Australian Government

Department of Sustainability, Environment,
Water, Population and Communities
Supervising Scientist

internal report 580

Summary of presentations and key issues raised at the Biological Monitoring Review Workshop, October 2006, with status as of August 2010

Buckle D, Humphrey C &
Turner K (eds)

October 2010

(Release status – internal use only – non-sensitive)

This page has been left blank intentionally.

Summary of presentations and key issues raised at the Biological Monitoring Review Workshop, October 2006, with status as of August 2010

Edited by D Buckle, C Humphrey & K Turner

Supervising Scientist Division

GPO Box 461, Darwin NT 0801

October 2010

Registry File SG2002/0220

(Release status: internal use only – non-sensitive)

Australian Government

Department of Sustainability, Environment, Water, Population and Communities

Supervising Scientist

How to cite this report:

Buckle D, Humphrey C & Turner K (eds) 2010. Summary of presentations and key issues raised at the Biological Monitoring Review Workshop, October 2006, with status as of August 2010. Internal Report 580, October, Supervising Scientist, Darwin.

Location of final PDF file in SSDX Sharepoint:

Supervising Scientist Division > PublicationWork > Publications and Productions > Internal Reports (IRs) > Nos 500 to 599 > IR580_Biological Monitoring Review (Buckle et al)

Location of all key data files for this report in SSDX Sharepoint:

Please see page iv for list

Authors of this report:

Duncan Buckle – Environmental Research Institute of the Supervising Scientist, GPO Box 461, Darwin NT 0801, Australia

Chris Humphrey – Environmental Research Institute of the Supervising Scientist, GPO Box 461, Darwin NT 0801, Australia

Kate Turner – Environmental Research Institute of the Supervising Scientist, GPO Box 461, Darwin NT 0801, Australia

The Supervising Scientist is part of the Australian Government Department of Sustainability, Environment, Water, Population and Communities.

© Commonwealth of Australia 2010

Supervising Scientist
Department of Sustainability, Environment, Water, Population and Communities
GPO Box 461, Darwin NT 0801 Australia

This work is copyright. Apart from any use as permitted under the Copyright Act 1968, no part may be reproduced by any process without prior written permission from the Supervising Scientist. Requests and inquiries concerning reproduction and rights should be addressed to Publications Inquiries, Supervising Scientist, GPO Box 461, Darwin NT 0801.

e-mail: publications_ssd@environment.gov.au

Internet: www.environment.gov.au/ssd (www.environment.gov.au/ssd/publications)

The views and opinions expressed in this report do not necessarily reflect those of the Commonwealth of Australia. While reasonable efforts have been made to ensure that the contents of this report are factually correct, some essential data rely on references cited and/or the data and/or information of other parties, and the Supervising Scientist and the Commonwealth of Australia do not accept responsibility for the accuracy, currency or completeness of the contents of this report, and shall not be liable for any loss or damage that may be occasioned directly or indirectly through the use of, or reliance on, the report. Readers should exercise their own skill and judgment with respect to their use of the material contained in this report.

Printed and bound in Darwin NT by Supervising Scientist Division

Contents

Data files

Data files relating to monitoring programs reported in this document are in the following sharepoint folders:

Early detection techniques

Toxicity monitoring (creekside and in-situ)

Supervising Scientist Division > SSDX > Environmental Impact of Mining - Monitoring and Assessment > Toxicity Monitoring

Bioaccumulation in fishes and mussels

Supervising Scientist Division > SSDX > Environmental Impact of Mining - Monitoring and Assessment > Bioaccumulation

Supervising Scientist Division > SSDX > Bioaccumulation > Aquatic Fauna > Mussels and Fish > Monitoring

Biodiversity assessment techniques

Macroinvertebrate communities in seasonally-flowing streams

Supervising Scientist Division > SSDX > Environmental Impact of Mining - Monitoring and Assessment > Macroinvertebrates > Ranger streams

Fish communities in channel billabongs

Supervising Scientist Division > SSDX > Environmental Impact of Mining - Monitoring and Assessment > Fish > Ranger > Channel Billabongs

Fish communities in shallow billabongs

Supervising Scientist Division > SSDX > Environmental Impact of Mining - Monitoring and Assessment > Fish > Ranger > Shallow Billabongs

Macroinvertebrate communities in shallow billabongs

Supervising Scientist Division > SSDX > Environmental Impact of Mining - Monitoring and Assessment > Macroinvertebrates > Ranger billabongs

Foreword

A review of the biological monitoring activities of the Environmental Research Institute of the Supervising Scientist (***eriss***) was conducted via a workshop that was held in October 2006. This report contains the presentations delivered at the workshop and the recommendations that were made for change. It documents the changes that were implemented over the following four years to August 2010, and provides a summary of the status of the stream biological monitoring program at that time.

Executive summary

Background

The Environmental Research Institute of the Supervising Scientist (***eriss***) has been conducting research into biological monitoring techniques for surface waters around Ranger mine since 1978 (Humphrey et al 1990, Humphrey & Dostine 1994). Initially ***eriss***'s role focused on the development and subsequent transfer of monitoring techniques to mining companies. However in 2000, additional resources were provided to the Supervising Scientist Division (SSD) to formalise a non-statutory independent monitoring program. A preliminary monitoring program was implemented during the 2000–01 wet season with the finalised program commencing during the 2001–02 wet season.

eriss's biological monitoring techniques have been subject to a number of reviews in the lead up to implementation of the monitoring program that was formalised during the 2001–02 wet season. By 2006, over a decade of biological monitoring data had accrued and the opportunity was taken at this point to review the program in light of current 'best practice'. The biological monitoring review conducted in October 2006 took into account:

1 Possible reduced sampling (frequency/effort) for components of the program, considering factors such as:
   a sensitivity of monitoring organisms to mine-related, water quality changes
   b adequacy of current datasets as a basis for monitoring during the operational and rehabilitation phase
   c competing resources insofar as possible increased intensity of new monitoring approaches and rehabilitation research
2 Optimisation of existing techniques (ie similar results with similar power, but with fewer samples/data)
3 Wishes of stakeholders, including local landowners

With the delayed reporting of the results of this review, an update of current status as of August 2010 has been provided.

Early detection techniques

Creekside and in situ monitoring

Creekside monitoring (CSM) has been the mainstay of toxicity monitoring since the 1991–92 wet season. A two-year comparison of the CSM and an in situ snail egg production test procedure was conducted during the 2006–07 and 2007–08 wet seasons. In short, there was no significant difference in the response variable between the in situ and creekside methods

over the two-year period of comparison. Consequently, the in situ 'once-only-feeding' regime has superseded the CSM method and was first implemented during the 2008–09 wet season, along with cessation of the creekside program. Advantages of the in situ monitoring include improved water flow-through the test containers, portable infrastructure and reduced resourcing compared to the creekside program.

Toxicity monitoring using the in situ reproduction test (egg number) for freshwater snails will continue to be conducted over the standard four-day exposure periods every other week in Magela Creek, and building towards a similar frequency of testing in Gulungul Creek. In a given wet season, testing will commence after deployment of continuous monitoring sondes in both creeks.

Based on the low sensitivity of larval black-banded rainbowfish to mine-derived contaminants, the large amount of staff resources required to maintain broodstock and poor fish survival using in situ testing conditions (akin to creekside monitoring conditions), toxicity monitoring using survival of black-banded rainbowfish larvae was removed from the Ranger mine stream monitoring program following the 2005–06 wet season.

Bioaccumulation in fishes and mussels

Monitoring of bioaccumulation in freshwater mussels will continue with annual sampling conducted in Mudginberri Billabong. Sampling effort has been substantially reduced from sampling individual age classes to collecting a bulk sample of mussels for re-assurance purposes. Collections using this approach was implemented in 2009 and is currently planned for the 2010 dry season. The option for analysing individual age classes in more detail every third year (commencing in 2011 and as per approach used to 2008) will be considered at a later date. The suitability of Sandy Billabong as a control site for this study remains to be assessed given that concentrations of most analytes in the tissues of mussels from this site are much lower than for mussels collected from Mudginberri Billabong. The results of a longitudinal study of mussels along Magela Creek in May 2007 showed that, like Sandy Billabong, analyte concentrations in mussels collected from sites in Magela Creek upstream of the mine are unlike those recorded in Mudginberri Billabong. Hence these upstream Magela sites may not serve as suitable control sites either. The identification of, or indeed necessity for, a suitable control site for this study remains under review.

Fish bioaccumulation has been removed from the routine monitoring program because measured concentrations are very low (ie fish do not bioaccumulate Ranger mine-derived contaminants) and so this measure is not sensitive to changes in mine-derived contaminant concentrations. In the event that local Aboriginal residents become concerned about fish in Mudginberri Billabong, or a significant mine-site influence is detected from other monitoring programs, a specific sampling program may be considered.

Biodiversity assessment techniques

Macroinvertebrate communities of seasonally-flowing streams

Macroinvertebrate sampling in seasonally-flowing streams will continue with the same sampling effort and design. A level of pooling of within-site replicates was investigated to explore reduced costs/time for macroinvertebrate sorting. However, during this investigation it was shown that utilisation of the replicate data in the statistical analysis has much greater statistical power and also provides the ability to assess impacts after a particular (typically preceding) wet season, which pooled data cannot provide.

Fish communities in channel billabongs

Refinements to the experimental design of fish community monitoring in channel billabongs have been implemented. The monitoring technique now has reduced observer counts per transect, ie from 5 to 4 counts. The reduction in observer counts has not altered the statistical power of the impact detection test.

Fish communities in shallow lowland billabongs

Fish community sampling in shallow billabongs has historically been conducted annually in up to nine billabongs. A refined experimental design has been implemented since the 2007 sampling. Sampling will only be conducted every other year (biennially) from three exposure–control paired sites (Coonjimba vs Buba; Georgetown vs Sandy Shallow; Gulungul vs Wirnymurr). Impact detection analysis now incorporates duplicate data from within each of five crocodile exclosure areas to increase statistical power.

Macroinvertebrate communities of shallow lowland billabongs

Macroinvertebrate sampling in shallow billabongs may be conducted every (say) 5 years or otherwise may be initiated if mine-derived contaminant concentrations in the relevant wet season have noticeably increased in Coonjimba, or Georgetown Billabongs compared with analyte values associated with previous sampling (1995, 1996 and 2006). This (relative infrequency) is due to the large work load involved in sampling and sorting macroinvertebrates from shallow billabongs that is in addition to the routine annual macroinvertebrate monitoring. Further work is being conducted to understand the impoverished benthic macroinvertebrate communities in Georgetown Billabong.

In the event that macroinvertebrate sampling in shallow billabongs is scheduled due to a deterioration in water quality conditions in a mine-site-influenced billabong, macroinvertebrate sampling will take precedence over the fish community monitoring. This is due to the increased sensitivity of macroinvertebrates to mine site contamination.

This page has been left blank intentionally.

Summary of presentations and key issues raised at the Biological Monitoring Review Workshop, October 2006, with status as of August 2010

Edited by D Buckle, C Humphrey & K Turner

1 Introduction

The Environmental Research Institute of the Supervising Scientist (***eriss***) has been conducting research into biological monitoring techniques for surface waters around Ranger mine since 1978 (Humphrey et al 1990, Humphrey & Dostine 1994). Initially, ***eriss***'s role focused on the development and subsequent transfer of monitoring techniques to mining companies. However, in 2000, additional resources were provided to the Supervising Scientist Division (SSD) to formalise a non-statutory independent monitoring program. The formalisation was in response to a tailings water leak at Ranger during the 1999–00 wet season. The preliminary monitoring program was implemented during the 2000–01 wet season and the finalised program commenced during the 2001–02 wet season (Supervising Scientist Division 2002).

During the development of the biological monitoring techniques (prior to 2000), ***eriss***'s biological monitoring research was subject to a number of reviews. In 1993, an external review resulted in the implementation of a number of changes that ensured 'best practice' for the time. The program was re-assessed internally in 1997 (Humphrey & Pidgeon 1998) and again in 2000, just prior to commencement of SSD's independent monitoring program. The monitoring program constitutes an essential component of the SSD's 'multiple lines of evidence' approach to water quality assessment in the Alligator Rivers Region (van Dam et al 2002).

Since 2000, the biological monitoring program has not been further reviewed. With many of the biological monitoring techniques now having been conducted for over a decade, in October 2006 it was timely to review the program in light of current 'best practice' monitoring. Furthermore, with the introduction of new monitoring techniques and further research into the rehabilitation of Ranger Mine, there was, and will continue to be, increasing (and competing) demand on staff resources. Therefore, this review has taken into account and considered the following factors by way of study objectives:

1 Possible reduced sampling (frequency/effort) for components of the program, considering factors such as:

   a Sensitivity of monitoring organisms to mine-related, water quality changes

   b Adequacy of current datasets as a basis for monitoring during the operational and rehabilitation phase

   c Competing resources insofar as possible increased intensity of new monitoring approaches and rehabilitation research

2 Optimisation of existing techniques (ie similar results with similar power, but with fewer samples/data)

3 Wishes of stakeholders, including local landowners

Apart from identifying monitoring techniques for ongoing use in the program, it was hoped that the outcomes of this biological monitoring review would also be used to identify and direct key areas of research necessary to further develop and refine the program for assessment of current operations and future rehabilitation phases of Ranger mine.

1.1 Format of report

The layout of this report follows the order of the workshop, as summarised in the program below. The workshop comprised a series of PowerPoint presentations, with questions and discussion during and after each presentation. The key issues from each presentation, a summary of related discussion and the key outcomes, are summarised on the pages following the presentation. A brief summary of the project's status at August 2010 is also included.

2 Workshop program

Overview of the SSD's environmental monitoring program – Chris Humphrey

Early detection techniques

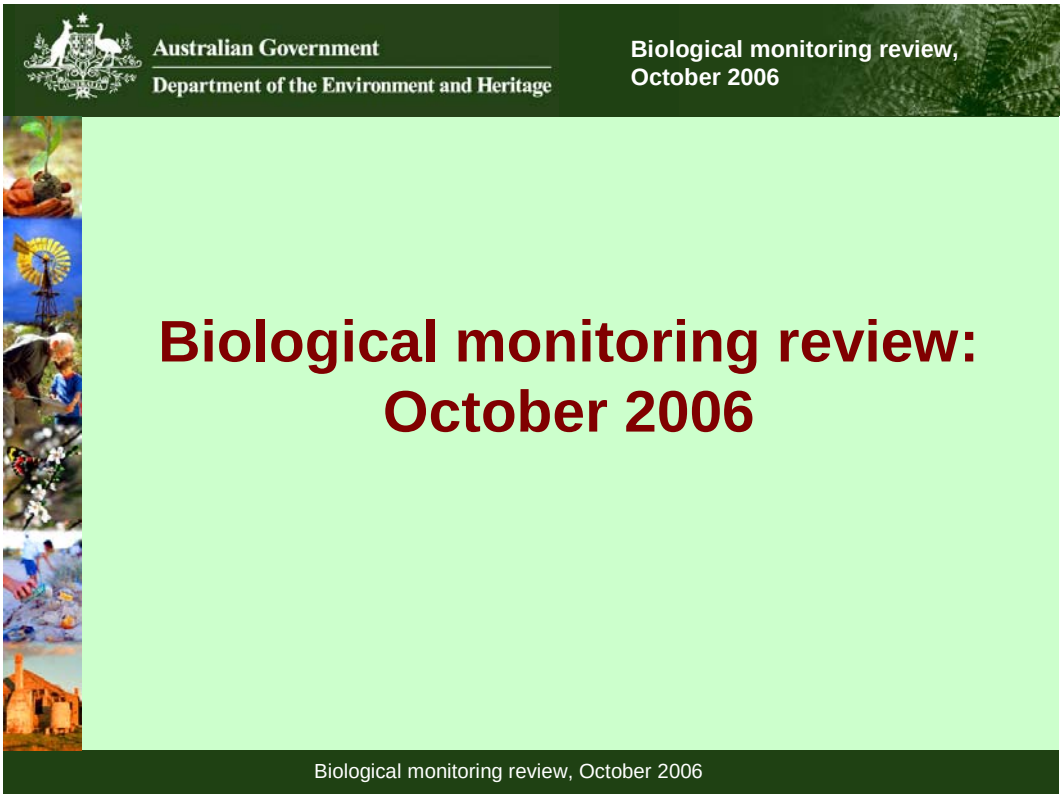
- Toxicity monitoring (creekside and in-situ) – C Humphrey
- Bioaccumulation in fishes and mussels – C Humphrey

Biodiversity assessment techniques

- Macroinvertebrate communities in seasonally-flowing streams – C Humphrey
- Fish communities in channel billabongs – D Buckle
- Fish communities in shallow billabongs – D Buckle
- Macroinvertebrate communities in shallow billabongs – C Humphrey

3 Overview of the SSD environmental monitoring program

C Humphrey



Review the SSD's routine biological monitoring program taking into account:

- Possible reduced sampling (frequencies/effort) for components of the program, considering factors such as:
  - Sensitivity of monitoring organisms to mine-related, water quality changes
  - Adequacy of current datasets as a basis for monitoring during Ranger rehabilitation phase
  - Competing resources insofar as possible increased intensity of new monitoring approaches and rehabilitation research
- Optimisation of existing techniques (the same result [with similar power] but with fewer samples/data)
- Wishes of stakeholders, including local landowners

Primary impetus for instigating an environmental monitoring program

- Additional resources provided to the SSD in 2000 to implement a routine environmental monitoring program focusing on human and ecosystem protection in KNP. A consequence of:
  - Tailings water leak (1999-2000) report; and
  - ISP/IUCN recommendations on monitoring requirements for Jabiluka
- Main elements of the program implemented during 2001–02 wet season
- Not statutory nor check monitoring, but independent monitoring program
  - Elements of which developed internally (and continually refined) after best-practice and according to national guidelines and protocols

Underpinning tenets of SSD's environmental monitoring program

Management of ecosystems of high conservation value (such as KNP) requires adherence to two important ESD tenets (ESD Steering Committee 1992): (i) precautionary management, and (ii) conserving and maintaining biological diversity. National and international recognition of this is provided in:

- The revised ANZECC & ARMCANZ (2000) Water Quality Guidelines
- The Ramsar obligation to maintain the ecological character of internationally important wetlands and to make wise use of such sites
- Aims of the World Heritage Convention, encouraging nations to protect and conserve natural and cultural heritage of worldwide importance

Ranger stream monitoring program: (i) Early detection

Prior to any wastewater release/dispersion

- Set conservative chemical standards (focus & action thresholds, max limit), derived from:
  - Reference site data
  - Ecotoxicity data
- Pre-release laboratory toxicity testing (if applicable).

During or after the wet season

- Water physico-chemistry,
- Toxicity monitoring (creekside, in situ),
- Bioaccumulation (far-field effects)

Ranger stream monitoring program:
(ii) Assessment of biodiversity

- **'Important' changes to the ecosystem, esp**:
  - changes to species richness, community composition and/or structure
- **In ARR**:
  - measurement of **fish** and **macroinvertebrate communities** in stream of interest and control streams
- **Greater complexity of design**

Why the need for biological assessment

- Continuous monitors, integrating effects of past and present exposure
- Key management indicators:
  - *Directly* assess progress towards achieving goals of ecosystem protection
  - Ultimate performance criterion in an integrated assessment program

Selection of indicators for ARR biological monitoring

- Macroinvertebrate and fish communities (or species therein), groups traditionally used in biological assessment programs
- Macroinvertebrates
  - Proven sensitivity (including ARR disturbances, lab & field), integrative, relatively sedentary etc
- Fish communities
  - Hold a high public profile: an important food resource for some communities, provide important social and cultural amenity
  - Useful for assessing landscape-level changes

The status of environmental controls

- Ability to employ optimal designs to infer strong inferences about possible mining impact in the ARR tempered and constrained:
  - Pre-mining (1980), baseline data are sparse for Magela Creek downstream of Ranger
  - R&D for macroinvertebrate and fish community monitoring techniques occurred over a relatively lengthy period
  - Designs using fish communities that are confined to a single stream (eg pre-1994 data for Ranger) may be confounded by fish movement along the stream
- In practice, extent of spatial and temporal control data in the ARR governed more by constraints of resources and availability/suitability (of controls) than by *a priori* power analysis that would assist in determining such matters.
- However, number of control sites and amount of control data are compatible with default recommendations provided in ANZECC and ARMCANZ (2000)

An appropriate level of monitoring?

- **Lack of pre-mining data places an onus on compensatory monitoring to provide a *weight of evidence* approach, to better enable correct inferences about potential impacts**

- **Emphasis on integrated assessment:**
  - Enhancing the monitoring battery (selecting and including additional sites, biological indicators)
  - Combining biological, chemical and toxicological results, seeking concordance is sought between field results and controlled experimental findings.

Some comments on data analysis procedures

- As far as possible, employ formal hypothesis testing using robust BACI-class designs, modified to account for any lack of 'Before' data. May include introduction of environmental covariates
- Other descriptive/complementary analytical techniques available:
  - Plots of the accruing time series of (raw or 'difference') data to portray any trends, control chart approaches
  - For community data, multivariate ordination of the full data-set. (Does the downstream 'impact' site lie within the space occupied by other sites?)
- Assessment of the biodiversity and conservation status of the fauna of 'exposed' sites (arising from the species-level analyses)
- In addition, undertake prospective power analyses of current fish and macroinvertebrate community data to ensure that the designs are optimised as far as possible

Overview of the SSD's environmental monitoring program – discussion and commentary

In 2000, as a result of the Ranger mine tailings dam leak, resources were allocated to the Supervising Scientist Division for the development of an independent, non-statutory biological monitoring program (Supervising Scientist Division 2002). The program's underpinning tenets of precautionary management and conservation of biodiversity are recognised nationally and internationally through the Australian & New Zealand Quality Guidelines, Ramsar guidelines for maintaining biological diversity of wetlands and the aims of World Heritage Convention (governing the management of Kakadu National Park) (Supervising Scientist Division 2002). These tenets are embedded through two key components of a multiple lines of evidence program:

1 Early Detection

a Setting standards prior to mine site discharge using ecotoxicological data and reference site data using the methods described in the national water quality guidelines.

b Physico-chemical, toxicological and bioaccumulation monitoring conducted throughout the wet season.

2 Assessment of biodiversity

a Identification of important changes to ecosystems (eg species richness, community structure) using fish and macroinvertebrate communities in control and impact sites throughout the Alligator Rivers Region (ARR).

Biological monitoring indicators include:

a Macroinvertebrates – sensitive, integrative, long life cycles and sedentary.

b Fish – high public profile, culturally and socially valuable, food resource for humans, diverse, intact (ie no exotic species), sensitive to landscape level changes.

There is a lack of true and comparative pre-mining/baseline spatial and temporal data for the Ranger mine as few studies were conducted prior to 1980. The lack of baseline data places onus on compensatory monitoring to better provide a weight of evidence approach. Statistical analyses of the biological monitoring data includes robust designs such as the MBACI (Multiple controls, Before, After, Control, Impact). Complementary approaches include control charts and multivariate ordination.

4 Early detection techniques

4.1 Creekside and in situ toxicity monitoring

C Humphrey

Creekside monitoring

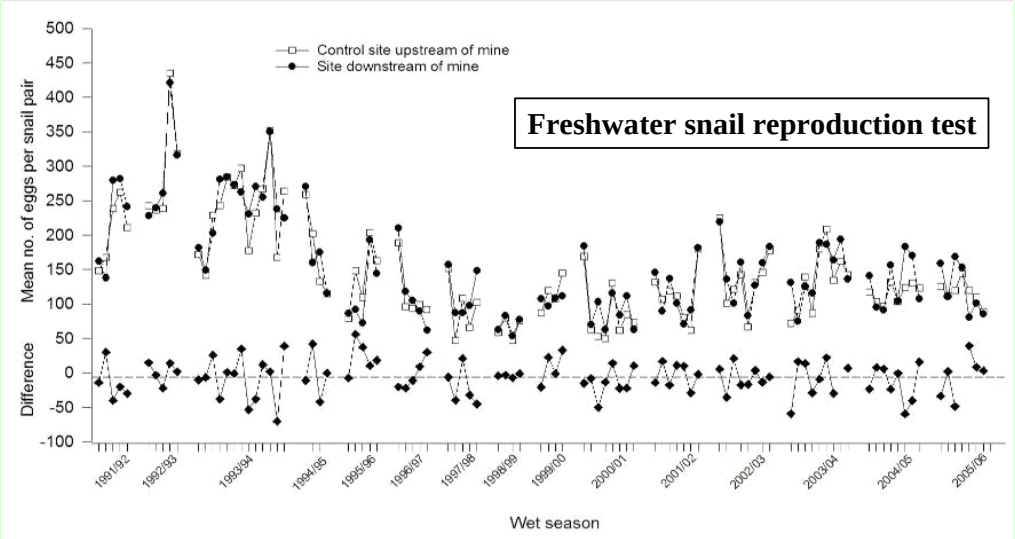
In situ monitoring

Biological monitoring review, October 2006

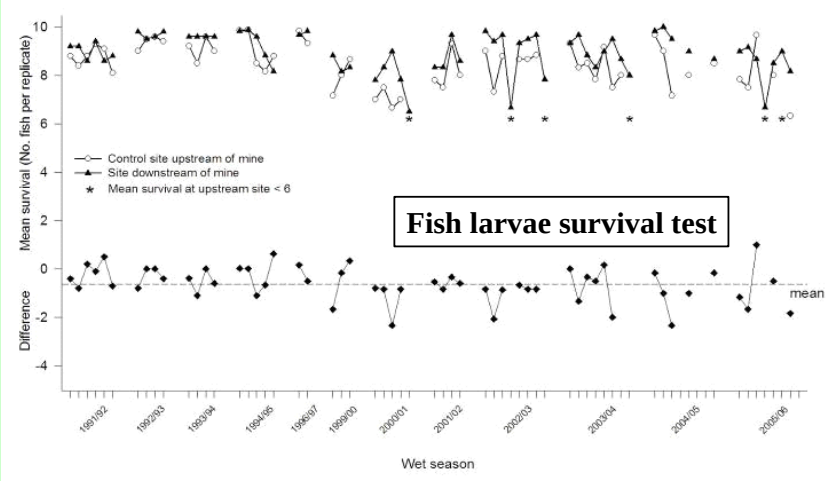
Early detection techniques:

(a) Toxicity monitoring – creekside and (to-be-developed) in-situ testing

- Several 4-day tests conducted throughout the wet season
- Survival of larval fishes and snail egg production
- Results often invoked to support 'no-observed-biological effects' arising from Ranger incidents
- In-situ testing (in place of creekside) has potential to reduce material and human resource costs substantially (amongst other advantages)



- Toxicological responses now well understood; amongst the most sensitive of organisms tested in the laboratory
- Close upstream-downstream concordance; data invariably meet statistical assumptions and test validity criteria
- Reliable 'early-detection' response; definitive conclusions reached at the end of the season



- Toxicological responses still being investigated but larval fishes generally not as sensitive to Ranger mine wastes as invertebrates
- Possibly as a consequence of a shift in upstream pump location in 1998, survival rates at upstream site decreased; approximately one test per season deemed to have failed because of poor control survival
- Some conferring of protection to larvae at the downstream site – presumably a consequence of emanating billabong waters with higher solute and/or 'nutrient' concentrations

Discussion points for toxicity monitoring

- **Creekside monitoring operationally expensive but alternative, improved testing conditions with reduced resources, in-situ procedures being tested**
- **Creekside fish and snail tests equivalent effort during the operation of a test (12 hours each per test) but over the 2-week block, fish husbandry, preparation time etc is more expensive – 20 hours versus 8 hours for snails**
- **Is the fish larvae test providing value?**
  1. Quite high (but usually acceptable) control mortality
  2. Higher survival downstream but such a 'difference' per se does not necessarily indicate the response is unsuitable for impact assessment
  3. Reliance on a single test species 'risky'
- **Possible way forward with fish test:**
  1. Is the poorer upstream survival, in part, an artefact of the creekside testing procedure? Test under in-situ conditions?
  2. Testing under in-situ conditions addresses the issue of the utility of the protocol to other locations in the ARR and NT
  3. Hormetic downstream responses warrant an explanation (experimentation) – mine solutes or natural billabong nutrients

Creekside and in situ toxicity monitoring – discussion and commentary

Organism sensitivity

It is widely accepted that the snail egg production response is sensitive to U and Mg (Hogan et al 2010, Van Dam et al 2010). However, the 96 hr Black-banded rainbowfish larval survival response is not particularly sensitive to Ranger mine waters and therefore may not be adding significant value to the toxicity monitoring program.

Rainbowfish larval test results could be compared to work completed by Bywater et al (1991) in an attempt to understand the high upstream mortality of larval fry. The high upstream mortality is, at this stage, assumed to be related to reduced solutes in the control waters possibly making test conditions at the upstream site unfavourable to larval fish, particularly during recessional flows.

- Should the larval fish test be removed from the program, the single remaining (snail) indicator may not be sufficient for assurance monitoring. Therefore it may be worth investigating an alternative organism to the fish. A less-sensitive, second test organism may be beneficial alongside the more sensitive snail in providing a graded sensitivity response in the event of a pulse of mine-derived constituents.

Resourcing

The comparison of creekside monitoring (CSM) and in situ monitoring will require additional resources to those currently available at the Jabiru Field Station (JFS). In order to free up resources for the snail egg production comparisons, CSM larval fish survival tests could temporarily cease with staff resources diverted to the snail egg production comparisons. Any additional resources available could focus on an assessment of upstream larval fish survival. This could be undertaken by way of larval fish comparative monitoring between in situ and CSM conditions, thereby providing some insight as to the cause of the higher mortality at the upstream CSM station. However, the risks associated with cessation of the CSM fish testing need to be carefully assessed.

- Stake holders may need reassurance that the resources made available by removing the fish tests will be redeployed to activities of equal, or greater, value in ensuring the environment remains protected.
- The potential advantages of in situ monitoring include improved water flow-through and contact conditions of the test organisms, portability, the ability to run an essentially continuous biological monitoring program and greatly reduced resourcing (staff infrastructure) and maintenance.

Key outcomes

- Highest priority task over the ensuing two wet seasons is the comparison of CSM and in situ snail egg production tests.
- Any additional resources that may be available should be directed at comparing fish larval survival between CSM and in situ test conditions at the upstream (only) site.

Status at August 2010

- A two year comparative evaluation of the CSM and in situ snail egg production tests was conducted over the 2006-07 and 2007-08 wet seasons. Results are reported in Humphrey et al (2009a). In summary, there was no significant difference in test responses measured between the in situ and creekside methods, and therefore, in situ monitoring will supersede the CSM method.

- A comparison of snail egg production between the CSM downstream duplicates (located on different sides of the western braid of Magela Creek) for the duration of testing has been completed. No significant difference in mean egg production per snail pair was observed between the locations. In line with the routine Ranger grab sampling and continuous monitoring for water quality, therefore, the duplicates for future in situ toxicity monitoring will be located on the western-most downstream site. The eastern site will be decommissioned.
- Toxicity monitoring using the in situ reproduction test (egg number) for freshwater snails will continue to be conducted over the standard four-day exposure periods every other week in Magela Creek, and building towards a similar frequency of testing in Gulungul Creek (trials commenced during 2009-10 wet season). In a given wet season, testing will commence after deployment of continuous monitoring sondes in both creeks.
- Comparative CSM versus in situ larval fish survival tests at the upstream site were conducted as time permitted. In total, six tests were completed (2 in 2006-07 and 4 in 2007-08). Survival of larval Black-banded rainbowfish was not enhanced under in situ conditions, suggesting that the upstream mortality of larval fish during CSM is due to natural water conditions and not related to the CSM infrastructure and conditions. Based on the known reduced (toxicological) sensitivity of larval Black-banded rainbowfish, the significant staff resources required to maintain broodstock, and the lack of noticeable improvement in organism survival under in situ testing conditions, toxicity monitoring using the Black-banded Rainbowfish larval survival test has been removed from the Ranger mine stream monitoring program.

4.2 Bioaccumulation in fishes and mussels

C Humphrey

Bioaccumulation in fishes and mussels

Biological monitoring review, October 2006

Early detection techniques:

(b) Bioaccumulation in mussels and fishes

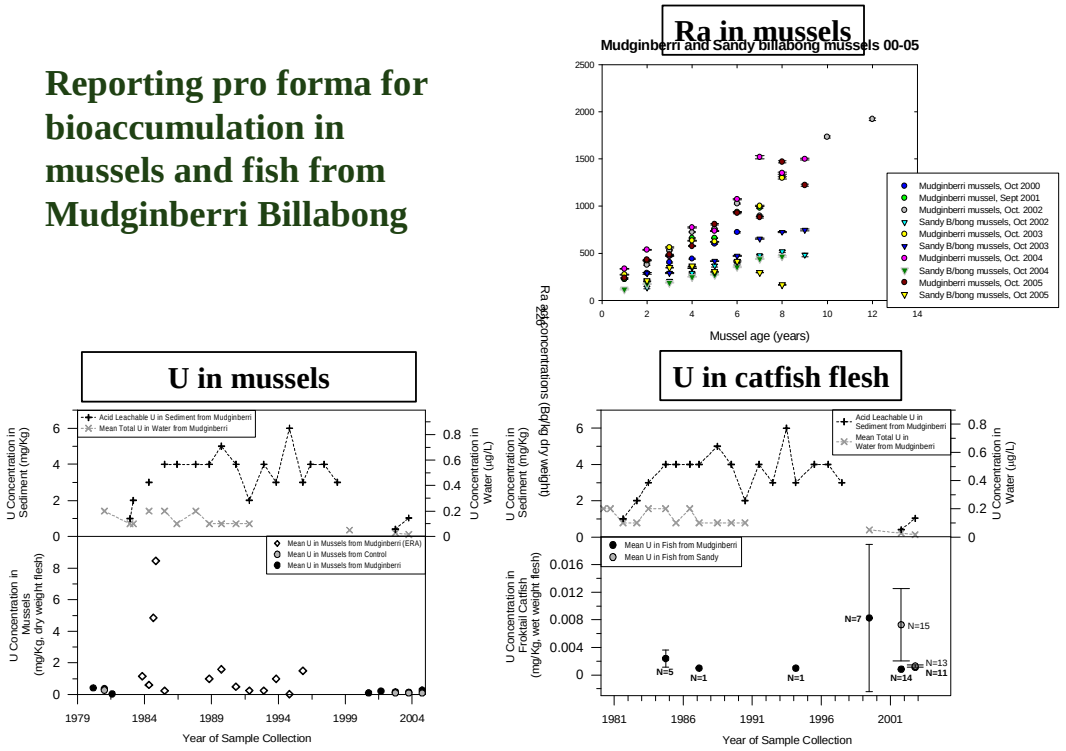
- Early detection of far-field effects arising from bioaccumulation
- Human and ecosystem health in a strategically important site
- Annual sampling of mussels from Mudginberri and Sandy Billabongs, biannual sampling of fishes from the same sites
- Program reviewed in October 2005

Biological monitoring review, October 2006

Progress with bioaccumulation studies

- Now have in place basic pro-forma plots for reporting: U and Ra in mussels, U in catfish flesh.
- Workshop recommendations being implemented:
  - Stream-line the bioaccumulation sampling program
  - Reassess analytes for measurement
  - Measure filter-feeder-relevant <63µm sediment fraction for analytes
  - Analyse mussels from Magela upstream of Ranger to assess the mine's contribution to Ra

Reporting pro forma for bioaccumulation in mussels and fish from Mudginberri Billabong



Discussion points for bioaccumulation

- October 2005 workshop largely addressed problems and major outstanding issues
- U & Ra in catfish liver should be added to pro-forma plots
- Fish samples still creating serious workload problems for radionuclide analysis: Further refinement necessary
- Dedicated project leader now required: environmental chemist and environmental radioactivity staff member?
- Dedicated effort required, as a matter of urgency, to complete protocols and summary bioaccumulation reports. Includes advice on data analysis procedures (e.g. regression comparisons for Ra-age relationships)
- Inter-catchment differences not necessarily a problem (Ra in mussels, Mudginberri vs Sandy). Interannual variations measured in Sandy still providing important inferential information

Are our early detection techniques meeting our requirements?

ESSENTIAL OR DESIRED ATTRIBUTE	Creekside snails	Creekside fish	Bioaccumulation mussels	Bioaccumulation fish
Sensitivity to the contaminant (diagnostic)	✓	Hormesis?	Analyte-dependant	Analyte-dependant
Respond and measure rapidly	✓	✓	? (have some exp'tal data)	?
High degree of constancy in time and space (precision)	✓	×?	Likely	Likely
Cost-effective	✓	?	N/A?	N/A?

Bioaccumulation in fishes and mussels – discussion and commentary

The bioaccumulation program has historically been split between Environmental Radioactivity (EnRad) and AEP (Aquatic Ecosystem Protection) programs with responsibilities changing over the years. The division in responsibility between AEP and EnRad has resulted in duplication of administration, duplication of metal analysis, lack of common data storage location and duplication of data without standardised quality control. The bioaccumulation program should be overseen by a single person to prevent duplication of resources as well as ensuring standardised analysis, and reporting on data analysis and assessment.

The scope of the bioaccumulation program should be reassessed with respect to the frequency of collections of biota, which could be determined based on changes to the water/sediment quality measured throughout the previous wet season. For example, if during the wet season no mine-related 'events' are detected in the routine surface water quality monitoring program then bioaccumulation monitoring could be omitted in the ensuing dry season. Alternatively, sampling efforts could be focused or increased following a mine-related event.

Mussels bioaccumulation

Mussels are sedentary filter feeders which make them good biomonitors that potentially accumulate contaminants over their (relatively long) life span (particularly Ra^{226}), thereby detecting changes that might not be detected in sediments and water alone. Due to the sedentary nature of mussels, inputs of Ra from Ranger Mine can be assessed by monitoring Ra^{226}:Ra^{228} ratio in mussel tissues. Any increase in the proportion of Ra^{226} indicates Ra inputs from Ranger Mine, whereas increased inputs from natural sources would increase the proportion of Ra^{228} in mussel flesh. Furthermore, mussel flesh is an important medium for monitoring Ra dose by the ingestion pathway (human health) as they are a valued food source for traditional owners living at Mudginberri community.

Sandy Billabong in Nourlangie Creek catchment has been selected as a control site for detecting natural changes in Ra concentrations accumulated by mussels. However this billabong is not well matched to Mudginberri Billabong as a control site because its catchment geology and general water chemistry is different to Mudginberri Billabong. Sandy Billabong substrate is predominantly clean sand, which has lower levels of associated Ra compared to the finer grained sediments typical of large sections of Mudginberri Billabong. Uptake of Ra by mussels in Sandy Billabong is further reduced by relatively higher concentrations of Ca in surface waters which competes with Ra for potential uptake within mussel tissues.

Fish bioaccumulation

Fish have a low accumulation capacity for the key Ranger mine contaminants and a high mobility within the catchment which makes them not a suitable organism to detect mine site contamination. However, monitoring of fish bioaccumulation does provide very good public assurance to local residents, particularly from Mudginberri community, as they regularly catch and consume fish from Mudginberri Billabong.

Processing of fish tissues (various organs) constitutes a large work load for ENRAD and Ra analysis can be time consuming, resulting in a large back-log of samples. Sometimes the sample integrity is compromised if samples are stored for too long before processing and analysis. Samples are usually stored frozen and are compromised if the freezer fails.

Given the fish bioaccumulation program does not provide a sensitive measure of mine site influence, this program could be reduced. If procedures are documented in protocols, sampling could be limited just to post mine-related events identified by other monitoring

techniques. Further consideration would need to be given to the human health assurance program, to ensure absence of fish bioaccumulation data does not weaken this assurance capacity. Furthermore, the option of compositing of samples would make the processing and analysis quicker.

A further review of the fish program would be required after the analysis and assessment of recent data. Points that may need to be considered in deciding upon the future of the fish bioaccumulation program include:

- Data from composite samples can have reduced reliability, as effects can be masked depending upon which organs/species are combined.
- The mass of various organs from an individual fish are typically sufficient for ICPMS analysis. However obtaining enough mass for radionuclide analysis is often difficult, thus composite samples would benefit radionuclide analysis.
- For ecosystem health (early detection) purposes, the signal to noise ratio for priority metals should be used to select which organs are analysed, only analysing the organs containing the highest accumulated concentrations.
- Laboratory uptake studies could be conducted using adult fish, defining the relationship between accumulated metal concentration and exposure concentration. This would help determine the bioaccumulation potential of various fish species.
- For many fish species, traditional owners consume the whole fish, including organs.

Key outcomes

- Jenny Brazier (AEP) is to be overall project leader for both mussels and fishes bioaccumulation work and should take carriage of data analysis and impact detection/assessment from the ecosystem protection perspective. Bruce Ryan (EnRad) will retain intellectual ownership of mussel radionuclide data for human health purposes.
- Mussel bioaccumulation program.
    a. October 2006 collection will be scaled back with the suite of analytes reduced as advised by Kate Turner based on the Ranger on-site water body study,
    b. April/May 2007 longitudinal study of Magela Creek with collections from upstream and downstream of Ranger in Magela creek channel (and Mudginberri if possible), and
    c. Usual sampling in October 2007, possibly modified on the basis of outcomes of part b (if data available).
- Fish bioaccumulation program.
    a. Jenny B to examine and assess existing metal and radionuclide data for fish,
    b. This assessment will determine whether (i) separate organ compositing is possible, and (ii) sufficient low-concentration data are available such that future sampling can be restricted to ‘reactive’ collections in response to significant mine events/incidents. If condition (ii) is not met, additional sampling should be conducted.

Status at August 2010

Jenny Brazier no longer works for the Department and her position within the Supervising Scientist Division (SSD) has moved from ***eriss*** to the Office of the Supervising Scientist

(OSS). Due to the change, her position will no longer be project leader for the bioaccumulation projects. Chris Humphrey will assume the position of project leader.

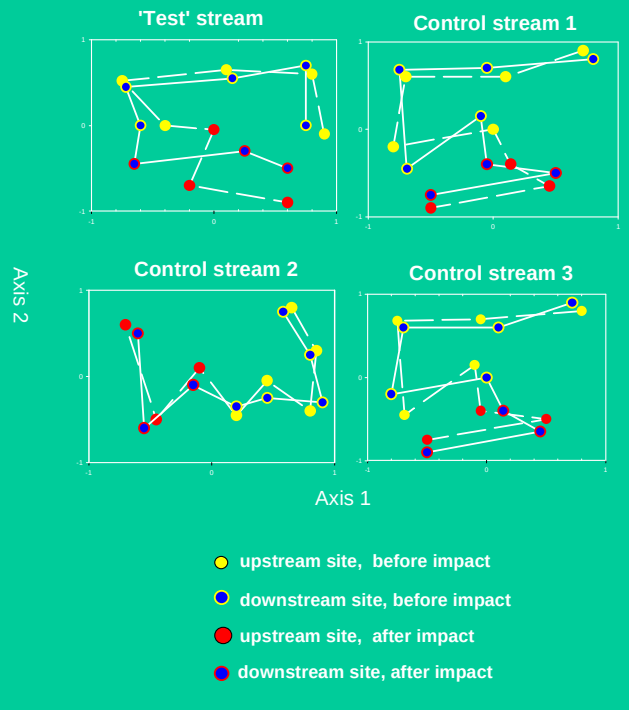
- A longitudinal study of mussels along Magela Creek was completed in May 2007 and has been reported (Brazier et al 2009a). The aim of this study was to assess results of metals and Ra^{226} and Ra^{228} activity concentrations from mussels, sediment and water from Magela Creek sites upstream and downstream of the mine, including Mudginberri Billabong. The study found that radium activity concentrations in mussels varied along the catchment and are driven by a range of factors unrelated to current mining activity at Ranger. The natural variation in geology, sediment and water quality along Magela Creek showed that sites upstream on the mine within the Magela catchment may not be useful as control sites for similar reasons discussed above. Selection of an ideal control site is still under review. Sandy Billabong is not an ideal control site and its role for this purpose requires further review for the Routine monitoring program.
- Mussel sampling in October 2007 was not conducted because the results from the May 2007 sampling provided adequate human health assurance for the consumption of mussels by the local community.
- A longitudinal study within Mudginberri Billabong was completed in October 2008 and has been reported (Bollhofer et al 2010). The aim of this project was to assess results of metal and Ra^{226} and Ra^{228} activity concentrations from mussels, sediment and water from three sites located along Mudginberri Billabong (billabong inlet, midway and outlet). In summary the study found subtle variations in the relative contribution of sources of lead and uranium in the tissue of the mussels collected from the three different sites. Importantly, Ra^{226} and ^{210}Pb activity concentrations in mussels (which determine most of the dose received via the ingestion of mussels) are not statistically different amongst sites. These results provide increased confidence that data from previous mussel collections conducted from several locations within the billabong over the years can be directly compared, provided factors that affect mussel condition (timing of mussel collection, duration of preceding wet season) are taken into account.
- Sampling effort on the mussel bioaccumulation monitoring program was substantially reduced in 2009 to collecting a bulk sample of mussels from Mudginberri Billabong for re-assurance purposes, this will be repeated in the 2010 dry season. The option for analysing individual age classes in more detail every third year (commencing in 2011 and as per approach used to 2008) will be considered at a later date.
- The reduced effort on the mussel bioaccumulation monitoring program is contingent on the current water quality being maintained in Magela Creek and will need to be reviewed in the event of a minesite incident, or change in water management practice, should that lead to substantial increases in the loads of metals and radionuclides input to Magela Creek. Other billabongs in the ARR may be investigated (depending on research priorities and staff resources) to further test the predictive power (and hence enhanced monitoring potential) of the [Ra]:[Ca] ratio in water to determine the radium load in mussels from various waterbodies.
- Fish sampling was carried out in October 2007 and the results have been reported (Brazier et al 2009b). Fish bioaccumulation has since been removed from the routine monitoring program because fish do not provide a sensitive measure of mine site contaminants. In the event that local traditional owners are concerned about fish in Mudginberri Billabong, or a significant mine site influence is detected from other monitoring programs, a specific sampling program will be considered.

5 Biodiversity assessment techniques

C Humphrey

Principle of monitoring using community structure

(MBACIP design: Hypothetical, idealised scenario)



Response variables used to assess changes to communities

- **Univariate parameters:** abundance (total or individual species), taxa number (so far not examined closely)
- **Multivariate:** dissimilarity indices – how 'dissimilar' community structure is between a pair of samples/sites (0-1)

  Dissimilarity indices:
  - Ecosystem-level response
  - Sensitive to community change (but do not indicate which site is responsible for the change)
  - In ARR experimental manipulations (Rockhole Mine Creek) dissimilarity measure was more sensitive to impact than the most sensitive macroinvertebrate species examined
  - Scope for additional sensitivity analyses

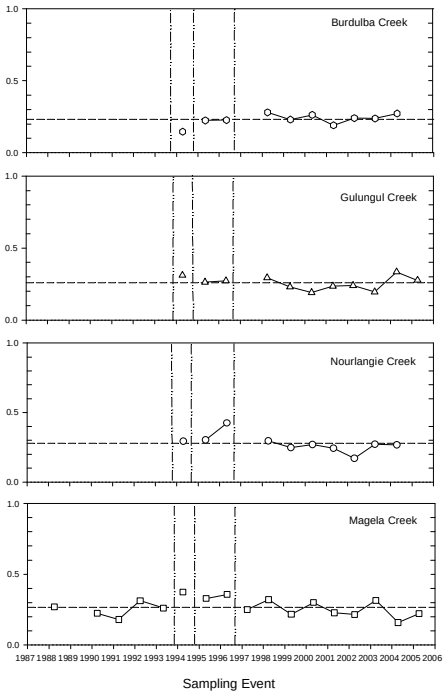
5.1 Seasonal streams

5.1.1 Macroinvertebrate communities in streams

C Humphrey

Sampling of macro-invertebrate communities from aquatic plant habitat in streams

Biological monitoring review, October 2006



Biodiversity assessment:
(b) Macroinvertebrate communities from stream sites

- Different methodologies used over time. Current method applied consistently since 1998
- Balanced MBACIP design (pairs of sites in two ‘exposed’, two reference streams)
- Data analysis and assessment focused on graphical comparison of paired-site data amongst streams – constancy and similar magnitude of the dissimilarity values
- Replicate data from within a site pooled for graphical ‘assessment’ and hence are under-utilised (e.g. no formal tests of impact in a wet season of interest)

ber 2006

Biodiversity assessment:
(b) Macroinvertebrate communities from stream sites

Progress with stream macroinvertebrate technique

- Underpinning R&D well established:
  - General sensitivity of macroinvertebrates to water quality in the ARR proven
  - Behaviour and sensitivity of dissimilarity measures as response variable have been investigated
- Not enough information is being gleaned from the data. Statistical advice is to utilise within-site replicate data to address hypotheses associated with particular streams and particular seasons
- Analysis of pooled and unpooled within-site replicate data has shown the design has implicit high power
- *How much* replicate data from within a site is necessary (in a power optimisation sense), however, should be investigated, as the sample processing time for macroinvertebrates is quite considerable
- Sample processing has focused mostly on family-level data. Strategic sample processing at the species-level would provide assurance that real (and subtle) impacts are not passing undetected

Discussion points for macroinvertebrate communities from stream sites

- Continue an annual sampling program: Ensuing data and results provide the most reliable assurance of the level of protection being met for downstream ecosystems
- Incorporate additional and formal statistical testing procedures in annual assessment of impact
- Examine the effect of some level of replicate pooling that would optimise the design noting (i) some level of within-site replication is necessary, and (ii) once pooling commences, there is ‘no turning back’!
- Undertake some strategic species-level sample processing for assurance that subtle changes in community structure are not occurring downstream of Ranger

Macroinvertebrate communities from stream sites – discussion and commentary

Macroinvertebrates are collected annually from two control and two exposed seasonally-flowing streams during recessional flows. At each of upstream-downstream (paired) sites in each stream, five replicate samples are collected from macrophyte beds in shallow runs. Various methods have been used for collection over the monitoring period. The current Surber sampler (0.0625 m^2 quadrat) has been used since 1997. Additional control streams were included in 1998 that then enabled the Multiple-Before-After-Control-Impact –Paired sites (MBACIP) impact detection method.

Each of the 5 replicate samples from each site are preserved in 90% ethanol for sorting in the laboratory. With each replicate taking four hours to sort, plus preparation time, the possibility of pooling replicates to reduce sorting time was discussed.

Whilst pooling of replicates will save a lot of staff resources it will also result in an irreversible loss of data. Before a decision can be made on this it needs to be determined if the use of replicate data in the MBACIP design provides substantially increased statistical power or greater interpretative capabilities.

Key outcomes

- Annual sampling should continue (per current sampling design and configuration of sites).
- A level of pooling of within-site replicates should be examined to optimise sample processing costs and statistical power.

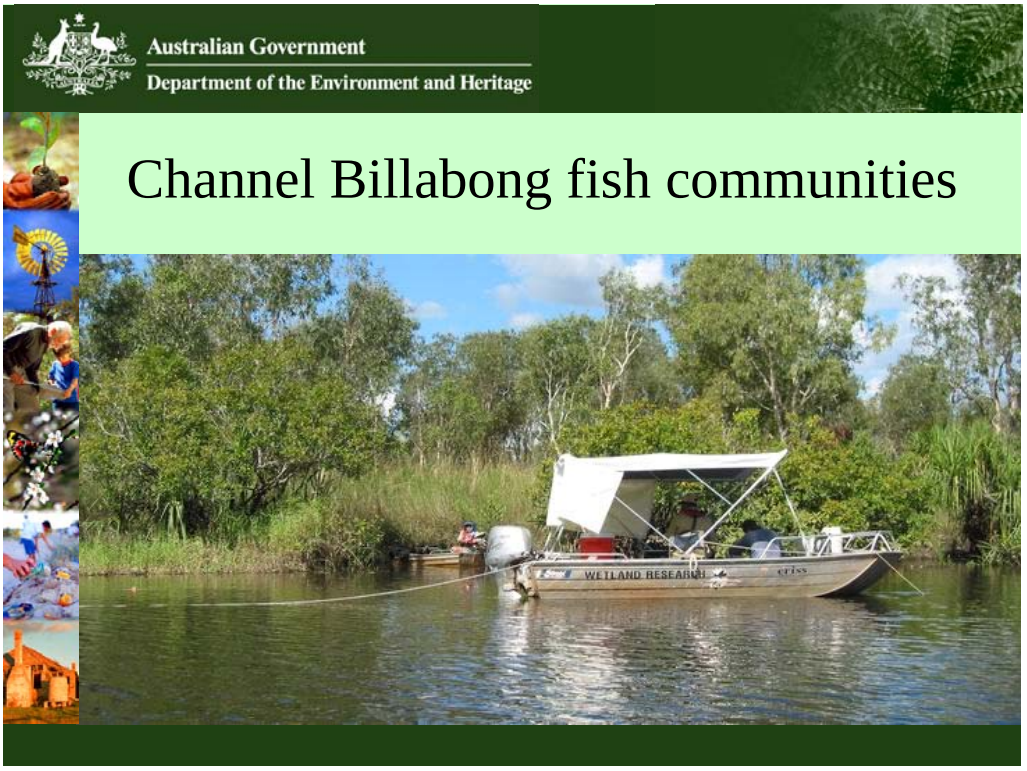
Status at August 2010

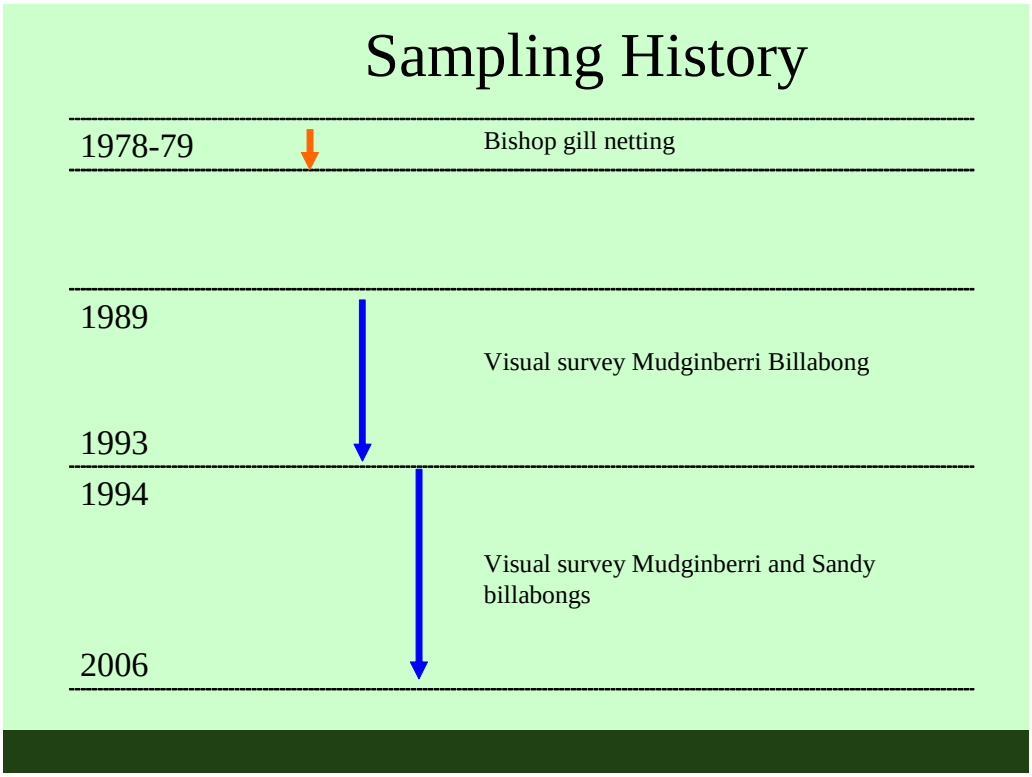
- Annual sampling has continued for 2007 (Humphrey et al 2008a), 2008 (Humphrey et al 2009b), 2009 (Humphrey et al 2010) and 2010 (currently being reported).
- Statistical advice from Keith McGuinness (CDU) has shown that the MBACIP model (ANOVA) has increased power, and an ability to assess impact in a particular year of concern, when replicates are included. The improved impact detection model has been adopted and the suggestion of pooling replicates prior to sorting dismissed.
- In-line with the MBACIP ANOVA impact detection test, the graphical presentation of the dissimilarity between paired upstream-downstream sites has been updated to include the replicate data and associated error bars.
- A portion of the macroinvertebrate samples collected and archived since 1999 are being identified to species level to determine if macroinvertebrate community structure derived from family level data is as sensitive, or adequate, as species level identification.

5.2 Channel billabongs

5.2.1 Fish communities in channel billabongs

D Buckle





Why monitor fish in Mudginberri Billabong

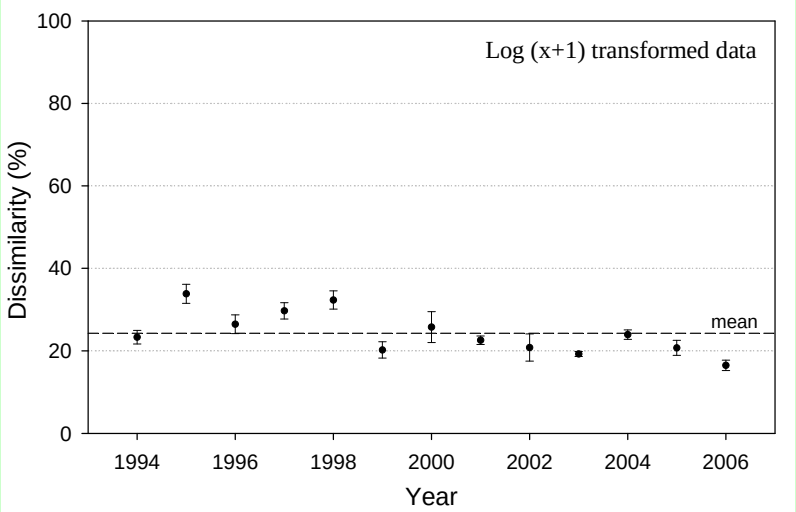
- Mudginberri Billabong is the first large permanent water body downstream from the mine.
- Important refuge habitat during the dry season (channel billabong).
- Important fishing location (resource) for locals at Mudginberri Community.
- Fish communities are an indicator of ecosystem health (long term monitoring)

Current experimental design

- Two billabongs – Mudginberri (exposed) and Sandy (control)
- Five sites (50 m transects) in each billabong
- Five visual counts at each transect by alternating observers
- Using a visual census method
  - non destructive and repeatable method.
- Impact detection is based on BACIP design (ANOVA)
- Bray Curtis dissimilarity value between billabong transects is the primary end point for analysis.

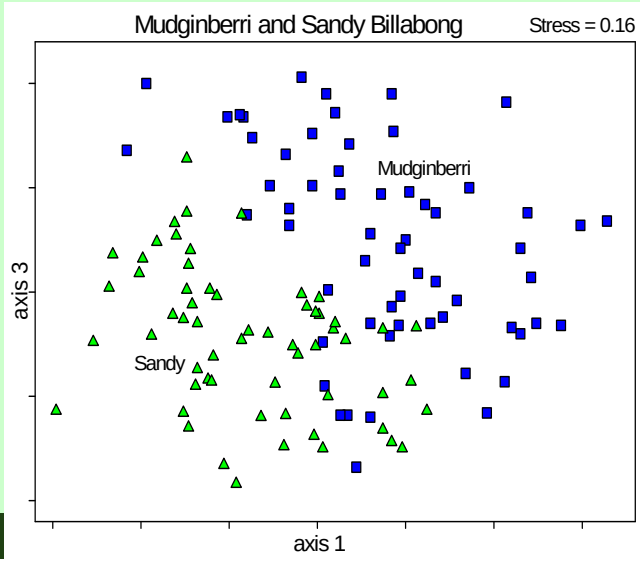
Current results

- Billabongs are becoming increasingly similar over time
  - Primarily influenced by Chequered rainbow fish and Glassfish
  - Rainbow fish numbers have declined in Mudginberri Billabong since 1989



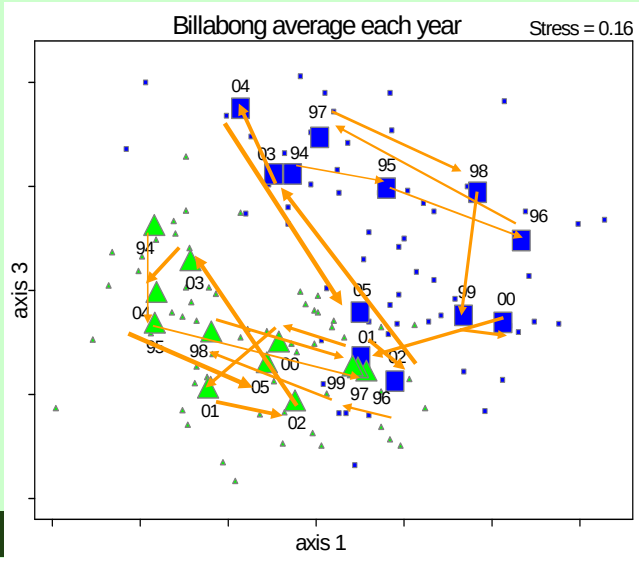
Community patterns

- Consistent differences shown by clear separation of sites in ordination space
- Temporal variation is large but difference between sites is consistent
- Sites have some semblance of tracking



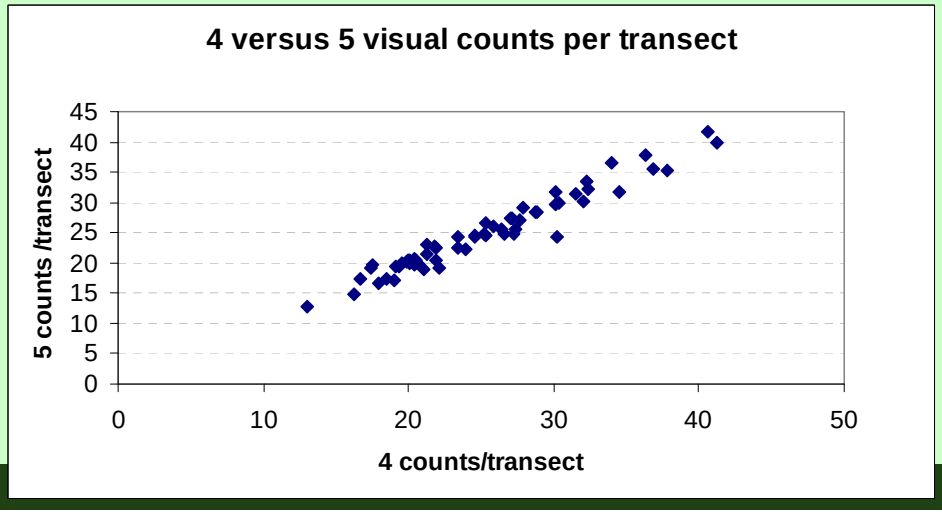
Community patterns

- Consistent differences shown by clear separation of sites in ordination space
- Temporal variation is large but difference between sites is consistent
- Sites have some semblance of tracking



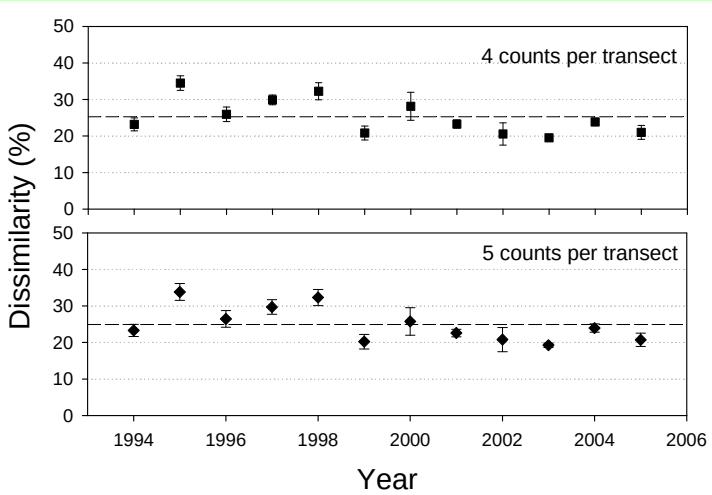
Sampling optimisation

- Reduce from 5 to 4 counts per transect?
  - Dissimilarity measure between 4 and 5 counts per transect is very highly correlated (r=0.97) (based on 1994 – 2005 data)



Sampling optimisation

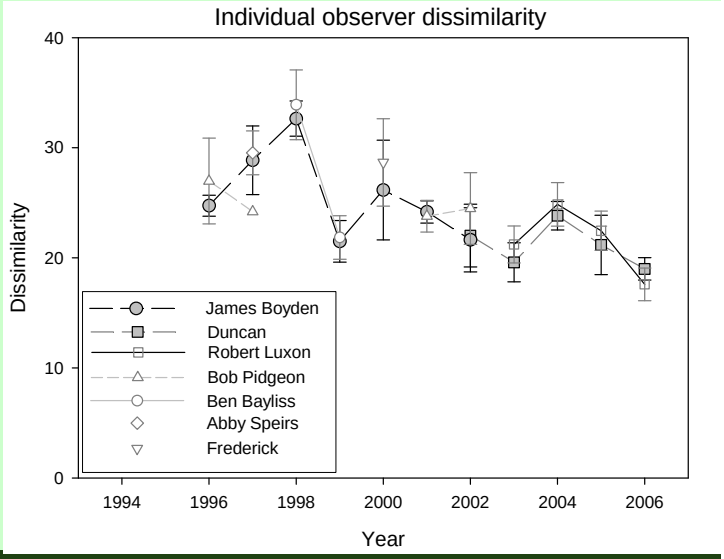
- Reduction from 5 to 4 counts per transect has no visual impact on dissimilarity value



For further information relating to sampling optimisation see Appendix 1

Sampling optimisation

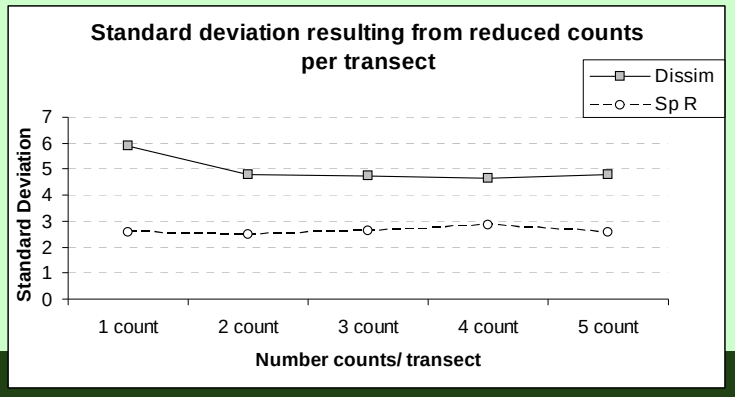
- 4 counts maintains balanced analysis between observers (two observations per transect)
- Will enable the inclusion of two training counts per transect. Training results can then be compared statistically with the experienced observers (ANOVA design).



Sampling optimisation

Estimation of power

- Reducing number of transect counts will not alter statistical model (same number of replicate site pairs)
  - –Thus reduction in power would be due to an increase in the Standard deviation of the response variables (Dissimilarity, Species richness difference)
- No changes in either – Power of test unaffected

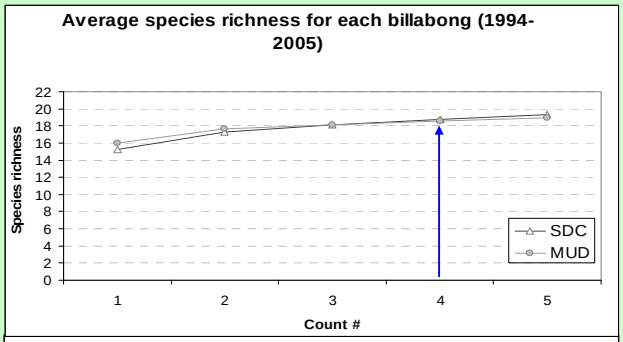


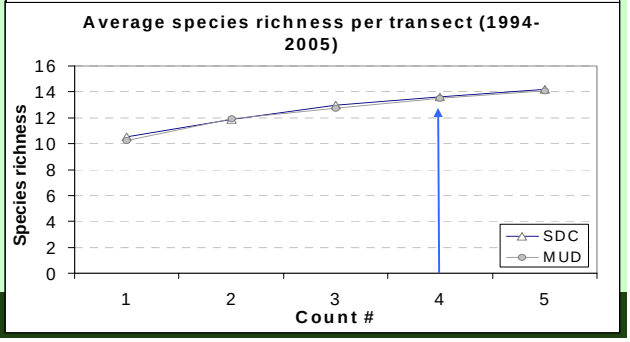
For further information relating to sampling optimisation see Appendix 1

Sampling optimisation

some loss of species data

- Some species data will be lost for each transect
- Species lost are rare species, but likely open water species which are infrequently encountered





Sampling effort
(people hrs)

task	People hours 5 counts/transect	People hours 4 counts/transect
Pre sampling hours	44	44
Sampling hours (MUD)	89	79
Sampling Hours (SDC)	98	88
Post sampling	20	20
Total	251	231

For further information relating to sampling optimisation see Appendix 1

Where to from here

- Continue sampling each year to monitor the declining dissimilarity.
  - To provide assurance to the local community and stakeholders, potentially perceived mine site influence on Mudginberri Billabong.
  - Increased time series data will help elucidate the driving factors behind the declining dissimilarity value (& discount mining influence).
- Refine sampling effort by reducing to 4 counts per transect (two counts per observer)
  - No reduction to the power of the impact detection test (dissimilarity, Sp Richness difference or Abundance difference)
  - No observable change to the dissimilarity value between MUD and SDC
  - Reduction will make training time available and more cost effective for new observers
  - Will save the equivalent of 3 working days

Fish communities in channel billabongs – discussion and commentary

Fish communities in channel billabongs were first sampled, using semi quantitative methods, in 1978–79 by Keith Bishop. The 1978–79 study used gill nets and is not directly comparable to the, non-destructive, visual survey method implemented from 1989 onward. Between 1989 and 1993 only Mudginberri Billabong was sampled. From 1994, after an external review of the ***eriss*** biological monitoring program, Sandy Billabong was introduced as a control site. Thus only data from 1994 onwards are used in the BACIP (control-impact) impact detection model.

The paired-site fish community dissimilarity between Mudginberri and Sandy Billabongs has shown an apparent decline between 1994 and 2006, ie the fish communities from Mudginberri and Sandy Billabong have become more similar over time. This decline is primarily influenced by Chequered rainbow fish, a species that conducts seasonal upstream migrations towards the end of the wet season. Rainbowfish abundances have significantly declined since sampling commenced in 1989. Whilst the decline in rainbowfish abundance, and the associated decline in dissimilarity, do not appear to be related to influences from Ranger mine, the decline needs to be studied to elucidate the causal mechanisms (Humphrey et al 2006).

Currently the visual assessment technique requires five visual counts by two alternating observers along each 50 m transect. There is scope to reduce the number of counts per transect from five to four without reducing community data collected or the statistical power of the impact detection model. Further details on the methods used to determine the appropriateness of a refinement from five to four visual counts per transect are described in Appendix 1.

Key outcomes

- Annual sampling should continue (per current sampling design and configuration of sites) to elucidate the potential causal mechanisms behind the declining dissimilarity for Mudginberri and Sandy Billabongs. Continuing the time series data (annual sampling) will be important to understand natural causes for community shifts.
- The number of observations made per site-transect should be reduced from 5 to 4. The extra time gained can be focused on training additional observers.
- The desk-top study on the statistical procedures should continue. Of particular interest is the approach to random pairing of transects between billabongs. Is the current random without replacement approach suitable or can all possible pairs be utilised?

Status at August 2010

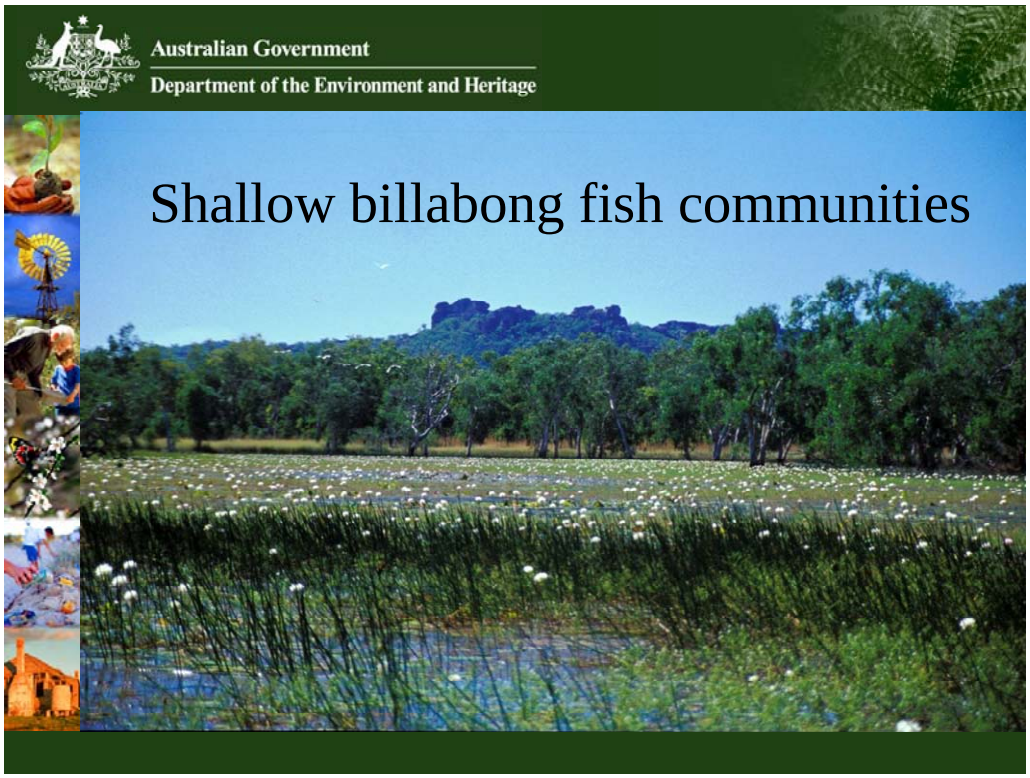
- The number of transect counts has been reduced to 4 and results for 2007, 2008 and 2009 have been reported (Buckle and Humphrey 2008, Buckle et al 2009 and Buckle et al 2010). Data has been collected in 2010 and is currently being reported.
- Statistical advice from Keith McGuinness (CDU) has shown that the random without replacement procedure is an appropriate method of random pairing. This approach has been adopted for this protocol.
- The most recent status on the decline in dissimilarity and chequered rainbow fish has been reported by Buckle et al (2010). In summary, the decline in dissimilarity remains significant over the full dataset (1994 to 2009, $P < 0.001$), despite an increase in dissimilarity that has occurred from 2006 to 2009. Humphrey and Buckle (2009) noted that a change in method procedure between the visual canoe (1989–2000) to the visual boat (2001–present) was an issue that required closer scrutiny in context to the

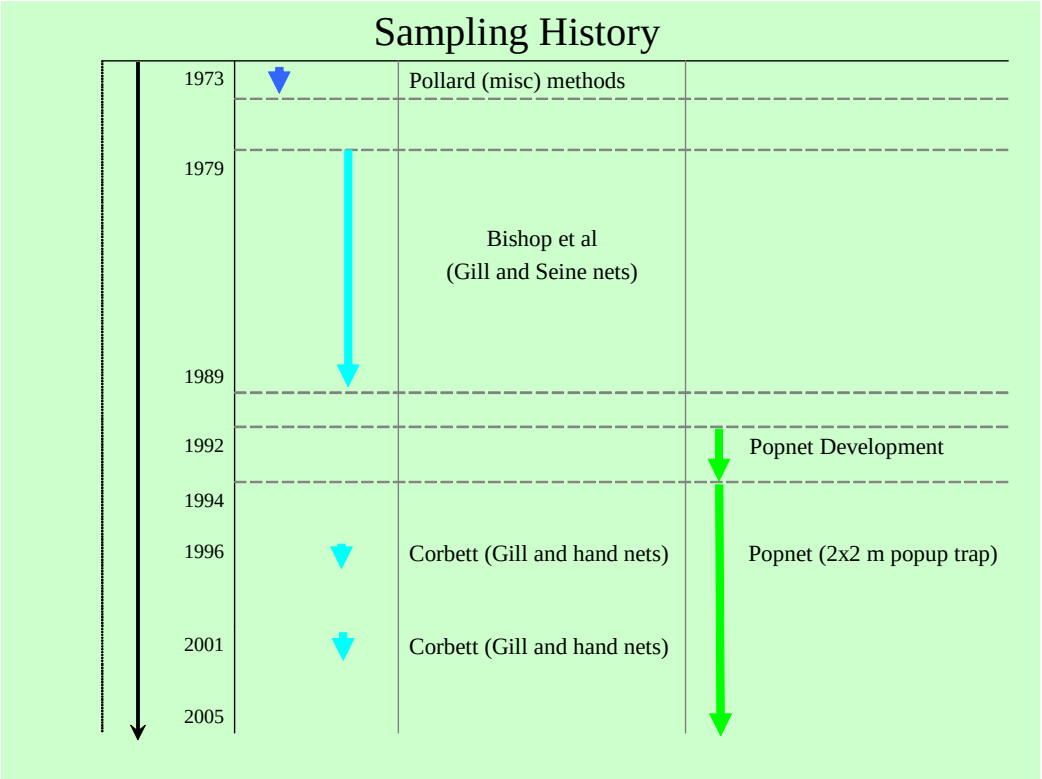
corresponding step-down in the dissimilarity values. The change potentially confounds the observed decline over time due to a significant increase in the time taken to complete each replicate visual count since 2001. However, the high dissimilarity value observed in 2009 (the highest recorded since the introduction of the visual boat in 2001) has occurred without change in sampling method, suggesting that transect times accompanying the change in observation method may not be so influential in determining dissimilarity values. The paired-site fish community dissimilarity value has increased since 2006 and may suggest that natural shifts in community structure over time are occurring. If this is the case, the nature of the community shift should become more evident over the next few years, leading to a possible explanation for the previously-identified decline or step down over time in community dissimilarity values.

5.3 Shallow lowland billabongs

5.3.1 Fish communities in shallow billabongs

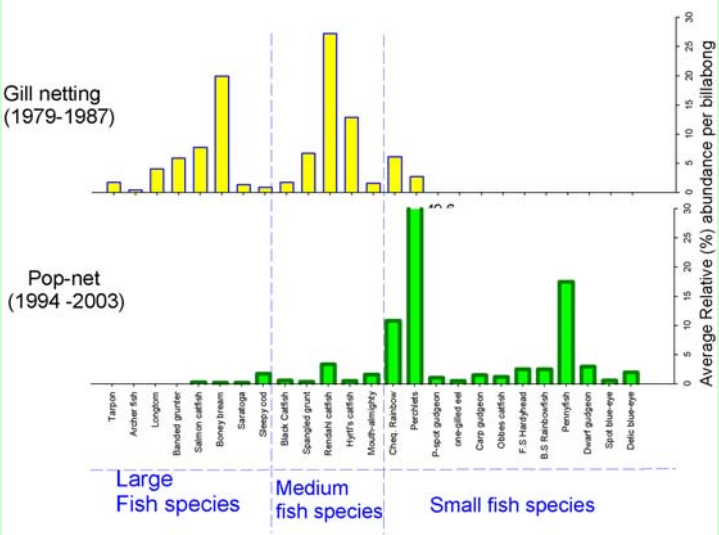
D Buckle





Historical method comparison

- Gill netting and pop-netting are not comparable for community based analysis
- Methods bias towards different fish communities



Why use fish in shallow billabongs

- Important recruitment areas in Wet season and, being depositional sites, at risk from contaminant accumulation.
- Considered best habitat for detecting effects of mining on fish communities.
- Advantage of multiple exposed and control sites
  - Important when understanding natural community shifts
- Fish are good indicators of the broader ecosystem health.

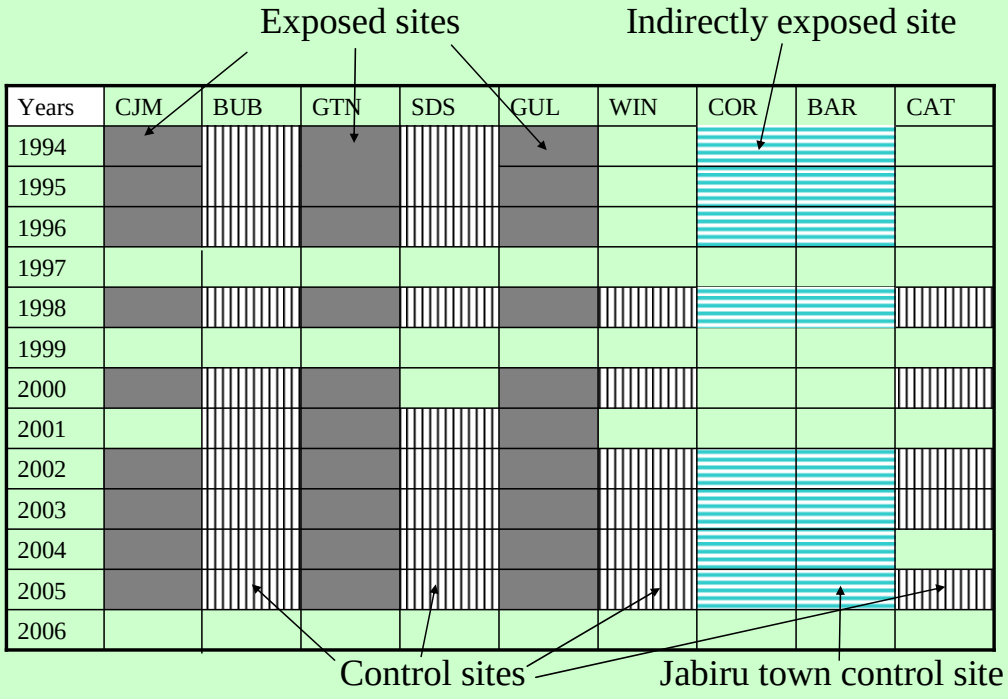
Experimental design

- 10 popnet traps (2x2 m) per billabong.
  - Focuses on shallow macrophyte margins (<1 m depth)
  - Increased effort in year 2000 with introduction of crocodile safety enclosures.
- Impact detection is based on BACIP design (ANOVA).
  - Bray-Curtis dissimilarity (control vs exposed) as the primary response.
- Three identified directly exposed vs. control site pairs.
  - Georgetown Billabong (GTN)– Sandy Billabong (SDS)
  - Coonjimba Billabong (CJM)– Buba Billabong (BUB)
  - Gulungul Billabong (GUL)– Wirnmuyurr Billabong (WIN)
- Sites are paired according to habitat similarities (depth, susceptibility to drying out)

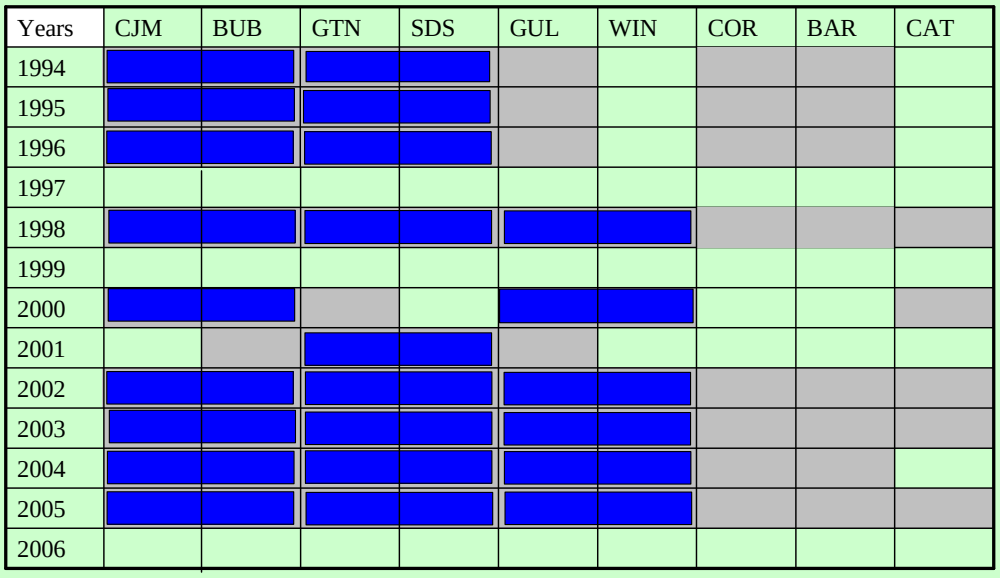
Impact assessment - limitations

- **Limited data available** - Statistical analysis requires a consistent data set across years – the current dataset is not balanced due to:
  - OH&S
  - Introduction of extra control sites in 1998
  - Cost cutting
- **Under utilising data available** – Improved statistical power by using individual popnets or crocodile enclosures as replicate site pairs
  - Further investigation and understanding required

Data available for all sites

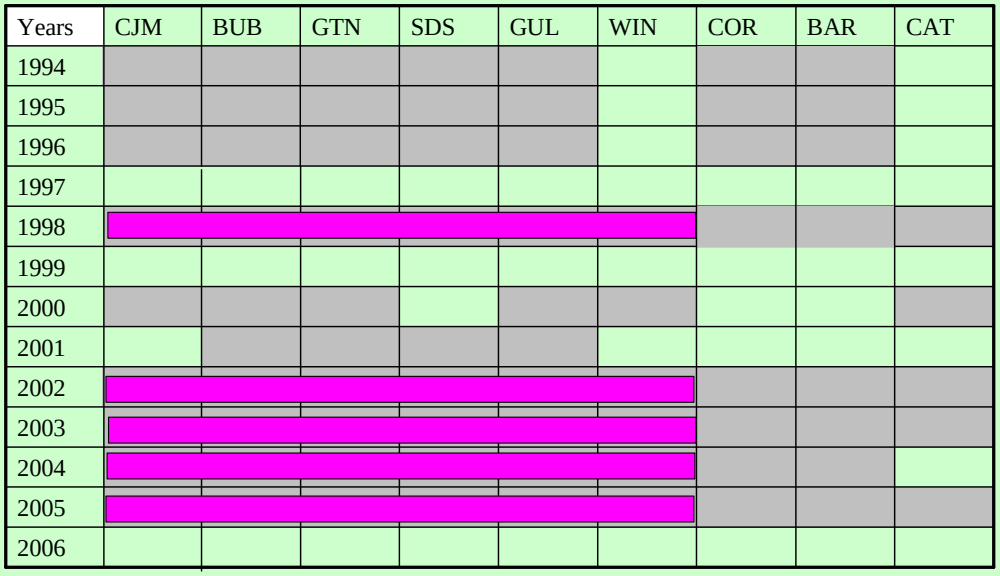


Data available for site pairs



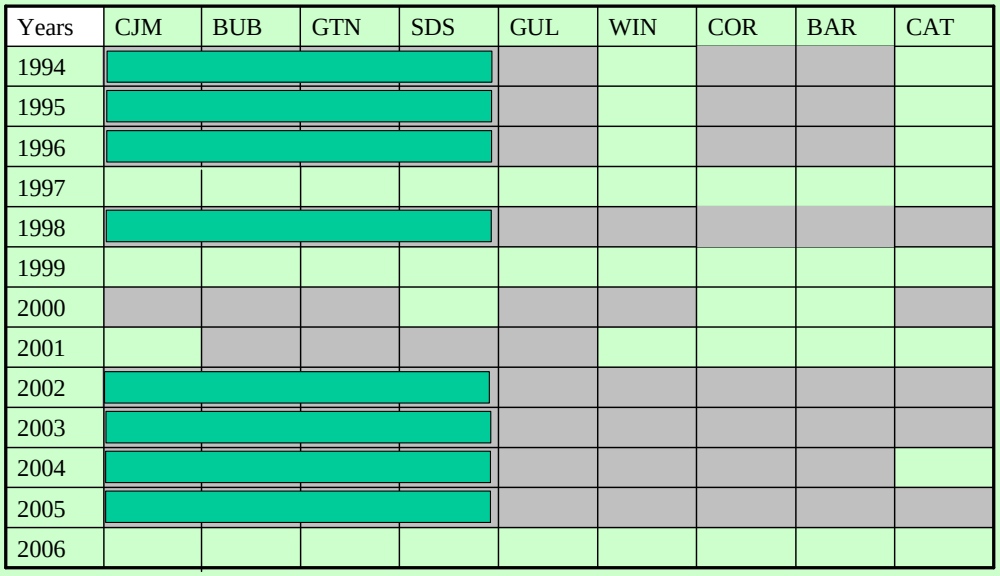
Data available for site pairs

Using preferred model of 3 paired sites



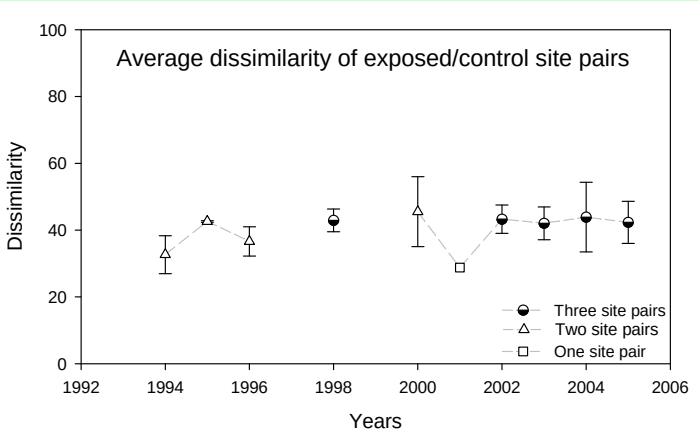
Data available for site pairs

Maximise time series data



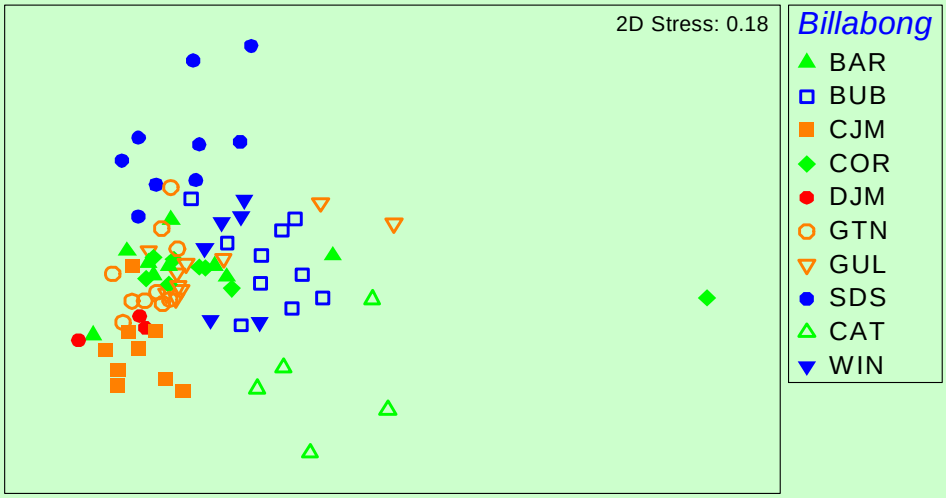
Site pair dissimilarity values

- Average dissimilarity of the three site pairs has remained similar over time.
  - Exposed billabong fish communities are not shifting from natural variations (control sites)

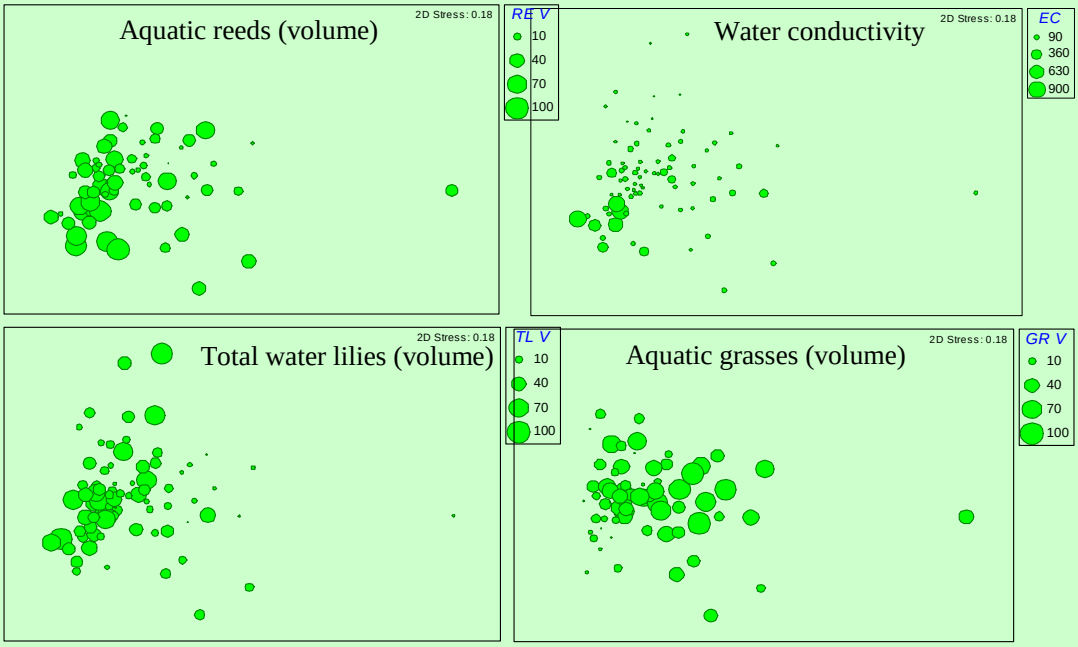


Community structures

- Fish communities are similar between billabongs (cluster together)
- Subtle differences in billabong fish communities evident by the individual billabong groups.

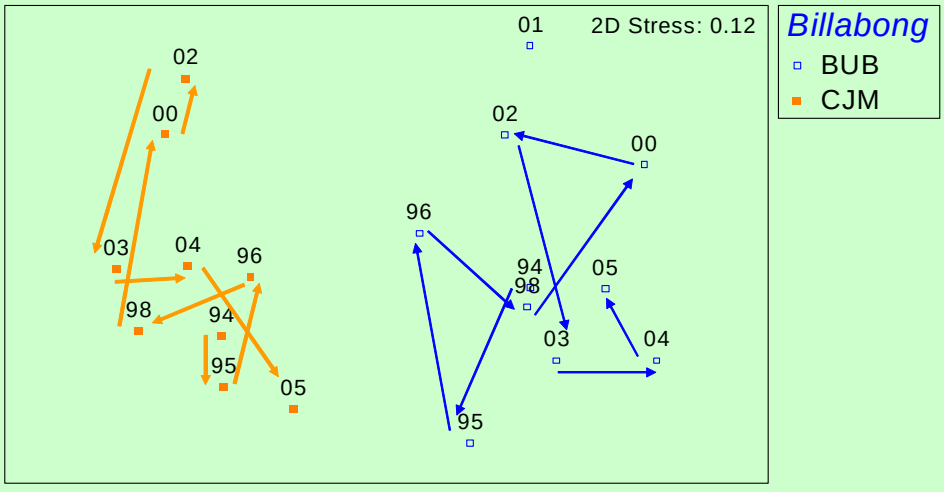


Community structures with habitat variables



Community structures of Coonjimba and Buba Billabongs

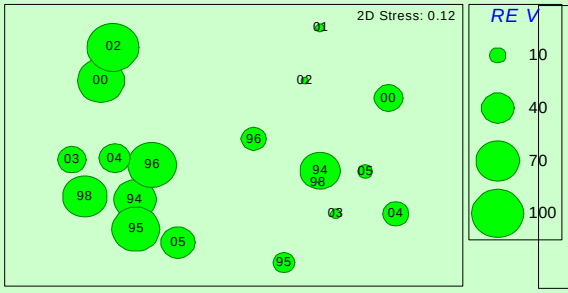
Some semblance of tracking



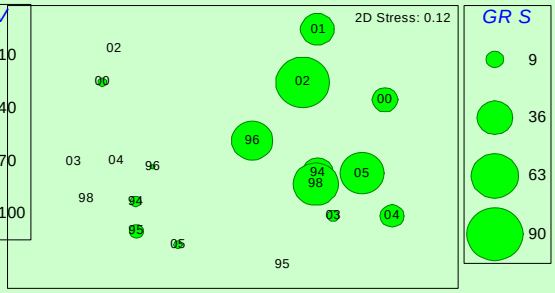
Coonjimba and Buba Billabongs

Habitat differences

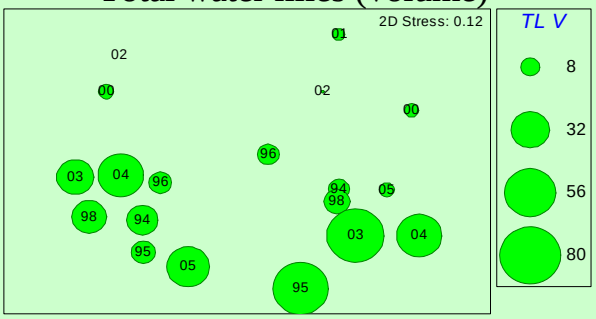
Aquatic reeds (volume)



Aquatic grasses (volume)

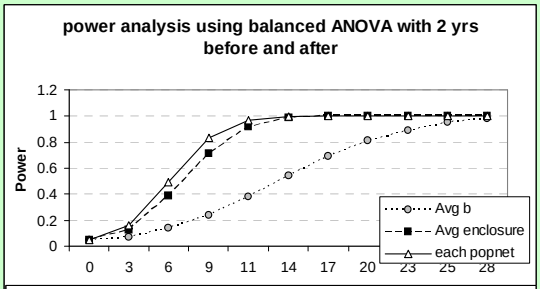


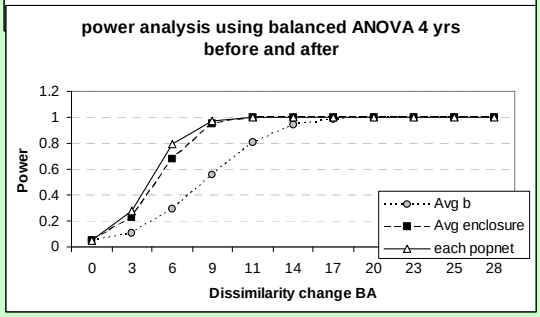
Total water lilies (volume)



Optimisation of current design data

- Statistical power can be increased substantially by using three factor ANOVA
  - Requires use of replicates from each billabong (popnets or croc enclosures)





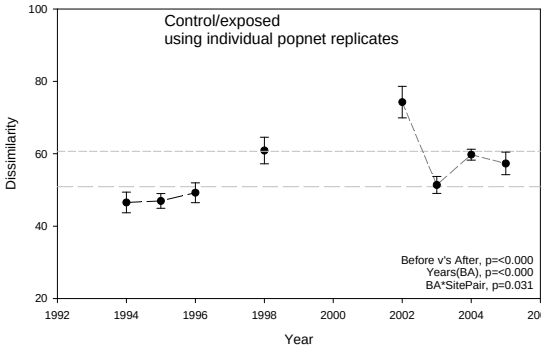
Results from ANOVA tests using

CJM VS BUB & GTN vs. SDS

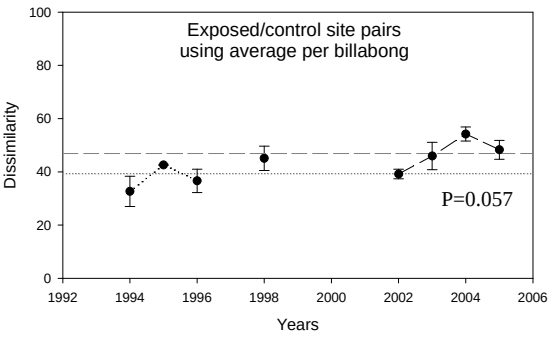
BEFORE = 1994 – 1998, After = 2002 - 2005

- Including replication (three factor ANOVA) provides a different BA result than no replication (current two factor ANOVA)
- Including replication provides greater statistical power – also provides information of within billabong variation (reflected in paired-site dissimilarity) –further understanding of interpretation required

Three factor ANOVA



Two factor ANOVA



For further information on the power analysis methods and results see Appendix 2

Shallow billabongs sampling effort

(people hrs)

task	1 Billabong	6 billabongs	8 Billabongs
Pre sampling hrs	92	92	92
Sampling hrs (per/billabong)	94	564	752
Post sampling hrs	70	70	70
Total	256hrs	726hrs	914hrs

Recommendations

- Possibly reduce sampling to every second year.
- Always include the three exposed/control site pairs in experimental design (min. 6 billabongs).
- Optional extra billabongs favourable because:
  - **Baralil Billabong** (BAR): monitors potential influence from Jabiru town on Gulungul Billabong.
  - **Corndorl Billabong** (COR): May detect indirect influences from the mine. Monitors potential effects of salvinia – a potential risk to mine site water bodies (rehabilitation).
- Remove Cathedral Billabong (CAT) from design - access makes it difficult and more costly.

Fish communities in shallow billabongs – discussion and commentary

Fish have been sampled in shallow billabongs since 1973. However, the various methods that have been used for sampling are not directly comparable. Studies by Pollard in 1973 (Pollard 1974) used a range of miscellaneous qualitative methods. Bishop *et al* (1986) used semi-quantitative methods combining the use of gill and seine nets from 1979 to 1989. These methods were repeated in 1995 (presented incorrectly on slide 2) and 2001 (Corbett et al 2004). Due to the removal of the water buffalo in the mid to late 1980s, many of these shallow billabongs experienced an increase in aquatic macrophytes preventing the effective use of gill and seine nets. In response to the increase in aquatic macrophytes, ***eriss*** trialled the popnet sampling method in 1992 and 1993. In 1994 this program was implemented with the inclusion of control sites. Unfortunately fish communities sampled using the popnet method are not directly comparable to the gill and seine net methods as there is a bias towards different community structures. As a result the impact detection model used by ***eriss*** (BACIP, control-impact) uses popnet data only, collected from 1994 onward.

Within the popnet dataset there are many years when not all billabongs have been sampled due to OH&S issues (crocodile safety), budget restrictions and the addition of a control site in 1998 (Wirnmuyurr Billabong). This has resulted in an unbalanced data set. There are only five out of ten years of collected data (1994 to 2005), when the three paired sites have been sampled concurrently. To maximise the number of years used in the impact detection model currently, it is recommended that only the two paired sites (Georgetown – Sandy and Coonjimba – Buba) be used as they provide eight years of data. However, it is recommended that future monitoring of fish communities in shallow billabongs includes the three paired sites (six billabongs) each year to increase the amount of baseline data to be used in the preferred impact detection model of three exposed-control site pairs.

The current impact detection model uses the average of all popnet replicates from each billabong. This model underutilises the data available from this program. By including replication within each billabong (ie individual popnet traps or crocodile enclosure data) the statistical power of this monitoring technique can be potentially increased and a test and assessment of impact made for each particular year of sampling. Further information on the power analysis methods using the three different analysis options has been provided in Appendix 2.

Key outcomes

- The default position is sampling every other year in a 3 exposure-reference paired site configuration (from 2007) unless it is decided that in the due year, macroinvertebrate sampling should be conducted (in which case, no fish sampling will be conducted).
- Desktop study of the statistical analysis procedure (no replication versus replication) should continue to determine the design which best optimises the statistical power.

Status at August 2010

- Sampling using three exposure – control paired sites was completed in 2007 and 2009, with the results reported (Buckle and Humphrey 2008, Buckle et al 2010).
- After further statistical advice from Keith McGuinness (Charles Darwin University), BACIP Impact detection using crocodile enclosure replicate data has been adopted, providing increased statistical power. The random without replacement pairing method is used to generate the paired site dissimilarity values.

5.3.2 Macroinvertebrate communities in billabongs

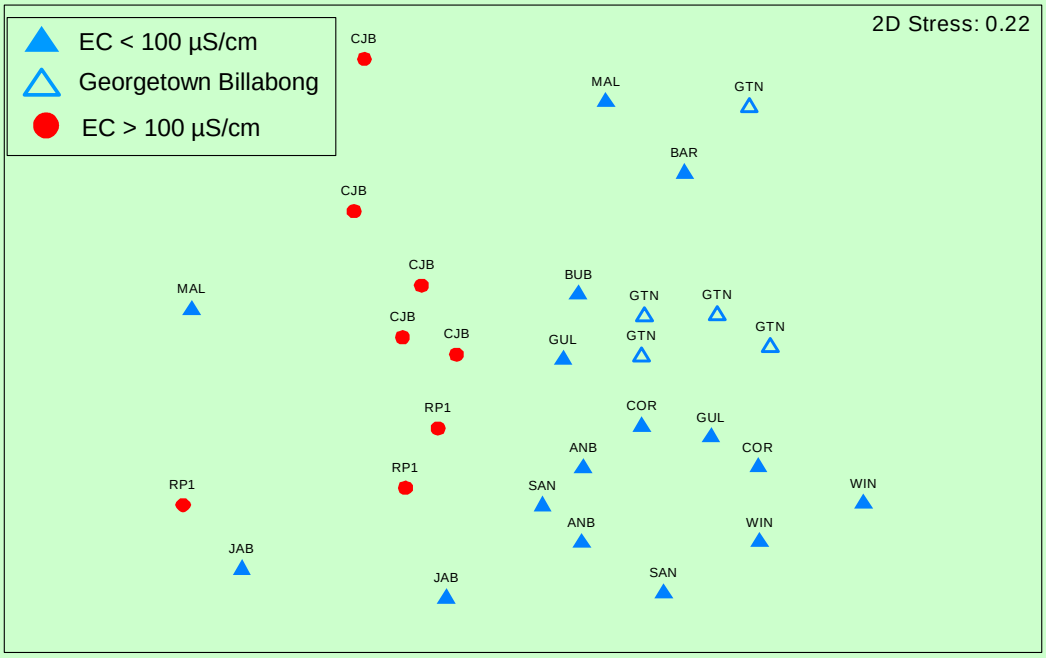
C Humphrey

Biodiversity assessment in shallow lowland billabongs:
(ii) Macroinvertebrate communities

- Serve a rehabilitation role (developing water quality closure criteria) but also useful monitoring role
- Design comprises sampling of 5 replicate samples in each of a number of different 'exposed' and reference waterbodies (~12 total).
- Because a number of mine-contaminated sites are included, a gradient response in macroinvertebrates across a contaminant gradient is possible (powerful inferential information)
- Common sampling approach has been applied in 1995, 1996 & 2006
- Mine-related effects appear to be evident in 1995 and 2006 but subdued in 1996 (different wet season intensities and exposures to inputs of water from the mine lease)
- Sample processing reasonably labour-intensive. Currently detracting from completion of samples from routine stream monitoring study

Biological monitoring review, October 2006

Preliminary macroinvertebrate data from shallow waterbodies, 2006



Biological monitoring review, October 2006

Discussion points for community studies in shallow lowland billabongs

- Strategically important sites: Because some contaminated sites included, provision of early warning information and 'dose-response' data from the real world
- Fish communities responding in a similar manner between catchments to non-mining-related, environmental changes (landscape-level), macroinvertebrates responding to mine contaminants at least
- Alternating fish and macroinvertebrate sampling each year a sensible sampling strategy, ensuring continuity of valuable time-series data
- Fish community design refined to just 3 pairs of exposed-reference sites, with potential to improve power of the statistical design through use of unpooled, replicate data

Biological monitoring review, October 2006

Macroinvertebrate communities in lowland billabongs – commentary and discussion

Ideally, lowland billabongs should have alternate fish and macroinvertebrate sampling each year. This ensures adequate resources are available for each program whilst ensuring continuity of time-series data.

To provide an adequate baseline of data before rehabilitation of Ranger minesite, it is recommended that macroinvertebrate communities in shallow billabongs be sampled for at least another few years, comparing benthic and macrophyte macroinvertebrate communities independently.

With new land application areas in the vicinity of Coonjimba Billabong (Jabiru East and Djalkmara extension) there is the potential for increased contamination of surface waters with mine-derived constituents. If an increase in contamination is detected from Coonjimba Billabong then macroinvertebrate community sampling should be conducted to capture any EC gradient effects upon macroinvertebrate communities.

Key outcomes

- The 3-year dataset (1995, 1996 and 2006) should be analysed to determine (i) whether there are logical exposure-reference waterbody pairs, and (ii) whether there are sufficient existing data to restrict further sampling to a smaller configuration of billabong pairs – as now done with fish in lowland billabongs.
- For the purposes of deriving water quality closure criteria and monitoring macroinvertebrates in shallow billabongs for assessing ecosystem health, macroinvertebrate sampling should be triggered after wet seasons where billabong water quality has been notably impacted by discharges of water from the minesite. This would provide the potential for water quality closure criteria to be reviewed and enable the SSD to assess the extent, if any, of potential impacts.

Status at August 2010

- Due to the large work load involved in sampling macroinvertebrates from shallow billabongs, future sampling will only occur if an increase in contamination is detected in Coonjimba or Georgetown Billabongs. In such an event, macroinvertebrate sampling will take precedence over the fish community work due to the increased sensitivity of macroinvertebrates to mine site contamination.
- Results from the 2006 macroinvertebrate survey have been reported in Humphrey et al (2008b). In summary, the data indicate that macroinvertebrate communities from the sediments of Georgetown Billabong are relatively impoverished and resemble those from the sediments of higher EC mine-influenced waterbodies. The report listed a number of possible reasons for this, including (i) contamination of sediments from mine-derived constituents, and/or (ii) physical and chemical attributes of the sediments that are unrelated to mining. The littoral macroinvertebrates from macrophyte habitat, however, continue to indicate impact from mining (ie community structure similar to that shown in control billabongs). The impoverished benthic macroinvertebrates in Georgetown Billabong are being further investigated.
- Additional sediment samples for chemical analysis were collected in August 2007 from the same billabongs sampled for macroinvertebrate communities in 2006. The results suggest that the uranium concentration in sediments have increased since 2002. However, a number of potentially confounding effects diminish the ability to infer mining-related change to the benthic communities of water bodies in this study. Future investigations of this aspect have been identified in Humphrey et al (2009c).

The recommendations from Humphrey et al (2009c) are reprinted below;

- better quantifying and describing the physical nature of sediments from the various waterbodies by way of particle size distribution (to confirm the fine-grained nature of sediments in Georgetown Billabong in particular);
- collecting a limited number of littoral and corresponding deeper-water sediment samples from Georgetown for chemical analysis. (The littoral samples collected in 2007 may be unrepresentative of the more central billabong samples collected by other agencies in the past);
- examine the extent of metal extraction from sediments using different digest techniques on different size fractions. The results would be used to assess the degree to which historical sediment quality data, often derived using different digest methods and size fractions, may be validly compared; and
- using data from dot-points 1 and 2, re-analyse and model environmental and biological data to better assess the degree and extent, if any, of possible mine-related change to benthic communities of Georgetown Billabong.

The outcome from this more detailed assessment will indicate if billabong closure criteria may be needed for sediments as well as water.

5.4 Are our biodiversity techniques meeting our requirements

C Humphrey

Are our biodiversity assessment techniques meeting our requirements?

ESSENTIAL OR DESIRED ATTRIBUTE	Macroinv'brates streams	Macroinv'brates billabongs	Fish Channel b'bongs	Fish Shallow b'bongs
Sensitivity to the contaminant (diagnostic)	✓	✓	? (but landscape change, likely)	? (but landscape change, likely)
Direct measures of biodiversity	✓	✓	✓	✓
High degree of constancy in time and space (precision)	✓	Likely	✓ (but trend evident)	Likely
Staff resources (PD)	52 (4 streams, 8 sites)	120 (12 water-bodies)	33 (2 b'bongs)	97 (6 b'bongs)

Biological monitoring review, October 2006

6 References

Bishop KA, Allen SA, Pollard DA & Cook MG 1986. *Ecological studies on the freshwater fishes of the Alligator Rivers Region, Northern Territory. Volume I Outline of the study, summary, conclusions and recommendations*. Research report 4 (i), Supervising Scientist for the Alligator Rivers Region, AGPS, Canberra.

Bollhoefer A, Humphrey C, Ryan B & Buckle 2010. A study of radionuclide and metal uptake in mussels from Mudginberri Billabong. In ***eriss*** *research summary 2008–2009*. eds Jones DR & Webb AL, Supervising Scientist Report 201, Supervising Scientist, Darwin NT, 99–104.

Brazier J, Bollhoefer A, Humphrey C & Ryan B 2009a. A longitudinal study of radionuclide and metal uptake in mussels from Magela Creek and Mudginberri Billabong. In ***eriss*** *research summary 2007–2008*. eds Jones DR & Webb AL, Supervising Scientist Report 200, Supervising Scientist, Darwin NT, 91–99.

Brazier J, Bollhoefer A, Ryan B & Humphrey C 2009b. Bioaccumulation in fish and freshwater mussels from Mudginberri Billabong. In ***eriss*** *research summary 2007–2008*. eds Jones DR & Webb AL, Supervising Scientist Report 200, Supervising Scientist, Darwin NT, 53–56.

Buckle D and Humphrey C 2008. Monitoring of Ranger mine using fish community structure. In ***eriss*** *research Summary 2006–2007,* eds Jones DR, Humphrey C, Van Dam R & Webb A, Supervising Scientist Report 196, Supervising Scientist, Darwin NT, 56–59.

Buckle D, Humphrey C & Davies C 2009. Monitoring using fish community structure. In ***eriss*** *research summary 2007–2008*. eds Jones DR & Webb AL, Supervising Scientist Report 200, Supervising Scientist, Darwin NT, 60–64.

Buckle D, Humphrey C & Davies C 2010. Monitoring using fish community structure. In ***eriss*** *research summary 2008–2009*. eds Jones DR & Webb AL, Supervising Scientist Report 201, Supervising Scientist, Darwin NT, 88–92.

Bywater JF, Banaczkowski R & Bailey M 1991. Sensitivity of tropical freshwater fishes and four species of cladocerans from Northern Australian uranium. *Environmental Toxicology and Chemistry* 10 (11), 1449–1458.

Corbett L, Batterham R & Zimmermann A 2004. Aquatic studies at Ranger Mine: Whole-Ecosystem Monitoring – Round 2. Unpublished report to ERA Ranger Mine by EWL Sciences Pty Ltd.

Hogan AC, van Dam RA, Houston MA, Harford AJ & Nou S 2010. Uranium exposure to the tropical duckweed, *Lemna aequinoctialis*, and pulmonate snail, *Amerianna cumingi*: fate and toxicity. *Archives of Environmental Contamination and Toxicology* 59, 204–215.

Humphrey CL, Bishop KA & Brown VM 1990. Use of biological monitoring in the assessment of effects of mining wastes on aquatic ecosystems of the Alligator Rivers Region, Tropical Northern Australia. *Environmental Monitoring and Assessment* 14, 139–181.

Humphrey C, Buckle D & Davies C 2009a. Development of in situ toxicity monitoring methods for Magela Creek. In ***eriss*** *research summary 2007–2008*. eds Jones DR & Webb AL, Supervising Scientist Report 200, Supervising Scientist, Darwin NT, 86–90.

Humphrey C, Chandler L, Camilleri C & Hanley J 2008a. Monitoring of Ranger mine using macroinvertebrate community structure. In ***eriss*** *research Summary 2006–2007*, eds Jones DR, Humphrey C, Van Dam R & Webb A, Supervising Scientist Report 196, Supervising Scientist, Darwin NT, 52–55.

Humphrey C, Chandler L & Hanley J 2009b. Monitoring using macroinvertebrate community structure. In ***eriss*** *research summary 2007–2008*. eds Jones DR & Webb AL, Supervising Scientist Report 200, Supervising Scientist, Darwin NT, 57–59.

Humphrey CL and Dostine PL 1994. Development of biological monitoring programs to detect mining-waste impacts upon aquatic ecosystems of the Alligator Rivers Region, Northern Territory, Australia. *Mitteilungen Internationalis Vereinigung Limnologiae* 24, 293–314.

Humphrey C, Jones D, Hanley J, Chandler L & Camilleri C 2008b. Developing water quality closure criteria for Ranger Billabongs using macroinvertebrate community data. In ***eriss*** *research summary 2006–2007*. eds Jones DR, Humphrey C, Van Dam R & Webb A, Supervising Scientist Report 196, Supervising Scientist, Darwin NT, 84–89.

Humphrey C & Pidgeon R 1998. Development of biological monitoring programs used to assess mining impact in the ARR: A report to the Alligator Rivers Region Technical Committee, December 1997. Internal report 286, Supervising Scientist, Canberra. Unpublished paper.

Humphrey C, Pidgeon B & Buckle D 2006. Monitoring of fish communities in deep channel billabongs associated with Ranger uranium mine, Northern Territory. In *A guide to monitoring fish stocks and aquatic ecosystems*. Australian Society for Fish Biology Workshop Proceedings, Darwin, Northern Territory, 11–15 July 2005, eds Phelan MJ & Bajhau H, Fisheries Incidental Publication 25, NT Department of Primary Industry, Fisheries and Mines, Darwin, 86–103.

Humphrey C, Turner K & Jones D 2009c. Developing water quality closure criteria for Ranger billabongs using macroinvertebrate community data. In ***eriss*** *research summary 2007–2008*. eds Jones DR & Webb AL, Supervising Scientist Report 200, Supervising Scientist, Darwin NT,130–135.

Pollard DA 1974. The freshwater fishes of the Alligator Rivers 'Uranium Province' area (Top End, Northern Territory) with particular reference to the Magela Creek catchment (East Alligator River System). Report 1. In *Alligator Rivers Area Fact Finding Study Four AAEC Reports*, ed PJF Newton, Report E305, Australian Atomic Energy Commission, AAEC, Sydney.

Supervising Scientist 2008. Annual Report 2007–2008. Supervising Scientist, Darwin, 1–71.

Supervising Scientist Division 2002. Investigating an environmental monitoring program to protect aquatic ecosystems and humans from possible mining impacts in the Alligator Rivers Region. Environment Australia. www.environment.gov.au/ssd/monitoring/pubs/env-mon-prog-background.pdf

Van Dam RA, Hogan AC, McCullough CD, Houston MA, Humphrey CL & Harford AJ 2010. Aquatic Toxicity of Magnesium Sulfate, and the influence of Calcium, in very low oinic concentration water. *Environmental Toxicology and Chemistry* 29(2), 410–421.

van Dam RA, Humphrey CL & Martin P 2002. Mining in the Alligator Rivers Region, northern Australia: Assessing potential and actual effects on ecosystem and human health. *Toxicology* Vols 181 & 182, December 2002, 505–515.

Appendix 1 Further information on the analysis presented for Channel billabong fish community structure project

Files relating to the analysis contained in this document can be found in SharePoint at the directory location:

\\Environmental Impact of Mining - Monitoring and Assessment\Fish\Ranger\Channel Billabongs\Experimental design optimisation\Transect counts required

Acknowledgments

Dr Keith McGuinness, Charles Darwin University, for developing the ANOVA models used for impact detection assessments and valuable advice on the power analysis procedures used in this document.

1 Introduction and objective

1.1 Aim

To determine if sampling effort in the monitoring of fish communities using visual census methods can be reduced without reducing statistical power or community information.

1.2 Background

This report reviews the resources required to conduct fish community monitoring using the visual census method. This review forms part of the biological monitoring program review, which aims to optimise resources whilst maintain best practice monitoring.

The report explores a reduction in the number of visual counts completed along each transect within each billabong. Currently five visual counts are completed along each transect by two alternating observers. The counts at each transect are then averaged to give an average transect count. The average transect counts are then used to generate the paired-site community dissimilarity values which are used in the Impact detection analysis (ANOVA model see section 1.1.2). A reduction in the number of visual counts per transect will result in a change to the experimental design as less data per transect is included/collected. However, the proposed change would not alter the statistical impact detection test used (ANOVA model see section 1.1.2). For a reduction in the number of counts per transect to be implemented it is important that the statistical power of the impact detection test (and other commonly used analysis) is not reduced and that future community data collected is directly comparable with historic data using the impact detection ANOVA model.

This document does not consider a reduction in sampling frequency. The observed decline in fish communities (dissimilarity) between Mudginberri and Sandy billabongs is not fully understood. Annual sampling is considered necessary to elucidate the causal factors behind the declining dissimilarity. The current sampling method is outlined below.

1.2.1 Current sampling methods

The current visual sampling procedure is conducted annually, in the early Dry season, and measures the abundance of different fish species near the margins of the two channel billabongs: Mudginberri Billabong on Magela Creek downstream of Ranger (‘exposed’ site) and Sandy Billabong on Nourlangie Creek (control site). Observations are made using a small boat with a transparent observation bow. This is necessary because diving is prevented by the presence of large crocodiles (*Crocodylus porosus)*. The fish species and abundances are recorded from five 50 m transects located along the bank in each billabong. Two trained observers are used to make five counts in each of the five transects to minimise effects of observer bias and maximise the number of species detected. The structure of fish communities – species and relative numbers – are compared between the two sites and with values obtained in previous years.

1.2.2 Current impact detection model

The monitoring technique is based on principles of a Before-After-Control-Impact-Paired differences (BACIP) design described by Stewart-Oaten et al (1986, 1992). The monitoring objective is attained by comparison of data from a ‘baseline’ time series collected *before*, with data obtained *after* suspected contamination by mine waste water or some other ‘event’ or a particular period of interest.

The BACIP design uses a form of temporal replication. The *difference* between sampled responses at the *paired* sites (control-minus-exposed) at any one time is regarded as a replicate observation. Where community data are derived, multivariate *dissimilarity* values may be used as the measure of difference between the sites at each time of sampling (Faith et al 1991, 1995); these measures reduce the differences between the two communities over many different species to a single value.

The means of sets of randomly paired differences or dissimilarity measures (see below for details) between the two sites; Mudginberri Billabong on Magela Creek downstream of Ranger (exposed site) and Sandy Billabong on Nourlangie Creek (control site), before and after (BA) an event or period of interest are compared using a nested ANOVA test.

The model for the ANOVA is:

Model: BA + Year(BA)

In this model, BA is a fixed factor (or effect), testing for differences from before to after the event. Year(BA) is also a fixed factor but is nested in the BA factor with the different years sampled before and after the event. The replicate observations are the Bray Curtis dissimilarities, species richness and abundance differences from the randomly generated pairs of transects: these are used to derive an estimate of error (or residual) variation.

The ANOVA table is as follows (b = years before; a = years after; n = number of replicates)

Source	**df**	**F**
BA	1	MS_{BA}/MS_{Error}
Years(BA)	$(b-1)+(a-1)$	$MS_{Years(BA)}/MS_{Error}$
Error	$(b+a)(n-1)$	
Total	$n(b+a)-1$	

The monitoring technique is designed to evaluate the primary null hypothesis that there has been no change in fish community structure at the Magela Creek 'test' site, relative to the Nourlangie Creek control site, between two time periods of interest, eg before and after a possible impact event, between the current wet season results and those from previous wet seasons, before and after mine rehabilitation etc. Specifically, the null hypothesis of primary interest is:

H_0: Mean dissimilarity before event (or the period of interest) equals mean dissimilarity after event (or the period of interest).

If the test for BA effect (source) is significant, then the null hypothesis (1) is rejected: mean dissimilarity after the event differs (is either smaller or larger) from that before the event.

2 Methods to assess sampling effort

Data from Mudginberri and Sandy billabongs from 1994 to 2005 have been used in this analysis. Training observations or observations extra to the first five counts at each transect have not been included.

Replicate counts for each transect are averaged for each year to give one representative count for each transect in each billabong. Average transect counts have been compiled with one, two, three, four and five replicate counts per transect.

To ensure consistency of the data set across all years the following fish species (in code) have been grouped into the following;

TAM = AMM, AMA, AMJ, AM

TCS = CS, FRY

TMI = MI, MJ

All other species observed during the sampling period have been included, rare species have been included.

The effects of reduced replicate counts for each transect and its influence on statistical inference of mining impact, has been looked at in terms of;

- Community indices (Species richness and Total abundance)
- Community structure (multivariate Bray-Curtis Dissimilarity value using Log(X+1) transformation)
- Impact detection responses (dissimilarity, Species richness difference and total abundance (log) difference)
- Statistical power of Impact detection responses.

2.1 Community indices

2.1.1 Species richness

Species richness summation curves have been derived for one, two, three, four and five replicate counts for each transect in each billabong. Two curves have been generated; 1) Average species richness for each billabong - the average species richness from the 1994 to 2005 period; and 2) Average species richness per transect - the average species richness of each transect from the 1994 to 2005 period.

One way ANOVA has been used to compare the mean species richness between one, two, three, four and five replicate counts for each transect in Mudginberri billabong. Analysis has been conducted in Minitab.

Correlation analysis has been used to compare species richness from four and five replicate counts per transect in Mudginberri Billabong, using average transect data. Analysis has been conducted in Minitab.

2.1.2 Abundance

One way ANOVA has been used to compare the total abundance between one, two, three, four and five replicate counts for each transect in Mudginberri billabong. Analysis has been conducted in Minitab.

Correlation analysis has been used to compare Total abundance from four and five replicate counts per transect in Mudginberri Billabong, using average transect data. Analysis has been conducted in Minitab.

2.2 Community structure

Assessment of change in community structure using reduced counts per transect has been determined using the Bray-Curtis dissimilarity measure between Mudginberri and Sandy Billabongs (Clarke and Warwick 2001). Dissimilarity values were derived using Log (X+1) transformed abundance data for one, two, three, four and five counts per transect in each billabong using randomly paired transect values.

For each year of sampling, Bray-Curtis dissimilarity values (Clarke and Warwick 2001), or differences (species richness or abundances) were calculated for five independent pairs of transects between Mudginberri and Sandy billabongs. The transect 'pairs' were selected at random (using random without replacement) for each year from the 25 possible comparisons using the RAND function in Excel 2003. The random pairs selected remained the same for all comparative analysis.

One way ANOVA has been used to compare the mean dissimilarity between one, two, three, four and five counts per transect in each billabong. Analysis has been conducted in Minitab.

Correlation analysis has been used to compare dissimilarity values derived from four and five replicate counts per transect, using average transect data. Analysis has been conducted in Minitab.

2.3 Impact detection model

2.3.1 Statistical power

Assessment of reduction in statistical power of the impact detection model was done by comparing the standard deviation (sigma, see below) from each ANOVA test using one, two, three, four and five replicate counts for each transect in each billabong. The number of counts per transect does not alter the ANOVA model thus any reduction in statistical power would be caused by an increase in sigma.

Statistical power has been assessed for the primary community indicy (dissimilarity). Species richness difference and total abundance difference have also been analysed. Dissimilarity and difference values have been generated using the random without replacement pairing method discussed above (within community structure, section 2.2). Analysis is based on average transect data.

Sigma value has been derived from running an ANOVA with the treatment data.

Sigma = √(mean square (MS) error (residual) value from ANOVA result table)

2.3.2 Impact detection tests using ANOVA model

Impact detection tests using the ANOVA model (see section 1.1.2) were conducted for four and five counts per transect.

Before and After years were taken as:

Years before event (1994, 1995, 1996, 1997, 1998, 1999 and 2000)

Years after event (2001, 2002, 2003, 2004 and 2005).

Tests were conducted using dissimilarity values, abundance (log) difference (control – exposed) values and species richness difference values.

Correlation analysis comparing dissimilarity, species richness difference (control-exposed) and Total abundance difference from four and five replicate counts per transect has been completed in Minitab. Data used is average transect data randomly paired using the random without replacement method (see above).

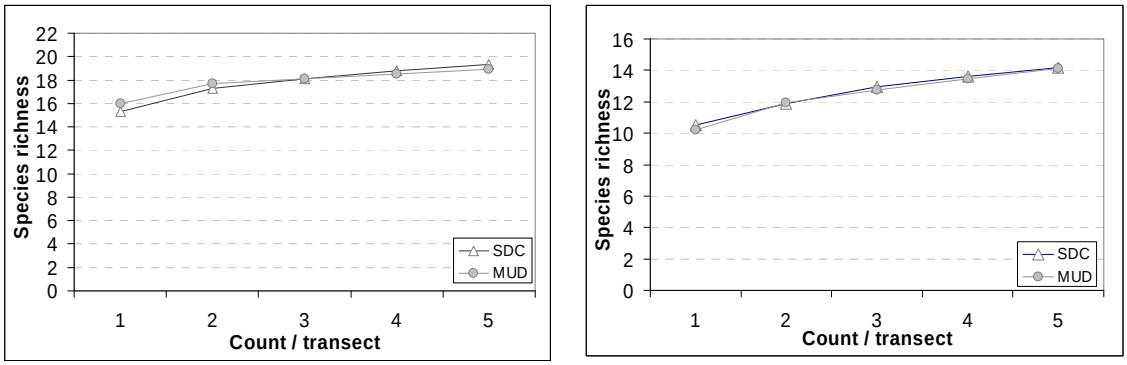
3 Results/discussion

3.1 Community indices

3.1.1 Species richness

From the species summation curves some species information is lost with reducing counts per transect. Across all counts per transect (1, 2, 3, 4 and 5) there is a significant difference in the mean species richness per transect (ANOVA, p=<0.001). There is no significant difference between four and five counts per transect (ANOVA, p=0.233), however there is a significant difference between three and five counts per transect (ANOVA, p = 0.012).

Correlation analysis between average transect species richness (4 counts per transect versus 5 counts per transect) in Mudginberri is highly significant (correlation, r = 0.961, p=<0.000). Indicating very little change in data derived from four or five counts per transect.



(A) Average species richness for each billabong

(B) Average species richness per transect for each billabong

Figure 1 Species summation curves for each billabong using 1994 to 2005 data. SDC = Sandy Billabong, MUD = Mudginberri Billabong

3.1.2 Total abundance

There is no significant difference between mean abundance per transect between 1, 2, 3, 4 or 5 counts per transect (ANOVA, p=0.994).

Correlation analysis between average transect total abundance (4 counts per transect versus 5 counts per transect) in Mudginberri is highly significant (correlation, r = 0.995, p=<0.000). Indicating very little change in data derived from four or five counts per transect.

3.2 Community structure

Plotting dissimilarity plots over time for one, two, three, four, and five counts per transect show some differences with reducing counts per transect, however very little visual difference in three, four and five counts (Figure 2). The dissimilarity decrement curve, displaying the mean dissimilarity (with 95% confidence limits) for each treatment (1,2,3,4 and 5 counts per transect), also show little difference between three, four and five counts per transect (Figure 3).

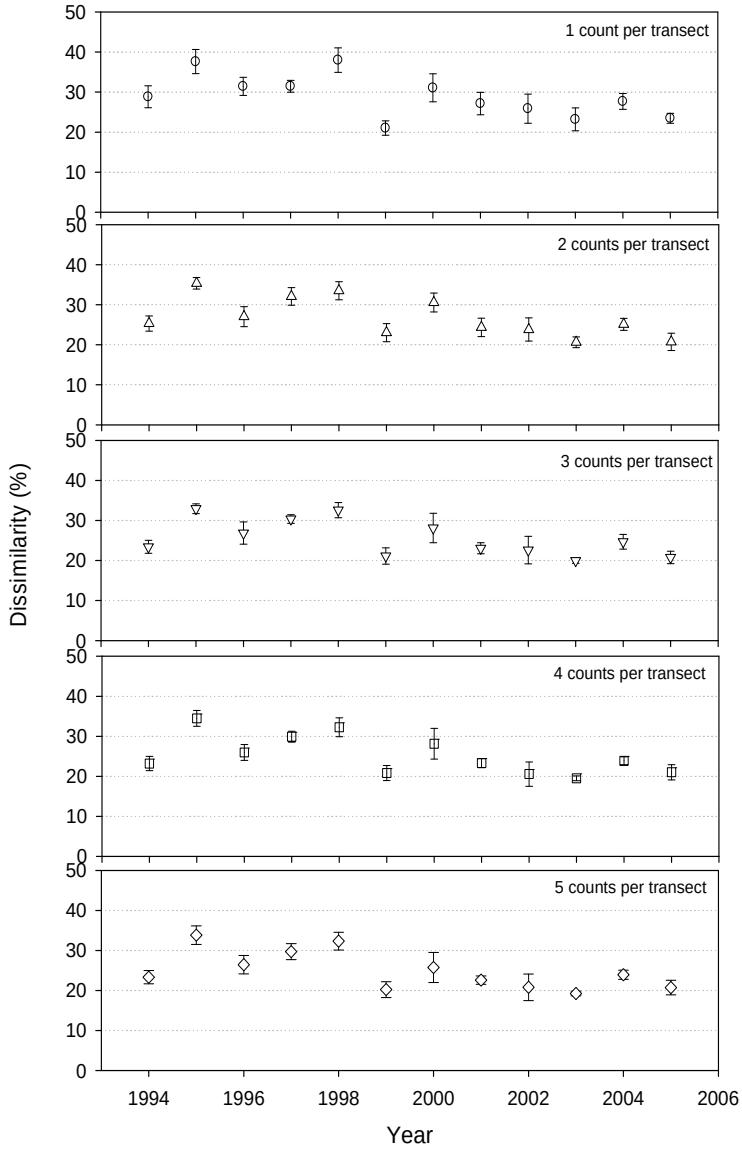


Figure 2 Bray-Curtis dissimilarity plot between Mudginberri and Sandy Billabong with data from one, two, three, four and five counts per transect

Analysis using one way ANOVA shows a significant difference between the five different counts per transect treatments (ANOVA p=0.006). However, no significant difference between 3, 4, and five counts per transect (ANOVA p=0.851), supporting the visual observation in Figure 2 and 3. Further more, the dissimilarity values for four and five counts per transect are highly correlated (correlation r=0.974, p=<0.000). Indicating very little change between the data derived from four and five counts per transect.

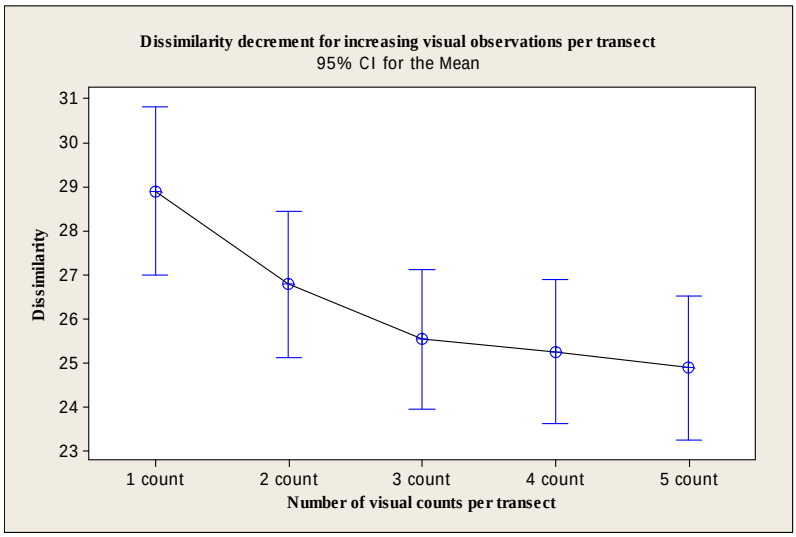


Figure 3 Mudginberri versus Sandy billabongs paired-site dissimilarity decrement curve using Log(X+1) transformation for one, two, three, four and five visual counts per transect using data from 1994 to 2005

3.3 Impact detection model

3.3.1 Statistical power

The standard deviation derived from dissimilarity values, species richness difference and total abundance differences are displayed in Figure 4. The Standard deviation of the community structure dissimilarity, species richness and total abundance does not change between four and five counts per transect. Based on this there should be no reduction in the statistical power of the impact detection model if the number of counts per transect are reduced to four.

Further more, correlation analysis between four and five counts per transect for dissimilarity, species richness difference and total abundance difference are all highly correlated (r=0.974 p=<0.000, r =0.886 p=<0.000 and r=0.953 p=<0.000 respectively) indicating good concordance of these data sets.

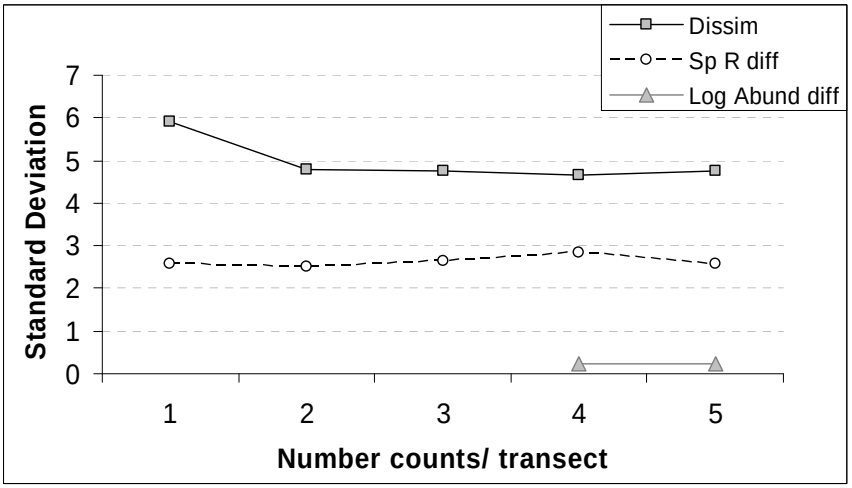


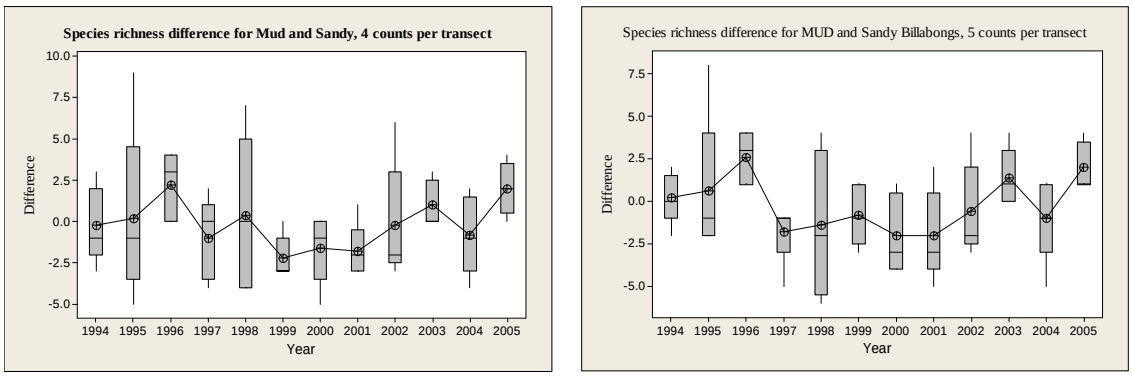
Figure 4 Standard deviation in data generated by one, two, three, four and five counts per transect. Standard deviation is sigma derived from ANOVA analysis. Only four and five counts plotted for log abundance difference.

3.3.2 Impact detection tests using ANOVA model

Analysis of the impact detection ANOVA using data derived from both four and five counts per transect shows no difference in the significance of the primary hypothesis (Before versus After in Table A1). However there is a change in 'year(BA)' for species richness difference values. It shows that years are significantly different with data derived from five counts per transect and not with data from four counts. From Figure 5 we can visualise the change between four and five counts per transect. Interestingly if you further reduce the number of counts to three counts per transect the Year(BA) is, once again, significant (p=0.022). This could be suggesting some differences between observer counts in some years.

Table 1 Comparison of significance values from four and five counts per transect using the ANOVA model: 'BA + Year(BA)'

	Dissimilarity		Species R diff		Abundance diff (log)	
	4 counts/trn	**5 counts/trn**	**4 counts/trn**	**5 counts/trn**	**4 counts/trn**	**5 counts/trn**
BA	P= 0.000	P=0.000	P=0.637	P=0.623	P=0.009	P=0.012
Year(BA)	P=0.001	P=0.002	P=0.239	P=0.044	P=0.000	P=0.000



(A) Four counts per transect (B) Five counts per transect

Figure 5 Species richness difference between Mudginberri and Sandy billabongs 1994 – 2005

4 Conclusion

A reduction from five to four counts is recommended. A reduction from five to four counts will have minimal influence on the data derived for this monitoring protocol. The primary analysis endpoint, dissimilarity values (community structure based on Log(X+1) transformed data), shows virtually no change. The dissimilarity values derived from four counts are highly correlated with dissimilarity values derived from five counts. No change to statistical power is detected as the standard deviation of data from the two data sets is almost identical. Further more no difference is detected when running a hypothetical impact detection test using the ANOVA model.

Changes to species richness are detected when reducing from five to four counts per transect. A reduction in species richness at each transect and at each billabong can be expected, however missing species are likely to be rarely observed open water species (not target species) having little influence on community structure (as indicated by no detectable change in dissimilarity values).

Observed differences in species richness difference values in the impact detection significance values (ANOVA) is interesting, but not of concern because this endpoint is highly variable and not of primary interest. The results between four and five counts (1994 – 2005 data) per transect are significantly correlated and highly variable between years. The differences observed in the ANOVA test between the number of counts per transect could possibly indicate biases within the data set. During data collection observers alternate counts at each transect, interestingly so do the ANOVA significance results (3 counts, p=0.022; 4 counts, p =0.239; 5 counts, p=0.044). The potential of observer bias within the data set will be looked at separately, however this result supports a reduction to four counts per transect so observers have an even number of counts in each transect and that the averaged data is not weighted by the observer with the most counts.

The visual observations require two alternating observers at each transect each year. The two observers counts are then assessed for any bias using the average of each observer's counts at each transect. The averaging of counts helps to reduce the natural variation (as fish swim in and out of the transect) between counts. It is not advisable to reduce the number of counts per transect below four (two counts per observer). A reduction to four counts is advantageous as it enables new observers to be trained during the monitoring program. A training observer can conduct two observations at the completion of monitoring counts (ie counts 5 and 6).

5 References

Clarke KR & Warwick RM 2001. *Change in Marine communities: an approach to statistical analysis and interpretation*. 2nd edition, Primer-E, Plymouth.

Faith DP, Dostine PL & Humphrey CL 1995. Detection of mining impacts on aquatic macroinvertebrate communities: Results of a disturbance experiment and the design of a multivariate BACIP monitoring program at Coronation Hill, Northern Territory. In *The use of biota to assess water quality*, *Australian Journal of Ecology* 20, 167–180.

Faith DP, Humphrey CL & Dostine PL 1991. Statistical power and BACI designs in biological monitoring: comparative evaluation of measures of community dissimilarity based on benthic macroinvertebrate communities in Rockhole Mine Creek, Northern Territory, Australia. *Australian Journal of Marine and Freshwater Research* 42, 589–602.

Stewart-Oaten A, Bence JR & Osenberg CW 1992. Assessing effects of unreplicated perturbations: no simple solutions. *Ecology* 73, 1396–1404.

Stewart-Oaten A, Murdoch W & Parker K 1986. Environmental impact assessment: 'Pseudoreplication' in time? *Ecology* 67, 929–940.

Appendix 2 Further information on the power analysis presented for shallow billabong fish community structure project

Files relating to the analysis contained in this document can be found in SharePoint directory, location:

> **\\Environmental Impact of Mining - Monitoring and Assessment\Fish\Ranger\Shallow Billabongs\Statistical methods assessment**

Acknowledgments

Dr Keith McGuinness, Charles Darwin University, for developing the ANOVA models used for impact detection assessments and valuable advice on the power analysis procedures used in this document.

1 Introduction

1.1 Aim

To determine if the statistical power for impact detection testing using shallow billabong fish communities can be increased by using replicate data from each billabong (average enclosure area or individual popnets).

1.1 Background

During the development of the statistical analysis for this project, by Keith McGuiness (CDU), it has been noted that the current Impact detection test (section 1.1.2) underutilises the data available within the experimental design (section 1.1.1). The current impact detection test uses the average of ten popnet traps from each billabong to derive paired site dissimilarity values that are used in a two factor ANOVA. This document explores the possibility of including replication from each billabong within a three factor ANOVA (section 2.1) to increase the statistical power to detect mining related change in shallow billabong fish communities. Outcomes of this review would not result in any change to the current experimental design – this is purely a desk top exercise.

1.1.1 Current experimental design

The experimental design here is based on six sites (3 control-impact paired sites), as recommended in the biological monitoring review workshop (this report). The three other discontinued sites included two indirectly exposed sites and one difficult to access control site. These sites have been removed to reduce the amount of resources required to sample shallow billabong fish communities.

This sampling procedure is conducted biennially (as recommended in this report), in the early dry season, and measures the abundance of different fish species in the shallow margins of 6 shallow billabongs. This includes three control sites; Buba, Sandy and Wirnmuyurr billabongs and three exposed sites; Coonjimba, Georgetown and Gulungul billabongs. Pop-net traps (2x2m pop up traps) are used to capture fish from dense vegetation in 10 quadrats, paired within five crocodile enclosure areas, in each billabong. Trained fish observers identify and count the numbers of each fish species captured. The ten replicate samples are required to adequately sample the array of species present.

The structure of fish communities – species and relative numbers – are compared between billabongs exposed to contaminants from Ranger uranium mine (RUM) and control billabongs on different catchments and with values obtained in previous years. Paired-sites have been determined based on physical similarities between billabongs (size, depth – susceptibility to drying out). Paired-sites are as follows;

Coonjimba-Buba (CJM/BUB)

Georgetown-Sandy Shallow (GTN/SDS)

Gulungul-Wirnmuyurr (GUL/WIN)

1.1.2 Current impact detection model

This monitoring technique uses three control-impact paired-sites that are based on principles of a Before-After-Control-Impact-Paired differences (BACIP) design described by Stewart-Oaten et al (1986, 1992). The monitoring objective is attained by comparison of data from a ‘baseline’ time series collected *before*, with data obtained *after* suspected contamination by

mine waste water or some other 'event' or a particular period of interest. Fish community structure data are collected at the same time of year from 10 quadrates from within each of six billabongs.

The BACIP design uses a form of temporal replication. The *difference* between sampled responses at the *paired* sites (Control-minus-impact) at any one time is regarded as a replicate observation (Stewart–Oaten et al 1986, 1992). Where community data are derived, multivariate *dissimilarity* values may be used as the measure of difference between the sites at each time of sampling (Faith et al 1991, 1995); these measures reduce the differences between the two communities over many different species to a single value.

The means of differences or dissimilarity measures between the paired-sites before and after an event or period of interest, are compared using a two factor ANOVA, with factors BA (before-after the event), and Site pair.

The model for the ANOVA is:

Model: BA+Sitepair+BA*Sitepair

In this model, BA is a fixed factor (or effect), testing the differences from before to after the event. Sitepair is also a fixed factor, testing for differences between the Sitepairs.

The ANOVA table is as follows (*b* = years before; a=years after; *sp* = number of sitepairs.

Source	**df**	**F**
BA	1	MS_{BA}/MS_{Error}
Site-Pair	(*sp*-1)	$MS_{Site\text{-}Pair}/MS_{Error}$
BA x Site-Pair	*(sp-1)*	$MS_{BA \times Site\text{-}Pair}/MS_{Error}$
Error	Sp((b-1)+(a-1))	
Total	sp(b+a) - 1	

The monitoring technique is designed to evaluate the primary null hypothesis that there has been no change in fish community structure at the exposed sites, relative to the control sites, between two time periods of interest, eg before and after a possible impact event, between the current wet season results and those from previous wet seasons, before and after mine rehabilitation etc. Specifically, the null hypothesis of primary interest is:

H_0: Mean dissimilarity before event (or the period of interest) equals mean dissimilarity after event (or the period of interest).

If the test for BA effect (source) is significant, then the null hypothesis (1) is rejected: mean dissimilarity after the event differs (is either smaller or larger) from that before the event.

2 Methods

Statistical power has been compared using three treatments of the shallow billabong fish community data; (A) using the current impact detection model (section 1.1.2) which is an average for each billabong and year – resulting in no replication for each paired-site year; (B) using the average data from each crocodile exclusion area – resulting in replication of five for each paired-site year; and (C) using the ten individual popnet trap data for each billabong – resulting in replication of ten for each paired-site year. Treatment 2 and 3 require the introduction of a three factor ANOVA (see section 2.1).

Data from years 1994, 1995, 1996, 1998, 2002, 2003, 2004, 2005 from paired-sites Coonjimba-Buba (CJM/BUB) and Georgetown-Sandy Shallow (GTN/SDS) have been used

in this analysis. Missing years are due to data not collected or data missing for one or both paired-sites. Paired-site Gulungul-Wirnmuyurr has not been included due to Wirnmuyurr only being sampled from 1998.

For the ANOVA models years before and after (BA) have been hypothetically assigned. Years before are 1994, 1995, 1996 and 1998, years after are 2002, 2003, 2004 and 2005.

The community structure has been determined using the Bray-Curtis dissimilarity measure between paired-sites (Clarke and Warwick 2001) identified above (section 1.1.1). Dissimilarity values were derived using Log (X+1) transformed abundance data between the paired-sites. Replicates created using treatments B and C have been randomly paired using a random without replacement procedure for each paired-site and year using the RAND function in Excel. The dissimilarity values have been plotted as scatter plots with error bars (standard error) over time using SigmaPlot graphics program.

All fish species observed over the sampling period have been included in the analysis.

ANOVA analysis has been conducted in MINITAB using a general linear model.

Power analysis has been conducted using Piface (Lenth 2006), see section 2.2 for more details.

2.1 Three factor ANOVA proposed

The model for the ANOVA is:

Model: BA+Years(BA)+Sitepair+BA*Sitepair+Year*Sitepair

In this model, BA is a fixed factor (or effect), testing the differences from before to after the event. Year(BA) is also a fixed factor but is nested in the BA factor with the different years sampled before and after the event. Sitepair is also a fixed factor, testing for differences between the Sitepairs.

The ANOVA table is as follows (b = years before; a=years after; sp = number of sitepairs; n = number of replicates)

Source	**df**	**F**
BA	1	MS_{BA}/MS_{Error}
Years(BA)	$(b-1)+(a-1)$	$MS_{Years(BA)}/MS_{Error}$
Sitepair	$(sp-1)$	$MS_{sitepair}/MS_{Error}$
BA x Site-Pair	$(sp-1)$	$MS_{BA \times Site\text{-}Pair}/MS_{Error}$
Year x Site-Pair	$(Sp-1)\ ((b-1)+(a-1))$	$MS_{Year \times Site\text{-}Pair}/MS_{Error}$
Error	$Sp\ (b+a)(n-1)$	
Total	$Spn(b+a)-1$	

2.2 Power analysis approach

Power analysis has been conducted using Piface (Lenth 2006) which uses a balanced number of years before and after (BA). It is unlikely this monitoring technique will be balanced, but more likely to have many years before an event with only a few, one initially, after. In this situation the analysis in Piface (based on a balanced ANOVA) is likely to be conservative, particularly when analysing only a few years before and after (ie 2 years before and after), due to the reduced number of years (n). Determining the power of unbalanced years before

and after is more complex and not really necessary for this exercise, particularly when Piface provides such a simple power analysis tool. The Piface analysis conducted here has used two, three, four and five years before and after perturbation.

Treatment 1 uses the two factor ANOVA described in section 1.1.2.

Treatments 2 and 3 use the three factor ANOVA described in section 2.1.

The requested standard deviation values used within Piface are as follows;

SD(within) = The sigma value generated from the respective treatments and ANOVA models

Sigma = √(mean square (MS) error (residual) value from ANOVA result table)

SD(ba) = The standard deviation of the difference between before and after dissimilarity values (ie dissimilarity from 50 before to 40 after, STDEV = 7.07) (see figure 1).

The significance level of all tests has been set to p=0.05.

Results of the power analysis (BA) were entered into Excel and plotted as line graphs. Piface results have been back calculated to show the statistical power of detecting the dissimilarity difference before and after, rather than the standard deviation of the difference (as produced by Piface).

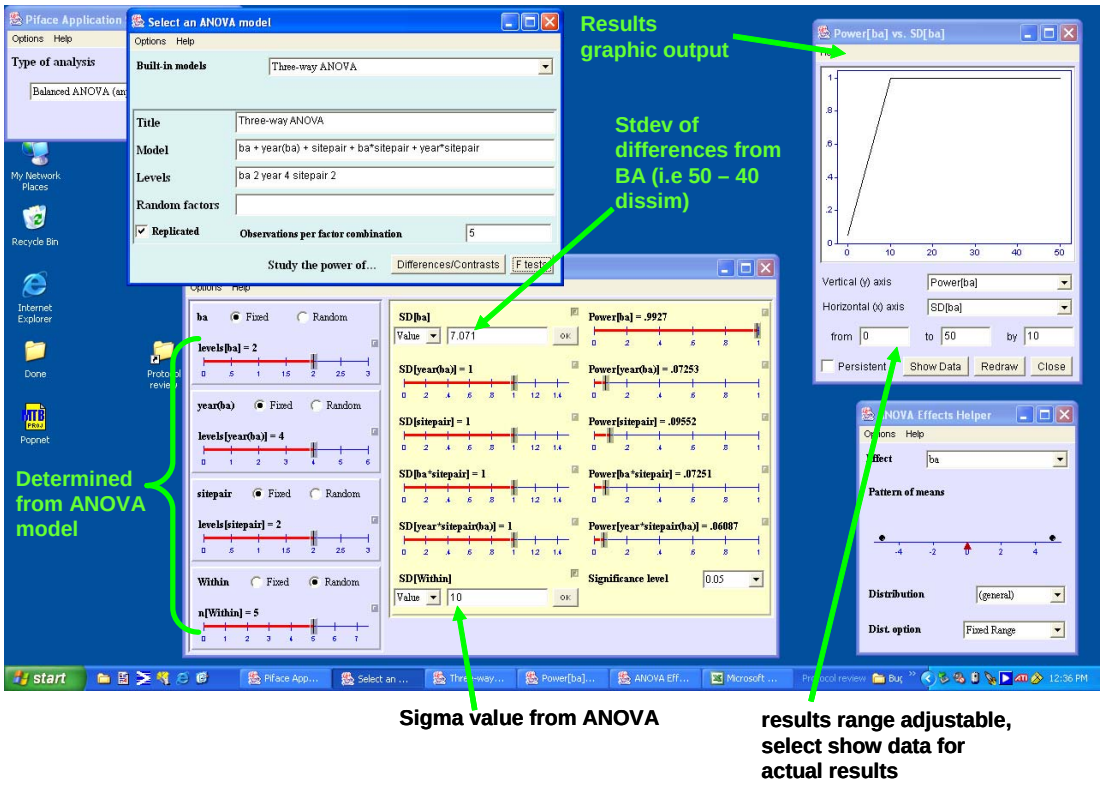


Figure 1 Screen capture of Piface imputing three factor ANOVA to determine statistical power to detect a change from 50 to 40 dissimilarity (STDEV =7.071)

3 Results

Results for the BA power analysis have been plotted for two, three, four and five years before and after perturbation (figure 2A, 2B, 2C, 2D respectively). These figures show that increasing replication within billabongs increases the power of impact detection before and after perturbation. This is most evident with fewer years before and after perturbation (2 years

before and after). Indicating that increased replication will greatly increase the statistical power when analysing with only a few years after perturbation (ie detecting change in the year of interest).

The results also show that increasing the number of years before and after an event increases the power of all the statistical tests. Indicating that the more years collected after perturbation the greater the statistical power to detect change from before to after.

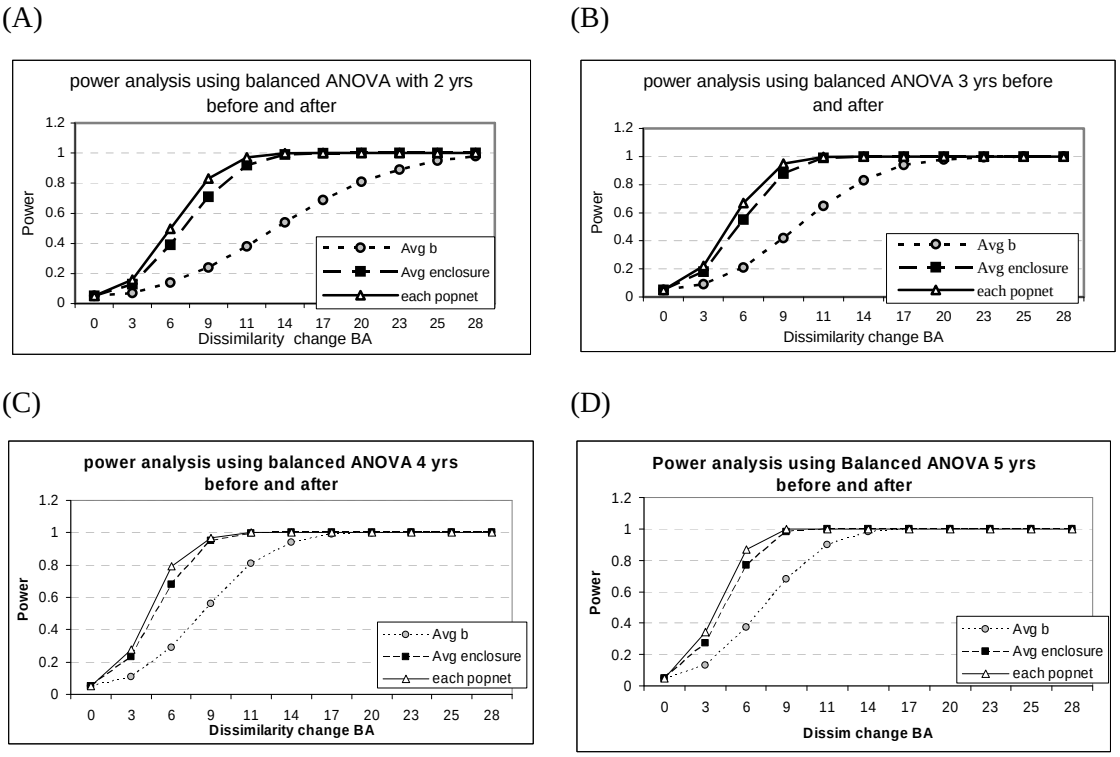
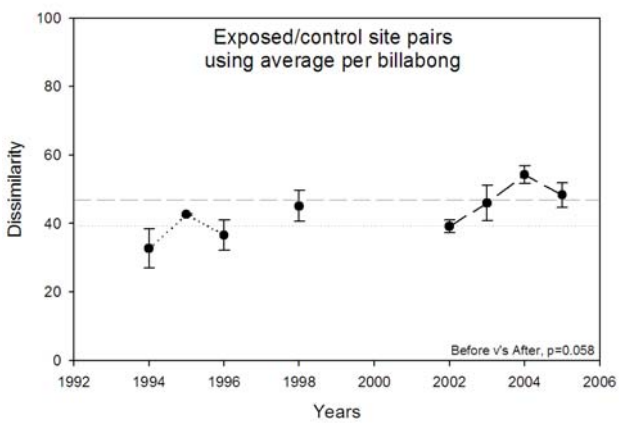


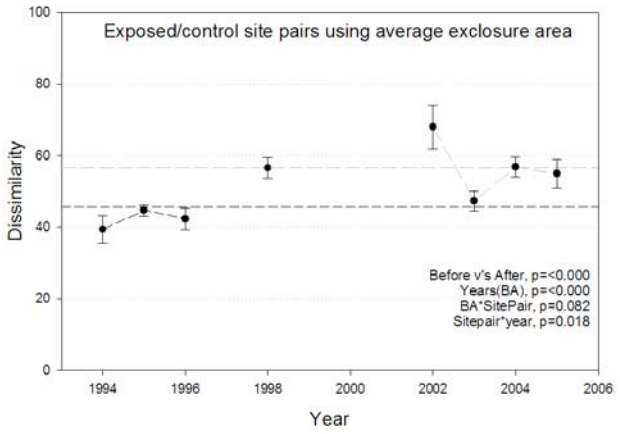
Figure 2 Statistical power analysis of three treatments of data from shallow billabong fish communities. A to D show different numbers of years before and after perturbation.

Dissimilarity results for the three treatments are presented in Figures 3A, 3B and 3C. ANOVA results from their respective models have been included in these figures. The results show that treatments B and C have greater variance in the years dissimilarity means from before to after perturbation. This is also reflected in the ANOVA tests BA results as treatment A is not significant and treatments B and C are highly significant. The results show a large discrepancy in the year 2002 between treatments A, B and C. The wet season preceding 2002 sampling had below average stream discharge with many sites recording increased aquatic vegetation within sample areas. The inclusion of replicates has resulted in increased statistical power to detect differences between paired sites each year. The average billabong (treatment A) has masked the increased variation in fish communities, possibly resulting from increase aquatic vegetation, between the site pairs.

(A)



(B)



(C)

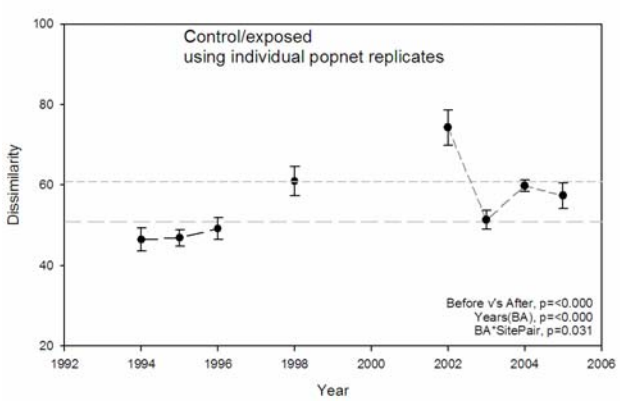


Figure 3 Paired control-exposed site (Coonjimba-Buba and Georgetown-Sandy Shallow) dissimilarity values (using Bray-Curtis measure) calculated for community structure in fish in shallow billabongs, values are means (± standard error) from three different replication treatments. (A) No replication for years within each pair-site (average value for each billabong); (B) Five replicates per paired-site (average value for each enclosure area in each billabong); (C) Ten replicates per paired-site (individual popnet trap result in each billabong).

4 Conclusion

Increased power can be obtained from including replicates in the statistical impact detection test. Although greatest power is achieved using individual pop net traps it would be preferable to use the average enclosure area values (five replicates, treatment 2). Enclosure areas are suggested because the setting of crocodile exclusion nets (which form the enclosure areas) groups popnets into five pairs, each pair typically having similar habitat. Thus this would seem a logical approach.

5 References

Clarke KR & Warwick RM 2001. *Change in marine communities: an approach to statistical analysis and interpretation*. 2nd edition. Primer-E, Plymouth.

Faith DP, Dostine PL & Humphrey CL 1995. Detection of mining impacts on aquatic macroinvertebrate communities: Results of a disturbance experiment and the design of a multivariate BACIP monitoring program at Coronation Hill, Northern Territory, In The use of biota to assess water quality, *Australian Journal of Ecology* 20, 167–180.

Faith DP, Humphrey CL & Dostine PL 1991. Statistical power and BACI designs in biological monitoring: comparative evaluation of measures of community dissimilarity based on benthic macroinvertebrate communities in Rockhole Mine Creek, Northern Territory, Australia. *Australian Journal of Marine and Freshwater Research* 42, 589–602.

Lenth RV 2006. Java Applets for Power and Sample Size [Computer software]. Retrieved October 2nd, 2006, from http://www.stat.uiowa.edu/~rlenth/Power.

Stewart-Oaten A, Bence JR & Osenberg CW 1992. Assessing effects of unreplicated perturbations: no simple solutions. *Ecology* 73, 1396–1404.

Stewart-Oaten A, Murdoch W & Parker K 1986. Environmental impact assessment: ‘Pseudoreplication’ in time? *Ecology* 67, 929–940.