

Australian Government

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





This page has been left blank intentionally.

Environmental monitoring protocols to assess potential impacts from Ranger minesite on aquatic ecosystems: Macroinvertebrate community structure in streams

Supervising Scientist Division

Supervising Scientist Division GPO Box 461, Darwin NT 0801

July 2013

Project Number: MON-1995-001

Registry File: SSD2010/0062

(Release status - Unrestricted)



Australian Government

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

How to cite this report:

Supervising Scientist Division 2013. Environmental monitoring protocols to assess potential impacts from Ranger minesite on aquatic ecosystems: Macroinvertebrate community structure in streams. Internal Report 591, July 2013, Supervising Scientist, Darwin.

Project number – MON-1995-001

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

<u>Supervising Scientist Divison>SSDX>Environmental Impact of Mining - Monitoring and</u> <u>Assessment>Macroinvertebrates>Ranger streams</u>

Authors of this report:

Supervising Scientist Division, 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 2013

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 enquiries concerning reproduction and rights should be addressed to Publications Enquiries, Supervising Scientist, GPO Box 461, Darwin NT 0801.

e-mail: enquiries_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

E	xecutive summary	v
Ρ	reamble	vi
A	cknowledgements	vii
1	Introduction	1
	1.1 Objective	1
	1.2 Background	1
	1.3 Principle of the monitoring technique	2
	1.4 Taxonomic resolution	4
2	Experimental design	5
	2.1 History of experimental design	5
	2.2 Current statistical design and analysis	5
3	Field procedures	14
	3.1 Occupational health and safety	14
	3.2 Consultations required for site access	14
	3.3 Timing of sampling	15
	3.4 Sampling sites	15
	3.5 Number of replicate samples and habitat sampled	16
	3.6 Measurement of environmental variables	17
	3.7 Macroinvertebrate sampling	18
	3.8 Field QA/AC	19
4	Laboratory procedures	20
	4.1 Order of processing of samples	20
	4.2 Sub-sampling, identification and enumeration	20
	4.3 QA/QC for laboratory procedures	21
5	Data entry, storage and associated QA/QC	22
	5.1 Data Storage	22
	5.2 Data entry and QA/QC	22
6	Data analysis	23
	6.1 Data preparation	23
	6.2 Impact detection and assessment	24
7	Reporting	34
	7.1 Overview	34
	7.2 Timing	34

7.3	Reporting results to Traditional Owners and Aboriginal residents	34
7.4	Supervising Scientist Annual Report	35
7.5	Internet	35
7.6	Alligator Rivers Region Technical Committee and Annual	
	Research Summary (Supervising Scientist Report)	35
7.7	Summary report for stakeholders	36
8 Ref	erences and additional reading	37
Appe	ndix 1 Comparison of Analysis Of Variance (ANOVA) with	
PE	Rmutational MANOVA for the Ranger macroinvertebrate	
CO	mmunity, stream monitoring program	39
A1	Background	39
A2	Comparison of PERMANOVA and MBACIP ANOVA analysis	
	approaches	40
A3	Comparison of PERMANOVA and MBACIP ANOVA results	45
A4	PERMANOVA and Minitab program functionality differences	49
A5	Conclusions	52
A6	References	53
Annex	C – PERMANOVA and MBACIP ANOVA results for mid	
ye	ar reporting: Assessment of 2010–11 wet season results	54
Ba	ckground	54
Im	pact detection and assessment	54

Executive summary

The SSD operates an integrated chemical (including radiological), physical and biological monitoring program to ensure protection of the aquatic ecosystems of the ARR from the operation of uranium mines in the region. This stream monitoring program is an independent, assurance program, unlike the compliance and check water chemistry monitoring programs of the mining company (Ranger) and the NT government regulator respectively.

The techniques and 'indicators' used in the monitoring program satisfy two important needs of environmental protection: (i) the early detection of significant changes in measured indicators to avoid short or longer term ecologically important impacts; and (ii) assessing ecological or ecosystem-level effects by way of measured changes to surrogate indicators of biodiversity.

The SSD has prepared protocols for the measurement programs required to implement each of these monitoring techniques. For each technique, two types of protocols have been prepared, high-level protocols and detailed operational manuals. This document is the high-level protocol, describing the science underpinning one of the biological biodiversity-assessment techniques, namely monitoring of macroinvertebrate communities in streams.

This protocol for the stream macroinvertebrate monitoring technique provides an overview of the monitoring principles and objectives, experimental and statistical design, field, laboratory, data analysis and impact assessment procedures and reporting requirements.

Preamble

Full descriptions of the field and laboratory procedures, methods for data collation and worked examples of data analysis required to implement the methods described in this protocol are contained in a companion "Operational manual". The manual is the working document used by monitoring staff at the Jabiru Field Station to conduct the stream macroinvertebrate monitoring technique. The Operational manual is in a loose-leaf, ring-bound form, allowing for ready revision and update. Revisions must be approved by the SSD Monitoring Support Unit (Darwin) before the Operational manual can be updated.

The specific material contained in the Operational manual includes:

- datasheet pro forma for field, laboratory and QA/QC results
- details of all field and laboratory procedures used for macroinvertebrate sampling and sample processing, including:
 - o instructions on use of meters, microscopes and other instruments
 - o photographs and maps of the location of sample sites
 - data codes for macroinvertebrates and environmental variables
 - o taxonomic keys and photographs of macroinvertebrate taxa
- details of all statistical procedures used for data analysis and worked examples of these procedures
- examples of all required corporate reports
- training pro-forma
- summary information on key references and supporting studies.

Acknowledgements

The following *eriss* personnel have been involved in the development of this protocol since it was conceived in 1988: Chris Humphrey, Lisa Chandler, Duncan Buckle, Alistair Cameron, Ruth O'Connor, Lisa Thurtell, Frederick Bouckaert, Peter Dostine, Abbie Spiers, Julie Hanley and Caroline Camilleri.

Many volunteers and temporary employees, including local Aboriginal people, have assisted with the fieldwork as data recorders, crocodile spotters and general field hands over this time.

Traditional owners of the country containing the sampling sites (Gagadju and Mirrar) have generously allowed access to the sites and assisted in the fieldwork.

Advice and assistance with access from staff of Kakadu National Park and ERA is gratefully acknowledged.

Dr Keith McGuinness from Charles Darwin University has provided valuable advice during recent years on the use of statistical procedures for data analysis. Drs David Jones and Amy George provided critical review of the draft protocol.

Contact officers:

Dr Chris Humphrey Environmental Research Institute of the Supervising Scientist, PO Box 461, Darwin NT 08 8920 1160 Chris.Humphrey@environment.gov.au

Lisa Chandler

Environmental Research Institute of the Supervising Scientist, Locked Bag 2, Jabiru NT, 0886 08 8979 9707 Lisa.Chandler@environment.gov.au

1 Introduction

1.1 Objective

The objective is the detection of any¹ changes to macroinvertebrate communities in streams downstream of the Ranger Uranium Minesite associated with dispersed mine waste-waters.

1.2 Background

1.2.1 General Introduction

Of the types of biological assessment available for aquatic ecosystem protection, biological monitoring using aquatic macroinvertebrate communities is regarded in the Australian and New Zealand guidelines for fresh and marine waters (ANZECC & ARMCANZ 2000) as a form of 'biodiversity' assessment, ie used to determine the ecological importance of impacts or the effects upon the ecosystem as a whole. Implicit in this assessment objective is the knowledge that no programs are sufficiently resourced that they are able to measure every facet of ecosystem 'health'. Hence, it is inevitable that biodiversity assessment will be limited to the measurement of ecosystem surrogates — communities/assemblages of organisms, or habitat or keystone-species indicators that have been closely linked to ecosystem-level effects. Information on the ecological importance of effects will best be met in programs that have regional coverage and encompass a full disturbance gradient (ANZECC & ARMCANZ 2000).

Benthic macroinvertebrates (operationally, > 0.5 mm in size) encompass a diverse assemblage of organisms that are common to all freshwater ecosystems. Benthic macroinvertebrates include insects, crustaceans, molluscs, flatworms and annelid and nematode worms. They play a major role in the functioning of these systems with a variety of functional groups represented, including shredders (feeding on large pieces of organic matter), collector-filter feeders (removing fine particulate matter (FPOM) from surrounding waters), collector-gatherers (gathering FPOM deposited in streams), scrapers (grazing or scraping attached algae and associated microbes from various surfaces), wood miners (burrowing into, and living in, submerged wood) and predators. A diverse and abundant macroinvertebrate community in freshwater ecosystems is also important for the health of organisms higher in the food chain that are dependent upon macroinvertebrates as their key food source.

Benthic macroinvertebrates are regarded as the most broadly applicable group for biological monitoring of freshwater ecosystems worldwide, including Australia. Included amongst their inherent advantages are: widespread distribution within most habitats; generally limited mobility; the considerable diversity of forms such that there are very few water quality stressors and habitat alterations to which macroinvertebrate community structure is unlikely to respond to; ease of collection using well-established sampling techniques; and widespread skill base including taxonomic and general life history knowledge (Hellawell 1986, ANZECC & ARMCANZ 2000).

The aquatic life stages of most benthic macroinvertebrates are relatively long-lived and are commonly resident in tropical streams for up to several weeks, though some taxa (eg bivalve molluscs) can live for several years or more. This attribute and the generally sedentary nature

¹ Enhanced (stimulatory) effects, eg increased number of species and abundances, may indicate the response of macroinvertebrate communities to low-level contaminant concentrations (such as solutes, nutrients) which could provide sufficient cause to trigger management action.

and environmental sensitivities of macroinvertebrates, when combined with sound sampling designs, provide compelling *prima facie* evidence of disturbance or impact when communities at exposed sites differ significantly from those recorded at representative reference sites and/or from the same sites sampled prior to disturbance. This principle underpins the objective of the protocol as stated above (section 1.1).

1.2.2 Some background relevant to ARR studies

Changes to the major streams downstream of Ranger that are associated with mining activities manifest predominately as changes to surface water quality (as opposed to changes in habitat). Hence mine-related changes to the biota of these receiving water ecosystems will potentially and predominately be associated with water quality alterations. For Northern Hemispheric waters and some systems in southern Australia, information about an organism's tolerances to certain types of contaminants are used to interpret and infer degree of water pollution (ANZECC & ARMCANZ 2000), and various biotic indices have been derived to quantify such biological contaminant effects. This supporting diagnostic information is not yet available for wet-dry tropical streams of Australia. Therefore mine-related changes in macroinvertebrate communities may be evident through changes to community structure, ie changes to the relative abundances of the different taxa comprising the community.

Monitoring of macroinvertebrate communities in the sand channels of streams downstream of Ranger commenced in 1988. Between 1988 and 1993, up to 10 sites in Magela Creek, upstream and downstream of Ranger, were sampled annually. From 1994, annual sampling was extended to an additional (mine-) 'exposed' stream, Gulungul Creek, as well as two control streams, Burdulba and Nourlangie Creeks. Further refinements to sampling were made between 1994 and 1997, until the final sampling program described in this protocol was settled upon in 1998. The history and rationale behind this annual sampling program is described in section 2.1 below.

1.3 Principle of the monitoring technique

Monitoring is conducted at the wet-dry transition (April-May period) (see rationale in section 0). Paired upstream-downstream sites are sampled in the two exposed streams, Magela and Gulungul Creeks, and in the two control streams, Burdulba and Nourlangie Creeks – see Figure 1. (The presence of a second upstream site in Magela Creek, MCS2, is explained in section 3.4 below.) Sites on the two exposed streams are upstream and downstream of the Ranger minesite respectively. Amongst streams, the distance between the paired sites is similar (Figure 1).

On each sampling occasion and at each site, five replicate benthic samples are collected from the stream bed in macrophyte habitat found along the creek edges and in flowing waters. Water quality and habitat data are gathered concurrently at each site. Samples are preserved in ethanol and sorted under stereo-microscope later in the laboratory to extract macroinvertebrates (> 0.5 mm). Animals are identified, at least to family-level, and enumerated.



Figure 1 Map showing location of paired upstream-downstream macroinvertebrate sampling sites in 'exposed' (Magela and Gulungul Creeks) and control (Burdulba and Nourlangie Creeks) streams in relation to Ranger minesite and lease.

In general terms, the structure of macroinvertebrate communities – families and their relative abundances – is compared between sites exposed to minewater contaminants and control sites upstream (in the same creek). The behaviour or concordance of that paired site comparison in community structure is in turn compared with (i) the same comparison of paired (upstream-downstream) sites measured at the same time in control streams unaffected by mining, and (ii) paired site comparisons obtained in previous years from both exposed and control streams.

1.4 Taxonomic resolution

Where biological assessment involves study of macroinvertebrate communities, it is often assumed that identifications at lower taxonomic levels (eg genus or species) will provide greater sensitivity to impact detection than where identifications are based upon higher taxonomic levels (eg family-level). Where impact detection is the issue, the assumption of greater sensitivity at lower, especially species, level is based upon the reasonable argument that congeners may vary in their sensitivities to different stressors (eg Cranston 1990). Hence by summarising a response at a higher taxonomic level, critical information may be lost. However, statistical sensitivity is also a key factor in determining appropriate taxonomic resolution because precision of the community summaries (eg dissimilarity values, taxa richness) used to measure spatial and temporal change may vary according to taxonomic level.

Humphrey and Chandler (2009) undertook an analysis of macroinvertebrate data from three streams in the Alligator Rivers Region that aimed to inform decision-making on the appropriate level of taxonomic resolution to apply to biological monitoring in the region. They predicted that: (i) data summarised at lower taxonomic levels would be more variable over time (within either the 'before' or 'after' impact period) that those summarised at higher taxonomic levels, but (ii) the magnitude of change from 'before' to 'after' impact for data summarised at lower taxonomic levels would be greater than that summarised at higher taxonomic levels. These predictions might imply that the advantage of greater magnitude of change detected at lower taxonomic levels after impact (aspect (ii) from above) may be lost or diminished by loss in statistical power associated with the higher variability in these data.

In the analysis by Humphrey and Chandler (2009), the predictions were confirmed. In particular, while the magnitude of change before and after impact (in Rockhole Mine Creek) was observed to be greater at species than family level, the difference in sensitivity was small; a conclusion that supported an earlier and independent (and different) analysis of the same data by Faith et al (1995).

For macroinvertebrate studies in the ARR, species-level identifications are about three times more costly in staff resources than family-level identifications, and require a more specialist skill base (Humphrey & Chandler 2009). On the other hand, only identifications at lower taxonomic (especially species) levels can provide information for biodiversity and conservation assessment. This may be particularly important for sites of high conservation value such as in the ARR to address scenarios such as: (1) an impact has been detected but what is the ecological importance of this; have any species been 'lost' from the system?; or (2) assurance that subtle change is not occurring (eg gradual species replacement). For now, a dual family- and species-level analysis is being applied. Species-level data are being sought for (upstream and downstream) sites in Magela Creek, the receiving waters of main concern for Ranger, while only family-level data are being sought from all other sites.

2 Experimental design

2.1 History of experimental design

The current protocol for assessing potential impacts from mine wastewater discharge using macroinvertebrate communities has undergone considerable review and evolution since its initial implementation in 1988. Research focussed on developing methods that could be used for monitoring and followed peer review and best-practice at the time. The changing sampling approaches and significant review points are summarised in Table 1. The sampling design evolved over time from a simple within-stream (Magela-only) design to increasingly more sophisticated variants of the BACI (Before-After-Control-Impact) class of design. These revisions in sampling approach are explained in more detail in Humphrey & Pidgeon (1998). In 2001, SSD took on the responsibility and role of independent off-site monitoring for the Ranger Uranium Mine and since then has reported the results of the current monitoring program on an annual basis in the Supervising Scientist's annual reports to Parliament.

Years	Method	BACI-design type
1988-1993	Sampling of 7-10 sites in Magela Creek (only) (2 upstream of Ranger, 5-8 downstream).	BACI; gradient
1994	Sampling of three sites in Magela Creek (two upstream and one downstream of Ranger), and similar paired (up-/downstream) sites in 5 other control streams. Relocation of the downstream site in Magela Creek to SSD's monitoring site, MCDW.	Unbalanced MBACIP
1995-96	Same as 1994, but rapid assessment approaches used for sampling and sample processing.	Unbalanced MBACIP
1997	Sampling and sample processing reverted to pre-1995 traditional quantitative approaches at three Magela Creek sites. Magela-only study focussed on the required amount of within-site replication and sampling area per replicate to apply in future years.	N/A
1998 to present	Quantitative sampling approach derived from 1997 study applied to paired (up-/downstream) sites in 'exposed' streams, Magela and Gulungul Creeks (two upstream sites in Magela Creek), and 'control' streams, Burdulba and Nourlangie Creeks.	Balanced MBACIP

Table 1 Changes to the sampling design of the Ranger macroinvertebrate monitoring study from 1988to the present (see also Humphrey & Pidgeon 1998).

2.2 Current statistical design and analysis

The current monitoring technique (last revised in 1998) is based on principles of a Before-After-Control-Impact-Paired differences (BACIP) design described by Stewart-Oaten et al (1986, 1992). As applied to a single stream where an adjacent mine site is located, macroinvertebrate data are collected simultaneously at two sites, one located upstream (*Control*) and the other downstream (*Impact* or '*exposed*' site) of the mine, both *Before* suspected contamination by mine waste water (or some other 'event' or a particular period of interest) and *After*. The BACIP design uses a form of temporal replication. The *difference* between sampled responses at the *Paired* sites (upstream-minus-downstream) at any one time is regarded as a replicate observation with the replicate differences gathered at multiple times before and after impact.

The basic BACIP design in this study has two important embellishments:

- 1. Because community data are acquired, multivariate *dissimilarity* values substitute as the measure of difference between the sites at each time of sampling (Faith & Cranston 1991, Faith et al 1995); these measures reduce the differences between the two communities over many different species to a single value. A value of 0 indicates macroinvertebrate communities identical in structure, while a value of 100 (percent) indicates totally dissimilar communities, sharing no common taxa.
- Paired sites in an additional impact stream and in two control streams are added to the basic single-stream model with the extended design being termed M-BACI (Keough & Mapstone 1995) where 'M' denotes the inclusion of multiple control sites (or 'streams' in this case). Tests for impact and interaction are conducted within a multi-factor ANOVA.

MBACI(P) designs are generally assumed to be asymmetric, meaning they have replicated control sites but only a single impact site (eg Underwood 1993, Keough & Mapstone 1995). For example, an asymmetric model is applied to the macroinvertebrate study designed for proposed mining at Jabiluka where Ngarradj was the only 'exposed' stream of several streams sampled (McGuinness 2003). However, the Ranger macroinvertebrate study is not asymmetric, but instead the design is spatially (but not temporally) balanced, with two 'exposed' streams, Magela and Gulungul, and two control streams, Burdulba and Nourlangie – as depicted in Figure 2.



Figure 2 Schematic of the monitoring design used for macroinvertebrate community structure in the seasonal streams around Ranger Mine.

Replicate observations in the MBACIP model are Bray-Curtis dissimilarity values derived for each of the randomly-paired, upstream-downstream replicates (from section 6.1.3 and Figure 2); the replicate observations provide an estimate of error (or residual) variation. For the community data collected for each stream, five dissimilarity values are derived, each representing one of the five possible randomly-paired upstream and downstream replicates. Mean dissimilarity values are calculated Before and After (BA) an event or period of interest.

In general terms, the expectation is that any impact which occurs should predominantly affect the downstream site of either, or both, exposed streams, causing macroinvertebrate communities at the site(s) to diverge from those at the upstream site(s) of the same stream(s). Thus, an impact should be detectable by observing that, after the impact commences, the upstream–downstream dissimilarities on the exposed stream(s) increase, relative to changes on the control streams.

A key caveat to be applied to this biological monitoring technique is that the hypotheses tested in formal analyses of change detection (section 2.2.1.2) consider 'Before' data as commencing no earlier than 1998, when sampling and processing methods were last standardised across sites from exposed and control streams. It should be noted generally that for this and other biological monitoring techniques used to assess potential impact associated with mining at Ranger, inferences that may be drawn from the time series data are, a priori, weakened because there are no similar pre-mining baseline (pre-1980) data upon which to compare. Nevertheless and despite the lack of pre-mining baseline data, comprehensive and detailed chemical and biological monitoring, as well as ecotoxicological, assessments (or multiple lines of evidence) have been conducted since 1980. The results of the monitoring and assessments are reported upon annually by the Supervising Scientist. Annual reporting has concluded that any mine-related changes to water quality that may occur at the compliance site in receiving waters downstream of Ranger from year to year have been insufficient to have altered the structure and function of resident biological communities.

2.2.1 MBACIP design with ANOVA

2.2.1.1 MBACIP design and ANOVA model

The BA values are statistically tested (compared) using a nested four factor ANOVA model.

The model for the ANOVA is:

BA + Year(BA) + Exposure + Stream(Exposure) + BA*Exposure + BA*Stream(Exposure) + Year(BA)*Exposure + Year(BA)* Stream(Exposure)

In this model, BA is a *fixed* factor (or effect), testing for differences between 'before' and 'after' periods. Exposure is also a *fixed* factor, testing for differences between exposures. However, Year(BA) is a random factor, nested in the BA factor because different years are sampled before and after the event (or between the periods of interest). Stream(Exposure) is also a *random* factor, nested in the exposure factor because streams are unique and only allocated to one exposure condition. Some comment is required on the choice of Years(BA) as a *random* factor since there is debate in the literature about whether 'time' in ANOVA designs should be assigned as 'fixed' or 'random' (see Quinn & Keough 2002). Although the years are a sequential (and not random) sample, the primary aim of this analysis is to answer the question: Does the magnitude of any change from before to after exceed that expected, given the natural variation from year to year? Treating Year (BA) as a random factor, addresses this question because years are not selected on the basis of particular characteristics they possess that may influence the response. While the assignment of time as a random factor would likely lead to a more conservative (less powerful) test (compared to fixed-factor designation), it is also worth noting that the BA test performed in this situation is comparable to the Student t-test method recommended by Stewart-Oaten et al (1986), authors of the original BACIP analysis method, and subsequently adopted routinely in numerous subsequent studies, where 'Seasons' (or time) are also treated as random.

Community summaries such as species richness and abundances are not used in analyses of the data gathered in the current protocol. However, they could be included as response measures, using corresponding (univariate) difference values between the randomly-paired, control-exposed (upstream-downstream) sites as input data to the ANOVA model.

The ANOVA results table is described in Table 2.

Where significant differences are found to occur between a time-series of stream dissimilarity values, further investigation is required to assess whether or not the change is associated with inputs from the mine. Water quality variables measured over the duration of the wet season in Magela and Gulungul creeks upstream (SSD and Energy Resources of Australia water chemistry monitoring programs) and at each site at the time of macroinvertebrate sampling, are used to assist in determining the potential influence of mine waste-water inputs upon macroinvertebrate communities. Flow dependant taxa are used to help determine the influence of any change in magnitude of wet season from year to year (ie more flow dependant taxa might be present after a larger wet season that has resulted in higher flow velocities).

Table 2ANOVA results table used for monitoring macroinvertebrate communities in seasonal streamsaround Ranger Mine. BA and Exposure are fixed factors, Year(BA) and Stream are regarded as randomfactors.(b = years before; a=years after; St = number of streams; n = number of replicates).

Source	df	F
BA	1	No exact F test
Year(BA)	(b-1)+(a-1)	$\text{MS}_{\text{Year(BA)}} \ / \ \text{MS}_{\text{Year(BA)}^*\text{Stream(Exposure)}}$
Exposure	1	No exact F test
Stream(Exposure)	2(St -1)	$MS_{Stream(BA)} \ / \ MS_{Year(BA)^*Stream(Exposure)}$
BA* Exposure	1	No exact F test
BA* Stream(Exposure)	2(St -1)	$MS_{BA^*Stream(BA)} / MS_{Year(BA)^*Stream(Exposure)}$
Year(BA)* Exposure	(b-1)+(a-1)	$MS_{Year(BA)^{*}Exposure} \ / \ MS_{Year(BA)^{*}Stream(Exposure)}$
Year(BA)* Stream(Exposure)	[(b-1)+(a-1)]*[2(St-1)]	$\rm MS_{\rm Year(BA)^*Stream(Exposure)} \ / \ \rm MS_{\rm Error}$
Error	2St(b+a)(n-1)	
Total	2Stn(b+a) -1	

2.2.1.2 Hypotheses

The monitoring technique based upon the ANOVA model is designed to evaluate the primary null hypothesis that there has been no change in macroinvertebrate community structures in the exposed streams, relative to the control streams:

- 1. between two time periods of interest, eg before and after a possible impact event, or
- 2. between the current wet season and all previous wet seasons, or
- 3. before and after mine rehabilitation etc.

Specifically, the null hypothesis of primary interest is:

1. H₀: Mean dissimilarity before event (or the period of interest) is consistent to mean dissimilarity after event (or the period of interest) between exposure type.

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

Whether significant or not, results of the BA*Exposure interaction must always be interpreted together with those from the 'BA*Stream(Exposure)' interaction which considers whether or not the change in magnitude of paired-site dissimilarity within either, or both, exposure types

(exposed and control streams) is consistent within or between the before and after periods. The null hypothesis for this second interaction is:

2. H₀: Mean dissimilarity before event (or the period of interest) is consistent to mean dissimilarity after event (or the period of interest) for streams within or between exposure type.

If the test for BA*Stream(Exposure) effect is significant, then the null hypothesis (2) is rejected. In this monitoring design, the BA*Stream(Exposure) interaction is likely to be the most important because it will indicate if, within one or both of the two exposure types, one of a stream's upstream-downstream dissimilarity is responding differently to that of the other stream (eg exposed creek Magela has changed over time compared with corresponding exposed creek Gulungul by virtue of the higher concentrations of contaminants it receives from the minesite). In this situation, further investigation of the aberrant stream upstream-downstream dissimilarity would be required.

Concomitant examination of the results of testing of both the primary null hypothesis (BA*Exposure interaction) and the secondary hypothesis (BA*Stream(Exposure)' interaction) is important in the event of one of three important outcomes and interpretations:

- 1. The Primary hypothesis is accepted but the secondary hypothesis is rejected: indicates within one or both exposure types, one of a stream's upstream-downstream dissimilarity is responding differently to that of the other stream.
- 2. The Primary hypothesis is rejected and the secondary hypothesis is rejected: same interpretation as for (1) but the intensity of the significant response is greater than in (1)
- 3. The Primary hypothesis is rejected and the secondary hypothesis is accepted: indicates within one or both exposure types, both of the stream's upstream-downstream dissimilarity are responding similarly and have changed.

A number of other hypotheses are tested in this ANOVA model, most relating to temporal or spatial variation that is not relevant to impact detection. For example, an interaction that is consistently significant after each annual analysis and assessment is the Year(BA)*Stream(Exposure) interaction. This simply indicates that dissimilarity values for the different streams – regardless of their status (BA, exposure) – show differences through time. Interpretation of results arising from all factors and interactions in this model are detailed in Table 3.

Possible causes of any observed changes or trend are assessed and/or investigated further to ensure accurate inference about mining impact.

2.2.1.3 Hypotheses for mid-year annual reporting

Results of the sampling program are reported on an annual basis in the Supervising Scientist's annual reports to Parliament (section 2.1 and 7.4). At the time the annual report is drafted, in July of each year, not all replicate samples collected from the previous (just-completed) wet season have been processed (samples collected in the recessional flow period, April or May). As a compromise for reporting, so that an interim assessment of potential mining impact may be made, priority is given to processing the replicate samples from just the exposed streams, Magela and Gulungul Creeks (section 4.1).

In the absence of data from the control streams, it is not possible to run the full ANOVA as described above (section 2.2.1.1) and instead, a modified ANOVA model is run using the factors Before/After (BA; fixed), Year (nested within BA; random) and Stream (upstream vs downstream paired dissimilarities; random), examining just the exposed creeks, Magela and

Gulungul, to determine if any change in these streams has occurred. (Thus all 'Before' data employed in this analysis also comprise just Magela and Gulungul data. The approach is effectively a paired BACIP design.)

Factors and interactions from Table 2 that are now *not* relevant to this modified analysis include Exposure, BA*Exposure and Year(BA)* Exposure. Modified hypotheses for the factor and interaction of relevance (from Table 2) are:

1. BA: When averaged over both exposed streams, the magnitude of upstream-downstream dissimilarity does not differ between the year of interest and previous years.

2. BA*Stream: Any change in magnitude of upstream-downstream dissimilarity between the year of interest and previous years is the same for both exposed streams.

If the BA test is significant, then the null hypothesis 1 is rejected. That is, the mean dissimilarity after the event differs (is either smaller or larger) from that before the event. Whether significant or not, results of the BA test must always be interpreted with that of the 'BA*Stream' interaction which considers whether or not the change in magnitude of paired-site dissimilarity within either, or both, exposed streams is consistent within or between the before and after periods. If the test for BA*Stream effect is significant, then the null hypothesis 2 is rejected. In this modified analysis, the BA*Stream interaction is likely to be the most important because it will indicate if, within one or both of the two exposed streams, one of a stream's upstream-downstream dissimilarity is responding differently to that of the other stream (as discussed in section 2.2.1.1 for example, exposed creek Magela has changed over time compared with corresponding exposed creek Gulungul by virtue of the higher concentrations of contaminants it receives from the minesite). In this situation, further investigation of the aberrant stream upstream-downstream dissimilarity would be required.

Concomitant examination of the results of testing of both the primary null hypothesis (BA) and the secondary hypothesis (BA*Stream interaction) is important in the event of one of three important outcomes and interpretations:

- 1. The Primary hypothesis is accepted but the secondary hypothesis is rejected: indicates one of the stream's upstream-downstream dissimilarity is responding differently to that of the other stream.
- 2. The Primary hypothesis is rejected and the secondary hypothesis is rejected: same interpretation as for 1. but the intensity of the significant response is greater than in 1.
- 3. The Primary hypothesis is rejected and the secondary hypothesis is accepted: indicates that both of the stream's upstream-downstream dissimilarity are responding similarly and have changed.

The Year(BA)*Stream interaction is also tested for in this modified ANOVA model. As discussed in section 2.2.1.1 above, however, interpretation of this interaction is not relevant to impact detection.

Possible causes of any observed changes or trend are assessed and/or investigated further to ensure accurate inference about mining impact. See also section 6.2.3 and the Annex of the Appendix for further discussion of the analysis of this modified (mid-year) dataset.

2.2.2 MBACI design and PERMANOVA model

PERMANOVA (PERmutational Multivariate ANalysis Of Variance) (Anderson 2001, McArdle & Anderson 2001, Anderson et al 2008), an add-on function of PRIMER multivariate software (Clarke & Gorley 2006), is an analytical method that has become

available since the inception of the original BACIP design. PERMANOVA can use any distance measure appropriate to the data (including Bray-Curtis), and uses permutations to perform hypothesis tests which are largely, but not entirely, free of distribution type. As such, and by adopting an approach to partitioning of variation like that employed in ANOVA, it can perform analyses of multivariate (or univariate) data in the same manner as the more complex experimental designs and models associated with BACIP and ANOVA that are only applicable to univariate data and that are used in the current protocol.

 Table 3
 Interpretation of each factor and interaction for the BACIP ANOVA analyses on macroinvertebrate community structure data in seasonal streams. Important factors/interactions for interpreting impact detection are identified.

Factors	Interpretation
ВА	A significant result Indicates change from the before to after periods in the magnitude of upstream-downstream dissimilarity across both streams and exposures. While this factor is not directly interpreted for impact detection, it can be interpreted for non significant BA*Exposure and BA*stream(Exposure) interactions to indicate if overall change from before to after has occurred
Exposure	A significant result indicates differences in the magnitude of upstream-downstream dissimilarity between the two exposure types (across both streams within each exposure type). May indicate mining impact, though care in the interpretation is needed because any differences could be due to natural variation between catchments. Consideration of the stream(exposure) factor is needed to determine if natural variation between streams occurs (indicated if this factor is significant). Exploration of correlates of macroinvertebrate community structure using BIOENV analysis may be required to aid interpretation (ie natural habitat variation versus mine-related water quality)
Year(BA)	A significant result indicates the magnitude in dissimilarity across all streams differs amongst years within the before or after period. This factor is not important for impact detection.
Stream(Exposure)	A significant result Indicates streams differ in the magnitude of upstream-downstream dissimilarity within either or both exposure type. Pairwise comparisons are required to explore a significant result to test whether differences occur amongst streams within control or exposed designations, or both. Differences amongst streams within both exposure types most likely indicate natural differences amongst streams.
BA*Exposure	Key interaction for impact detection. A significant result indicates that across streams within each exposure type the change in the magnitude of upstream-downstream dissimilarity from before to after periods is not consistent between the control and exposed streams (suggesting minesite influence). A non-significant result for this factor needs to be interpreted carefully because a significant BA*stream(Exposure) interaction may indicate streams within either, or both, exposure type are not consistent within or between the before and after periods. A significant BA*stream(Exposure) interaction requires further investigation to understand the nature of the differences.
BA*Stream(Exposure)	Key interaction for impact detection. A significant result indicates the change in magnitude of upstream-downstream dissimilarity between streams within either, or both, exposures are not consistent within or between the before and after periods. When significant, this interaction is the most important for impact detection because, for example, it might reflect that one exposed stream (closest to the minesite) is responding differently to all other streams. A significant result for this interaction requires further investigation to determine the nature of the inconsistencies and whether or not they indicate mine impact. Pairwise comparisons can be used to explore a significant result.
Exposure*Year(BA)	A significant result indicates that the change in magnitude of upstream-downstream dissimilarity across streams within each exposure type are not consistent amongst Years(BA) between the two exposures. This interaction can be useful for impact detection if, for example, variation in the upstream-downstream dissimilarity over time was only detected across exposed streams. Further consideration of the Year(BA)*Stream(Exposure) interaction is required to determine if variation overtime within streams for either, or both, exposures is consistent. Variation overtime within all streams, regardless of exposure would suggest temporal variation is due to natural causes.
Year(BA)*Stream(Exposure)	A significant result indicates that the variation in magnitude of upstream-downstream dissimilarity amongst Years(BA) is not consistent amongst streams within either, or both, exposure type. This interaction can be useful for impact detection if, for example, variation in the upstream-downstream dissimilarity over time was only detected amongst exposed streams. Pairwise comparisons can be used to explore a significant result.

Unlike the (M)BACIP approach where sitepair dissimilarities are employed to meet data assumptions of independence of temporal replicates, PERMANOVA is not so constrained. It uses data from the individual sites and as such offers increased partitioning of data variation (and hence increased factors). Its ability to use the complete multivariate dissimilarity matrix enables PERMANOVA to better detect changes in *direction* in multivariate space that might otherwise be missed when using the simple sitepair dissimilarity metric data. (See Appendix for further information.)

Using the current design configuration, PERMANOVA is employed: 1) to better interpret the Multi-Dimensional Scaling (MDS) ordination graphic (section 6.2.3.2) as a comparative analysis to the ANOVA design (see Appendix).

2.2.3 Multivariate analyses to assist in impact assessment

Inferences about mining impact are not solely reliant upon formal statistical testing techniques that underpin the approaches described above. Macroinvertebrate community data are also displayed according to the underlying Bray-Curtis dissimilarity matrix by Multi-Dimensional Scaling (MDS) ordination (Clarke & Gorley 2006). The resulting ordination graphic depicts the patterns in macroinvertebrate community structure for each stream site over time and the relationship of these site communities to each other. Points on the MDS with greatest separation represent sites with greatest differences in macroinvertebrate communities shown in the MDS may be formally tested (using PERMANOVA, see section 2.2.2 above), the graphical representation of the sample points alone may assist in determining the significance of year to year changes in the structure of macroinvertebrate communities from the control and exposed streams.

The multivariate SIMPER and BIOENV routines from PRIMER multivariate software (Clarke & Gorley 2006) can also assist in identifying possible causal mechanisms of changes in assemblage structure by providing a greater understanding of the macroinvertebrate taxa and environmental (including habitat) factors contributing to the separation of stream sites, respectively.

The suite of multivariate analyses described in this section is not simply an add-on to the formal (PERM)ANOVA statistical tests for impact detection described above (sections 2.2.1 and 2.2.2). Rather, the two approaches are run side-by-side and each informs the other to increase the power of the approach for not only detecting impact, but also inferring a cause.

3 Field procedures

3.1 Occupational health and safety

eriss has established project and field safety approvals processes, guidelines and procedures that must be followed prior to and during all field work. These include completion of a risk-based field safety analysis (FSA) of the required work program. All participants are made aware of potential dangers and the procedures implemented to minimise risks by communicating and understanding the FSA before field work commences. Special arrangements are in place to facilitate communication of this information to indigenous personnel engaged as day labour.

The main health and safety risks in this project are transport to and from, and around, sampling sites, heat stress, dehydration, manual handling strains, mosquito bites and working around water, including crocodile attack. Due to shallow water depth and high water clarity at the time of sampling, potential for crocodile attack is invariably rated as 'extremely low'. A reconnaissance for signs of crocodiles in the area is made prior to entering the site and dedicated 'crocodile spotters' are used at all times a staff member is in, or close to, the water.

Extra precautions are taken if: water level at a particular site is above 70 cm; the site is adjacent to a large waterbody; and/or crocodiles are known to frequent the location (eg Nourlangie Creek upstream site (NOUC1) for the latter two criteria). In these cases, crocodile exclusion nets may be deployed for added assurance and safe operation within sample areas. These heavy gauge nets (12 mm, cargo netting, mesh 200 mm) are either set upstream and downstream of sampling area across the stream (where possible), or in a half circle from shoreline upstream and downstream of the sampling area. For correct procedures for deploying exclusion nets see SSD Shallow Billabong Fish Operational Manual. A decision on whether safety nets are deployed at sites is made at the time of sampling.

Details on mitigation measures for the identified health and safety risks for this work are discussed in the Operational manual for this protocol.

3.2 Consultations required for site access

Sampling is conducted within Kakadu National Park (KNP) and the Ranger uranium mine lease. The Aboriginal people within KNP maintain strong cultural ties with their lands and take responsibility in its management. The following stakeholders need to be consulted, via SSD's community liaison officer, to ensure necessary protocols are followed, well in advance of sampling:

- Local traditional owners should be consulted directly through their community organisations for approval to access their land and to conduct the sampling. The Northern Land Council should be advised of intended activities. This should be carried out well in advance of sampling.
- For access to Magela Creek and Gulungul Creek sites, approval from Energy Resources of Australia Pty Ltd, environmental management branch, should be obtained prior to accessing the Ranger mineral lease. ERA security should be notified prior to start of work at both upstream Magela Creek sites.
- For all sites, Parks Operation and Tourism Branch project officer and local district offices should be notified to ensure that park management activities do not conflict with sampling dates (eg feral animal control, fire management).

Formal approval to use non-recreational fishing methods, including collection of macroinvertebrates, is required from the Northern Territory Government, Department of Primary Industry, Fisheries and Mines – Fisheries Licencing. A special permit to collect fish and aquatic life should be held by the project leader and other *eriss* staff members regularly involved in this work. The Director of Fisheries issues permits for a one-year term.

3.3 Timing of sampling

For macroinvertebrate community-based monitoring in the Alligator Rivers Region, sampling is conducted annually, during post wet season recessional flows which usually coincide with the April-May period.

Tripodi (1996) examined temporal changes in macroinvertebrate diversity (taxa number and abundances) at sites in the permanent and seasonally-flowing sections of Magela Creek over the 1995–96 wet season. In the seasonally-flowing section of the creek near Ranger, diversity was highest in macrophyte edge habitat at the end of the wet season (see also supporting results in Garcia et al 2011). Further, at this time, monitoring is most likely to integrate the full effects of wet season inputs of mine runoff waters (Humphrey et al 1990). Accessibility to the sites, coupled with Tripodi's findings, reinforced the decision to sample at the end of the wet season.

Similarity in macroinvertebrate community structure amongst wet-dry tropical stream sites, in the absence of human disturbance, is highly dependent upon stream discharge (Humphrey et al 2009) and to this end, it is very important that sampling from year to year is standardised to the same water levels in the creeks to avoid temporal confounding. Hence, the commencement of the stream macroinvertebrate sampling program is triggered by the water level of Magela Creek, as recorded from the Georgetown gauge board (near the upstream water quality monitoring pontoon). The average gauge height recorded at the time of sampling from previous years is 11.8 m (1995–2011 field data). This level indicates conditions suitable for sampling and also conditions that are dry enough to enable overland access to the sampling sites. Sampling of all sites is accomplished over a fortnight period, and the order of sampling, Magela, Gulungul, Burdulba then Nourlangie, is the same every year.

An alternative trigger level is a stage height reading of less than or equal to 1.7 m from gauging station G8210009 (near the Magela downstream monitoring point). G8210009 stage height data are accessed online.

3.4 Sampling sites

Four streams are sampled in the current monitoring design (Table 4; Figure 1). On each stream there are two sampling sites, except in the case of Magela Creek where there are three. The addition of the third (upstream, control) site on Magela provides additional assurance that the upstream reference condition for this stream has been adequately characterised in the event that impact is inferred at the downstream exposed site. For example, a significant shift in the paired-site dissimilarity measure depicting similarity of community structure between upstream and downstream sites in Magela (section 2.2.1) could be a consequence of an unusual change at the *upstream* site. Additional data from another upstream site can assist in elucidating the nature of the change.

Site	Site Code	Longitude	Latitude	UTM (Easting, Northing)	Creek type	Up- /D'stream
Magela Creek Site 1	MCS1	-12.679103	132.937145	275973, 8597465	Exposed	Up
Magela Creek Site 2	MCS2	-12.674222	132.930722	275271, 8598000	Exposed	Up
Magela Creek Site 3	MCS3	-12.638557	132.899337	271830, 8601919	Exposed	Down
Gulungul Creek Site 1	GULC1	-12.700922	132.887252	270572, 8595008	Exposed	Up
Gulungul Creek Site 2	GULC2	-12.666725	132.884572	270250, 8598789	Exposed	Down
Nourlangie Creek Site 1	NOUC1	-12.889562	132.780528	259157, 8574036	Control	Up
Nourlangie Creek Site 2	NOUC2	-12.856113	132.777339	258779, 8577735	Control	Down
Burdulba Creek Site 1	BURC1	-12.776317	132.788654	259932, 8586575	Control	Up
Burdulba Creek Site 2 A	BURC2	-12.779996	132.777217	258693, 8586158	Control	N/A
Burdulba Creek Site 3	BURC3	-12.778156	132.756794	256473, 8586342	Control	Down

Table 4 Sampling sites in Ranger macroinvertebrate monitoring program (datum WGS84).

A: Sampling at this site was discontinued in 2005. The data from the site do not contribute to formal change detection analysis.

3.4.1 Exposed sites

The Magela and Gulungul creek sites are potentially directly exposed to contaminants from Ranger Mine at their downstream sampling locations. Even though the upstream sites are not exposed to minewaters, in the event of impact, the paired (upstream-downstream) dissimilarity response variable would potentially change (section 2.2.1) and so the site-pairs within Magela and Gulungul Creeks are regarded as 'exposed'.

3.4.2 Control sites

Burdulba and Nourlangie Creeks are located in the Nourlangie Creek catchment of the South Alligator River system. There is no mining activity or urban development in this catchment, and so Burdulba and Nourlangie Creeks are regarded as control streams.

3.5 Number of replicate samples and habitat sampled

3.5.1 Habitat sampled

A comparison of the macroinvertebrate diversity in Magela Creek during the 1995–96 wet season, from sandy substrates, backwater leaf packs and macrophyte habitats, indicated that the highest taxa richness and abundances by the late wet season and recessional flow period occurred in macrophyte habitat (Tripodi 1996). Diversity was hypothesised to be higher in the macrophyte habitats due to the greater stability of the macrophyte beds, compared to the sandy substrates and leaf packs that are easily disturbed during spates. From 1997, macrophyte habitat was therefore reinstated as the preferred habitat to be sampled in this protocol.

All sites are located in shallow macrophyte 'riffle' zones, in water depths mostly less than 50 cm. The macrophyte areas are dominated by submerged plants such as *Eriocaulon* sp., *Blyxa* sp. and *Xyris* sp. which are exposed to moderate to high stream currents during recessional flows (O'Connor et al 1997).

3.5.2 Within-site replication

At each site and on each sampling occasion, replicate samples are collected, each replicate sample comprising five 0.0625 m² Surber samples. Taxon accretion curves for macroinvertebrates collected from Magela Creek sites in 1997 and identified to the lowest possible taxonomic resolution (usually species-level), indicated that greater than 80% of taxa are represented in five replicate samples (*eriss*, unpublished data). As a consequence, for the

current protocol, five replicate samples are collected from each site and are processed separately.

The replicate samples are collected from suitable areas of homogeneous macrophyte habitat of similar flow conditions, both within and amongst sites, to minimise confounding in the detection of impact due to differing habitat conditions. Percentage cover of macrophytes at the sampling site should be greater than 50% when assessed visually. Flow velocities vary considerably among sites and the most suitable guidance is that samples must be collected from visually-flowing waters (ie not from backwaters) that are not excessively fast-flowing as to result in the collection of high abundances of flow-dependant taxa such as simuliids and hydropsychids. When selecting the sampling locations, the five sub-replicates from each replicate should be collected within 2 m of each other.

Replicate samples are generally collected from the same locations from within each site each year. However, macrophyte beds may shift from year to year, or their percentage cover may decrease, and so it may be necessary to move the sampling location according to habitat availability and thereby meet the plant cover and flow velocity criteria from above. It should be noted that Magela Creek Site 3 (MCS3) must always have three replicates located on the western bank, and two replicates on the eastern bank, of the western-most channel.

3.6 Measurement of environmental variables

Change in macroinvertebrate communities may be associated with changes in water quality or habitat structure among sites and years. Changes in environmental variables may arise from either natural environmental perturbations or through human-induced change. Consequently, at each site variables relating to water chemistry, depth, surface flow and structure of the habitat created by the aquatic macrophytes are measured. These may serve as potential environmental correlates of the macroinvertebrate community structure data for that year.

In situ water chemistry parameters are measured and water samples collected before the site is disturbed by sampling. A water sample is collected for later laboratory determination of ions and trace metals. Measured variables are described in Table 5.

Habitat variables (Table 5) are recorded for each site replicate at the same time as macroinvertebrate sampling is undertaken. Variables such as depth and surface flow are generally measured once the macroinvertebrate sampling has been completed.

Category	Feature/analyte	Units
In situ water	Electrical Conductivity	µS/cm
physico-	Turbidity	NTU
chemistry	рН	Units
	Temperature	°C
	Dissolved Oxygen	mg/L & % saturation
Laboratory	Sulfate (SO ₄ ²⁻), Calcium (Ca ²⁺), Magnesium (Mg ²⁺)	mg/L (filtered)
measured water physico- chemistry ^B	Aluminium (Al), Copper (Cu), Iron (Fe), Manganese (Mn), Lead (Pb), Uranium (U) and Zinc (Zn)	µg/L (filtered)
Habitat Structure ^c	Macrophyte genera/species – Identified using macrophyte taxonomic keys and eriss common plant photograph guide. A visual estimate of the percentage cover of the total macrophyte cover for each taxon is determined for each replicate.	% contribution

 Table 5
 Site physico-chemistry and habitat description measured for stream macroinvertebrate monitoring.

Macrophyte cover – A visual e	stimate of the percentage cover of the total	% total covor
macrophytes occupying the area	a of each replicate	
Substratum characteristics – clay (0 – 0.063 mm), Fine sand 2 mm) for each replicate.	A percentage estimate of the proportion of (0.64 - 0.5 mm) and coarse sand (0.51 –	% contribution
Detritus cover – A visual estim replicate	ate of the percentage cover for each	% contribution
Root masses – Terrestrial origi estimate of the percentage cover	n, primarily fine <i>Melaleuca</i> roots. A visual r for each replicate	0 = absent 1 = present
Surface flow – Measured by tin along the area that the macroiny completed three times per replic recorded.	ning a float travelling over a 2 m distance vertebrate sample was taken. This is ate and an average for each replicate is	m/s
Average water depth – Measu replicate sample was taken. Dep replicates are then undertaken a samples.	red over the 2 m stretch from where the oth measurements from three of the sub- and then averaged for each of the replicate	m

Superscripts A = water quality variables measured in the field at each sampling site, B = samples collected from each and sent to a NATA accredited laboratory, required filtration is done in the field, C = measurements taken for each replicate within a site on the day of sampling.

3.7 Macroinvertebrate sampling

Aquatic macroinvertebrates are collected with a Surber sampler. It consists of a 0.0625 m² quadrat to which a frame of a collecting net is permanently attached as an open-ended sleeve of netting for sampling.



Figure 3 Surber sampler for collecting macroinvertebrate samples in ARR seasonally-flowing streams (measurements in centimetres).

The Surber sampler is placed over a chosen area of macrophyte bed (selected according to habitat criteria described in section 3.5.1) with the collection net facing downstream and pushed into the substrate to a depth of approximately 7 cm. All macrophyte and substrate material within the Surber quadrant area are removed to a depth of 5 cm. This method of sampling is completed five times for each replicate. When all five sub-replicates have been collected, the net is removed and washed in-stream to remove any fine particulate matter (eg fine sand and clay < 500 μ m) whilst ensuring no sample content can escape from the wide opening of the net. The sample is then placed into a series of buckets containing clean creek water and on the water's edge, where it is washed and elutriated to separate

macroinvertebrates and other fine organic matter from coarse sand and the bulk of the macrophyte material. The resulting replicate sample is preserved in plastic pots containing 100% ethanol ready for later sample processing (sub-sampling, sorting and identification) in the laboratory.

3.8 Field QA/AC

Sampling is conducted or supervised by trained personnel. Personnel with limited experience are accompanied by trained staff until deemed competent. Experienced operators must be present at all times throughout the sampling operation to minimise influence of bias. Further details are provided in the Operational Manual.

3.8.1 Water Chemistry

Water chemistry sampling must be performed by a competent staff member with working knowledge of the SSD water chemistry protocols. Personnel collecting field data and samples should be trained in these procedures or supervised by someone who is trained.

3.8.2 Habitat

Quality control and assurance of recording habitat variables are ensured by the most senior field sampler who maintains an appropriate knowledge of aquatic macrophyte taxonomy. Use of standardised field datasheets helps to ensure standardised habitat assessments. For novice field samplers, training is obtained 'on-the-job' from the senior sampler. Senior and novice samplers should complete the habitat assessments together and results compared until the values from five replicate assessments are within 10% of one another.

4 Laboratory procedures

4.1 Order of processing of samples

The laboratory processing of macroinvertebrate samples is labour intensive and in order to meet reporting needs for sampling completed in the wet-dry transition period (April or May), a priority order of processing is required. In that way, results from key sites at least may be available in time for mid-year annual reporting (section 7.4). Thus from highest to lowest priority, samples are processed in the order:

- 1. Magela Creek site 3 then site 1
- 2. Gulungul Creek site 2 then 1
- 3. Nourlangie Creek sites 1 and 2
- 4. Burdulba Creek sites 1 and 3
- 5. Magela Creek site 2

4.2 Sub-sampling, identification and enumeration

Sub-sampling of replicate samples is necessary as they generally contain large amounts of organic material, and too many macroinvertebrates, to be managed efficiently in the laboratory. Sub-sampling is achieved using a box sub-sampler designed by Marchant (1989).

All macroinvertebrate specimens from the sub-sample are then removed, counted and identified under a stereo microscope until only debris remains in a channel tray. Sub-sampling continues until at least 200 animals have been recorded. Such a fixed-count approach to sample processing has been shown to be more effective at distinguishing macroinvertebrate samples from one another than proportional sample processing approaches, (Barbour & Gerritsen 1996, Walsh 1997). So that inordinate amounts of time are not spent on sorting, a maximum period of 4 hours (sorting only) is applied. This means that the sorting process is carried out until at least 200 organisms are collected, or for a period of 4 hours, whichever is arrived at first. The four hour limit does not apply to the 1-2 hours still required for the washing of the samples, and associated vial preparation and labelling. The entire sub-sample must be completely sorted, even if the 200 organism count or 4 hour time limit is reached, to enable estimates of total abundance.

Macroinvertebrate taxa are identified, using currently accepted keys (see Operational Manual), to family-level resolution where possible. However, worms (eg Nematoda, Oligochaeta) are only identified to class and mites are identified to the suborder Oribatida or sub-cohort Hydracarina.

One final step in the laboratory sorting process involves completing a coarse scan of the remaining unprocessed portion of the sample. This quick scan, without the aid of stereo microscope, picks up any large organisms that may potentially have been missed during the sub-sampling process.

As discussed in section 1.4, species-level data are sought for upstream and downstream sites in Magela Creek. Taxa collected during the family-level identification process are identified to the lowest possible taxonomic level, with a low power dissection microscope and a compound microscope, using available published keys and the voucher collection held by *eriss* at Jabiru Field Station. External specialist taxonomic experts are contacted to assist with identification of difficult taxa and for QA/QC purposes.

4.3 QA/QC for laboratory procedures

An introductory training schedule is used for staff completing sample processing for the first time. QA/QC criteria and thresholds for acceptance of results, against the independent (cross-checked) results of the senior macroinvertebrate specialist and trainer, include:

- 1. At least 95% of taxa are removed from the processed residue
- 2. At least 95% of the taxonomic identifications are correct
- 3. The estimate of total abundance of the sub-sample is within 5% of the value derived by the trainer.

QA/QC must be completed for the first five samples of any novice sorter and the standard must be met in all five samples. If the standard is not met, training is provided to overcome the deficiencies and the staff member stays in the training schedule until three consecutive samples meet the standard required. Once the staff member has passed schedule 1) it is no longer necessary for the staff to retain any processed residues.

The senior taxonomist must check a random 5% of all samples processed each annual sampling period. A pass level of 95% on the taxonomy and enumeration of a sample is required. Failure to meet the 95% pass level reinstates the taxonomist in the introductory training schedule until a minimum of three consecutive samples meet the standard.

5 Data entry, storage and associated QA/QC

5.1 Data Storage

All original datasheets are archived in SEWPAC/SSD registry files that are currently kept at the Jabiru Field Station.

Macroinvertebrate counts, water chemistry and habitat data are stored electronically in Microsoft Excel spreadsheets located in folders for each sampling season within the *eriss* computer network on SSIMS SharePoint.

e.g. \\Environmental Impact of Mining - Monitoring and Assessment > Macroinvertebrates > Ranger streams > 2011-12

All relevant site photographs are stored electronically on the SSD PhotoD database and copies are generally made to a CD and stored in the relevant registry file for the sampling period.

Backup, CD/external hard drive, copies are made annually and stored at the Jabiru Field Station.

5.2 Data entry and QA/QC

Field environmental data and macroinvertebrate count data are entered into Excel spreadsheets and verified by an independent person. Both staff members entering and validating the data must sign and date the datasheets when they have completed their tasks.

A commercial NATA accredited laboratory conducts water chemistry analysis. The results of these analyses are scrutinised before being entered into the Excel database. Water chemistry QAQC involves the following:

- Check field blank samples for contamination.
- Compare replicate sample results looking for discrepancy that would indicate contamination (< 20%).
- Check results for unusually high measurements.

Further information is available in the SSD's Surface water interpretation and reporting Operational manual.

No data should be used until the validation procedure is completed.

6 Data analysis

For greater detail of the statistical methods for data analysis, including that of the worked example, refer to the corresponding Operational manual.

6.1 Data preparation

6.1.1 Rejection of data

Data are only rejected if they are not representative of the experimental design and other procedural requirements described in this protocol. Data not included in the MBACIP statistical analysis (ANOVA) are macroinvertebrate data gathered prior to 1998 which were gathered using different sampling and sample processing methods (section 2.1). Magela Creek data from 1997, however, were gathered using the current protocol and may be combined with Magela data from 1998 to the present if specific Magela-only analyses were to be considered in future. Note as well, that samples and data gathered from all streams prior to 1997 (for Magela dating back to 1988) may be used for other impact assessment purposes, including an assessment of whether or not there have been compositional changes in macroinvertebrate communities over time (eg at species or genus level).

Apart from pre-1998 data and those not collected according to the current protocol, consideration may also be required about the validity of (and hence possible need to omit) other types of data and unusual circumstances, including, for example:

- Annual data that do not constitute the full complement of five replicates per site (eg samples may be lost or missing) where there are indications that the remaining replicates may be unrepresentative.
- Unusual habitat or macroinvertebrate compositional changes to a stream site relative to the corresponding paired site in the same stream. For example, severe flooding may alter or remove sampling habitat at one site but not the other site, or may significantly delay sampling of one site relative to the other site which was sampled prior to flooding or stream rise.

Professional judgement is required about which data may need to be omitted from analyses and there are no a priori rules for this.

6.1.2 Standardising macroinvertebrate data

For final analysis of macroinvertebrate data, taxa abundance data for each replicate sample are standardised (scaled-up) to 1 m² area of the stream.

As the macroinvertebrate data recorded on the data sheets are only a sub-sample of the sample collected in the field, it is necessary to standardise the sub-sample abundances to reflect expected densities for 100% of the sample prior to data analysis. The coarse scan data are also added to this figure.

To achieve standardisation of the data the following equation is applied in Excel:

(Subsample data * percentage subsampled + coarse scan data) x 3.2

The number '3.2' is the multiplier used to standardise the sample to 1 m^2 area. This was arrived at using the following steps:

1 x Surber Sample quadrat = 0.25 m x 0.2 m = 0.0625 m^2

5 x Surber Sample sub-replicates = 5 x 0.0625 = 0.3125 m²

Sample area = 1/0.3125 = 3.2

6.1.3 Random pairing of replicates

For each year of sampling, Bray-Curtis dissimilarity values are calculated for the five possible independent pairs of replicate data between upstream and downstream sites for each stream. The replicate 'pairs' are selected at random (using a random-without-replacement approach) for each year from the 25 possible pairwise comparisons. The randomly-selected replicate pairs for each year remain the same in all ensuing analyses relevant to that year of study.

6.2 Impact detection and assessment

The important change detection hypotheses evaluated in this protocol using the MBACIP ANOVA model were listed in section 2.2.1.2. The primary null hypothesis states that there has been no change in macroinvertebrate community structures in the exposed streams, relative to the control streams:

- 1. between two time periods of interest, eg before and after a possible impact event, or
- 2. between the current wet season and all previous wet seasons, or
- 3. before and after mine rehabilitation etc.

Hypothesis scenario b is tested on an annual basis and to date, has been the only hypothesis tested using this protocol. The scenario represents an unbalanced design in the 'Year' (time) factor. As described in section 2.2.1, impact detection and assessment for any of the scenarios are conducted on the multivariate sitepair or broader dissimilarity matrix using ANOVA. Supportive evidence is provided from PERMANOVA (section 2.2.2). Trends or unusual observations in the time series of dissimilarities are investigated using other multivariate approaches, including ordination, SIMPER and/or BIOENV from the Primer software package (Clarke & Gorley 2006, section 2.2.3).

6.2.1 Caveat

Section 2.2 draws attention to the fact that inferences that may be drawn from the time series of macroinvertebrate data – dating back to 1988 for Magela and 1994 for others streams – are, *a priori*, weakened because there are no similar pre-mining baseline (pre-1980) data upon which to compare. The formal MBACIP ANOVA itself is only applied to data gathered from 1998 to the present. Section 6.1.1 notes, however, that samples and data gathered from all streams prior to 1997 (for Magela dating back to 1988) may be used for other impact assessment purposes, including an assessment of possible shifts in macroinvertebrate community composition over time.

Even using the current and advanced MBACIP design with its balanced multiple exposed and control streams, making a correct inference about mining impacts should always rely upon detailed examination of mine and non-mine-related environmental factors that may explain the observed responses, with reference to other lines of monitoring or experimental evidence (including water chemistry, bioaccumulation, laboratory and field-based ecotoxicology, and community studies of macroinvertebrates and fish in shallow billabongs).

6.2.2 Testing of statistical assumptions and other test criteria

Statistical analyses are conducted on the dissimilarity values arising from the observations from each year, with five values for each year being derived from the randomly-paired

replicates between the upstream and downstream sites for each stream. It is important to check on an annual basis that the full set of dissimilarity values to be analysed conforms to the underlying assumptions required for application of ANOVA, including normality, homogeneity of variance and independence (eg Stewart-Oaten et al 1986, Sokal & Rohlf 1995).

6.2.2.1 Data assumptions of normality and equal variances

Assumptions of **normality** and **equal variances** are checked graphically, as recommended by McGuinness (2002), using plots of the *residuals* or *errors* (ie the difference between a dissimilarity observation and the mean dissimilarity for the group). A worked example of this procedure is provided in the Operational manual using Minitab software. Both assumptions have invariably been met in macroinvertebrate community data obtained since 1998.

6.2.2.2 Data assumption of independence

If the residuals are arranged in time order of data collection, they should succeed each other in a random sequence. If this is observed, then data sets from each year meet the assumption of **independence**. Departure from independence (often termed serial correlation, or autocorrelation) would be indicated in the event of an extended sequence of positive residual values followed by an equally long sequence of negative values, ie *positive autocorrelation*, or regular periodicity of positive and negative values, ie *negative autocorrelation*.

The plot of residuals versus observation order may be used as an initial screening assessment to check for lack of independence, with formal testing conducted using the von Neumann test as detailed in Sokal & Rohlf (1995: pp 394-396). An example of the use of the latter test is demonstrated in the Operational manual. Using replicate dissimilarity data from 1998 to 2011, the following results were observed:

Stream	Significance	Type of autocorrelation
Magela	P < 0.01	Positive
Gulungul	NS	N/A
Nourlangie	P < 0.01	Positive
Burdulba	NS	N/A
Streams combined	P < 0.05	Positive

Plots of dissimilarity over time (five replicate dissimilarity values per stream per sampling occasion, commencing 1998, graphs not provided here) indicated a random pattern of dissimilarity for the smaller Gulungul and Burdulba creeks (ie no autocorrelation), and typically, sequences of higher dissimilarity values followed by sequences of lower dissimilarity values for the larger Magela and Nourlangie creeks, as well as all streams combined (ie significant positive autocorrelation).

Humphrey et al (2012) observed that water quality and biological responses of captive organisms (freshwater snails) held in situ were more variable in Gulungul Creek compared to the same responses measured in Magela Creek. They attributed the greater variability in the smaller (Gulungul) creek to the greater proportional influence of runoff to the stream from catchment sources between the upstream and downstream sites in this relatively small drainage basin. The greater responsiveness of flow in the smaller stream to rainfall events would also result in greater peaking in the hydrograph and hence potential for greater between-site variability in flow. It is not unexpected therefore, that between-site dissimilarity in smaller streams is more strongly influenced by stochastic processes and greater site independence that lead to more random sequences of dissimilarity over time. Conversely,

greater between-site dependence can be expected in water quality and quantity in larger streams and because the wet-dry tropical climate is characterised by sequences of alternating wetter and drier periods, sequences of similar and less similar community structure among sites would be expected. For example, higher flow years may lead to greater homogeneity of habitat and more even dispersion of organisms throughout the stream, and hence lower between-site dissimilarity. Similarly, the reduced flow-related disturbance to streams in lower flow years may also promote greater homogeneity of habitat and hence also lower betweensite dissimilarity.

The influence of climate generally on paired-site dissimilarity is illustrated in Figure 4 where, since 1998, there has been a pattern of greater between-site similarity at low and high rainfall (and hence stream discharge) seasons (as suggested above), and less between-site similarity in seasons of intermediate or closer-to-average wet season rainfall and discharge.



Figure 4 Mean (±SE) of pooled annual paired-site dissimilarity for the four streams sampled in the Ranger macroinvertebrate study in relation to annual rainfall categories (Jabiru airport). Significant two-order polynomial equation describing the fit to mean dissimilarity and rainfall data is provided.

Positive autocorrelation in this multi-factor ANOVA, even though (i) its presence in some streams but not others would dilute the effect, and (ii) it is distributed evenly between exposed and control streams, has the potential to inflate the Type I error rate (K McGuinness, Charles Darwin University, pers comm) – a significant change is detected when in fact there is none ('false positive'). In environmental protection studies, this error is more acceptable than Type II error (an impact passes undetected). Rather than abandoning ANOVA because of this violation in independence assumption, the potential for increased Type I error is acknowledged and the results carefully scrutinised in the event of a significant outcome in annual statistical analysis. An array of accompanying environmental data and complementary multivariate techniques is available to correctly interpret results – see case study in section 6.2.3.

6.2.2.3 Trend analysis

In an extensive investigation detecting long-term change in marine intertidal communities using the BACIP design, Steinbeck et al (2005) removed from ANOVA analysis any data that

had a significant trend (ie with a regression R^2 of > 0.5) in the 'Before' period. This criterion was applied to the current dataset (1998 to 2011) using the two regressions, (i) mean (pooled) dissimilarity for each stream (Y) against year (X), and (ii) each of the 5 replicate dissimilarity values for each stream and from each year (Y) against year (X). Because of significant autocorrelation in a number of the dissimilarity time series (from above), regressions were corrected using the Hildreth-Lu procedure, available as a macro in Minitab software. (The Hildreth-Lu procedure uses nonlinear least squares to jointly estimate the regression parameters with a first-order autoregressive process model.)

Significant regressions were found for cases (i) and (ii) from above (see table below) and these were confined to Nourlangie Creek, with marginal significance for 'replicate' Magela analysis.

Stream	Significance	Negative (–) or positive (+) slope	R ² value
Pooled replicates			
Magela	NS	_	N/A
Gulungul	NS	+	N/A
Nourlangie	P<0.05	_	0.24
Burdulba	NS	+	N/A
Streams combined	NS	_	N/A
Replicates			
Magela	NS (P= 0.053)	_	0.04
Gulungul	NS	+	N/A
Nourlangie	P < 0.05	_	0.06
Burdulba	NS	+	N/A
Streams combined	NS	_	N/A

The two smaller streams, Gulungul and Burdulba, display increasing (but not significant) between-site dissimilarity over time, while the two larger streams, Magela and Nourlangie, display decreasing (only Nourlangie significant) between-site dissimilarity over time. For the significant regressions tabulated above, R^2 values are very low (< 0.25) and hence there is insufficient trend in the time series of sitepair dissimilarity values to discount use of ANOVA.

6.2.3 Impact detection and assessment

For the Supervising Scientist's Annual Report, prepared mid-year, two assessments of potential mining impact are provided: (i) a full assessment using the complete dataset for exposed and control streams, but relevant only to the wet season *prior* to the just-completed wet season; and (ii) an interim assessment for the just-completed wet season using data from just the exposed streams, Magela and Gulungul Creeks. The rationale for using just a subset of data for the interim assessment, together with relevant hypotheses for testing, are provided in section 2.2.1.2 above. In the Annex to the Appendix of this protocol, results of statistical testing and impact assessment for interim assessment scenario (ii) from above, for the 2010–11 wet season, are provided, and the reader is referred to that description as the template for making that assessment. All data preparatory steps for this interim analysis and assessment are provided in the preceding sections.

The descriptions hereafter refer to data analysis and assessment using the complete dataset for exposed and control streams, ie assessment scenario (i) from above.

An impact resulting from a change in composition of creek water (which may result from minesite inputs) may be inferred from a statistically significant change or trend in the upstream-downstream paired-site dissimilarity values for macroinvertebrate community structure on exposed streams between any two time periods. This assessment of possible

mine-related change and its significance is based on comparison of macroinvertebrate communities in exposed streams with those in control streams. Other environmental information, especially water chemistry and hydrology data, is used to interpret a significant result.

For routine monitoring, the most useful statistical tests to apply are the BA*Exposure and BA*Stream(Exposure) factors of the four-factor ANOVA, comparing the current (wet) season's data with data from all previous seasons (section 2.2.1). Assessment of the environmental performance of the Ranger Mine, based on macroinvertebrate community structure monitoring results, is reported each year in the Supervising Scientist's Annual Report.

6.2.3.1 Detecting potential impact using the MBACIP (ANOVA) test

The MBACIP ANOVA analysis has been conducted using data from 1998 to 2011, with 2011 being the year of interest (the 'After' period). Analysis for potential or suspected minesite impact is conducted using the randomly-paired (section 6.1.3) multivariate dissimilarity values, based upon Log [x+1] transformed data (time series of dissimilarity values illustrated in Figure 1). The dissimilarity values for the two time periods of interest are formally analysed using a General Linear Model (GLM) (see section 2.2.1.1 for ANOVA model).



Figure 5 Paired upstream-downstream dissimilarity values (using the Bray-Curtis measure) calculated for community structure of macroinvertebrates in seasonally-flowing streams around Ranger uranium mine. Magela and Gulungul creeks are 'exposed' streams and Nourlangie and Burdulba creeks are 'control' streams. Values are means (± standard error) of the 5 possible (randomly-selected) pairwise comparisons of replicate data between the up and downstream sites.

Analyses of the replicate upstream-downstream difference data for species richness or total abundance can be performed in a similar manner. However, these analyses have not been conducted as part of this protocol.

The MINITAB output for the ANOVA results is shown in Table 6. The details of how to set up the ANOVA model and run the analysis are provided in the Operational manual. The ANOVA results show no significant difference between the control and exposed streams in the before (1998 to 2010) and after (2011) comparison (ie the BA*Exposure interaction is not significant). However, the BA*Stream(Exposure) interaction for the same before-after comparisons is significant (p = 0.01), which indicates the change in magnitude of paired-site dissimilarity in either, or both, exposure types (exposed and control streams) is not consistent within or between the before and after periods. Further investigation using pairwise tests showed that the significant difference in the change in dissimilarity was associated with the Gulungul sites (p=0.002). The result indicates the exposed stream (Gulungul Creek) has responded differently, from the before to after periods, from all other streams. From Figure 5, a sharp rise in dissimilarity for Gulungul Creek can be observed. Closer examination of the data is required to assess whether or not this result may be associated with the Gulungul downstream site, and thereby indicate possible mining impact (see section 6.2.3.2 for further exploration).

Table 6ANOVA results for stream macroinvertebrate community upstream-downstream dissimilarityvalues from two exposed streams and two control streams. Years 'Before' are 1998–2010, year 'After' is2011.

Source	DF	Seq SS	Adj SS	Adj MS	F	Р
BA	1	46.72	46.72	46.72	0.10	0.784*
Exposure	1	28.60	202.61	202.61	0.79	0.457*
Year(BA)	12	1256.83	1256.83	104.74	0.99	0.507
Stream(Exposure)	2	289.16	467.65	233.82	0.50	0.666
BA*Exposure	1	392.84	392.84	392.84	0.80	0.458*
BA*Stream(Exposure)	2	934.30	934.30	467.15	5.66	0.010
Exposure*Year(BA)	12	1269.43	1269.43	105.79	1.28	0.291
Year(BA)*Stream(Exposure)	24	1981.94	1981.94	82.58	1.71	0.024
Error	224	10794.00	10794.00	48.19		
Total	279	16993.83				

* Not an exact F-test

The Year(BA)*Stream(Exposure) interaction was also significant in the same analysis (p = 0.024); this interaction has been shown to be significant in previous years as well and simply indicates that dissimilarity values for the different streams – regardless of their status (Before, After, Control, Impact) – show natural differences through time. This is evident in the upward and downward trends over time (eg a recent downward trend in the Nourlangie paired-site dissimilarity, Figure 5).

6.2.3.2 Impact assessment

At the time of publication of this protocol, and for the first time, a significant difference has been observed between upstream-downstream dissimilarity in Gulungul Creek from a particular wet season (2010–11) and previous wet seasons' results (from section 6.2.3.1 above). This result reflected an increase in magnitude of the paired-site dissimilarity for the year of interest (2010–11) (Figure 5). A limitation of the BACIP analysis is that the influence of individual sites that comprise the paired-site dissimilarity cannot be determined, and hence

it is not possible to determine whether the result is associated with the Gulungul downstream site (downstream of minesite), and thereby indicate possible mining impact. Further analyses to attribute the cause of this change are described in the following sections.

6.2.3.2.1 Multivariate analyses

Analysis of the multivariate data using Multivariate Dimensional Scaling (MDS) ordination and PERmutational Multivariate Analysis Of Variance (PERMANOVA) (Anderson et al 2008) has been conducted to depict the relationship of the community sampled at any one site and sampling occasion with all other possible samples. The ordination can assist in determining whether the upstream and/or downstream Gulungul communities have changed over time or are aberrant compared to the other communities sampled over time.

To support multivariate analysis, abundances of the numerically-dominant taxa were compared between the upstream and downstream sites over time to determine what types of shifts in taxa abundances may have occurred recently (2011 in particular).

Figure 6 depicts the ordination derived using replicate within-site macroinvertebrate data (ie data described in section 0.2, without random pairing of site replicates). Data points are displayed in terms of the sites sampled in Magela and Gulungul creeks downstream of Ranger for each year of study (to 2011), relative to Magela and Gulungul creeks upstream (control) sites for 2011, and all other control sites (Magela and Gulungul upstream sites, all sites in Burdulba and Nourlangie). Samples close to one another in the ordination indicate a similar community structure.



Figure 6 Ordination plot of macroinvertebrate community structure data from sites sampled in several streams in the vicinity of Ranger mine for the period 1988 to 2011. Data from Magela and Gulungul creeks (Exposed streams) for 2011 are indicated by the enlarged symbols.

The ordination depicted in Figure 6 indicates that Gulungul Creek communities collected from the upstream site in 2011 differ from communities from other sites and times (Figure 6). Conversely, data-points associated with the 2011 Gulungul and Magela downstream sites are generally interspersed among the points representing the control sites, indicating that these

'exposed' sites have macroinvertebrate communities that are similar to those occurring at control sites.

The aberrant 2011 Gulungul upstream result was further examined using PERMANOVA (section 2.2.2) to determine if a priori groups, exposure type ('exposed' Magela and Gulungul creeks versus control Burdulba and Nourlangie creeks) and site location (upstream versus downstream), and the interaction between these two factors, show significant differences.

PERMANOVA conducted on all replicate data from all available years and sites showed a significant difference for BA*Upstream/Downstream*Stream (Exposure) (p = 0.022, see Table A3 in the Appendix). This indicates that differences between sites for streams within either, or both, exposure types are not consistent within or between the before and after periods. This may indicate that one exposed downstream site (Magela or Gulungul) is responding differently to all other downstream sites, and thus requires further investigation. A pairwise comparison was undertaken to determine the nature of the significant difference. It indicated that the upstream Gulungul Creek site was significantly different (p = 0.0122) from the before to after period. This result supported the above interpretation of Figure 6.

Interestingly, and not influential over the upstream-downstream dissimilarity, was a significant difference at the upstream Burdulba Creek site (p = 0.0046) (see Appendix for further discussion on the discrepancy between BACIP ANOVA and PERMANOVA results). As neither Gulungul nor Burdulba Creek upstream sites are influenced by minesite activity, the changes at these sites must be associated with natural or non-mine-related conditions. This result was further supported by examination of the taxa abundance information, as discussed below.

6.2.3.2.2 Patterns in dominant macroinvertebrate taxa

Abundances of numerically-dominant taxa were examined between Gulungul and Burdulba upstream and downstream sites over time.

For Gulungul Creek, the analysis found that, historically and typically, there are a greater proportion of taxa at the upstream site with a preference for high velocity waters associated with this location in the creek (ie so-termed 'flow-dependent' taxa). While this remained the pattern in 2011, the abundances of these taxa at the upstream site in 2011 were unusually high compared with values found in previous years and were about three times the abundances observed at the downstream site in 2011 (Figure 7). Given that dissimilarity values are sensitive to taxa abundances, the discrepancy in macroinvertebrate abundances between the Gulungul sites in 2011 can explain the increase in mean dissimilarity observed in the paired-site dissimilarity plot (Figure 5) and the separation of Gulungul upstream sample points observed in the ordination (Figure 6, see also Figure A2a in Appendix).

The habitat and flow conditions prevailing at the upstream Gulungul site in 2011 have yet to be examined closely to better interpret these results. However, given that rainfall for the 2010–11 wet season was the second highest on record (2291 mm, Figure 7), it would appear that the flow characteristics at the upstream Gulungul site in 2011 reflected correspondingly high flows favouring flow-dependent taxa, relative to both the downstream site and previous years.



Figure 7 Total abundance of flow-dependent taxa collected from upstream and downstream Gulungul Creek sites over time with total annual wet season rainfall for Jabiru Airport (data from Bureau of Meteorology). Flow dependent taxa include Simuliidae, Leptophlebiidae, Crambidae, Hydropsychidae and Philopotamidae.

Burdulba Creek, also had unusually high abundances of taxa with a preference for high velocity waters at the upstream site in 2011 compared to previous years (Figure 8). Similar to Gulungul Creek, it is likely that the increased flow velocities associated with the second highest annual rainfall on record resulted in the increased abundances of flow dependant taxa at this site in 2011. Providing further support to the influence of wet season intensity on flow dependant taxa is the increased abundance of these taxa observed at the Burdulba downstream site in 2007, the antecedent wet season being the largest on record (2540 mm, Figure 8).



Figure 8 Total abundance of flow-dependent taxa collected from upstream and downstream Burdulba Creek sites over time with total annual wet season rainfall for Jabiru Airport (data from Bureau of Meteorology). Flow dependent taxa include Simuliidae, Leptophlebiidae, Crambidae, Hydropsychidae, Philopotamidae and Polycentropodidae.

The influence of large wet seasons on flow-dependant taxa is not consistent between creek sites and years (Figures 7 and 8). This is because high wet season stream flows will affect stream geomorphological conditions, including bed load movement and associated shifts in flow velocities, at local scales that cannot be predicted. Macroinvertebrate community structure and associated multivariate analyses are influenced by the most abundant taxa. To this end, the shift in the assemblage for the Burdulba Creek upstream site in 2011 (Figure 8, also illustrated in Figure A2b in Appendix) would be explained by the increased abundances of flow-dependant taxa at this site for this year.

6.2.3.2.3 Conclusions

The increase in magnitude of the paired upstream-downstream dissimilarity for the exposed Gulungul Creek in 2011 was not associated with mining impact. Rather, it was due to changes in macroinvertebrate community structure at the upstream site, associated with increased abundances at this site of flow dependant taxa as a result of unusually large wet seasons. Similar changes for 2011 were observed at the upstream site of the control stream Burdulba Creek.

Macroinvertebrate community structure at downstream sites in the exposed streams (Magela and Gulungul) showed no change in 2011 when compared to previous years. This provides good evidence that changes to water quality downstream of Ranger as a consequence of mining activities between 1994 and 2011 have not adversely affected macroinvertebrate communities.

7 Reporting

7.1 Overview

Different reporting mechanisms are required for different forums and stakeholder groups. Summarised below is a (more or less) chronological sequence of corporate and other reporting during the calendar year:

- Reporting to Traditional Owners
- Supervising Scientist Annual Report (statutory requirement)
- Updating of the Internet monitoring pages following analysis of the collected data
- Review of ERISS science program outputs by the Alligator Rivers Region Technical Committee (ARRTC)
- Report of SSD wet season monitoring program results to the Alligator Rivers Region Advisory Committee (ARRAC)
- Annual Research Summary (Supervising Scientist Report)
- Additional summary reports for stakeholders as required

7.2 Timing

While sampling of macroinvertebrate communities is typically completed by early May each year, laboratory processing of samples may ensue for two to three months, depending upon technical staff availability. This will mean that not all samples will have been processed and data analysed in time for mid-year reporting (ie Supervising Scientist Annual Report). However, the order of priority for sample processing (from section 4.1), should ensure results for paired sites in 'exposed' streams at least (Magela and Gulungul), are available in time for this reporting. Results for the complete dataset may be reported later in the calendar year and/or be provided for annual reports submitted in the following year.

7.3 Reporting results to Traditional Owners and Aboriginal residents

There are two components involved in communicating the work and outcomes of the monitoring program (including the macroinvertebrate community assessment) to Aboriginal people:

Informing people of what tasks are to be undertaken, when, by whom and why.

Providing feedback to people on the results of the work and providing assurance that the environment and their lifestyle has been protected.

Communication occurs through a variety of mechanisms including:

- Involvement of Aboriginal people in the actual monitoring program, especially through employment.
- Regular updates and reports of monitoring results presented by the Community Liaison Officer at meetings and associations. Larger meetings or Open days may also be planned for this purpose. Monitoring staff (and more senior Darwin based staff) are available to people (particularly Traditional Owners and Aboriginal residents) to

answer questions or provide additional information as requested. Information is provided on what programs are to be undertaken and their timetable. Feedback is also sought on any key questions and needs.

• Illustrated report of monitoring results for Traditional Owners and Aboriginal residents.

7.4 Supervising Scientist Annual Report

This statutory report is tabled in Parliament in the latter part of each year. A summary of the macroinvertebrate community monitoring results (which may be an abbreviated version of the summary reports described in section 7.6 below), is included in the Report. For this report, two assessments of potential mining impact are provided: (i) a full assessment using the complete dataset for exposed and control streams, but relevant only to the wet season *prior* to the just-completed wet season (see section 6.2.3); and (ii) an interim assessment for the just-completed wet season using data from just the exposed streams, Magela and Gulungul Creeks (see the Annex to the Appendix of this protocol).

7.5 Internet

Once the macroinvertebrate community monitoring data have been analysed and reported in the Supervising Scientist Annual Report the text and figures are adjusted appropriately for presentation on the SSD website:

http://www.environment.gov.au/ssd/monitoring/magela-bio.html

Papers and reports produced to address the other communication requirements for the monitoring and science programs of SSD (listed in Section 7.1 above) are also posted to the SSD Website as these become available.

7.6 Alligator Rivers Region Technical Committee and Annual Research Summary (Supervising Scientist Report)

A verbal summary of results-to-date is reported to the first meeting of the Alligator Rivers Region Technical Committee (ARRTC) that occurs in the mid–late wet season (typically February-April period). A full summary report of work conducted in the wet season prior is provided to ARRTC for their late dry season review meeting. While the Supervising Scientist Annual Report drafted mid-year reports on only a subset of the macroinvertebrate data for the recent wet season of interest (from exposed streams, see the Annex of the Appendix), this late dry season ARRTC report includes a full impact assessment, using data from both exposed and control streams (see section 6.2.3). The ARRTC summary provides the basis for reporting in the *eriss* Annual Research Summary (a Supervising Scientist Report), compiled late in the calendar year, together with results from other stream monitoring programs.

The Annual Research Summary is circulated to a wide audience, including the key stakeholders, Energy Resources Australia, the Northern Land Council and the NT Department of Resources. A full list of recipients is available from the SSD Publications Officer.

The technical reports should contain the following information, and adhere to the required layout proforma:

1. Brief description and background of the monitoring program.

- 2. Details of the just-completed wet season, noting any specific or unusual issues of relevance. This includes water flow timing and period of controlled or accidental discharge events, and may include unusual weather or hydrological events, etc.
- 3. Brief description of methods with reference to the protocols. Any variations from the accepted Operational protocols and reasons for the variations should be reported.
- 4. Current wet season's results and comparisons to past wet seasons' trends and findings. This would include summary statistics for the data collected in the current season, BACIP (ANOVA) analysis of macroinvertebrate community dissimilarity values, and the relationship, if any, of these biological data to environmental conditions and variables.
- 5. Evaluation of results in the context of any impact being detected.
- 6. Recommendations based on conclusions drawn from the evaluation.

7.7 Summary report for stakeholders

Consistent with the reporting to ARRTC and with similar timing, two reports and presentations are provided each calendar year to the Alligator Rivers Region Advisory Committee ARRAC, representing a wide range of stakeholders for the ARR (not necessarily with technical backgrounds). The reports contain a summary of major results and conclusions, and should be in a more plain-English form to those reports described in section 7.6 above, given the broader range of stakeholder participation in ARRAC.

8 References and additional reading

- Anderson MJ 2001. A new method for non-parametric multivariate analysis of variance. *Austral Ecology* 26, 32-46.
- Anderson MJ, Gorley RN & Clarke KR 2008. PERMANOVA+ for PRIMER: Guide to software and statistical methods, Primer-E:, Plymouth, UK.
- ANZECC/ARMCANZ 2000. Australian and New Zealand guidelines for fresh and marine water quality, <u>http://www.deh.gov.au/water/quality/nwqms/index.html#quality</u>. Last accessed: July 2013.
- Barbour MT & Gerritsen J 1996. Subsampling of benthic samples: a defence of the fixedcount method. *Journal of the North American Benthological Society* 15 (3), 386-391.
- Clarke KR & Gorley RN 2006. *PRIMER v6: User Manual/Tutorial.* 6 edn, Primer-E Ltd, Plymouth.
- Cranston PS 1990. Biomonitoring and invertebrate taxonomy. *Environmental Monitoring and Assessment* 14 (2/3), 265-273.
- Faith DP & Cranston PS 1991. Could a cladogram this short have arisen by chance alone? On permutation tests for cladistic structure. *Cladistics* 7 (1), 1-28.
- 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 programme at Coronation Hill, Northern Territory. *Australian Journal of Ecology* 20 (1), 167-180.
- Garcia E, Dostine P, Humphrey C, Douglas M, Pusey B & Cook B 2011. Aquatic invertebrates. In *Aquatic biodiversity in northern Australia: patterns, threats and future,* ed BJ Pusey, Charles Darwin University Press, Darwin NT, 51-70.
- Hellawell JM 1986. Biological Indicators of Freshwater Pollution. Elsevier, London.
- Humphrey C & Chandler L 2009. Taxonomic resolution for biological assessment of mining impacts in tropical streams of the Northern Territory. Paper presented at *Combined Australian Entomological Society's 40th AGM & Scientific Conference and Society of Australian Systematic Biologists and 9th Invertebrate Biodiversity & Conservation Conference*. Darwin 25–28 September 2009.
- Humphrey C, McGuinness K & Douglas M 2009. Experimental design considerations for monitoring of macroinvertebrate communities in tropical seasonally-flowing streams. Paper presented at *Australian Society for Limnology*. Alice Springs Convention Centre 28 September – 2nd October 2009.
- 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 Division, Canberra.
- 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 (2-3), 139-181.

- Humphrey CL, Buckle D & Davies C 2012. Ranger stream monitoring research: Further analysis of toxicity monitoring data for Magela and Gulungul creeks. *eriss* research summary 2010-2011. Supervising Scientist Report 203, Supervising Scientist, Darwin, NT.
- Keough MJ & Mapstone BD 1995. Protocols for designing marine ecological monitoring programs associated with BEK mills. Technical Report 11, National Pulp Mills Research Program, CSIRO, Canberra.
- Marchant R 1989. A subsampler for samples of benthic macroinvertebrates. *Australian Society for Limnology Bulletin* 12, 49-52.
- McArdle BH & Anderson MJ 2001. Fitting multivariate models to community data: a comment on distance-based redundancy analysis. *Ecology* 82, 290-297.
- McGuinness K 2003. Estimating the power of the ARR monitoring protocols: The MBACI(P) design. Internal Report 423, Supervising Scientist, Darwin.
- McGuinness KA 2002. Of rowing boats, ocean liners and tests of the ANOVA homogeneity of variance assumption. *Austral Ecology* 27, 681-688.
- O'Connor R, Humphrey CL & Lynch C 1997. Macroinvertebrate community structure in Magela Creek between 1988 and 1996: Preliminary analysis of monitoring data. Internal Report 261, Supervising Scientist, Canberra.
- Quinn GP & Keough MJ 2002. *Experimental design and data analysis for biologists*. Cambridge University Press, Melbourne.
- Sokal RR & Rohlf FJ 1995. Biometry: the principles and practice of statistics in biological research. 3rd edn, WH Freeman and Co, New York.
- Steinbeck J, Schiel D & Foster M 2005. Detecting long-term change in complex communities: a case study from the rocky intertidal zone. *Ecological Applications* 15 (5), 1813-1832.
- 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.
- Tripodi E 1996. Temporal and spatial variations in the macroinvertebrate communities of the seasonally flowing portions of Magela Creek, Northern Territory. Internal Report 240, Environmental Research Institute of the Supervising Scientist, Jabiru, NT.
- Underwood AJ 1993. The mechanics of spatially replicated sampling programmes to detect environmental impacts in a variable world. *Australian Journal of Ecology* 18, 99-116.
- Walsh CJ 1997. A multivariate method for determining optimal subsample size in the analysis of macroinvertebrate samples. *Marine and Freshwater Research* 47 (3), 241-248.

Appendix 1 Comparison of Analysis Of Variance (ANOVA) with PERmutational MANOVA for the Ranger macroinvertebrate community, stream monitoring program

A1 Background

PERMANOVA (PERmutational Multivariate ANalysis Of Variance) (Anderson et al 2008), an add-on function of the PRIMER software (Clarke & Gorley 2006), represents an alternative and potentially superior data analysis method for impact assessment using the current seasonal stream macroinvertebrate community dataset. The method of non-parametric multivariate analysis of variance was introduced by McArdle & Anderson (2001) and its application to community data by Anderson (2001). There are two main technical advantages of the PERMANOVA method that may ultimately prove superior to the ANOVA method that is currently used for impact detection:

- 1. Whereas ANOVA or the multivariate (and computationally difficult) equivalent, MANOVA, assume normal distributions and, implicitly, Euclidean distance (MANOVA), PERMANOVA can use any distance measure appropriate to the data (including Bray-Curtis), and uses permutations to perform hypothesis tests which are largely, but not entirely, free of distribution type. As such, and by adopting an approach to partitioning of variation like that employed in ANOVA, it can perform analyses of multivariate (or univariate) data in the same manner as the more complex experimental designs and models associated with MBACIP and ANOVA that are restricted to univariate data such as used in the current protocol.
- 2. Unlike the MBACIP approach used in the present protocol where sitepair (upstream vs downstream) dissimilarities are employed to meet data assumptions of independence of temporal replicates, PERMANOVA is not so constrained and offers increased partitioning of data variation (and hence increased factors) by way of its ability to use the complete multivariate dissimilarity matrix. Use of the complete data matrix enables PERMANOVA to better detect changes in direction in multivariate space that might otherwise be missed when using the simple sitepair dissimilarity metric data (See Figure A1 for hypothetical illustration of how the sitepair dissimilarity value might be similar for the before and after periods yet mask a real change that occurs in multivariate direction). (Note that this is not an issue when using univariate data as the difference can only be positive or negative.)

This appendix compares the results from the MBACIP ANOVA analysis (section 6.2.3 from the main report) to results from PERMANOVA analysis on the equivalent dataset. This comparison has been carried out for two reasons:

- 1. To provide complementary results to the MBACIP ANOVA conducted in section 6.2.3 (particularly in the context of the greater information available from PERMANOVA, viz advantage item 2 above)
- 2. To compare and assess any differences in results between the two analysis approaches on the same monitoring dataset.

At this stage, the PERMANOVA method is not afforded more prominence in this protocol because it is relatively new (when compared with MBACIP and associated ANOVA methods) and as such has not been assessed with the same statistical scrutiny as the MBACIP ANOVA. Because of this, the univariate ANOVA analysis technique using sitepair dissimilarity remains the formal statistical testing procedure used to analyse the stream macroinvertebrate data.



Figure A1 Hypothetical scenario showing that an analysis using a control-impact sitepair dissimilarity value will not detect all changes that occur in multivariate direction. In this case, while a change has occurred at the exposed site after impact, the sitepair dissimilarity remains similar for the before and after periods.

A2 Comparison of PERMANOVA and MBACIP ANOVA analysis approaches

While PERMANOVA and the MBACIP ANOVA use the same original dataset, the actual analysis is undertaken on different manipulations of the same dataset, thus:

- PERMANOVA uses the complete dissimilarity matrix which enables the analysis to detect not only changes across streams, but also within each up- or downstream site.
- MBACIP ANOVA uses control-impact sitepair dissimilarity values for each year in order to eliminate or reduce spatial and temporal variability. (As such, the approach assumes that, in the absence of human disturbance, natural variation between the upstream-downstream pair will be consistent from year to year.) This approach is an established design for impact detection, particularly with paired upstream and downstream comparisons (ANZECC & ARMCANZ 2000). However, when using multivariate data, the MBACIP sitepair approach reduces the multi-dimensional data to just one dimension (a metric scale from 0-100). As illustrated above (Figure A1), this results in a potential loss of information relating to *direction* of change in multivariate space. This constraint on the sitepair dissimilarity data limits interpretation of results as change in macroinvertebrate community structure at the exposed downstream site could occur in different directions in multivariate space. Thus, a non-significant before to after dissimilarity value could mask real change that is occurring at the impact site.

Given the different analysis approaches, the hypotheses being tested with each analysis are somewhat different. Because the MBACIP sitepair dissimilarity aims to eliminate or reduce temporal and spatial variation, interpretation is based upon the changes between sitepairs (dissimilarity value) for each year. PERMANOVA retains the spatial and temporal variation but partitions these sources of variation by an additional factor, thus interpretation is based upon changes amongst years for each site (downstream).

To this end, the factors used to analyse the data differ slightly between the two models and are outlined in Table A1. While both models include the 'Before vs After' (BA), 'Years', Exposure and Stream factors, PERMANOVA includes an additional factor, upstream vs downstream. The MBACIP ANOVA uses the 'stream' factor to partition the replicate upstream-downstream dissimilarity values within each of the four streams. PERMANOVA, based on the complete dissimilarity matrix, requires the additional fixed factor (upstream vs downstream) to partition the equivalent data by the up and downstream sites. For PERMANOVA, the replicates are values (from the complete dissimilarity matrix) that represent the macroinvertebrate community structure within each of the up- and downstream sites, enabling analysis of the before and after periods within each site.

Table A1	Description of factors	used for PERMA	NOVA and	MBACIP	ANOVA	analysis with	1 factor
designatio	n included (nested, ar	nd fixed or randor	m) .				

Factors	Nested in	Fixed or Random	Analysis that includes the factor
Before vs. After (BA)		Fixed	PERMANOVA, ANOVA
Years	ВА	Random	PERMANOVA, ANOVA
Exposure		Fixed	PERMANOVA, ANOVA
Stream	Exposure	Random	PERMANOVA, ANOVA
Upstream vs downstream		Fixed	PERMANOVA

Furthermore, because the two analysis models use different datasets (and one additional factor for PERMANOVA) the approach to interpreting the two sets of results also differs. The interpretation of results arising from each model is detailed in Table A2. In summary, impact detection viz the MBACIP ANOVA model is assessed using the 'BA*Exposure' interaction if the 'BA*Stream(Exposure)' interaction is not significant (ie that both streams upstreamdownstream dissimilarity show a consistent change from the before to after period). However, in this monitoring design, the BA*Stream(Exposure) interaction is likely to be the most important because it will indicate if a stream's upstream-downstream dissimilarity is responding differently to the other streams (ie a stream that is closer to the minesite is more likely to show mine-related change). In this situation, further investigation of any aberrant stream upstream-downstream dissimilarity would be required. In PERMANOVA, the 'BA*Exposure*Us/Ds' interaction (not available in MBACIP ANOVA) is the important source of variation to interpret for impact detection after ensuring the 'BA*Us/Ds*Stream(Exposure)' interaction is not significant (ie that all changes between sites and streams within each exposure type show a consistent change from the before to after period). In this monitoring design, equivalent to the MBACIP ANOVA model, the 'BA*Us/Ds*Stream(Exposure)' interaction is likely to be the most important because it will indicate if stream sites within exposure types are responding differently to the other stream sites (ie an exposed downstream site that is closer to the minesite is more likely to show mine related change).

 Table A2
 Interpretation of each factor and interaction for the PERMANOVA and MBACIP ANOVA analyses on macroinvertebrate community structure data in seasonal streams. Important factors/interactions for interpreting impact detection are identified.

Factors	Analysis type	Interpretation
ВА	PERMANOV A	A significant result indicates change from the before to after periods across all streams and exposures. While this factor is not directly interpreted for impact detection, it can be interpreted for non significant BA*Exposure*Up/Ds and BA*Up/Ds*stream(Exposure) interactions to indicate if overall change from before to after has occurred
	ANOVA	A significant result indicates change from the before to after periods in the magnitude of upstream-downstream dissimilarity across both streams and exposures. While this factor is not directly interpreted for impact detection, it can be interpreted for non significant BA*Exposure and BA*stream(Exposure) interactions to indicate if overall change from before to after has occurred
Exposure	PERMANOV A	A significant result indicates differences between the two exposure types (across all streams and sites within each exposure type). Not important for impact detection because exposure is assigned to each stream (ie Magela and Gulungul are Exposed streams: Nourlangie and Burdulba are Control streams) and each stream has both an upstream (not influenced by the mine) and a downstream (influenced by mine on exposed streams) site.
	ANOVA	A significant result indicates differences in the magnitude of upstream-downstream dissimilarity between the two exposure types (across both streams within each exposure type). May indicate mining impact, though care is needed in the interpretation because any differences could be due to natural variation between catchments. Consideration of the stream(exposure) factor is needed to determine if natural variation between streams occurs (indicated if this factor is significant). Seeking correlates of macroinvertebrate community structure using BIOENV analysis may be required to aid interpretation (ie natural habitat variation versus mine-related water quality)
Up/Ds	PERMANOV A	A significant result indicates differences between the upstream and downstream sites (across all streams, exposures and years). Not important for impact detection because the control streams include downstream sites that are in reference condition.
Year(BA)	PERMANOV A	A significant result indicates if years differ within the before or after period across all sites, streams and exposures. This factor is not important for impact detection.
	ANOVA	A significant result indicates the magnitude in dissimilarity across all streams differs amongst years within the before or after period. This factor is not important for impact detection.
Stream(Exposure)	PERMANOV A	A significant result indicates streams (across both sites) differ within either or both exposure types. Not important for impact detection because this test combines upstream (reference at all sites) and downstream sites (impacted only at the two exposed sites). This factor useful in understanding spatial variation amongst streams.
	ANOVA	A significant result indicates streams differ in the magnitude of upstream-downstream dissimilarity within either or both exposure types. Pairwise comparisons are required to explore a significant result to test whether differences occur amongst streams within control or exposed designations, or both. Differences amongst streams within both exposure types most likely indicate natural differences amongst streams.
BA*Exposure	PERMANOV A	A significant result indicates that across sites and streams within each exposure type the change from before to after periods is not consistent between the control and exposed streams. Not important for impact detection because exposure is assigned to each stream (ie Magela and Gulungul are Exposed streams: Nourlangie and Burdulba are Control streams) and each stream has both an upstream (not influenced by the mine) and a downstream (influenced by mine on exposed streams only) site.
	ANOVA	Key interaction for impact detection. A significant result indicates that across streams within each exposure type the change in the magnitude of upstream-downstream

Factors	Analysis type	Interpretation
		dissimilarity from before to after periods is not consistent between the control and exposed streams (suggesting minesite influence). A non-significant result for this factor needs to be interpreted carefully because a significant BA*stream(Exposure) interaction may indicate streams within either, or both, exposure type are not consistent within or between the before and after periods. A significant BA*stream(Exposure) interaction requires further investigation to understand the nature of the differences.
BA*Up/Ds	PERMANOV A	A significant result indicates that across streams and exposures the change from before to after periods is not consistent between the up- and downstream sites. Not important for impact detection.
Exposure*Up/Ds	PERMANOV A	A significant result indicates that across years and streams within each exposure type the change from upstream to downstream is not consistent between the control and exposed streams. May indicate mining impact, though care in the interpretation is needed because any differences could be due to natural variation between catchments. Consideration of the stream(exposure)*Up/Ds factor is needed to determine if natural variation between streams occurs (indicated if this factor is significant). Seeking correlates of macroinvertebrate community structure using BIOENV analysis may be required to aid interpretation (ie natural habitat variation versus mine-related water quality).
BA*Stream(Exposure)	PERMANOV A	A significant result indicates that across sites for streams within either, or both, exposure type, the changes are not consistent within or between the before and after periods. Not important for impact detection because factor, streams, includes both the up- and downstream sites.
	ANOVA	Key interaction for impact detection. A significant result indicates the change in magnitude of upstream-downstream dissimilarity between streams within either, or both, exposure types is not consistent within or between the before and after periods. If significant, this interaction is the most important for impact detection because, for example, it might reflect that one exposed stream (e.g. closest to the minesite) is responding differently to all other streams. A significant result for this interaction requires further investigation to determine the nature of the inconsistencies and whether or not they indicate mine impact. Pairwise comparisons can be used to explore a significant result.
Exposure*Year(BA)	PERMANOV A	A significant result indicates that across sites and streams within each exposure type the changes amongst Years(BA) are not consistent between the two exposures. This interaction is not important for impact detection.
	ANOVA	A significant result indicates that the change in magnitude of upstream-downstream dissimilarity across streams within each exposure type is not consistent amongst Years(BA) between the two exposures. This interaction can be useful for impact detection if, for example, variation in the upstream-downstream dissimilarity over time was only detected across exposed streams. Further consideration of the Year(BA)*Stream(Exposure) interaction is required to determine if variation over time within streams for either, or both, exposures is consistent. Variation over time within all streams, regardless of exposure, would suggest temporal variation is due to natural causes.
Up/Ds*Year(BA)	PERMANOV A	A significant result indicates that across streams and exposure type the changes amongst Years(BA) are not consistent between the up- and downstream sites. This interaction is not important for impact detection.
Up/Ds*Stream(Exposure)	PERMANOV A	A significant result indicates that across years for streams within either, or both, exposure types, the changes are not consistent between the up- and downstream sites.
BA*Exposure* Up/Ds	PERMANOV A	Key interaction for impact detection. A significant result indicates that differences between sites across streams within exposure types are not consistent within or between the before and after periods (suggesting minesite influence). A non-significant result for this factor needs to be interpreted carefully because a significant BA*Up/Ds*stream(Exposure) interaction may indicate differences between up- and downstream sites at streams within either, or both, exposure types are not consistent within or between the before and after periods. A significant BA*Up/Ds*stream(Exposure) interaction to understand the nature of the

Factors	Analysis type	Interpretation
		differences.
Year(BA)*Stream(Exposure)	PERMANOV A	A significant result indicates that across sites for streams within either, or both, exposure types, the variation amongst Years(BA) is not consistent. This interaction is not important for impact detection.
	ANOVA	A significant result indicates that the variation in magnitude of upstream-downstream dissimilarity amongst Years(BA) is not consistent amongst streams within either, or both, exposure types. This interaction can be useful for impact detection if, for example, variation in the upstream-downstream dissimilarity over time was only detected amongst exposed streams. Pairwise comparisons can be used to explore a significant result.
BA* Up/Ds*Stream(Exposure)	PERMANOV A	Key interaction for impact detection. A significant result indicates that differences between sites for streams within either, or both, exposures are not consistent within or between the before and after periods. When significant, this interaction is the most important for impact detection because, for example, it might indicate that one exposed downstream site (from the stream closest to the minesite) is responding differently to all other downstream sites. A significant result for this interaction requires further investigation to determine the nature of the inconsistencies and whether or not they indicate mine impact. Pairwise comparisons can be used to explore a significant result.
Exposure* Up/Ds*Year(BA)	PERMANOV A	A significant result indicates that across streams within each exposure type the changes amongst Years(BA) are not consistent between sites and/or the two exposures.
Up/Ds*Year(BA)*Stream(Exposure)	PERMANOV A	A significant result indicates that the variation amongst Years(BA) is not consistent between sites and streams within either, or both, exposure type. This interaction can be useful for impact detection if, for example, variation over time was only detected amongst exposed downstream sites. Pairwise comparisons can be used to explore a significant result.

The analysis methods for the MBACIP ANOVA are described in section 2 of the main report. For PERMANOVA, the analysis has followed the MBACIP ANOVA data preparation procedures with the exception of the upstream-downstream pairing stages (unnecessary). The default settings in the PERMANOVA analysis package (Anderson et al 2008) have been used with the following exceptions:

- 1. The number of permutations was increased to 9999.
- 2. An unrestricted model was selected in order to match the approach used by Minitab's General Linear Model (GLM, required for unbalanced designs) ANOVA. Selecting an unrestricted model in PERMANOVA is achieved by un-selecting the 'Fixed effects sum to zero'. See below (section A4) for a brief discussion on the different models.

A3 Comparison of PERMANOVA and MBACIP ANOVA results

For this comparison and by way of example, data for the paired (up/downstream) sites from the two exposed and two control streams have been analysed in order to compare the most recent 'after' year (2011) with the 13 'before' years (1998 to 2010), and so determine whether a potential impact has occurred in the exposed stream(s) in the 'after' period (year of interest).

The complete results for PERMANOVA are shown in Table A3. The complete results for the MBACIP ANOVA analysis are available in section 6.2.3 (Table 6) of the main report. For both PERMANOVA and the MBACIP ANOVA analyses, a summary interpretation is provided in Table A4.

From Tables 6 (main report) and A3, the PERMANOVA analysis has a greater total degrees of freedom (515) than the MBACIP ANOVA (279) so, theoretically, PERMANOVA should provide greater statistical power due to the inclusion of more data.

A3.1 Impact detection

The results of the PERMANOVA analysis showed that upstream and downstream sites across streams within each exposure are consistently similar between the 'before' and 'after' periods (BA*Exposure*Up/downstream, p = 0.6902). However, sites (upstream or downstream) differ within either. or both. exposures from before to after (BA*Up/downstream*Stream(Exposure), p = 0.0221) (Table A3). Further investigation using pairwise comparisons indicates that differences from Before to After occur at the upstream sites for both Gulungul (p = 0.0122) and Burdulba creeks (p = 0.0046) (illustrated in Figure A2). Both these (upstream) sites are not influenced by minesite activity hence the changes from before to after are due to natural shifts in the macroinvertebrate communities (see section 6.2.3.2) in main report for further information). Furthermore, these results indicate that the macroinvertebrate communities at exposed downstream site have not changed from the before to after period and hence there is no indication of a mine-site influence at the exposed sites in the after period. These results highlight the advantage of PERMANOVA in enabling assessment and interpretation of BA periods at each site (upstream and downstream) within each stream.



Figure A2 Ordination plots (three dimensional) of macroinvertebrate community structure for (A) Gulungul Creek (axis 1 and 2) and (B) Burdulba Creek (axis 1 and 3). Upstream and downstream data are identified for the years Before (1998-2010) and After (2011).

The results of the sitepair MBACIP ANOVA analysis showed that across streams, within each exposure, the change from 'before' to 'after' periods is consistent (BA*Exposure interaction (p = 0.458). However, streams within either, or both, exposures differ from before to after (BA*Stream(Exposure), p = 0.010). Examination of this interaction using pairwise comparisons in the PERMANOVA² package indicates that the dissimilarity between upstream-downstream in Gulungul Creek (exposed site) differs from the Before to After period. (For this same analysis, no difference is detected for Burdulba Creek). Without further investigation, this result could indicate minesite impact. However, for the MBACIP ANOVA analysis, interpretation is not available for each up or downstream site and as a result the

² The PERMANOVA package has operational advantages over the MINTAB program. In PERMANOVA pairwise comparisons are possible on interactions that involve random factors (cautions are provided).

analysis does not (nor can it) include information on within-site changes over time. To this end, interpretation for impact detection is limited because it is not possible to determine the nature of the change in magnitude of dissimilarity in Gulungul Creek from Before to After (ie change could be occurring at either the up- or downstream, or both, sites).

Table A3 PERMANOVA (unrestricted model) results for stream macroinvertebrate community structureusing Bray-Curtis dissimilarity matrix. Analysis includes two Control (Burdulba & Nourlangie) and twoExposed (Magela & Gulungul) streams. Years Before (B) are 1998-2010, year after (A) is 2011.

Source	df	SS	MS	Pseudo-F ¹	P(perm) ²	Unique Perms ³
BA	1	8972.90	8972.90	1.8955	0.0135	9892
Exposure	1	4967.70	4967.70	0.7769	0.8306	9885
Up/Ds	1	4978.60	4978.60	1.2114	0.2261	9893
Year(BA)	12	50362.00	4196.80	1.9625	0.0001	9809
Stream(Exposure)	2	13122.00	6561.10	1.6182	0.2041	9949
BA*Exposure	1	978.98	978.98	0.8383	0.7654	9888
BA*Up/Ds	1	1230.30	1230.30	0.8080	0.7693	9904
Exposure*Up/Ds	1	3140.60	3140.60	0.8807	0.6605	9891
BA*Stream(Exposure)	2	2420.90	1210.50	0.7148	0.9333	9871
Exposure*Year(BA)	12	18260.00	1521.70	1.1933	0.0408	9724
Up/Ds*Year(BA)	12	11912.00	992.69	1.3456	0.0425	9853
Up/Ds*Stream(Exposure)	2	7756.00	3878.00	2.3178	0.1917	9939
BA*Exposure* Up/Ds	1	1169.60	1169.60	0.8686	0.6902	9904
Year(BA)*Stream(Exposure)	24	31505.00	1312.70	1.4208	0.0016	9732
BA* Up/Ds*Stream(Exposure)	2	3346.30	1673.10	1.8061	0.0221	9905
Exposure* Up/Ds*Year(BA)	12	8852.80	737.73	0.7992	0.9387	9787
Up/Ds*Year(BA)*Stream(Exposure)	24	22174.00	923.93	2.2628	0.0001	9715
Res	448	182920.00	408.31			
Total	559	448660.00				

Superscripts 1 = Pseudo-F is the permutation equivalent to a standard F value, 2 = P(perm) is the permutation equivalent to a standard p (significance) value, 3 = Unique perms are the number of unique permutations used to determine P(perm).

Comparing the results between MBACIP ANOVA and PERMANOVA analyses for this dataset highlights the limitations when interpreting multi-dimensional data (analysed via PERMANOVA) reduced to one dimension (dissimilarity, analysed via MBACIP ANOVA) (limitations discussed in sections A1 and A2). Specifically for this comparison, two limitations are encountered:

- The MBACIP ANOVA analysis identifies change for Gulungul Creek but is not able to determine the nature of this change. PERMANOVA, on the other hand, identifies that the macroinvertebrate community change occurs at the upstream, control site and not the exposed, downstream site (see Figure A2a). Figure A2a illustrates the increased separation, or dissimilarity, between the upstream and downstream sites for Gulungul Creek in 2011. However, the shift in direction from the before (1998-2010) to after period (2011) is greatest, and significantly different, for the upstream site only.
- 2. The MBACIP ANOVA does not detect change in macroinvertebrate (upstream downstream) dissimilarity data for Burdulba Creek. PERMANOVA, however, identifies a change in macroinvertebrate community structure from before to after at the upstream

site. In this case, the MBACIP ANOVA appears to mask real change that has occurred in multivariate direction. Figure A2b illustrates the separation, or dissimilarity, between the upstream and downstream sites for Burdulba Creek in 2011 has remained similar to all other years. However, a significant shift in direction from the before (1998-2010) to after period (2011) has occurred, for the upstream site only, a situation similar to the hypothetical scenario illustrated in Figure A1.

To this end, the inclusion of temporal and spatial variation in the PERMANOVA analysis (using the multivariate dissimilarity matrix) enables greater interpretation of the data and hence greater assurance that macroinvertebrate communities at the exposed downstream sites have not been influenced by minesite activity (pairwise comparison of the significant BA*Up/Ds*stream(Exposure) interaction shows changes from Before to After occur only at the upstream sites).

Of further note with the sitepair MBACIP ANOVA analysis, is that exact F tests could not be used to derive significance levels on all the factors and interactions, which could result in less reliable results. In this analysis, the BA*Exposure interaction, which is an important interaction for impact detection, has been generated using a "quasi F test". Re-analysis of the same sitepair dissimilarity data using the PERMANOVA package³, using the equivalent unrestricted model, shows that the BA*Exposure interaction remains non significant (p = 0.568) (Table A5). Importantly, in the MBACIP ANOVA analysis, the significant BA*Stream(Exposure) interaction (see above) is not influenced by the "quasi F test".

A3.2 Temporal variation

For PERMANOVA, variation over time is detected across all sites and streams in the before or after period (Year(BA), P = 0.0001). Because there is no replication (of years) in the after period, this significant variation must occur within the before period. Furthermore, the variation over time is not consistent amongst years between exposure conditions (Exposure*Year(BA), p = 0.0408), between upstream and downstream sites across all streams (Up/Ds* Year(BA), p = 0.0425), between streams within either exposure condition (Year(BA)*Stream(Exposure), p = 0.0016), and between upstream and downstream sites within each stream (Up/Ds*Year(BA)*Stream(Exposure, p=0.0001) (Table A3). Pairwise comparison indicates that significant differences amongst years in the before period occur at all sites within all streams (regardless of exposure), and that differences also vary between streams (results not shown). The results indicate that the temporal variation over time is natural for macroinvertebrate communities in these seasonal tropical streams.

Despite the MBACIP design approach (which aims to remove temporal variation), significant temporal variation is nonetheless detected by the MBACIP ANOVA (Year(BA)*Stream(Exposure), p = 0.024) (Table A4) which simply indicates that dissimilarity values for the different streams – regardless of their status (Before, After, Control, Impact) – show differences through time. Pairwise comparison of this factor using the PERMANOVA⁴ package shows that significant differences occur amongst years within control and impacted streams but vary between streams within exposures (results not shown).

³ Analysis using the PERMANOVA procedure has been completed on the sitepair dissimilarity data (same data derived for the ANOVA analysis) to determine if the quasi F tests used during ANOVA analysis has compromised the interpretation of these factors or interactions. PERMANOVA is more robust than ANOVA in this respect because it uses a random permutation procedure to generate a Pseudo-F value.

⁴ The PERMANOVA package has operational advantages over the MINTAB program. In PERMANOVA pairwise comparisons are possible on interactions that involve random factors (cautions are provided).

The retention of temporal and spatial information enables PERMANOVA to interpret the natural variations in macroinvertebrate communities amongst years within exposures, streams and sites (up- or downstream). While the BACIP ANOVA analysis does detect temporal change, the variation detected is just the change in magnitude of the between-site (up-/downstream) dissimilarity. Such changes in dissimilarity could be caused by changes at either, or both, sites.

A3.3 Spatial variation

Not surprisingly, the MBACIP ANOVA successfully removes spatial variation in macroinvertebrate communities by its upstream-downstream sitepair approach. Interestingly, PERMANOVA using the unrestricted model does not detect spatial variation either (Table A4).

A4 PERMANOVA and Minitab program functionality differences

There are three noticeable programming advantages to PERMANOVA over Minitab in this comparison. These are:

- PERMANOVA allows pairwise comparison with random factors (after a suitable warning). This is useful when exploring significant differences in factors that are considered random, but in reality are not truly random (ie the factors 'years' and 'streams' same streams used over time after the initial selection). Minitab will not conduct pairwise comparisons on random factors.
- 2. PERMANOVA allows the choice of analysis model type (restricted or unrestricted models), but defaults to a restricted model to overcome the intrinsic over-parameterisation of the ANOVA model (see Anderson et al 2008, p. 45). The GLM ANOVA required in Minitab uses an unrestricted model. The choice of model (restricted versus unrestricted) appears to be still debated amongst statisticians (Quinn & Keough 2002, box 9.7 p. 233; Anderson et al 2008) and is not discussed further here. However, use of either of the two models does not influence the interpretation of results for impact detection when applied to the current dataset.⁵ For other comparisons, however, differences can occur and this requires further consideration/advice to determine the most appropriate model.
- 3. The random permutation procedures used by PERMANOVA are less restricted by data assumptions and are not constrained by the F test procedure used by GLM ANOVA in Minitab. Minitab resorts to 'quasi F tests' (potentially less reliable methods to calculate an F statistic) when an exact F test calculation is not available for more complex ANOVA interactions. The calculation of a pseudo F statistic in PERMANOVA is not restricted by the model complexity. To this end, the PERMANOVA package could be useful for analysing sitepair dissimilarity data to confirm test statistics generated by the quasi-F tests in ANOVA or for provide assurance of the statistical robustness of the ANOVA model when data assumptions are not met.

⁵ Comparison of restricted versus unrestricted models with the seasonal stream macroinvertebrate community data has been conducted. Results for the BACIP sitepair dissimilarity data analysed using PERMANOVA and using restricted and unrestricted models remained similar for interpretation. However, the Stream(Exposure) factor has marginal non-significance, p = 0.072, using the restricted model. Analysis of the multivariate dissimilarity data (analysed using PERMANOVA) using both models gave different significance results for: the Stream(Exposure) factor; Exposure*year(BA) interaction; Up/Ds*year(BA) interaction; and Up/Ds*Stream(Exposure) interaction.

Table A4 Comparison of change-detection analyses (based on macroinvertebrate community structure data), between a five-factor PERMANOVA using the entire Bray-Curtis dissimilarity matrix and a four-factor ANOVA (MINITAB) and PERMANOVA using the upstream-downstream sitepair dissimilarity values, from two Control (Burdulba & Nourlangie) and two Exposed streams (Magela & Gulungul). Change detection based on years after (A) (2011) versus years before (B) (1998 – 2010).

Factors	MBACIP ANOVA	PERMANOVA	Interpretation
BA	0.784 x	0.0135	PERMANOVA ; Across all streams (control and exposed) there is significant change in the macroinvertebrate community structure from before and after (2011 versus earlier years)
			ANOVA; Across all streams (control and exposed) there is no significant difference in the mean dissimilarity measures calculated before and after (2011 versus earlier years)
Exposure	0.457 x	0.8306	PERMANOVA; Across all time periods (before and after) there is no significant difference in the macroinvertebrate community structures between control and exposed sites.
			ANOVA; Across all time periods (before and after) there is no significant difference in the mean dissimilarity measures calculated between control and exposed streams
Up/Ds	N/A	0.2261	Across all time periods (before and after) there is no significant difference in the macroinvertebrate community structures between upstream and downstream sites
Year(BA)	0.507	0.0001	PERMANOVA; Within either of the two time periods (B and A), there is significant change in macroinvertebrate community structure amongst years across control and exposed streams
			ANOVA; Within either of the two time periods (B and A), there is no significant difference in dissimilarity amongst years across control and exposed streams
Stream(Exposure)	0.666	0.2041	PERMANOVA ; Within the two exposure groups there is no significant differences in macroinvertebrate community structure between the two 'duplicate' streams across all times (B and A).
			ANOVA; Within the two exposure groups there is no significant difference in mean dissimilarity of the two 'duplicate' streams across all times (B and A)
BA*Exposure	0.458 x	0.7654	PERMANOVA; There is no significant difference between the exposed and control streams in the change (in macroinvertebrate community structure) from before to after
			ANOVA; There is no significant difference between the exposed and control streams in the change (in dissimilarity) from before to after
BA*Up/Ds	N/A	0.7693	Across all streams there is no significant difference between the upstream and downstream sites in the change (in macroinvertebrate community structure) from before to after
Exposure*Up/Ds	N/A	0.6605	There is no significant difference between the exposed and control streams with respect to differences between upstream and downstream sites (across streams within each exposure).
BA*Stream(Exposure)	0.010	0.9333	PERMANOVA; Within either of the two exposure groups, there is no significant difference in macroinvertebrate community structure of the two respective 'duplicate' streams in the change from before to after.
			ANOVA; Within either of the two exposure groups, there is a significant difference in mean dissimilarity of the two respective 'duplicate' streams between times B and A. Pairwise comparison using the PERMANOVA program shows the exposed stream, Gulungul Creek, is significantly different (p=0.002) from the before to after period.

Factors	MBACIP ANOVA	PERMANOVA	Interpretation
Exposure*Year(BA)	0.291	0.0408	PERMANOVA; Within either of the two time periods, B and A, there is significant change in macroinvertebrate community structure amongst years within either of the two exposures, or within both.
			ANOVA; Within either of the two time periods, B and A, there is no significant difference in mean dissimilarity amongst years within, or between, either of the two exposures
Up/Ds*Year(BA)	N/A	0.0425	Within either of the two time periods, B and A, there is significant change in macroinvertebrate community structure amongst years within, or between, either of the upstream or downstream sites (across all streams)
Up/Ds*Stream(Exposure)	N/A	0.1917	Within either of the two exposure groups, there is no significant difference in macroinvertebrate community structure between the two 'duplicate' streams within, or between, either of the up- or downstream sites, across all times (B and A).
BA*Exposure* Up/Ds	N/A	0.6902	Within either of the upstream or downstream sites across streams, within either exposure type, there are no significant differences in macroinvertebrate community structures from the before to after period.
Year(BA)*Stream(Exposure)	0.024	0.0016	PERMANOVA; Within status categories Control, Exposed, Before, After, macroinvertebrate community structure for at least one year differed from that for other years. Pairwise comparisons show significant difference amongst years for all streams
			ANOVA; Within status categories Control, Exposed, Before, After, mean dissimilarity for at least one year differed from that for other years
BA* Up/Ds*Stream(Exposure)	N/A	0.0221	Within either the upstream or downstream sites at either 'duplicate' stream within the exposure conditions, a significant difference has occurred between macroinvertebrate community structures in the change from before to after. Pairwise comparison shows that upstream sites for Burdulba (p=0.0046) and Gulungul (p=0.0122) differ from the before to after period. In both cases, these are control sites and the changes are not due to minesite influence.
Exposure* Up/Ds*Year(BA)	N/A	0.9387	Within either of the two time periods, B and A, and within the upstream or downstream sites, no significant difference in the macroinvertebrate community structures has occurred between either of the exposure (Exposed or Control) conditions.
Up/Ds*Year(BA)*Stream(Exposure)	N/A	0.0001	Within streams in either of the exposure conditions and amongst years in ether the B or A period, there are significant differences in the macroinvertebrate community structures between the upstream and downstream sites. Pairwise comparisons show significant differences amongst years for upstream and downstream sites within all streams (results not provided)

x Not an exact F-test

 $\overline{\Omega}$

Table A5 Sitepair dissimilarity (univariate) results using PERMANOVA (unrestricted model) for stream
macroinvertebrate community sitepair (up-/downstream) dissimilarity values using two Exposed streams
(Magela & Gulungul,) and two Control streams (Burdulba & Nourlangie). Years Before (B) are 1998-
2010, year After (A) is 2011.

Source	df	SS	MS	Pseudo-F ¹	P(perm) ²	Unique perms ³
BA	1	46.73	46.73	0.2261	0.9800	996
Exposure	1	202.61	202.61	0.8398	0.5290	999
Year(BA)	12	1256.80	104.73	0.9901	0.5280	999
Stream(Exposure)	2	467.65	233.82	0.5005	0.6613	490
BA*Exposure	1	392.84	392.84	0.8298	0.5680	999
BA*Stream (Exposure)	2	934.30	467.15	5.6569	0.0080	999
Year (BA) *Exposure	12	1269.40	105.79	1.2810	0.2970	999
Year (BA)*Stream (Exposure)	24	1981.90	82.58	1.7137	0.0210	996
Res	224	10794.00	48.19			
Total	279	16994.00				

Superscripts 1 = Pseudo-F is the permutation equivalent to a standard F value, 2 = P(perm) is the permutation equivalent to a standard p (significance) value, 3 = Unique perms are the number of unique permutations used to determine P(perm).

A5 Conclusions

The BACIP ANOVA result has limitations for impact detection due to the loss of withinstream (upstream or downstream) information. This, a significant change from the Before to After period is identified for the Exposed Gulungul Creek, but with information available only on the magnitude in dissimilarity between upstream-downstream sitepair, it is not possible to determine whether the change is occurring at either, or both, sites. PERMANOVA, on the other hand, utilises the temporal and spatial variation (using the multivariate matrix) which enables further partitioning within the dataset and hence greater interpretation for *each* upstream and downstream macroinvertebrate community. To this end, PERMANOVA identifies that significant change from before to after has occurred at the upstream Gulungul Creek site, not at the downstream site, and hence the change is not minerelated.

Furthermore, PERMANOVA has detected change from before to after at the upstream Burdulba Creek site that has remained undetected in the BACIP ANOVA sitepair dissimilarity data. In this instance, it appears PERMANOVA has detected change in macroinvertebrate community structure that has occurred in a different direction in multivariate space but which has passed unnoticed in the one-dimensional sitepair dissimilarity data analysed by ANOVA (see Figure A1 for hypothetical example of this scenario).

From a programming perspective, the PERMANOVA package provides benefits over the Minitab program in relation to:

- 1. Minitab does not allow pairwise comparisons on random factors. Pairwise comparisons on random factors need to be done with caution, as comparisons between terms which are truly randomised (ie sites randomly selected on each sampling occasion) are meaningless.
- 2. The ability to select the analysis model type (restricted or unrestricted model).
- 3. The ability to analysis data and design models that may not fit well with traditional ANOVA methods in Minitab. (ie PERMANOVA can provide supportive evidence for

significance values generated by quasi-F tests, or support analysis conducted on data not strictly conforming to ANOVA data assumptions).

A6 References

- Anderson MJ 2001. A new method for non-parametric multivariate analysis of variance. *Austral Ecology* 26, 32-46.
- Anderson MJ, Gorley RN and Clarke KR 2008. *PERMANOVA+ for PRIMER: Guide to software and statistical methods*. Primer-E: Plymouth, UK.
- ANZECC & ARMCANZ 2000. Australian and New Zealand guidelines for fresh and marine water quality. National Water Quality Management Strategy Paper No 4, Australian and New Zealand Environment Conservation Council & Agriculture and Resource Management Council of Australia and New Zealand, Canberra.
- Clarke KR & Gorley RN 2006. Primer v6: User manual/Tutorial. Primer-E: Plymouth.
- McArdle BH & Anderson MJ 2001. Fitting multivariate models to community data: a comment on distance-based redundancy analysis. *Ecology* 82, 290-297.

Annex – PERMANOVA and MBACIP ANOVA results for mid year reporting: Assessment of 2010–11 wet season results

Background

At the time of preparing the Supervising Scientist's annual report, only samples from Magela and Gulungul Creeks from the just-completed wet season are available for analysis (see section 2.2.1.2). Without comparable data from the two control streams, it is not possible to run the full ANOVA testing for the wet season of interest. Instead, modifications to data analysis are performed, as described in the following sections, where Magela and Gulungul data from 2011 sampling are analysed as a worked example.

Impact detection and assessment

A modified ANOVA model is performed using the factors Before/After (BA; fixed), Year (nested within BA; random) and Stream (upstream vs downstream paired dissimilarities; random) examining just the exposed creeks, Magela and Gulungul, to determine if any change in these streams has occurred (see associated hypotheses described in section 2.2.1.2).

The MINITAB output for the ANOVA results using 2011 data is shown below (Table AA1). The ANOVA showed no significant change from the before (pre 2011–10) to after (2011–10) periods in the magnitude of upstream-downstream dissimilarity across both 'exposed' streams (BA not significant). However, the BA*Stream interaction for the same before-after comparison is significant (p < 0.05, Table AA1), which indicates the change in magnitude of paired-site dissimilarity in either, or both, exposed streams is not consistent within or between the before and after periods.

Source	DF	Seq SS	Adj SS	Adj MS	F	Р	
BA	1	355.26	355.26	355.26	0.59	0.585 *	
Year(BA)	12	995.08	995.08	82.92	1.00	0.499	
Stream	1	1.26	421.83	421.83	0.7	0.558	
BA*Stream	1	606.95	606.95	606.95	7.33	0.019	
Year(BA)*Stream	12	993.12	993.12	82.76	1.48	0.141	
Error	112	6256.42	6256.42	55.86			
Total	139	9208.09					
							-

 Table AA1
 ANOVA results for stream macroinvertebrate community upstream-downstream dissimilarity values for the two exposed streams, Magela and Gulungul Creeks. Years 'Before' are 1998–2010, year 'After' is 2011.

* Not an exact F-test

Further investigation using pairwise tests showed that the significant difference in the change in dissimilarity was associated with the Gulungul sites (p=0.002) – Gulungul Creek has responded differently from the before to after periods. From Figure 5 of the main protocol, a sharp rise in dissimilarity for Gulungul Creek can be observed. Closer examination of the data is required to assess whether or not this result may be associated with the Gulungul downstream site, and thereby indicate possible mining impact. The steps described in section 6.2.3.2 of the protocol are followed for further analysis and assessment. Steps entail:

- 1. Multi-Dimensional Scaling (MDS) ordination to assist in determining whether the upstream and/or downstream Gulungul communities have changed or are aberrant compared with the other communities sampled over time;
- 2. PERMANOVA to examine which site(s) in the significant sitepair (in this case Gulungul Creek) are aberrant in the before to after period; and
- 3. To support the ordination from step 1, abundances of the numerically-dominant taxa are compared between the upstream and downstream sites over time to determine what types of shifts in taxa abundances may have occurred in 2011.

The ordination was conducted using sites sampled in Magela and Gulungul Creeks downstream of Ranger for each year of study (to 2011), relative to Magela and Gulungul Creek upstream (control) sites for 2011, and all other control sites (Magela and Gulungul upstream sites, all sites in Burdulba and Nourlangie except 2011) (ordination plot not shown here though it closely resembled the the ordination shown in Figure 6 of the main protocol which included Burdulba and Nourlangie sites from 2011).

The ordination indicated that Gulungul Creek communities collected from the upstream site in 2011 differed from communities from other sites and times (see also section 6.2.3.2). Conversely, data-points associated with the 2011 Gulungul and Magela downstream sites were generally interspersed among the points representing the control sites, indicating that these 'exposed' sites have macroinvertebrate communities that are similar to those occurring at control sites.

Using 2011 versus previous years' data from just the exposed creek sites (Magela and Gulungul), the aberrant 2011 Gulungul upstream result (evident from the ordination) was further examined using PERMANOVA. From Table A2, the key factors to examine in the PERMANOVA are the BA and BA x Stream x Upstream/Downstream interaction. The latter interaction is used to confirm whether there is inconsistency in the differences between sites of the exposed streams within or between the before and after periods. A significant interaction indicates that one of the sites (in Magela or Gulungul) is responding differently between the two time periods, and thus requires further investigation. This was confirmed in the PERMANOVA (Table AA2, BA x Stream x Upstream/Downstream, p = 0.006) where the effect is also sufficient to result in a significant Before/After effect averaged across the two streams (Table AA2, BA, p = 0.02).

Table AA2 PERMANOVA (unrestricted model) results for stream macroinvertebrate communitystructure using Bray-Curtis dissimilarity matrix. Analysis is conducted for the two exposed streams,Magela and Gulungul Creeks. Years Before (B) are 1998-2010, year after (A) is 2011.

Source	df	SS	MS	Pseudo-F ¹	P(perm) ²	Unique Perms ³
BA	1	5190.6	5190.6	1.9118	0.0101	9901
Up/Ds	1	5264	5264	1.3368	0.2339	9926
Year(BA)	12	31578	2631.5	1.8028	0.0001	9813
Stream	1	6728.1	6728.1	4.6093	0.0019	9919
BA*Up/Ds	1	1184.1	1184.1	0.69172	0.8135	9926
BA*Stream	1	847.12	847.12	0.58035	0.8658	9928
Up/Ds*Year(BA)	12	9967.6	830.63	1.0758	0.3529	9864
Up/Ds*Stream	1	3684.6	3684.6	4.7723	0.0051	9940
Year(BA)*Stream	12	17516	1459.7	3.5653	0.0001	9772

BA* Up/Ds*Stream	1	1997.4	1997.4	2.587	0.033	9936
Up/Ds*Year(BA)*Stream	12	9264.8	772.07	1.8858	0.0001	9785
Res	224	91709	409.41			
Total	279	2057400				

Superscripts 1 = Pseudo-F is the permutation equivalent to a standard F value, 2 = P(perm) is the permutation equivalent to a standard p (significance) value, 3 = Unique perms are the number of unique permutations used to determine P(perm).

A pair-wise comparison was undertaken to determine the nature of the significant difference, ie whether an upstream or downstream site (in Magela or Gulungul) is responding differently over time. This test indicated that of the exposed stream sites, Gulungul Creek upstream had the only significant difference from the before to after periods and hence this supported the MDS plot described above. As Gulungul Creek upstream is not influenced by minesite activity, the changes at this site must be associated with natural or non-mine-related conditions. This result was further supported by examination of the taxa abundance information, as discussed in section 6.2.3.2, 'Patterns in dominant macroinvertebrate taxa' of the main protocol.

Collectively, the graphical and statistical results provide good evidence that changes to water quality downstream of Ranger as a consequence of mining during the period 1994 to 2011 have not adversely affected macroinvertebrate communities.