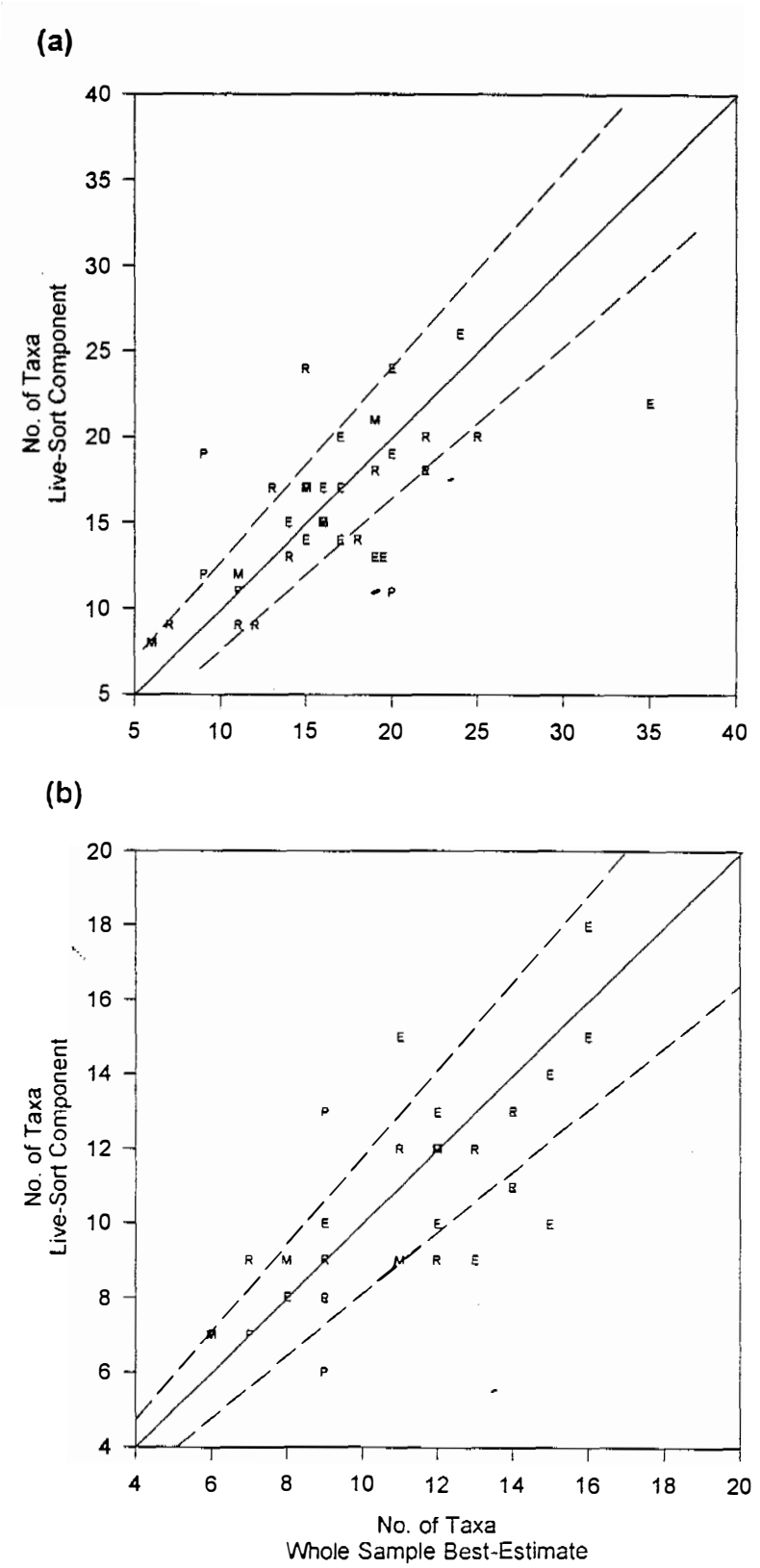


Figure 1: Relationship between number of taxa encountered in agency 'live-sort' and 'best proportional estimate of the whole sample' components, according to habitat with (a) all taxa and (b) rare taxa removed. Codes to habitats: R = riffle; E = Edge; P = Pool; M = macrophyte. Solid line indicates 1:1 ratio while samples with taxa in the live-sort component lying within 20 percent of the number of taxa recorded in the best proportional estimate of the whole sample are bound within the dashed lines.

Appendix 1 Requested Residues from MRHI Agencies - Current Status

State	Round	Site	Habitat	Received	% Subsampled	Sorted & ID	Livepick ID
TAS	3	D14 North Esk	Edge	yes	10	14-Feb	29-Jun
		A18 Staneley	Riffle	yes	30	26-Feb	29-Jun
		Dmon4 Great Forester	Riffle	yes	30	27-Feb	29-Jun
		Bmon4 Emu	Edge	yes	30	22-Feb	29-Jun
		B7 Keith	Edge	yes	20	29-Jun	29-Jun
		C15 Quamby	Riffle	yes	10	24-Feb	29-Jun
		Cmon4 Meander/Birralee	Riffle	yes	10	24-Feb	29-Jun
		D13 Nile	Riffle	yes	30	16-Feb	29-Jun
		B19 Leven/Blackmarsh	Edge	yes	10	13-Feb	29-Jun
		% habitat requested		**5**			
	4	Amon2 Little Henty	Edge	yes			
		Bmon2 Wilsons Ck	Riffle	yes			
		Cmon2 Western Ck	Edge	yes			
		Dmon1 Sth Esk	Edge	yes			
		A12 Ewart Ck	Riffle	yes			
		A15 Farrell Rt	Edge	yes			
		A15 Farrell RT	Riffle	yes			
		B10 Black Ck	Edge	yes			
		B10 Black Ck	Riffle	yes			
		B16 Flowerdale	Riffle	yes			
		C3 Lobster Rt	Riffle	yes			
		C13 Meander	Edge	yes			
		D5 North George	Edge	yes			
		D23 Little Forester	Riffle	yes			
		% habitat requested		**8**			
VIC	3	003809 Moleside Ck	Riffle	yes	10	19-Apr	May-96
EPA*		032900 Yackandandah Ck	Edge	yes	10	24-Apr	May-96
		033300 Kiewa R @ Bonegilla	Riffle	yes	20	18-Apr	May-96
		037100 Victoria R	Riffle	yes	5	17-Apr	May-96
		036400 Mitta Mitta R	Edge	yes	10	22-Apr	May-96
		00501O Shannassy R	Edge	yes	10	16-Apr	May-96
		297100 Cement Ck	Riffle	yes	20	26-Apr	May-96
		% habitat requested		**3**			
EPA**	4	003701 Eumerella R	Edge	no			
		032800 Running Ck	Riffle	yes			
		003811Glenelg R @ Dartmoor	Edge	yes			
		005012 Armstrong Ck	Edge	yes			
		003436 Woady Yaloak	Edge	yes			
		033600 West Kiewa R	Riffle	yes			
		033600 West Kiewa R	Edge	yes			
		003819 Grange Burn Ck	Riffle	yes			
		004982 Yarra R @ Big Peninsula	Edge	yes			
		004982 Yarra R @ Big Peninsula	Riffle	yes			
		% habitat requested		**13**			
Water	3	408202 Avoca R @ Ampitheatre	Pool	yes			
Eco		942668 Wonnangatta R @ Maroaka	Pool	yes			
		224206 Wonnangatta R @ Crooked	Riffle	yes			
		232211 Moorabool R W. Branch	Riffle	yes			
		223997 Tambo R ds Bark Sheds Ck	Riffle	yes			

		223998 Timbarra R @ Timbarra	Riffle	yes				
		231213 Lerderberg R	Riffle	yes				
		231230 Parwan Ck @ Rowsley	Pool	yes				
		230202 Jackson Ck @ Sunbury	Riffle	yes				
		230209 Barringo Ck @ Barringo	Pool	yes				
		230205 Maribyrong R @ Bulla	Pool	yes				
		404996 Ryans Ck	Riffle	yes				
		404214 Broken Ck @ Katamatite	Pool	yes				
		% habitat requested			6			
	4	224206 Wonnangatta R @ Crooked	Riffle	yes				
		223214 Tambo R us Smith Ck	Riffle	yes				
		404998 Moonee Ck us Lima East	Riffle	yes				
		231998 Werribee R @ Cobbledick F	Riffle	yes				
		224995 Wentworth R @ Jones Rd	Pool	yes				
		224995 Wentworth R @ Jones Rd	Riffle	yes				
		223210 Nicholson R @ Deptford	Pool	yes				
		232211 Moorabool R @ Mt. Doran	Pool	yes				
		230209 Barringo Ck @ Barringo	Pool	yes				
		230209 Barringo Ck @ Barringo	Riffle	yes				
		% habitat requested			7			
NSW	3	Hast 22	Edge	yes				
		Mann06	Riffle	yes				
		Hunt04	Riffle	yes				
		Shoa24	Edge	yes				
		Clyd31	Riffle	yes				
		Hawk08	Edge	yes				
		Towa05	Riffle	yes				
		Snow05	Riffle	yes				
		Lach01	Macro	yes				
		Bega07	Riffle	yes				
		Murr11	Edge	yes				
		Murr15	Edge	yes				
		Murr22	Riffle	yes				
		Clar19	Edge	yes				
		Clar30	Riffle	yes				
		Macq10	Macro	yes		20	08-Feb	01-Mar
		Rich01	Pool	yes		100	08-Feb	04-Mar
		Rich01	Edge	yes		30	07-Feb	04-Mar
		% habitat requested			4			
	4	Clar03	Edge	No				
		Bell15	Edge	No				
		Hast18	Edge	No				
		Hunt10	Riffle	No				
		Shoa05	Riffle	No				
		Bega04	Riffle	No				
		Snow08	Riffle	No				
		Murr27	Riffle	No				
		Murr27	Edge	No				
		Bidg10	Log	No				
		Gwyd10	Macro	No				
		Clar10	Riffle	No				
		Macl06	Riffle	No				
		Mann10	Edge	No				

		Hawk01	Edge	No				
		Clyd18	Edge	No				
		Towa06	Edge	No				
		Murrm5	Edge	No				
		Lach09	Edge	No				
		Darl02	Edge	No				
		Rich06	Riffle	No				
		% habitat requested			5			
ACT	3	Murrumbidgee 531	Edge	No				
		Limestone 32br	Riffle	No				
		Goodradigbee 30br	Riffle	No				
		Ryries 15m	Edge	No				
		Murrumbidgee 551	Edge	No				
		Murrumbidgee 16m	Riffle	No				
		% habitat requested			5			
	4	Blue Bull Ck 36co	Riffle	No				
		Oaks Ck 23br	Edge	No				
		Kybeyan R 40co	Riffle	No				
		Numeralla R 38co	Edge	No				
		Murrumbidgee 8c	Riffle	No				
		Murrumbidgee 631	Edge	No				
		% habitat requested			5			
QLD	3	Palmer R @ Drumduff Rd	Riffle	No				
		Rifle Ck @ Font Hills	Pool	No				
		Hann R @ Cape York Rd	Riffle	No				
		E. Normanby @ Dev Rd	Bank	No				
		E. Normanby @ Dev Rd	Pool	No				
		Normanby R @ 12 Mile Hole	Macro	No				
		Peets Ck @ Causeway	Riffle	Yes				
		Babinda Ck @ Babinda	Pool	No				
		Babinda Ck @ Babinda	Macro	No				
		Sth Johnstone R @ Corsi's	Riffle	Yes				
		Taylors Ck @ Warraker	Bank	Yes				
		Nth Johnstone R @ Malanda Falls	Pool	No				
		Ithaca Ck @ Clarks Track	Macro	Yes				
		Gowie Ck @ Abergowrie	Pool	Yes				
		Millstream Ck @ Diversion Weir	Macro	Yes				
		Flinders R @ Walkers Bend	Bank	No				
		Corella R @ Lake Corella	Riffle	Yes				
		Corella R @ Lake Corella	Pool	Yes				
		Porcupine Ck @ Mt. Emu Plains	Macro	Yes				
		Wyandotte Ck @ Wyandotte	Riffle	Yes				
		Burdekin R @ Blue Range	Bank	Yes				09-May
		Pelican Ck @ Kerale	Pool	Yes				
		Chinaman Ck @ Hydrosite	Macro	Yes				
		Gregory R @ Collingvale	Riffle	No				
		Middle Lethe Brook	Bank	Yes		10	06-Mar	16-May
		Gregory R @ Collingvale	Pool	Yes				
		Middle Lethe Brook	Macro	No				
		O'Connell R @ Caping Siding	Bank	Yes				
		St Helens @ Brothwells	Pool	Yes				
		Cattle Ck @ Williams Rd	Macro	Yes				
		Funnel Ck @ Main Rd	Pool	Yes		5	01-Apr	20-May

	Dawson R @ Taroom	Macro	No			
	Burnett R @ Mt Lawless	Riffle	Yes	10	19-Mar	16-May
	Burnett R @ Mt Lawless	Pool	Yes	5	20-Mar	15-May
	Burnett R @ Eidsvold	Bank	No			
	Kolan R @ Bucca Xing	Macro	Yes	5	09-Apr	14-May
	Amamoor Ck @ Zachariah	Riffle	Yes	10	11-Apr	10-May
	Mary R @ Fishermans Pocket	Bank	Yes	10	11-Apr	13-May
	Mary R @ Home Park	Pool	Yes	100	10-Apr	14-May
	Obi Obi Ck @ Alpin Rd	Macro	Yes	5	25-Mar	16-May
	Coochin Ck @ Bruce Hwy	Macro	Yes	5	11-Apr	09-May
	Logan R @ Rathdowney	Riffle	Yes	10	01-Apr	20-May
	Logan R @ Yarrahapplnl	Bank	Yes	10	11-Mar	15-May
	Brisbane R @ Savages Xing	Riffle	Yes	10	14-Mar	20-May
	Brisbane R @ Xing 26	Bank	No			
	Brisbane R @ Xing 26	Macro	Yes	10	13-Mar	20-May
	Coomera Ck @ Tuckers Lane	Riffle	No			
	Tallebudgera Ck	Bank	Yes	50	21-Mar	13-May
	Coopers Ck @ Boolloo Boolloo	Pool	No			
	Warrego R @ Wyandra	Bank	Yes	5	22-Mar	16-May
	Macintyre R @ Goondiwindi	Macro	Yes	10	26-Mar	15-May
	Condamine R @ Chinchilla	Pool	Yes	5	26-Mar	14-May
	Condamine R @ Chinchilla	Bank	Yes	10	12-Mar	07-May
	% habitat requested			**7**		
4	Rifle Ck	Bank	No			
	Mitchell R	Macro	No			
	East Normanby R	Pool	No			
	Normanby R	Bank	No			
	Fishery Fall Ck	Bank	No			
	Nth Johnstone	Bank	No			
	Gowrie Ck	Riffle	No			
	East Barrata Ck	Pool	No			
	Broken R	Pool	No			
	Broken R	Macro	No			
	Proserpine R	Macro	No			
	Boulder Ck	Riffle	No			
	Pioneer R	Bank	No			
	Camarvon Ck	Riffle	No			
	Baffle Ck	Bank	No			
	Baffle Ck	Pool	No			
	Burnett R	Macro	No			
	Sth Maroochy R	Bank	No			
	Coochin Ck	Pool	No			
	Dumaresq R	Pool	No			
	Condamine R	Riffle	No			
	Warrego R	Pool	No			
	Warrill Ck	Pool	No			
	Albert R	Pool	No			
	Currumbin Ck	Pool	No			
	Mitchell R	Pool	No			
	Palmer R	Riffle	No			
	Normanby R	Macro	No			
	Hann R	Riffle	No			
	Thiaki Ck	Pool	No			
	Gowrie Ck	Riffle	No			
	Gowrie Ck	Bank	No			

		Haughton R	Bank	No	
		Haughton R	Bank	No	
		Proserpine R	Riffle	No	
		O'Connell R	Macro	No	
		Finch Hatton Ck	Riffle	No	
		Finch Hatton Ck	Pool	No	
		Crinum Ck	Macro	No	
		Dawson R	Bank	No	
		Burnett R	Bank	No	
		Tewah Ck	Pool	No	
		South Maroochy R	Riffle	No	
		MacIntyre R	Bank	No	
		Dumaresq R	Riffle	No	
		Balonne R	Macro	No	
		Brisbane R	Bank	No	
		Burnett Ck	Macro	No	
		Tallebudgera Ck	Riffle	No	
		% habitat requested			**10**
SA	3	Yankalilla Ck @ Main South Rd	Edge	No	
		Scotts Ck @ Scotts Bottom	Riffle	No	
		Sturt R @ Sturt Rd	Macro	No	
		Gawler R @ Gawler Junction	Pool	No	
		Light R @ Kapunda Bridge	Macro	No	
		Hill R nr Andrews	Pool	No	
		Alligator Ck @ Alligator Gorge	Edge	No	
		Rocky R nr NP Headquarters	Riffle	No	
		Middle River @ Western River Rd	Pool	No	
		Cooper Ck @ Embarka WH	Edge	No	
		Burra Ck @ World's End	Pool	No	
		Truro Ck	Edge	No	
		Mt. Barker Ck us Mt. Barker Springs	Riffle	No	
		Mt. Barker Ck us Mt. Barker Springs	Macro	No	
		Drain M @ Penola-Robe Rd	Edge	No	
		% habitat requested			**5**
	4	Myponga R	Macro	No	
		First Ck @ Waterfall Gully	Riffle	No	
		Torrens R @ Windsor Grove	Macro	No	
		North Parra R @ Rowland Flat	Riffle	No	
		Willochra Ck S of Partacoona	Edge	No	
		Oratunga Ck @ Third Spring	Pool	No	
		Todd R @ Koppio	Edge	No	
		Todd R @ Koppio	Riffle	No	
		Wilson R S of Penneshaw	Edge	No	
		Artimore Ck @ Nildottie Springs	Edge	No	
		Artimore Ck @ Nildottie Springs	Macro	No	
		Morambo Ck @ The Gap	Pool	No	
		Eight Mile Ck	Riffle	No	
		Eight Mile Ck	Macro	No	
		% habitat requested			**5**
WA	1	MUR06	Channel	Yes	
Murdoch		MUR17	Riffle	Yes	
Uni		MUR24	Macro	Yes	
		MUR24	Channel	Yes	

		MUR30	Channel	Yes		
		MUR37	Macro	Yes		
		MUR42	Riffle	Yes		
		MURM2	Channel	Yes		
		% habitat requested				
CaLM	3	CLM15	Macro	Yes		
		CLM15	Channel	Yes		
		CLM18	Riffle	Yes		
		CLM30	Channel	Yes		
		CLM35	Pool	Yes		
		CLM43	Macro	Yes		
		CLM46	Channel	Yes		
		CLMM3	Riffle	Yes		
		% habitat requested				
Uni WA		UWA2	Channel	Yes		
		UWA3	Channel	Yes		
		UWA9	Macro	Yes		
		UWA13	Macro	Yes		
		UWA17	Channel	Yes		
		UWA23	Channel	Yes		
		UWA24	Organic	Yes		
		UWA27	Macro	Yes	100	30-Apr
		UWA29	Macro	Yes		
		UWA34	Macro	Yes		
		UWA37	Riffle	Yes		
		UWA41	Channel	Yes		
		% habitat requested				
Edith	4	ECU06	Channel	Yes		
Cowan		ECU12	Channel	Yes		
Uni		ECU12	Macro	Yes		
		ECU12	Pool	Yes		
		ECU18	Channel	Yes		
		ECU18	Macro	Yes		
		ECU21	Channel	Yes		
		ECU30	Pool	Yes		
		ECU30	Channel	Yes		
		ECU37	Pool	Yes		
		ECU37	Channel	Yes		
		ECU37	Riffle	Yes		
		ECU42	Channel	Yes		
		ECU44	Channel	Yes		
		ECU46	Channel	Yes		
		ECUM5	Channel	Yes		
		ECUM6	Channel	Yes		
		% habitat requested				
NT		No Residues Requested to date				

* Envelopes not received in time for some catchments
** Only 5 catchments sampled in round 4 (78 habitats)