Conceptual and statistical framework for a water quality component of an integrated report card for the Great Barrier Reef catchments

Petra Kuhnert[1] Rebecca Bartley[2], Erin Peterson[1], Matthew Browne[1], Bronwyn Harch[1],
Andy Steven[2], Mark Gibbs[1], Anne Henderson[2] and Vittorio Brando[2]

[1] CSIRO Mathematical and Information Sciences
[2] CSIRO Land and Water

Australian Government
Department of the Environment
and Water Resources

Funded through the
Marine and Tropical Sciences Research Facility
Project 3.7.7 Conceptual and statistical framework for the
water quality component of the Integrated Report Card

Copyright and Disclaimer

© 2007 CSIRO To the extent permitted by law, all rights are reserved and no part of this publication covered by copyright may be reproduced or copied in any form or by any means except with the written permission of CSIRO.

Important Disclaimer

CSIRO advises that the information contained in this publication comprises general statements based on scientific research. The reader is advised and needs to be aware that such information may be incomplete or unable to be used in any specific situation. No reliance or actions must therefore be made on that information without seeking prior expert professional, scientific and technical advice. To the extent permitted by law, CSIRO (including its employees and consultants) excludes all liability to any person for any consequences, including but not limited to all losses, damages, costs, expenses and any other compensation, arising directly or indirectly from using this publication (in part or in whole) and any information or material contained in it.

This report should be cited as:
Kuhnert, P., Bartley, R., Peterson, E., Browne, M., Harch, B., Steven, A., Gibbs, M., Henderson, A. and Brando, V. (2007) *Conceptual and statistical framework for a water quality component of an integrated report card for the Great Barrier Reef catchments.* Unpublished report to the Marine and Tropical Sciences Research Facility. Reef and Rainforest Research Centre Limited, Cairns (124pp.).

Made available online by the Reef and Rainforest Research Centre Limited for the Australian Government's Marine and Tropical Sciences Research Facility.

The Marine and Tropical Sciences Research Facility (MTSRF) is part of the Australian Government's Commonwealth Environment Research Facilities programme. The MTSRF is represented in North Queensland by the Reef and Rainforest Research Centre Limited (RRRC). The aim of the MTSRF is to ensure the health of North Queensland's public environmental assets – particularly the Great Barrier Reef and its catchments, tropical rainforests including the Wet Tropics World Heritage Area, and the Torres Strait – through the generation and transfer of world class research and knowledge sharing.

The views and opinions expressed in this publication are those of the authors and do not necessarily reflect those of the Australian Government or the Minister for the Environment and Water Resources.

While reasonable efforts have been made to ensure that the contents of this publication are factually correct, the Commonwealth does not accept responsibility for the accuracy or completeness of the contents, 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 contents of this publication.

This report is available for download from the Reef and Rainforest Research Centre website.
Visit http://www.rrrc.org.au/mtsrf/theme_3/project_3_7_7.html

Executive Summary

This report presents the final deliverable from the project titled *Conceptual and statistical framework for a water quality component of an integrated report card'* funded by the Marine and Tropical Sciences Research Facility (MTSRF; Project 3.7.7).

The key management driver of this, and a number of other MTSRF projects concerned with indicator development, is the requirement for state and federal government authorities and other stakeholders to provide robust assessments of the present 'state' or 'health' of regional ecosystems in the Great Barrier Reef (GBR) catchments and adjacent marine waters. An integrated report card format, that encompasses both biophysical and socio-economic factors, is an appropriate framework through which to deliver these assessments and meet a variety of reporting requirements. It is now well recognised that a 'report card' format for environmental reporting is very effective for community and stakeholder communication and engagement, and can be a key driver in galvanising community and political commitment and action.

Although a report card it needs to be understandable by all levels of the community, it also needs to be underpinned by sound, quality-assured science. In this regard this project was to develop approaches to address the statistical issues that arise from amalgamation or integration of sets of discrete indicators into a final score or assessment of the state of the system. In brief, the two main issues are (1) selecting, measuring and interpreting specific indicators that vary both in space and time, and (2) integrating a range of indicators in such a way as to provide a succinct but robust overview of the state of the system. Although there is considerable research and knowledge of the use of indicators to inform the management of ecological, social and economic systems, methods on how to best to integrate multiple disparate indicators remain poorly developed. Therefore the objective of this project was to (i) focus on statistical approaches aimed at ensuring that estimates of individual indicators are as robust as possible, and (ii) present methods that can be used to report on the overall state of the system by integrating estimates of individual indicators.

It was agreed at the outset, that this project was to focus on developing methods for a water quality report card. This was driven largely by the requirements of *Reef Water Quality Protection Plan* (RWQPP) and led to strong partner engagement with the Reef Water Quality Partnership.

This project developed the following:

1. *A description of the various phases required for indicator development and the associated statistical monitoring design approaches for data collection.*

This report outlines 15 guidelines that should be used to evaluate each indicator before it is finally accepted in a report card for the GBR region. Broadly, the criteria encompass conceptual relevance, feasibility of implementation, response variability and interpretation and utility

2. *A list of mature indicators for the GBR catchment, freshwater and marine regions, as well as (limited) available information on thresholds of concern, potential data sets and indicators requiring development.*

In consultation with MTSRF research scientists, a total of 53 mature indicators were identified from an original 82 potential indicators that are suitable for use within a GBR water quality report card. It is envisaged that this initial list of 53 mature indicators will be reduced further as more results from the 'domain' research programs become available over the next 4 years. The next stage is now for the 'domain' research groups within the MTSRF program to evaluate each of the mature indicators according to the 15 guidelines

presented in Section 3.1 of this report. This step will need to be done in conjunction with all of the organisations responsible for monitoring catchment, freshwater and marine condition and trend in the GBR catchments, and in some cases this may require specific pilot programs to be set up to rigorously test the indicator of interest. The outcome of this process will be a final list of indicators that is suitable for use in a GBR report card.

3. Approaches for indicator assessment, integration and visualisation.

This report highlights that the process of indicator evaluation, assessment, and integration requires an understanding of the natural variability, measurement error and co-variability of each indicator with respect to ‘reference condition’. For an indicator to be useful, it should have a reasonably high responsiveness to human disturbance, must be shown to be a suitable predictor of the ecological resource of interest, and the level of natural variability and measurement error associated with the indicator must be relatively minimal. To facilitate the visualisation of a large number of potential indicators, the GBR water quality ‘data wheel’ has been put forward as the preferred visualisation approach for the GBR report card. After evaluating a range of options in previous MTSRF reports (see Browne et al., 2007), the data wheel approach is recommend for a number of reasons including (i) it is flexible and new indicators, or indicator groups (e.g. socio-economic indicators) can be added at any stage; (ii) this type of visualisation option removes the need for different visualisation approaches for the scientific technical reports versus community report cards as different layers of the data wheel can be removed to reduce the complexity of visualisation where required; (iii) variations in the colour scheme can be used to demonstrate uncertainty in the data sets and potentially highlight where more data are needed for certain indicators; and (iv) this approach promotes consistency between the catchment, freshwater and marine approaches, rather than having different visualisation options as done for other report cards.

4. A discussion of preferred statistical approaches that enable appropriate spatially focused reporting against targets or threshold of concern.

Spatial partitioning is the process of subdividing a geographical space into one or more homogeneous sub-regions based on data that reflects differences in environmental and/or biological characteristics. It represents a critical component of any environmental monitoring program because it helps to ensure that an indicator’s response to anthropogenic impacts will be interpreted correctly. *It is important to note that throughout this report ‘spatial partitioning’ refers to appropriate spatially focused reporting against targets or threshold of concern not the actual reporting zones.* In this report we present two statistically and ecologically valid methodologies that can be used for freshwater and marine spatial partitioning. The spatial framework for the freshwater spatial partitioning successfully accommodates hydrologic connectivity and nested catchments, which are a unique characteristic of stream networks. A range of cluster analysis techniques, distance metrics, and methods for determining cluster structure were explored. Choice of the number of clusters was determined using two quantitative techniques: the average silhouette width and the Gap statistic for each clustering solution examined. A fuzzy cluster analysis was also constructed to investigate the likely group membership at the edges of the clusters identified. This allowed us to inspect the boundaries of each cluster visually and to determine where uncertainties in classification arise.

5. Demonstration of the conceptual and statistical framework for a water quality component of an integrated report card using the Tully/Murray freshwater and marine regions of the GBR.

The Tully-Murray catchment and marine zone were chosen as a pilot region to showcase methods for spatial partitioning, indicator assessment, and indicator integration for use in an integrated reporting framework. Two indicator assessment methods were demonstrated for the freshwater and marine environments. In the first, thresholds of concern were used to test for compliance at each of the sites and a proportion of the region that met the threshold was provided. The second approach was based on a statistical technique known as ‘bootstrapping’, which was used to examine the variability around the indicator within a spatially relevant neighbourhood. More importantly, an estimate of the bootstrap confidence interval was produced for the value recorded at the site and compliance was assessed accordingly. There are a number of ways of integrating assessments to form an overall evaluation of ecosystem health in terms of the proportion of sites that have exceeded a threshold. Some of the methods described here included averaging within regions, a weighted average within regions, and averaging across regions were demonstrated in the report using data from the Tully/Murray catchment and marine zone.

Arising from this research, seven general recommendations are made for the further development of an integrated report card for the GBR region. It is emphasised that the development of such a report card will take several years and will require the strong engagement and coordination of the various science, operational and end user stakeholder groups.

Recommendation 1: Indicator development and Incorporation into an IRC

Dedicated working groups within each of the IRC disciplinary classes (catchment, freshwater and marine) need to be set up to facilitate the further development and finalisation of the indicators to be used in the report card. The working groups will need to consist of the research scientists testing the indicators, the government bodies undertaking the routine monitoring, as well the statisticians that will be involved in the final analysis and integration of the data. Where appropriate regional body stakeholders may also be involved to provide community feedback on indicator development. Each of the working groups needs to complete Table 2, Table 3 and Table 4 using the 15 points outlined in Section 3.1.

The outcome of the working group process will hopefully lead to a final list of indicators that are suitable for use in an IRC for the GBR, as well as match the reporting requirements of each of the agencies involved. It is important to note that while we are suggesting at least 3 working groups, it is vital that the different components (catchment, freshwater and marine) remain connected and continue to focus on the linkages at the catchment to reef scale. The connectivity between the groups could be facilitated via the Reef Partnership.

Recommendation 2: Set up pilot monitoring programs

As a follow on to Recommendation 1, there is the need to set up pilot monitoring programs to assist with indicator testing and selection. This will be crucial for developing the thresholds of concern which are expected to be very different between wet and dry catchments (and associated marine monitoring zones) and between upland and estuarine conditions. The pilot programs may build on existing monitoring programs (e.g. AIMS Long term Marine Monitoring Program, or QNRW’s Stream and Estuary Assessment Program), however, the programs may need to be re-focused to address specific indicators. It is acknowledged that monitoring across the entire GBR is not necessarily practical, therefore a preliminary spatial partitioning exercise, similar to the one conducted in this report for the Tully catchment, would be useful for the entire GBR region so that areas with different biophysical conditions can be identified and targeted for sampling and establishing targets and/or thresholds of concern. The pilot programs could also be set up along side other research programs such as CSIRO’s National Research Flagship on Water for a Healthy Country so that existing research can help inform the indicator development and testing.

Recommendation 3: Data management

There is a need for a central inventory of data sets that are available on the indicators suggested in this report for the whole of the GBR region. This will include both freshwater and marine systems, as well as the catchment based data (e.g. fertiliser use rates). The freshwater water quality aspect of this is being undertaken by staff at QNRW, and will be crucial for understanding data availability, quality, geographic extent and future needs. Similarly, staff at GBRMPA and AIMS have access to most of the current marine data. Information such as fertiliser and pesticide application rates in cane systems are available for ~25% of the Queensland cane region (pers. comm. Tim Wrigley), however, funding is required to generate the data into a suitable format. The equivalent should be done for other catchment indicators and agricultural systems such as grazing and horticulture.

Having a list of available data is only the first of a series of data management steps that need to be undertaken. There is an enormous amount of data available in report format for all aspects of the GBR system, however, very little of the data described in these reports are in a suitable format for use in a report card (i.e. in spreadsheet or GIS format). A web based data capture, storage and distribution system that can be used by all parties working on issues to do with GBR water quality (that may be accessed via a Reef Partnership web page for example) would be extremely useful.

The data inventory and storage processes will require data sharing agreements allowing data to be shared between organisations. The lack of data sharing agreements was a major impediment to accessing data for use in this report. Sensitivity may be required when attempting to obtain data on land use management practice, and specific agreements need to made between industry and government on the way in which this data will be used within a report card framework.

Recommendation 4: Using modelled data within a report card framework

Due to the size of the GBR catchments, a number of indicator variables will need to be modelled in some form, rather than directly measured. This may require the use of sediment transportation models, such as SedNet and associated modelling platforms such as E2, as well as approaches such as remote sensing. It is important to communicate this intention with model developers so that more effort can go into improving the accuracy and uncertainty estimates on specific parameters. This will facilitate the use of these types of data within a report card framework.

Recommendation 5: Responsibility for the GBR integrated report card

Within the natural resource management realm of the GBR, there are a large number of different reporting requirements for State, Federal and Regional bodies (e.g. State of Environment and State of Catchment reports). The organisation finally responsible for the GBR IRC needs to be very clear about how the GBR IRC is different (or similar) to other existing reporting frameworks used in the GBR to avoid duplication as well to help form synergies between different government departments.

Recommendation 6: Ownership and leadership of an integrated report card

For the GBR IRC to become a reality requires ownership and leadership by a single organisation. The Reef Water Quality Partnership is the best placed organisation to do this, however, it needs to be equipped with the appropriate resources and it will also need to expand its focus away from just reef ‘water quality’ to encapsulate other issues in the GBR (e.g. Rainforest, Torres Strait, etc).

Recommendation 7: Communication strategy

The communication of a final report card product is an enormous task, and a communication strategy should be developed early for the successful transfer of the final report card to the wider community.

Contents

1. Introduction 1
1.1 About this report 1
1.2 Background and rationale for this work 2
1.3 Intended audience 2
1.4 Report outline 3
2. Conceptual approach: developing an integrated report card 4
2.1 Components of a report card 4
2.2 Progress towards conceptual models of catchment to reef health 6
3. Indicator Development 10
3.1 Selecting relevant indicators: use and abuse 10
3.1.1 Phase 1: Conceptual relevance 11
3.1.2 Phase 2: Feasibility of Implementation 11
3.1.3 Phase 3: Response Variability 12
3.1.4 Phase 4: Interpretation and Utility 13
3.2 Establishing indicator validity and efficacy 14
3.2.1 Baseline data 14
3.2.2 Indicator evaluation 15
3.3 Indicator assessment and integration 22
3.3.1 Managing to a specific objective 22
3.4 Conclusions 28
4. Mature indicators for a Great Barrier Reef Integrated Report Card 29
4.1 Pressure-Vector-Response Framework for indicator selection 29
4.2 Data integration and visualisation options 37
5. Approaches to spatial partitioning for assessing compliance against targets and thresholds of concern 43
5.1 Introduction 43
5.2 Methods 45
5.2.1 Freshwater Spatial Partitioning 45
5.2.2 Marine Spatial Partitioning 49
5.2.3 Statistical Methods for Spatial Partitioning 49
5.3 Discussion 51
5.3.1 Application to Other Catchments in the GBR 52
5.3.2 Extensions and Future Work 52
6. Case Study: the Tully/Murray catchment 54
6.1 Background 54
6.1.1 Catchment and Freshwater Data 56
6.1.2 Marine Data 57
6.2 Spatial partitioning 58
6.2.1 Freshwater 58
6.2.2 Marine 66
6.3 Indicator Assessment and Integration 70
6.3.1 Methods for Assessment 70
6.3.2 Methods for Integration 77
6.4 Examples of data visualisation for the Tully catchment 83
6.5 Discussion 88

7. Summary and recommendations from this report 90
8. References 93

List of Tables

Table 1: Guide to the source of data that may be used for the GBR IRC. 31
Table 2: Table of potential mature pressure indicators for catchment condition 32
Table 3: Table of potential mature pressure indicators for freshwater condition 33
Table 4: List of potential mature response indicators for the Marine environment 35
Table 5: Example of how the individual indicator measurement scores (either as a score from 1-5, percentage compliance or other measure) are averaged up to cluster scores, groups scores and finally an Index 40
Table 6: Key research projects being conducted in the Tully as part of the WQIP and WfHC projects 55
Table 7: Summary of variables available for analysis in the Tully-Murray marine region. 58
Table 8. Mean length and area measurements for the freshwater spatial partitioning spatial framework. 59
Table 9. Thirty-two variables used as input into clustering algorithms. 62
Table 10: Thresholds of concern chosen for freshwater indicators within each of the defined regions shown in Figure 21 73
Table 11: Thresholds of concern chosen for marine indicators within each of the defined regions 77
Table 12: Raw Assessments of water quality for the Tully-Murray marine zone 81
Table 13: Bootstrap Assessments of water quality for the Tully-Murray marine zone using a neighbourhood consisting of a 7.5km radius. 81
Table 14: Raw Assessments for the Tully-Murray freshwater region for (a) habitat condition and (b) land use management 82
Table 15: Bootstrap assessments for the Tully-Murray freshwater region for (a) habitat condition and (b) land use management. 82
Table 16: Integrated raw assessments of habitat condition (riparian vegetation and percent macrophyte bed cover) and land use management (TSS). 83
Table 17: Integrated bootstrapped assessments of habitat condition (riparian vegetation and percent macrophyte bed cover) and land use management (TSS). 83
Table 18: List of potential catchment indicators for the IRC. 103
Table 19: List of potential freshwater indicators for the IRC. 104
Table 20: List of potential marine indicators for the IRC. 106
Table 21: Source, spatial scale, and pre-processing methods. 108
Table 22: Geologic mapping data reclass table. 109
Table 23: ASRIS soil texture reclass table. 110
Table 24: Reach, RCA (reach contributing area), and watershed variables used in the freshwater regionalisation. 113

List of Figures

Figure 1: A conceptual diagram showing the relationship of an Integrated Report Card (IRC) Framework to other key groups and programs........................4

Figure 2: Phases outlining the development of an integrated reporting framework. 6

Figure 3: DRAFT conceptual model providing examples of regional catchment and reef threats...7

Figure 4: Example of a DRAFT conceptual model describing the mechanisms of sediment generation and delivery from the GBR catchments...........................8

Figure 5: Example of a DRAFT conceptual model differentiating between the mechanisms of sediment generation and delivery from GBR catchments with poor versus good land management...8

Figure 6: Example of a DRAFT conceptual model describing the mechanisms of sediment generation and delivery in a dry tropical GBR catchment 9

Figure 7: Various relationships between a disturbance gradient (x-axis) and a hypothetical indicator score (y-axis) 18

Figure 8: Hypothetical example comparing indicator scores at two sites with the distribution at reference sites conditioned on a single environmental predictor. 21

Figure 9: Conceptual diagram of the relationship between a referential approach and management goals. 22

Figure 10: Change in exceedence factor score as the ratio of test score to objective increases. 26

Figure 11: Transformation of indicator point-measurements so as to represent exceedence over a time-space volume. 27

Figure 12: Pressure-Vector-Response Framework being used for indicator selection. 29

Figure 13: Example of GBR water quality data wheel approach that would be suitable for use in an integrated GBR report card 37

Figure 14: Example of the relationship used to build the index presented in the GBR water quality data wheel in Figure 13. 39

Figure 15: Examples of the data wheel with (A) the outer indicator layers removed and (B) showing only the Index layer..41

Figure 16: Example showing how different shaded colours could be used to demonstrate variation in data quality or uncertainty 42

Figure 17: The data wheel could also be used to present the results different catchments or regions within a catchment management boundary such as the Wet Tropics. 42

Figure 18: The three scales of freshwater spatial partitioning attributes: the reach (1a), the RCA (1b), and the catchment (1c)..48

Figure 19: Map of the Tully-Murray river catchments (indicated by the dotted black and yellow lines) and marine zone (indicated by the solid orange line)...57

Figure 20. A comparison of average silhouette widths produced by the FANNY algorithm using a Gower dissimilarity metric 63

Figure 21. Hard 4 cluster regions for the Tully-Murray watershed based on fuzzy clustering using Gower's dissimilarity metric. 64

Figure 22. The percent increase in mean square error (MSE) for each variable if it were omitted from the analysis 65

Figure 23. Uncertainty in region assignments found in the fuzzy four cluster solution. 66

Figure 24: Variables input into the spatial partitioning for the Tully-Murray marine extent. .. 67

Figure 25: Indicator variables used in the assessment for the Tully-Murray marine extent .. 68

Figure 26: Figures showing the (a) average silhouette width for each variable included in the cluster analysis and (b) the marine spatial partitioning with two distinct clusters representing a shallow and a deep zone. 70

Figure 27: Maps showing (a) the raw assessment and (b) the bootstrapped assessment for riparian vegetation for each region 74

Figure 28: Maps showing (a) the raw assessment and (b) the bootstrapped assessment for TSS for each region 75

Figure 29: Maps showing (a) the raw assessment and (b) the bootstrapped assessment for percent macrophyte bed cover for each region 76

Figure 30: Maps showing (a) the raw assessment and (b) the bootstrapped assessment for chlorophyll-a for each region 78

Figure 31: Maps showing (a) the raw assessment and (b) the bootstrapped assessment for vertical light attenuation for each region 79

Figure 32: Maps showing (a) the raw assessment and (b) the bootstrapped assessment for number of non-algal particles for each region 79

Figure 33: Example of how the data wheel approach can be linked to geographical maps of the catchment showing the different ratings for the catchment, freshwater and marine zones. 84

Figure 34: An alternative to the figure above showing how the data wheel can be used to show the score in different 'regions' or sub-catchments of a single catchment 85

Figure 35: Example of how data can be visualised at different levels within the data wheel depending on the audience 86

Figure 36: Example layout of a report card for the Tully catchment 87

1. Introduction

1.1 About this report

This report presents a conceptual and statistical framework for the development of an Integrated Report Card for the Great Barrier Reef (GBR) region in North Queensland, Australia. Integrated catchment management groups, as well as State and Federal Government bodies, now have an obligation to report on condition and trend in their catchments and marine waters as part of the response to the Reef Water Quality Protection Plan (Queensland Government, 2003). In this context there is a need for a consistent approach to reporting on catchment and water quality condition at each of these levels.

To facilitate the development of a Integrated Report Card (IRC) for the GBR region, the Marine and Tropical Sciences Research Foundation (MTSRF) funded this project (Project 3.7.7: *Conceptual and statistical framework for the water quality component of an integrated report card)*, with the broad objective of developing a scientifically robust framework to support the production of report card(s) that integrate biophysical and socio-economic data from indicators that represent the pressures, vectors and responses in tropical aquatic landscapes of the GBR and Torres Strait Regions. This will help support informed adaptive management of these landscapes.

The development of an Integrated Report Card (IRC) is also one of the primary tools for the integration of information resulting from the MTSRF program, as well as data from other regional, state and federal initiatives. The IRC will actively involve some 23 MTSRF projects, as well as have strong engagement with end users from a number of regional, state and federal natural resources management agencies and industry. The framework needs to be broad enough to accommodate inputs from a range of programs, however, it has been agreed that ***water quality*** will be the initial focus of the IRC program, and hence is the focus of this report. The priority of other programs will be determined in future years.

A wide range of resource condition and trend assessment frameworks are currently available within Australia and internationally. A summary of some of these can be found in (Auricht, 2004) and were reviewed and summarised in a previous milestone report for this project (see Browne et al., 2007). This report represents our final milestone and includes:

- A description of the various phases required for indicator development and the associated statistical and monitoring design approaches for data collection;
- A list of *mature* indicators for the GBR catchment, freshwater and marine regions, as well as (limited) available information on thresholds of concern, potential data sets and indicators requiring development;
- Approaches for indicator assessment, integration and visualisation are provided;
- Demonstration of a recommended method of data integration and visualisation;
- A discussion of a preferred statistical approaches that enable appropriate spatially focused reporting against targets or threshold of concern; and
- Demonstration of the conceptual and statistical framework for a water quality component of an integrated report card using the Tully/Murray freshwater and marine regions of the GBR.

1.2 Background and rationale for this work

There has been significant land management change in catchments draining to the Great Barrier Reef (GBR) since European settlement in 1850. As a result of this change, sediment, nutrient and contaminant loads from these catchments have increased (McCulloch et al., 2003; Neil et al., 2002), and recent research suggests that these increased loads are now having a detrimental effect on coral reef systems (Fabricius, 2005; Fabricius et al., 2005). In an effort to reduce this impact, the Reef Water Quality Protection Plan (RWQPP) was signed by the Queensland and Federal Governments in October 2003 with the overall aim of 'halting and reversing the decline in water quality entering the reef within 10 years'.

To facilitate the implementation of the RWQPP, a number of environmental monitoring projects have been initiated including a marine monitoring program (Haynes et al., 2005), estuarine monitoring (Haynes et al., 2005) and an assessment of the condition of rivers and streams (Negus and Marsh, 2006a). At present, however, no framework exists that allows for the integration and evaluation of the individual monitoring programs so that catchment and reef health can be systematically evaluated and reported across the entire GBR region (freshwater, estuarine and marine). In addition, many of the agencies responsible for collecting and analysing this data have often struggled with the issues and methods required for the appropriate integration of information across various indicators. As a result, there is a strong need for statistical methods that focus on the spatio-temporal design, collection, statistical analysis, integration and visualisation of the data which is generated from these programs. The information then needs to be put into a form that is appropriate for management authorities and the community to interpret (Vos et al., 2000). Examples of applicable statistical methods include (but are not restricted to); spatially balanced monitoring design optimising resources against information reported, validating indicators with respect to the human disturbance gradient, helping determine appropriate critical reference or target levels given the distribution of minimally impacted to highly impacted sites, multivariate clustering of sites for establishing appropriate spatial scale for reporting against targets (and/or thresholds of concern), modelling of natural and anthropogenic contributions to variation in indicators, methods for aggregating indicator information across space and time, and incorporating estimates of uncertainty into integrated water quality assessments .To date, there is no single publication that captures this information.

1.3 Intended audience

The intention of this report is to provide a general framework and demonstration of a range of robust statistical methodologies that may be used to develop an IRC for the GBR region, which will be lead by the Reef Partnership. This report does not provide all of the answers, nor does it provide a final report card product. Instead, it provides a range of technical options that may be used to develop a report card depending on the future investment in data collection, analysis and reporting within this program. Therefore the primary audience for this report is:

- The Reef Partnership
- Other MTSRF research programs involved in collecting and testing indicators to be included in a GBR IRC
- State and Federal agencies (e.g. QNRW and GBRMPA) responsible for routine monitoring and data collection.

This project is providing the statistical framework around the development of a report card and relies implicitly on inputs from other MTSRF projects (as well as other existing state and federally funded projects) in the development and trialling of an IRC.

1.4 Report outline

This report is presented in 6 main sections:

- Section 2 provides a re-cap on the conceptual framework for the development of report cards that was previously described in Browne et al. (2007);
- A description of a range of statistical methods that may be used for indicator selection and aspects for consideration in monitoring design are presented in Section 3;
- Section 4 then presents a list of mature indicators for the GBR region based on a Pressure-Vectors-Response framework, as well as a recommended approach for data integration and visualisation.
- Section 5 outlines a range of statistical approaches that enable appropriate spatially focused reporting against targets and/or thresholds of concern;
- Then in Section 6 these statistical approaches are demonstrated for the Tully/Murray catchment; and
- Section 7 provides final recommendations for on-going report card development.

2. Conceptual approach: developing an integrated report card

2.1 Components of a report card

Integrated ecosystem health report cards are increasingly being used around the world to define and measure progress towards environmental sustainability. Report cards can be an effective community and stakeholder communication and engagement tool and when used effectively, can be a key driver in galvanising community and political commitment and action. There are numerous examples of reporting frameworks which were summarised in an earlier MTSRF milestone report (see Browne et al., 2007).

Figure 1 illustrates the relationship of an IRC to other key groups and programs and highlights the importance of community and stakeholder engagement. Past experience has shown that government organisations are willing to devote time and resources to develop and implement a report card because of the central role that they are seen to play both in informing management decisions and in providing an interface between research scientists, environmental managers, stakeholders and the community.

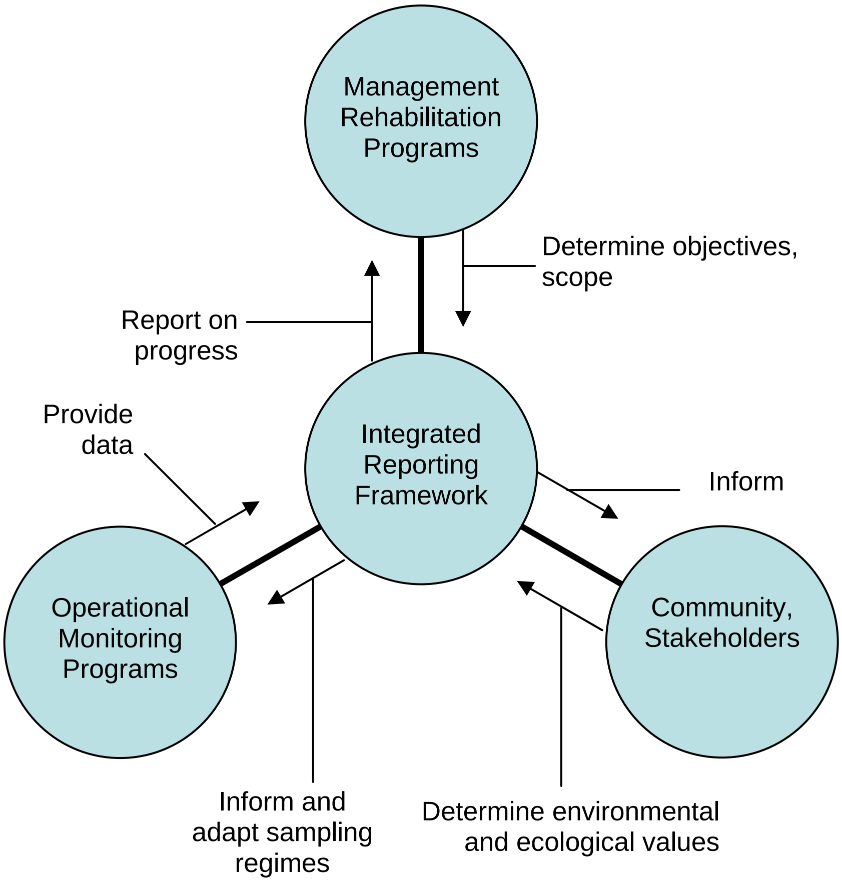
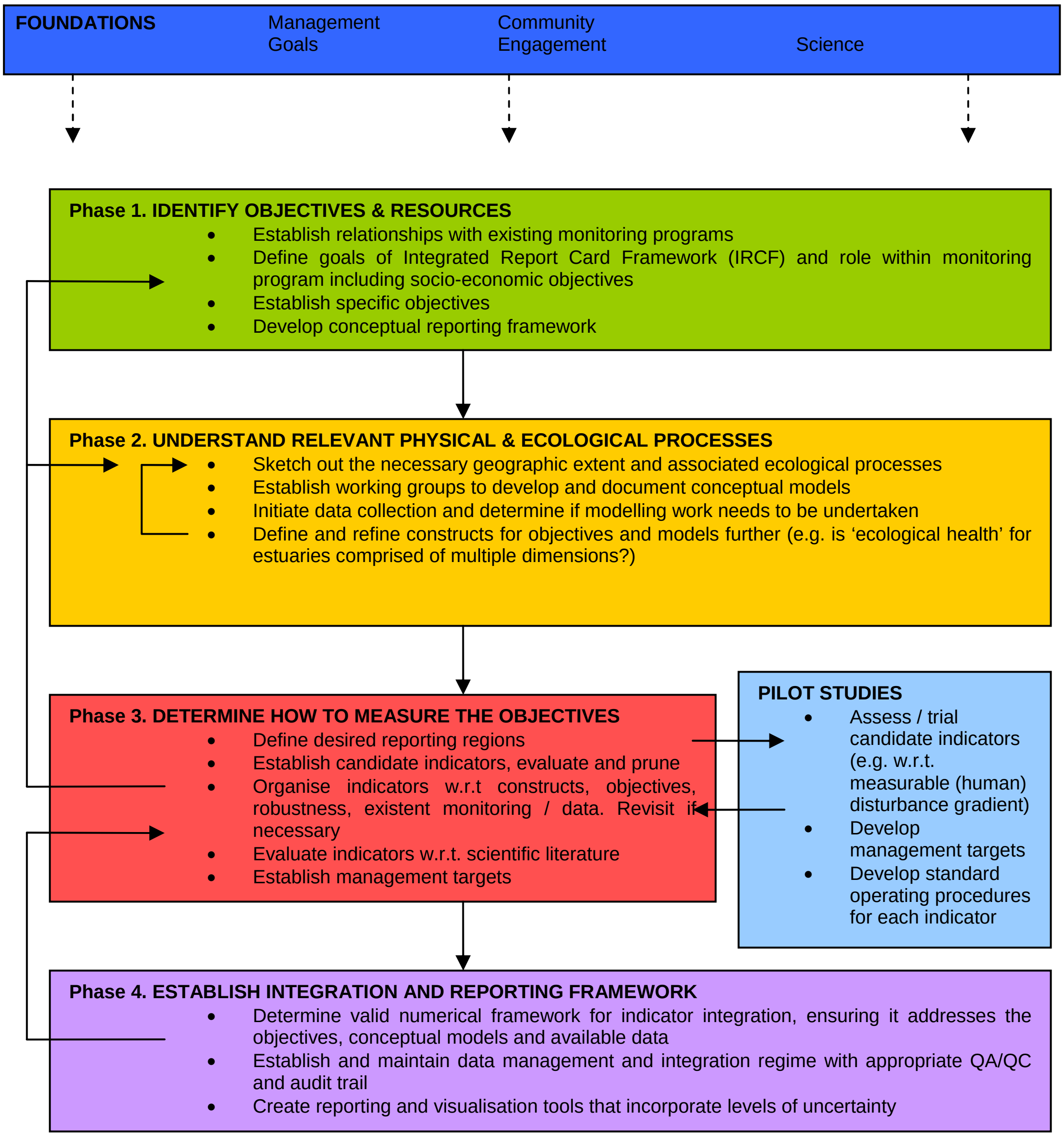


Figure 1: A conceptual diagram showing the relationship of an Integrated Report Card (IRC) Framework to other key groups and programs.

Figure 1 also serves to illustrate the relationships between the IRC, which we are primarily concerned with in this report, and the operational monitoring programs which collect the required data. The focus of a monitoring program is the design and implementation of an effective sampling regime in order to maximise the information generated, whereas an IRC is primarily concerned with the valid assessment, integration and presentation of a range of heterogenous data sources collected at various spatial and temporal scales. Furthermore, an integrated reporting framework will typically make use of existing monitoring data and/or will collaborate with concurrent efforts to adapt the sampling process as systems understanding is improved.

A general framework for report card development is captured in Figure 2, which outlines four sequential phases of the reporting process. The process is characterised by the close collaboration of research scientists and decision-makers (Harwell et al., 1999) and at each phase there is scope for refinement through feedback loops. Note, this framework for report card development incorporates detail from Harwell et al. (1999), Smith et al., (2001), Bunn and Smith (2007) and Negus & Marsh (2006). The tasks outlined within each phase represent a non-exhaustive list. This is intended to be indicative of the tasks that need to be accomplished for most IRCs. The phases in Figure 2 are relatively intuitive but we outline them briefly below for clarity.

1. **Define the scope and objectives** of reporting, and establish what existing monitoring resources are available.
2. **Understand the system**, drawing together and documenting all relevant theoretical and empirical knowledge. Conceptual frameworks and conceptual models play an important role here.
3. **Establish a measurement framework** that will address the constructs defined in Phase 1 in terms of the systems understanding identified in Phase 2.
4. **Establish an integration and reporting framework** that will integrate and present the data generated in Phase 3 in a valid spatio-temporal manner.



FOUNDATIONS Management Goals · Community Engagement · Science

Phase 1. IDENTIFY OBJECTIVES & RESOURCES

- Establish relationships with existing monitoring programs
- Define goals of Integrated Report Card Framework (IRCF) and role within monitoring program including socio-economic objectives
- Establish specific objectives
- Develop conceptual reporting framework

Phase 2. UNDERSTAND RELEVANT PHYSICAL & ECOLOGICAL PROCESSES

- Sketch out the necessary geographic extent and associated ecological processes
- Establish working groups to develop and document conceptual models
- Initiate data collection and determine if modelling work needs to be undertaken
- Define and refine constructs for objectives and models further (e.g. is ‘ecological health’ for estuaries comprised of multiple dimensions?)

Phase 3. DETERMINE HOW TO MEASURE THE OBJECTIVES

- Define desired reporting regions
- Establish candidate indicators, evaluate and prune
- Organise indicators w.r.t constructs, objectives, robustness, existent monitoring / data. Revisit if necessary
- Evaluate indicators w.r.t. scientific literature
- Establish management targets

PILOT STUDIES

- Assess / trial candidate indicators (e.g. w.r.t. measurable (human) disturbance gradient)
- Develop management targets
- Develop standard operating procedures for each indicator

Phase 4. ESTABLISH INTEGRATION AND REPORTING FRAMEWORK

- Determine valid numerical framework for indicator integration, ensuring it addresses the objectives, conceptual models and available data
- Establish and maintain data management and integration regime with appropriate QA/QC and audit trail
- Create reporting and visualisation tools that incorporate levels of uncertainty

Figure 2: Phases outlining the development of an integrated reporting framework.

2.2 Progress towards conceptual models of catchment to reef health

As described in Figure 2, the second phase in developing an IRC is to have a conceptual understanding of the system on which the report card is to be developed. Conceptual models play a critical role and aim to identify components, processes, factors and linkages and consequently are the foundation for the selection, assessment and integration of valid indicators. These types of models highlight how healthy ecosystems function and how processes respond to disturbance. They can also be useful for identifying possible management actions for rehabilitation (Bunn & Smith, 2007). Development of conceptual models for the GBR was not part of this project, however, conceptual model development is being undertaken by the Reef Partnership, and a range of models have been developed (presented in Figure 3 to Figure 6). These models were used as part of the process to identify indicators for the report card as described in Sections 3 and 4.

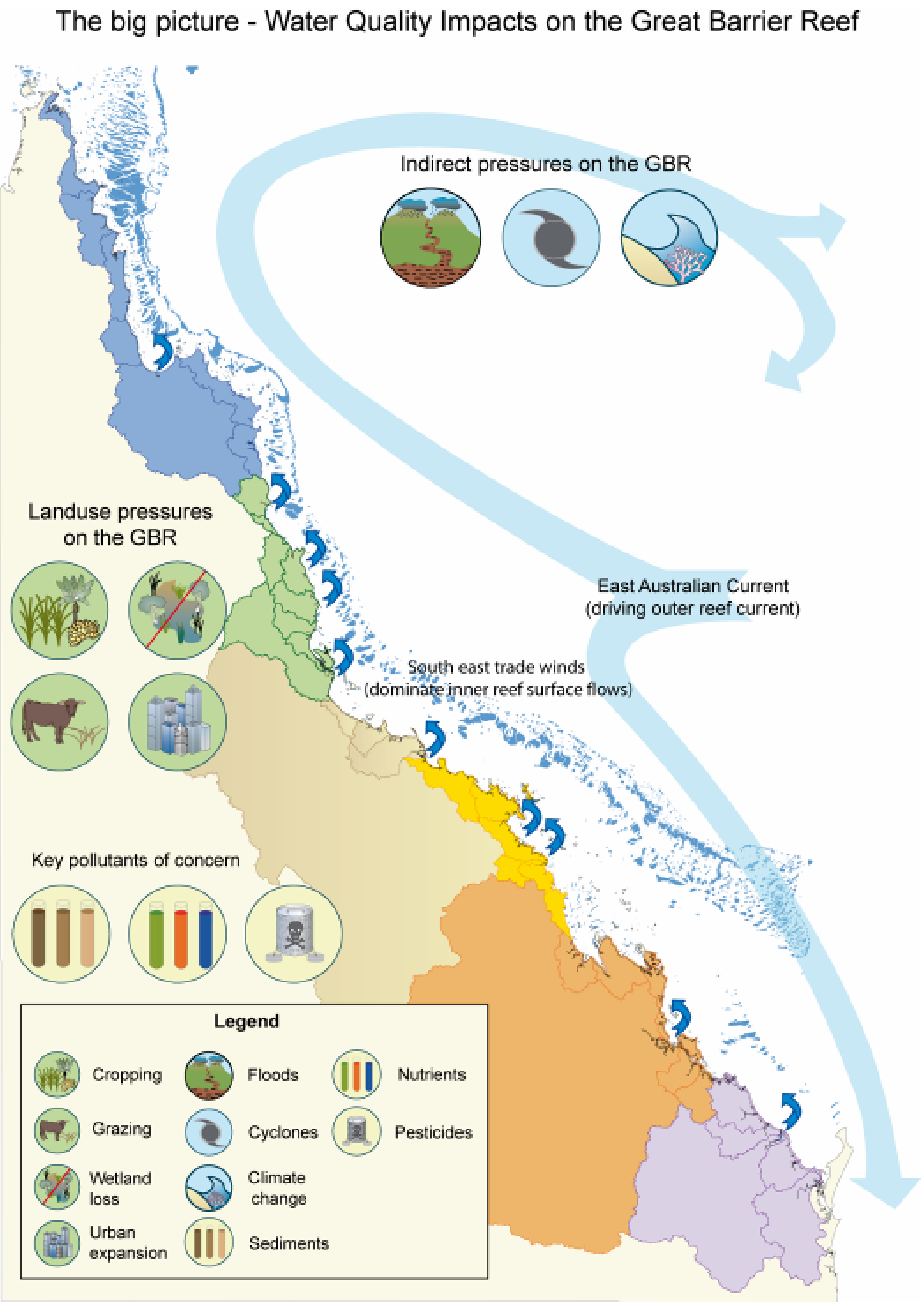


Figure 3: DRAFT conceptual model providing examples of regional catchment and reef threats (Image developed by Joelle Prange, GBRMPA, for the Reef Partnership).

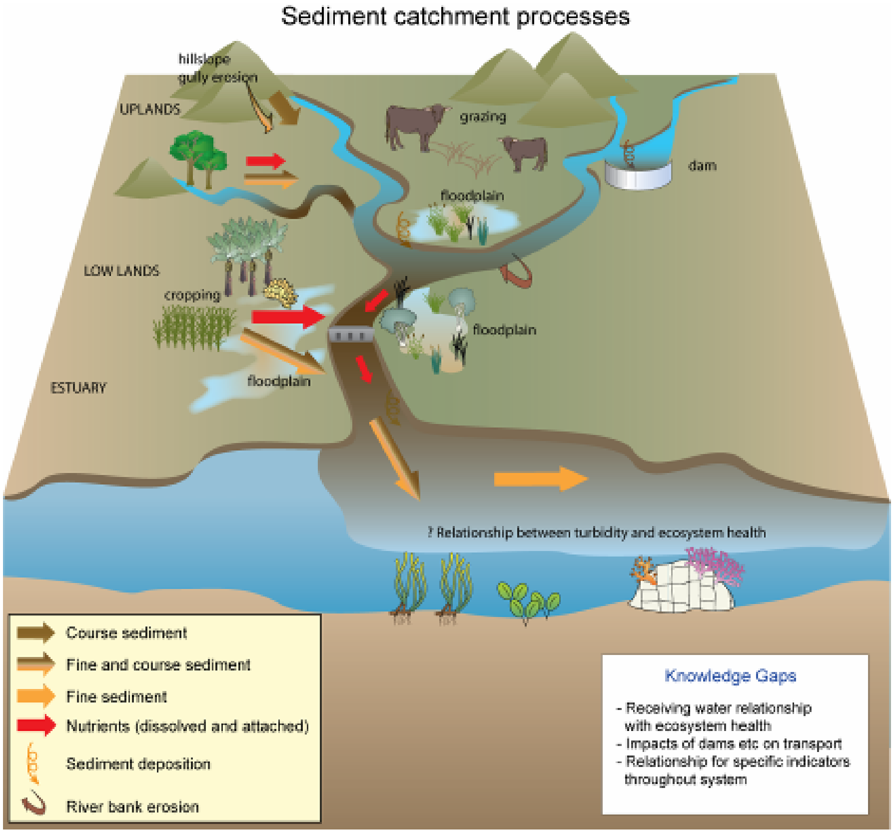


Figure 4: Example of a DRAFT conceptual model describing the mechanisms of sediment generation and delivery from the GBR catchments (Image developed by Joelle Prange, GBRMPA, for the Reef Partnership).

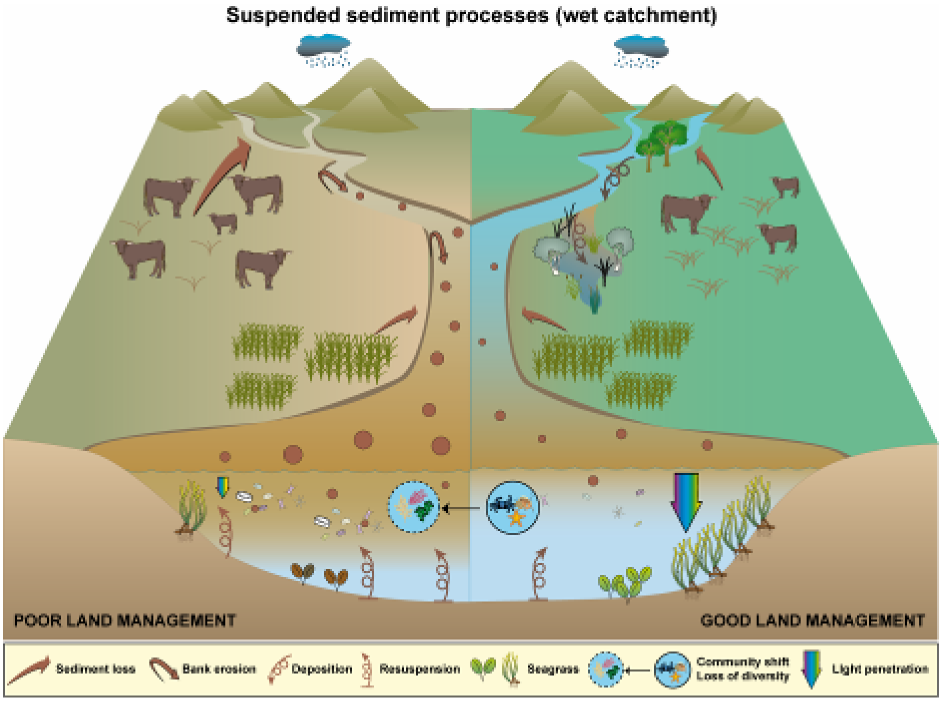


Figure 5: Example of a DRAFT conceptual model differentiating between the mechanisms of sediment generation and delivery from GBR catchments with poor versus good land management (Image developed by Joelle Prange, GBRMPA, for the Reef Partnership).

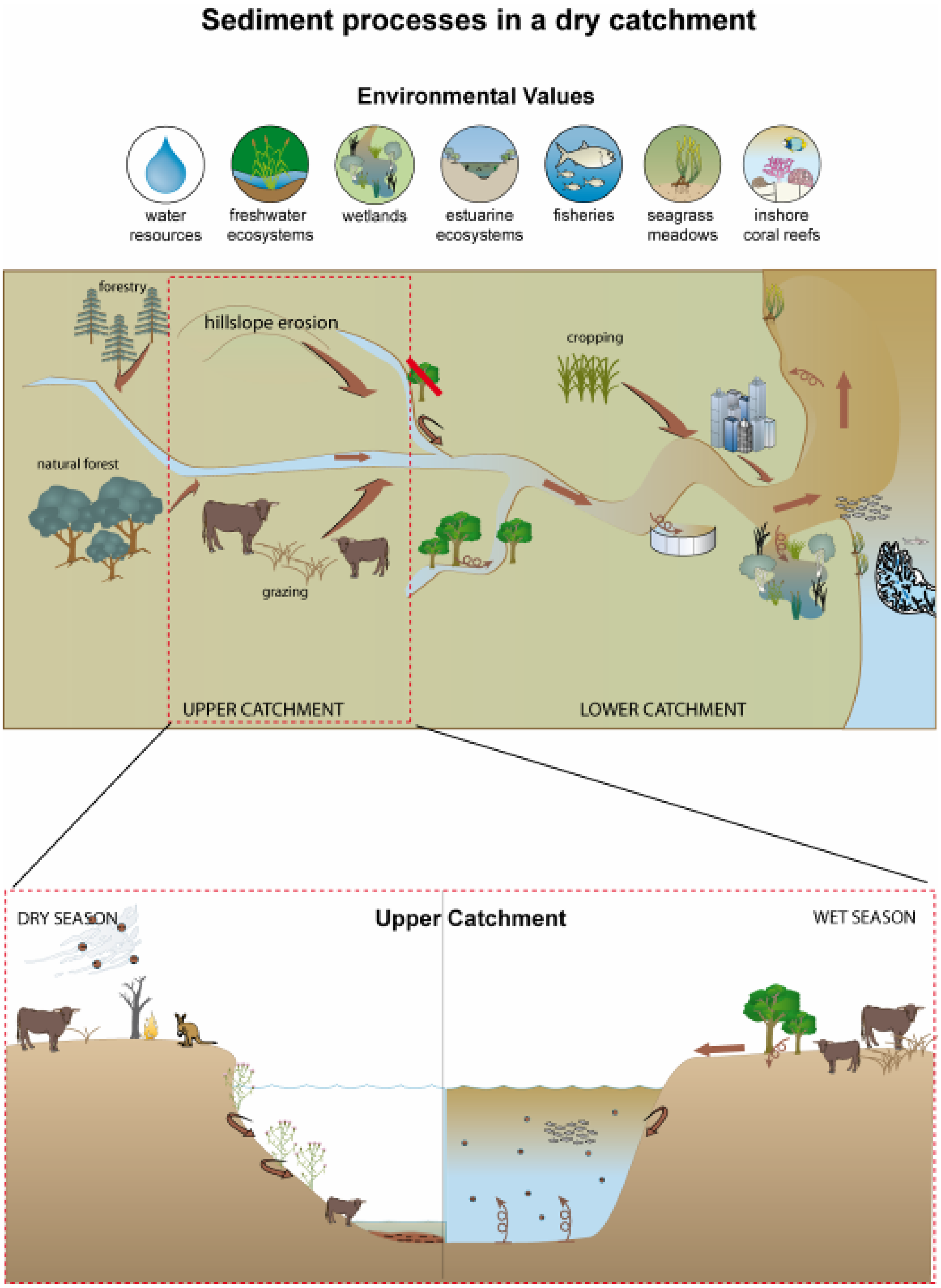


Figure 6: Example of a DRAFT conceptual model describing the mechanisms of sediment generation and delivery in a dry tropical GBR catchment (Image developed by Joelle Prange, GBRMPA, for the Reef Partnership).

3. Indicator Development

Coupled social-ecological systems are clearly complicated, and in fact are typically regarded as complex adaptive systems. The components of these systems are highly organised and feature dynamic structures and processes that often do not lend themselves to being reduced into tractable statistical and/or mathematical models. The inherent complexity of these coupled social-ecological systems means that it is particularly difficult to understand how these systems are behaving in a general sense, and to understand how human-induced impacts to particular components of the system are propagated through other parts of the system. Environmental managers often attempt to modify the behaviour of particular stakeholders in the human parts of the system in order to minimise impacts to the ecological parts of the system, and this relies on a qualitative, and where possible quantitative, understanding of the overall social-ecological system, and properties of particular sub-systems. However, the inherent complexity of the combined system means that information for management is always incomplete, and usually highly limited. Thus, a key objective in environmental management is to be able to direct limited environmental monitoring effort in order to detect and understand enough about the behaviour of at least the ecological components of the coupled system in order to underpin good decision making processes.

Reducing an ecological system to a set of variables that indicate the behaviour of key processes occurring within the system is difficult, and made more difficult in ecological systems given the present state of technology for measuring flows and transformations of mass and energy (the key indicators). Interestingly, in economic systems it can be more straightforward to measure key indicators (the flow of currency) as no special instrumentation are required. Invariably, the indicators that we end up monitoring in ecological systems are often not the most desirable variables for understanding how the system is working (e.g. the fluxes and flows of energy and material rather than simply abundances and concentrations of prominent organisms). Therefore, even when we have robust measurements of ecological indicators, the indicators themselves may not necessarily be providing much insight into how the ecological system is behaving. Therefore, given that it is almost a necessity for us to try and encapsulate very complicated dynamics into a small number of ‘indicators’ in order to manage coupled social-ecological systems, then we must be clearly aware of exactly what these indicators are telling us, and of equal importance how good our estimates of the true values of each parameter are.

3.1 Selecting relevant indicators: use and abuse

A tremendous amount of thinking has been directed towards the use and abuse of ecological indicators (e.g. there are multiple texts and a scientific journal dedicated to the use of ecological indicators). The abuse in particular, prompted the USA EPA to release a report on what constitutes a good ecological indicator and this is often regarded as a benchmark document (Jackson et al. 2000, p1-1). Hence, it is worthwhile reminding ourselves of the key results from this work. The EPA report identifies four key phases to indicator development, as follows:

- Phase 1: Conceptual relevance
- Phase 2: Feasibility of Implementation
- Phase 3: Response Variability
- Phase 4: Interpretation and Utility

The relevance and requirements of these phases will now be briefly discussed.

3.1.1 Phase 1: Conceptual relevance

Key outcome statement:

The indicator must provide information that is relevant to societal concerns about ecological condition. The indicator should clearly pertain to one or more identified assessment questions. These, in turn, should be germane to a management decision and clearly relate to ecological components or processes deemed important in ecological condition. Often, the selection of a relevant indicator is obvious from the assessment question and from professional judgement. However, a conceptual model can be helpful to demonstrate and ensure an indicator's ecological relevance, particularly if the indicator measurement is a surrogate for measurement of the valued resource. This phase of indicator evaluation does not require field activities or data analysis. Later in the process, however, information may come to light that necessitates re-evaluation of the conceptual relevance, and possibly indicator modification or replacement. Likewise, new information may lead to a refinement of the assessment question (Jackson et al. 2000, page 1-1).

Jackson et al. (2000) also suggest that guiding principles for this Phase should be:

Guideline 1: Relevance to the Assessment
Guideline 2: Relevance to Ecological Function

One of the key tensions in indicator development is between resource managers aspirations of having robust information supporting defensible decision making processes, and researchers aspirations of learning more about natural systems - often with the long-term objective of developing better management processes. As argued by researchers, often management-lead indicator development aims to address short-term management questions at the possible expense of a high opportunity cost of not learning about the system that may enable better long-term management. This often manifests itself into large numbers of short-term monitoring programs that feature different sampling protocols. This can lead to a lack of consistency between programs and a lost opportunity to gain a long-term understanding of system behaviour. It can also be argued that natural systems are also changing at a faster pace than our ability to implement management initiatives. By contrast, resource managers can equally well argue that there is little point attempting to gain long term understanding if immediate management issues cannot be adequately resolved. **The key message here is that the purpose of the indicator program must be very clear.** If it is indeed to feed directly into management processes (Guideline 1) than it is perhaps worthwhile developing alternative longer-term programs aimed at gaining a more fundamental understanding of system behaviour. However, having said this, it is also clear that any indicator must be ecologically relevant (Guideline 2) although not necessarily the 'most' ecologically relevant as financial, technological and logistical constraints may act to override the level of relevance.

3.1.2 Phase 2: Feasibility of Implementation

Key outcome statement:

Adapting an indicator for use in a large or long-term monitoring program must be feasible and practical. Methods, logistics, cost, and other issues of implementation should be evaluated before routine data collection begins. Sampling, processing and analytical methods should be documented for all measurements that comprise the indicator. The logistics and costs associated with training, travel, equipment and field and laboratory work should be evaluated and plans for information management and quality assurance developed. Additionally, logistics and costs associated with the analysis of integration of indicator data is needed.
Jackson et al. (2000) also suggest that guiding principles for this Phase should be:

Guideline 3: Data collection Methods
Guideline 4: Logistics
Guideline 5: Information Management
Guideline 6: Quality Assurance
Guideline 7: Monetary Costs

Methods used for measuring indicators must be widely-accepted and their limitations understood. The use of novel methods for measuring indicators has the potential for compromising both indicator programs, and the environmental management process. Efforts should also be made to avoid further impacting the environment when undertaking measurements.

Clearly any effective indicator program must be financially and logistical tractable. A large number of ambitious indicator programs have failed as a result of a lack of long-term commitment of resources. Furthermore, a pilot program is sometimes required in order to understand the variability in the particular parameters measured, not only so that robust sampling schemes can be developed, but also to assess the level of logistical resources required. Extra resources are often directed towards initial monitoring that later become unavailable once the novelty of the program dissipates and new management issues arise. Environmental management is also exposed to the reactive side of political and institutional processes that often undervalue longer-term programs in favour of reactive responses to short-term issues. Therefore the temporal extent of monitoring required depends on the particular management questions, relevant time frames, and the program should also be informed by the possibility that scarce monitoring resources may well be redirected elsewhere at some stage. For example, if an indicator program cannot be effective unless a time series of a decade or more is required, then such a program cannot be expected to robustly underpin a management strategy that operates over a shorter timescale. This highlights a common problem in modern environmental management: that the rate of change in coupled social-ecological systems is often faster than our ability to measure and understand these changes, let alone effectively manage them.

3.1.3 Phase 3: Response Variability

Key outcome statement:

It is essential to understand the components of variability in indicator results to distinguish extraneous factors from a true environmental signal. Total variability includes both measurement error introduced during field and laboratory activities and natural variation, which includes influences of stressors. Natural variability can include temporal (within season and across years) and spatial (across sites and catchments) components. Depending on the context of the assessment question, some of these sources must be isolated and quantified in order to interpret indicator responses correctly. It may not be necessary or appropriate to address all components of natural variability. Ultimately, an indicator must exhibit significantly different responses at distinct points along a condition gradient. If an indicator is composed of multiple measurements, variability should be evaluated for each measurement as well as for the resulting indicator.

Jackson et al. (2000 p. 1-3) also suggest that guiding principles for this Phase should be:

Guideline 8: Estimation of Measurement Error
Guideline 9: Temporal Variability - Within the Season
Guideline 10: Temporal Variability - Across Years
Guideline 11: Spatial Variability
Guideline 12: Discriminatory Ability

Guideline 8 addresses a key aspect that is often overlooked in indicator programs; the propagation of measurement errors into the final estimate of the indicator. Management decisions based on ecological indicators must be risk-informed. Guidelines 9, 10, and 11 address generally-well understood variability issues that must be taken into account in sampling design processes- a process that often requires a pilot survey as a result of lack of prior information on variability. Guideline 12 is one of the most commonly overlooked and misunderstood criteria in the development of ecological indicator programs. Stakeholders in the resource management process need to have decided apriori how different indicator estimate values translate into management actions to ensure that the resolution of the indicator estimate program is high enough to provide contrast. This does not necessarily mean that formal decision rules and trigger levels need to be specified in advance, but rather that stakeholders are aware of the performance criteria of the program. For example if a management objective is to reduce the annual mean concentration of chlorophyll a in a marine water-column by say 30%, then the indicator program must be able to firstly detect such a 30% change (sometimes not easy given the variability in chlorophyll a, e.g. De'ath, (2007)). Secondly if it requires say an 80% reduction in the outflow of terrestrially-derived nutrients to achieve this outcome (in some cases anything less would not invoke a measurable response in both the ecosystem and/or the indicator), than it may be more appropriate to attempt to measure the outflows from large nutrient point sources as using the chlorophyll a indicator may not be robust enough to recognise even large changes in nutrient inputs. In other words, ideally both an idea of how well the indicator tracks ecosystem response (not easy to achieve), in addition to knowledge of whether the indicator as measured actually has the ability to detect management level changes is required.

3.1.4 Phase 4: Interpretation and Utility

Key outcome statement:

A useful ecological indicator must produce results that are clearly understood and accepted by scientists, policy makers, and the public. The statistical limitations of the indicator's performance should be documented. A range of values should be established that defines ecological condition as acceptable, marginal, and unacceptable in relation to indicator results. Finally, the presentation of indicator results should highlight their relevance for specific management decisions and public acceptability.

Jackson et al. (2000 p.1-4) also suggest that guiding principles for this Phase should be:

Guideline 13: Data Quality Objectives
Guideline 14: Assessment Thresholds
Guideline 15: Linkage to Management Action

The utility of an indicator must be clearly communicated and signed off on by relevant managers and stakeholders (including managers and scientists) *before* the program is established. This means that estimates of uncertainty, ecological relevance, and resolution in terms of meeting management objectives must be clearly agreed to before the results of an indicators program are used for resource management. Once again, this does not necessarily imply that specific trigger levels and corresponding management actions should be specified *apriori*. Rather, it means that that the limitations and utility of indicators are understood before they are used for active management.

In summary, ecological indicators are one of the foundations of modern environmental management. However it is also clear that a number of indicator programs produce estimates of indicators that are highly uncertain, and often fail to accurately reflect the behaviour of key processes, let alone system behaviour, and have failed to adequately support good decision-making processes. The challenge is therefore to identify, measure, and communicate indicators in a manner that has high management utility and is cost-effective.

3.2 Establishing indicator validity and efficacy

The principles outlined above emphasise the importance of detailed consideration of indicator validity and efficacy. Establishing the natural variability and error surrounding each indicator is critical to ensure that the assessment and integration methods are appropriate and effective. We first outline the principles, and then methods for indicator evaluation and analysis. The goal of such an analysis is to provide guidance for indicator selection, weighting, reference bounds (or targets or thresholds of concern), and transformation schemes that take into account unexplained error and natural variability. The validation and testing work may take place prior to, or in conjunction with, the establishment of an assessment and integration scheme (discussed in Section 4).

In the first milestone report (Browne et al., 2007), we emphasised the importance of developing a good-quality database that enables the application of statistical methods that can be used to resolve empirical issues such as indicator thresholds, weighting and integration. Firstly, we will provide some specification of the kind of baseline data necessary to implement a typical assessment framework.

3.2.1 Baseline data

We emphasise that methods for indicator assessment (discussed in ‘approach 3’, below) for indicator assessment and integration depend on the ability to establish a high-quality baseline data set. As will be argued, developing reporting frameworks using these data provides the a robust foundation from which to conduct indicator assessment. This is because indicator assessment provides a mechanism with which to model and account for the various sources of variation inherent in indicator measurements so that test measurements can be compared with expected values giving an estimation of varying levels of human impact. This is critical for determining appropriate weighting, transformation, thresholds of concern and integration.

The notion of ‘baseline monitoring data’ may be extended to include not only data from a purposely designed baseline study, but also assimilation from existing monitoring programs. Further, monitoring done in the course of a report-card program may be added to the ‘baseline database’, leading to progressively better inferential models of human impact, and increasing the validity of the report-card over time. For this purpose, three types of data need to be collected (or estimated) in parallel at each site: we use the following classification, notation and definition of data types.

INDICATOR DATA: Measurements that are considered susceptible to anthropogenic impact. These measure-ments are used to infer a level of ecological impairment due to human disturbance.

ENVIRONMENTAL DATA: Measurements or properties of sites (or the zone or region) not considered subject to anthropogenic impact. During modelling, these measurements are used to capture natural variability, and moderate the estimated relationship between indicators and ecological impact attributed to human activity. Temporal and seasonal variables, possibly along with broad-scale regional categorical variables, also fall into this category.

HUMAN DISTURBANCE DATA: Measurements that relate to the expected level of human impact on a particular site. As with Environmental Data, these measurements may often not be strictly local, and may for instance pertain to the influencing catchment or water body.

In this discussion the term ‘measurement’ is used loosely to refer to in-situ measurement, remote sensing (with possible interpolation), process model estimates from indicator development stages (sediment load estimates, for instance), and more complex properties of sites. Often these properties are calculated using complex spatial model incorporating hydrology and other forms of spatial influence. These ‘measurements’ may even take the form of a broad-scale categorisation of regions based on expert opinion and previous research.

3.2.2 Indicator evaluation

A usual approach to indicator evaluation is to relate differences in indicator summary scores to aspects of environmental stress (Kennard et al., 2006a). While certain indicators may have a reasonably clear interpretation with respect to their significance to ecological health and human impact, other indicators require careful consideration. Evaluation of indicators is necessary not only for deciding whether or not they should be included in the assessment, but also for establishing their relative weighting, error bounds, and patterns of natural variability. Thus, indicator evaluation is critical for all subsequent processes of assessment, integration and visualisation.

The reference condition approach

The reference condition approach is strongly related to the disturbance gradient method of assessment, but has a different emphasis. In the referential framework, sampling during baseline evaluation is purposely done at undisturbed (reference) sites. Measurement of environmental structuring variables is also usually done, and the baseline study may be extended to include impacted sites.

Indicator evaluation within the reference condition approach (Reynoldson et al., 1997) is done by analysing the ability of indicators to discriminate between disturbed and undisturbed sites, often conditioning the discriminate function on environmental covariates. This encourages an evaluation scheme that compares metric scores with reference to an expected natural state. Models built on a referential approach permit inference as to the probability of a site being impacted, given the indicator data. In a multivariate, as opposed to a multi-metric approach (Bowman and Somers, 2006), test sites are compared with the distribution of reference sites using all variables (indicators) simultaneously. A number of methods are applicable in a multivariate context, but in general, the distance of a test site from reference condition is akin to a Mahalanobis distance (Legendre and Legendre, 1983). That is, it takes into account the means, variances and co-variances of all variables, thus accounting for redundancy amongst metrics. Importantly, a multivariate approach can identify differences that cannot be detected by considering each indicator individually. Multivariate comparison with reference can greatly increase sensitivity of the test to impairment by implicitly weighting indicators optimally given the baseline data, so that indicators that contribute little unique information are down-weighted or removed (Bowman and Somers, 2006). Multivariate comparisons, such as the multivariate t-test (i.e. special case of MANOVA), or discriminant analysis are fundamentally parametric, and probabilities of a site being significantly different from reference may be evaluated (Legendre and Legendre, 1983). If conventional distribution models prove to be a poor model of the actual data, a strongly non-parametric method (e.g. Browne, 2006) can also be used to directly model the density of undisturbed / disturbed sites in indicator space. Tests may incorporate models of the distribution of undisturbed sites only, or both sets of disturbed and undisturbed sites.

Data requirements

For most basic implementation of reference based approach, indicator data is required at reference sites, the number of sites being a factor of at least 5-10 larger than the number of indicators used.

Inference, validity and uncertainty

A reference condition approach allows inferences to be made regarding the probability of a test site being in undisturbed condition, as determined by the reference data set. The output of such an integration and assessment stage would therefore have a natural interpretation, supported by the analysis of empirical data. Adopting a multivariate stance ensures that information from multiple indicators is combined appropriately, and encourages appropriate indicator weighting. For example, if two indicators provide non-unique information on separating reference and test sites, then a multivariate discriminate function will implicitly weight the variables appropriately, while univariate approaches would tend to over-weight.

A disadvantage of the referential approach is that the use of a dichotomous validation variable (disturbed versus undisturbed) results in a reporting framework that does not resolve graduations of human impact. This gradient of human impact is considered in the following approach – known as ‘the disturbance gradient approach’.

The disturbance gradient approach

After selecting a subset of possible indicators for inclusion in a pilot or baseline monitoring study (see Section 4 for a list of mature indicators for the GBR), indicators may be evaluated using standard statistical methods such as regression according to:

1. The strength of the relationship of the indicator to the disturbance gradient;
2. The amount of residual variability after known anthropogenic and non-anthropogenic effects have been accounted for; and
3. The sensitivity of the indicator to the disturbance gradient (or the shape of the relationship – e.g. linear versus nonlinear) or alternatively, it’s efficacy for discriminating disturbed versus pristine sites.

Figure 7 illustrates some issues surrounding indicator evaluation with respect to a disturbance gradient. Here we assume that a baseline study has been conducted, with concurrent sampling of an indicator along with environmental variables and variables that that comprise the estimate of human impact at each site. The two columns in Figure 7 show examples of high- and low- residual variability, respectively. Three possible sensitivity types are shown from top to bottom; low sensitivity to higher levels of impact, equal sensitivity across the human disturbance gradient, and low sensitivity to low levels of human impact. The nature of the relationship between the indicator and the disturbance gradient informs how the indicator should be treated when analysed at test sites, where the goal is to estimate the (unknown) degree of human impact. A particular indicator score at a test site (shown with a red circle on the y-axis in Figure 7) is considered with respect to data from the baseline study, which implies not only an expected level of human impact, but also the residual variability surrounding this estimate (solid red region). Use of the mean estimate of human impact, given a particular indicator value, without incorporating the residual variability associated with the estimate, clearly has the potential to introduce unknown amounts of error into the eventual report-card grade. Another argument for incorporating residual variability into the integration framework is that it provides one solid basis for indicator weighting: assuming independence, indicators should be weighted according to the inverse of the uncertainty associated with the inferred level of human impact (e.g. the freshwater EHMP in

SE Qld considered the utility of weighting variables according tho the R^2 value [Harch 2007, *pers. comm*.). Naturally, indicators are often correlated, and the weighting scheme should also reflect this. Detection of high correlations between indicator variables in the baseline study may provide scope for removing redundant indicators from the assessments, based on consideration of the cost / information ratio. Indicator weighting will be discussed more fully in the following section on indicator integration and assessment.

As Figure 7 clearly shows, the degree of uncertainty associated with the estimated human impact from an indicator score at a test site depends on the strength of the relationship and also the sensitivity in that local region of the disturbance gradient. Note that this illustration does not include effects from temporal or structuring environmental variables (natural variation) on indicator uncertainty. The strength of a disturbance gradient approach is that it leads not only to an inference of the most likely level of human impact that gave rise to a particular indicator measurement, but also the degree of uncertainty surrounding the estimate. The potential for relying on baseline models in order to transform raw indicator scores into inferences (with error bounds) will be considered in the next section on indicator assessment.

Kennard et al's (2005) study as part of SEQ Freshwater EHMP detailed in Box 1 illustrates the kind of scientific sampling, analysis and evaluation that needs to be done for each included indicator in a report-card, in order for the report-card to have a scientifically credible interpretation. For the SEQ freshwater EHMP, this form of research justified the sampling, index / metric calculation, and further implied reference levels (levels of concern) for the indicator related to alien fish (amongst others).

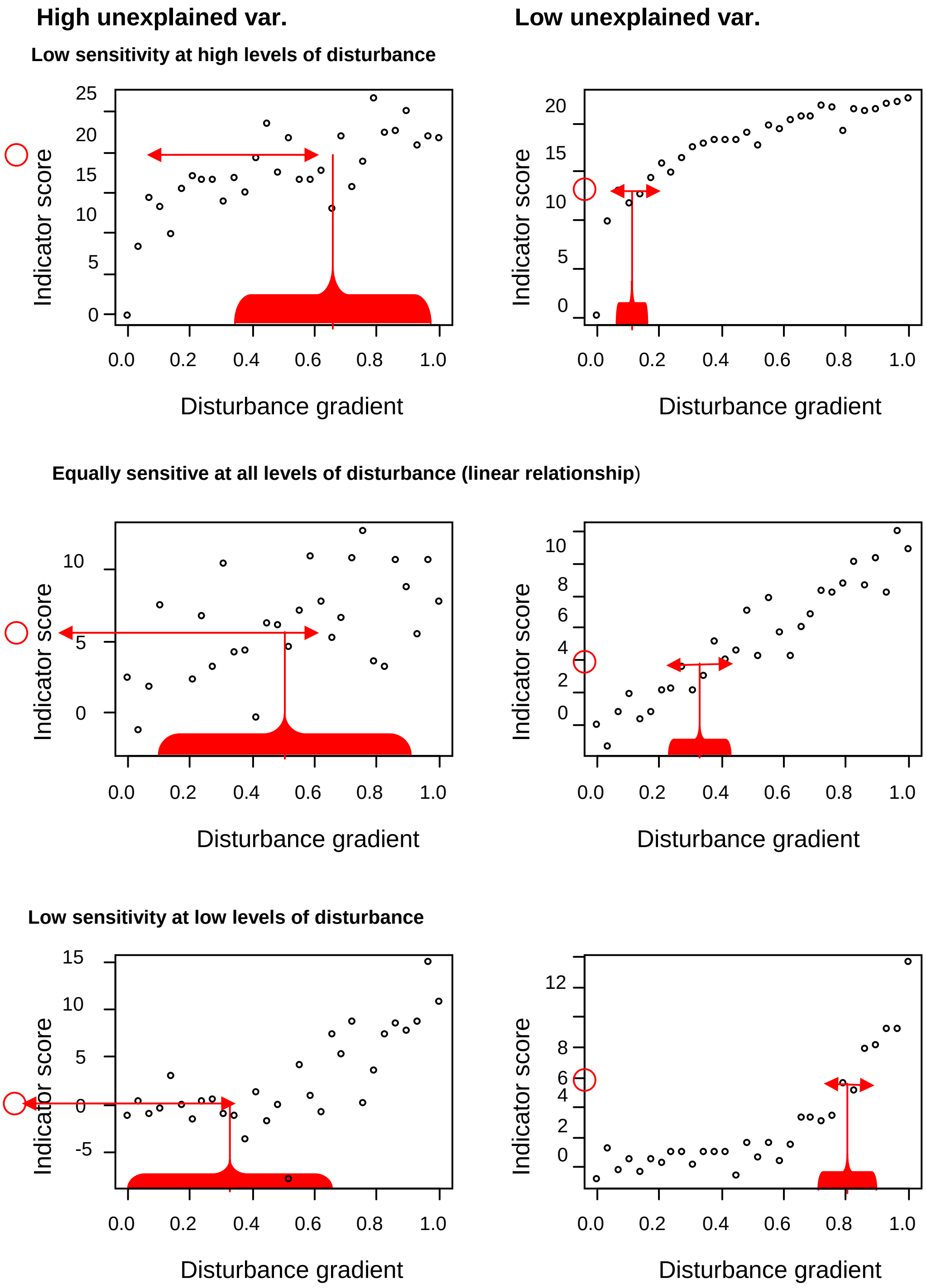


Figure 7: Various relationships between a disturbance gradient (x-axis) and a hypothetical indicator score (y-axis). After modelling, test site scores (red circles) imply a probability distribution of possible levels of human impact (shaded region, indicating the range of an unspecified error distribution). The left and right columns illustrate situations with high and low unexplained variability of the indicator, respectively.

Box 1. An example of indicator evaluation

There are very many instances of indicator validation in the ecological literature. Here is a single example which illustrates many of the standard statistical and methodological principles to providing evidence for justifying a candidate indicator. Kennard et al (2005) evaluated alien fish as an indicator of ecological health in freshwater rivers as part of the SEQ Freshwater EHMP development process. Forty-eight sites were selected that well-represented the human disturbance gradient in the region with an additional 72 reference sites with minimal human impact. Effort was made to sample so as to control for seasonal and hydrological effects. Statistical analyses included rank-order correlations, generalised linear modelling, and t-tests comparing ecological integrity variables at sites where alien fish were and were not present. Principlal Component Analysis was performed on a variety of measures of human disturbance which identified three dimensions that predicted alien species relative abundance and biomass relatively well. The authors concluded, with some caveats, that indices based on alien fish species can be a useful indicator of ecological health.

The DIBM3 (Bunn and Smith, 1997) study played a major role in the development of the freshwater component of the SEQ EHMP. The researchers adopted an approach applied previously developed by the GEEP (Group of Experts on Environmental Pollution) for detecting anthropogenic impacts in marine systems (Bayne et al., 1988; Stebbing and Dethlesfsen, 1992). In these studies, a broad range of indicators were evaluated against a known disturbance gradient. Research providing input to the DIBM3 (Fellows et al., 2006; Kennard et al., 2006b; Marshall et al., 2006b; Udy et al., 2006) considered 80+ descriptors of human disturbance in order to validate indicators. Selection of which aspects of the disturbance gradient variables to analyse was done with reference to conceptual models, and making use of dimension reduction methods such as Principal Component Analysis (PCA). The DIBM3 studies relied mostly on generalised linear models (GLMs) to establish the strength of the relationship between (often transformed) indicators and the disturbance gradient. The primary method to establish indicator validity was the proportion of variance explained (R^2). However, as noted by Fellows et al (2006), the criteria of what constituted a good indicator also extended to spanning a relatively large range of human impact, and to distinguish intermediate levels of disturbance as opposed to only between highly disturbed versus reference sites.

Data requirements

For basic implementation of a disturbance gradient approach, indicator data and indicators of human impact or disturbance are required, spanning the full range of human impact, from minimal (reference) to highly disturbed. The baseline data set will need to be large enough to detect the anticipated effect of human impact variables on indicators. The higher the ratio of natural variability / measurement error to human impact effect, the larger the baseline data set will be required (refer to Udy et al. 2006). More specific definition of data requirements ought to be accomplished via statistical power analysis. If the contribution of structuring (non-anthropogenic) environmental variables is to be accommodated in the analysis, then these must also be collated.

Inference, validity and uncertainty

The disturbance gradient approach to indicator evaluation permits inference to be made regarding the most likely level of human impact given the observed indicator data at test sites. Parametric modelling allows for the evaluation of uncertainty for these estimates, providing a known functional form to describe the relationship between the indicator and measures of the human impact (linear quadratic, cubic, etc). [Note non-parametric modelling approaches (including regression trees an their variants) are also applicable here, but usually require larger data sets.] Thus, the validity of a grade generated by the gradient approach is relatively high, and easily interpretable. The gradient approach distinguishes between degrees of human impact, not simply distinguishing between impacted and un-impacted sites. The development of a baseline dataset is also very important in this approach, with some attempt to stratify sampling with respect to environmental structuring variables. We also recommend including these variables as covariates when estimating indicator utility. Of course, the true validity of the approach is also dependent on the quality of the human disturbance validation variables. If these are a poor reflection of the true level of human impact, or incorporate spurious correlations with the indicator variables, then this will introduce bias and undefined uncertainty into the estimates.

Accommodating natural variability, uncertainty and establishing reference or target condition

Natural variability and uncertainty around indicator measurements are critical issues that must be addressed in the integration framework. Several forms of natural variability can be differentiated:

1. Seasonal variability
2. Random temporal variability due to climatic variation and weather events
3. Variability in space, including variability due to known physical and habitat differences, broad-scale differences in physical features and biological composition between regions
4. Unexplained variability and measurement error.

It is essential that the final scores of a report-card reflect human impact rather than other sources of variation. As discussed in the previous section, in the first instance, indicators should be selected that are expected to have as low as possible natural variability and low measurement error relative to the anticipated effect of human impact. Variability arising from small scale spatial variation and intrinsic sampling variability may be estimated and minimised by conducting standard procedures such as taking replicate samples at different spatial locations within a site. Variability arising from seasonal and non-anthropogenic environmental effects can be accounted for by co-measurement of these variables and factoring out these effects in an appropriate statistical modelling framework.

The Australia-wide Assessment of River Health (ARH), which is currently in development, is following the precedent set by such programs as the Victorian Index of Stream Condition and the SEQ EHMP in making assessments based on departures from reference conditions (NWC, 2006). There is a marked consensus that scientifically sound indicator assessment must be undertaken with respect to the presumed natural state of a site. As implied by the presence of natural variability, and as noted by Stoddard et al (2005), reference condition for a particular indicator is not a point, but a distribution consisting of both explained and unexplained variance components.

Figure 8 illustrates a distribution of a particular indicator - log(Choropphyll a) - at marine sites in the GBR, in relation to a single non-anthropogenic explanatory variable (cross-shore spatial location relative to the width of the GBR lagoon). The ‘threshold of concern’, for this indicator may be approached statistically as *values for that indicator which are unlikely to have been drawn from the undisturbed population, given non-anthropogenic explanatory information (covariates)*. This approach is essentially identical to the disturbance gradient approach discussed above, except that the disturbance gradient is considered as a categorical variable. In this figure, it is assumed that baseline data is only available from undisturbed sites. Figure 8 illustrates that the same indicator score, at measurements with different covariate scores, can be insignificant or significant, depending on covariate information. In this case, a Cl a reading of 1.35 (exp(0.3) on the y-axis of Figure 8) lies within the expected undisturbed distribution for coastal sites (green triangle), but represents a significant deviation if taken at a remote offshore location.

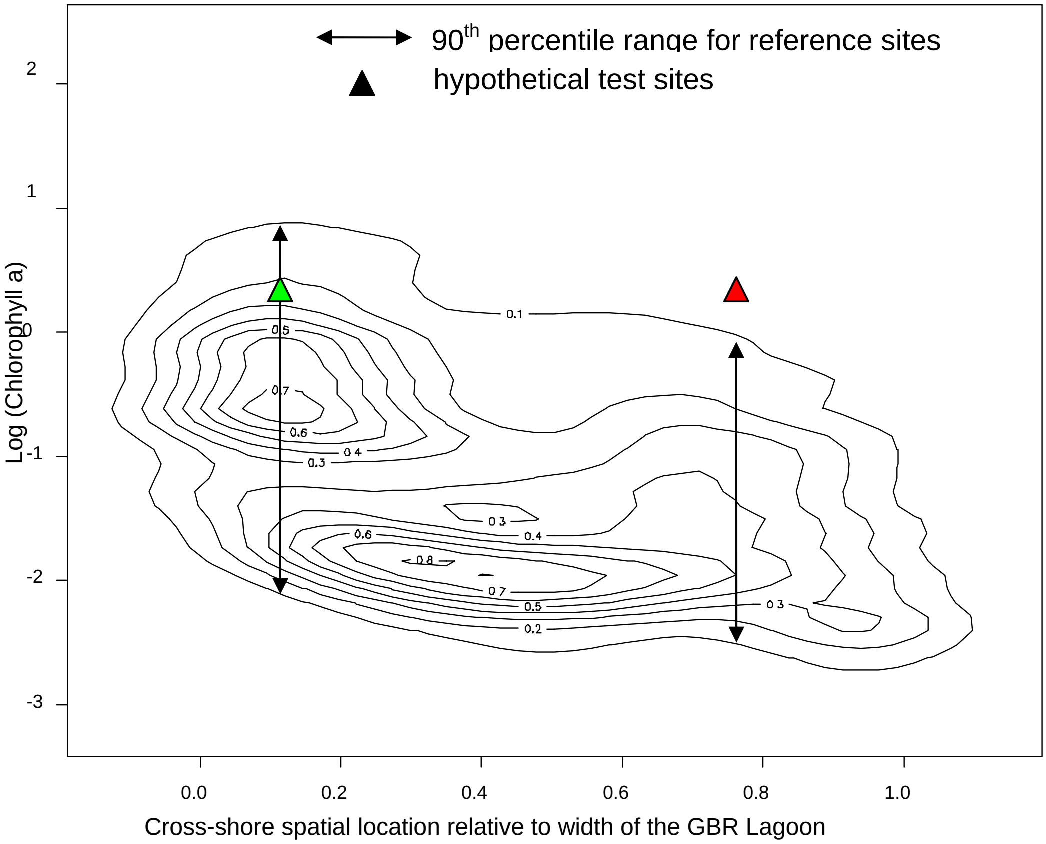


Figure 8: Hypothetical example comparing indicator scores at two sites with the distribution at reference sites conditioned on a single environmental predictor.

The approach detailed in Figure 8 illustrates that estimating a probability distribution of an indicator for minimally impacted sites, incorporating both explained and unexplained sources of variability, provides a basis for evaluating indicator scores at test sites. Incorporating the distribution of the indicator at impacted sites would also further improve the assignment of probabilities. On the other hand, if the natural variability and uncertainty associated with a particular indicator is not well understood, then it is very difficult to estimate with confidence the range of indicator scores that correspond to healthy ecosystem functioning, leading to uncertainty in how particular indicator scores should be evaluated in the assessment and integration framework.

There is often confusion regarding the relationship between reference levels and target levels (Stoddard et al., 2005). For instance, it is often thought, mistakenly, that adopting a referential approach entails that the management goals are to restore all sites to minimally disturbed, native or natural condition. We argue that the main role of the referential approach is to establish the *direction and distance* between present state and pristine state. The proportion of this distance that one aims to bring a disturbed site towards the pristine condition is a management decision. Figure 9 illustrates this idea.

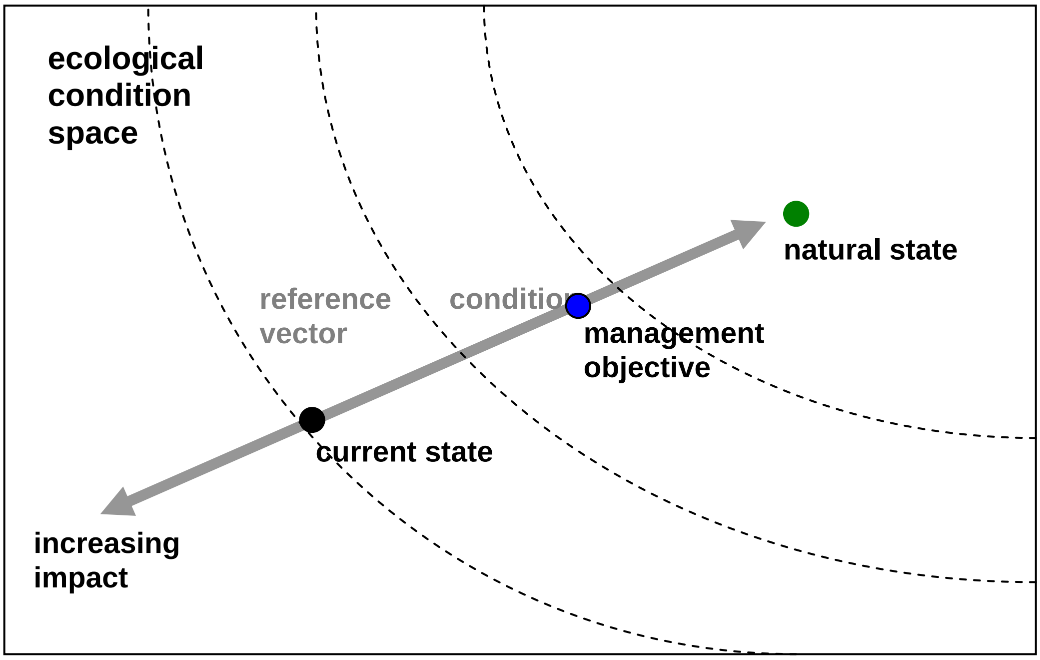


Figure 9: Conceptual diagram of the relationship between a referential approach and management goals. Establishing the reference gradient is necessary, even though management objectives may represent only partial restoration to natural ecological state.

3.3 Indicator assessment and integration

A robust IRCF must be founded on the 15 guiding principles outlined in Section 3.1. With the consideration of these guidelines, a robust framework for indicator selection, assessment and integration can be developed. It will be argued that the indicator assessment stage can and should influence the subsequent indicator integration process. Here we provide alternatives for indicator assessment for a water quality report card according to several objective-based approaches. They differ to the degree that objectives are determined empirically or apriori, and in the sophistication (complexity) of the assessment and integration method. In this context, the advantages of adopting a more sophisticated approach must be weighed against any loss of transparency, or ease of interpretation.

3.3.1 Managing to a specific objective

We consider three alternatives for indicator assessment. The first alternative, ‘compliance with objectives’ is based on a simple comparison of test-site indicator scores with given thresholds. The second alternative, ‘general compliance’, provides a more sophisticated method of assessing compliance, at the expense of some transparency. Both these assessment schemes amount to numerical rather than statistical operations: i.e. transformation and integration of raw data. Thus, they are strictly limited in terms of the level of inference they permit. Validity and uncertainty are also largely unquantified. However, if this has been accomplished in a prior indicator evaluation stage, in which case thought must be given on how to incorporate this validation information into the assessment. They also implicitly assume that an error free spatially and/or temporally explicit objectives or thresholds-of-concern have been developed: these inputs are necessary to take into account natural variation. The final alternative, ‘model based assessment’ incorporates validation of indicators, and estimation of natural variability. However, it is more demanding in terms of access to considerable baseline data (see the previous sections on baseline data and indicator assessment).

Compliance with objectives

A straight-forward compliance-based indicator assessment method involves comparing indicator measurements at test sites with the objective for that indicator. Indicator assessment in the SEQ Estuarine/Marine EHMP largely follows this approach (Pantus and Dennison, 2005). This assessment method assumes that objective levels for indicators are given. The work of Pantus and Dennison (2005) which contributed to the SEQ Estuarine/Marine EHMP details methods for integrating indicators with respect to management targets. Rather than attempting to determine pristine state empirically through the use of reference sites, the approach is to define the index as being *inversely proportional to the distance between the operational management objective and the measured state of the system* (Pantus and Dennison (2005). In practice, this is often done, for each indicator, by considering the number of sites (or proportion of area) that exceeds the threshold of concern. Thus, indicator readings at sites (or grid points, in the case of remotely-sensed data) are compared to thresholds. Calculating the proportion of sites that exceed the threshold is the most natural method for integrating this representation of the data. This yields the proportion of compliant sites for each region and each indicator. This assessment can be improved by incorporating an estimate of sampling reliability using bootstrap methods (illustrated in Section6.3.1).

Adopting an objective-based approach essentially places the responsibility for determining indicator weighting, validation, reference levels / thresholds and natural variation outside of the parameters of the current reporting process. There may be a number of reasons for doing so, including:

- The means and resources for doing so are not available in the current project;
- Significant previous work has been done to establish such ‘meta-parameters’ as thresholds of concern, indicator importance, etc; and
- Ecological objectives have been largely determined by an outside process (e.g. national standards, local management goals).

Averaging over indicators can take place over regions and indicators according to the indicator organisation scheme (a candidate indicator organisation scheme is described in Section 6). Weighted averages can also be used if it is desired to weight some indicators more than others. In general, indicators with greater unexplained variability and less relationship with the disturbance gradient should be weighted less. High co-variability between particular indicators would also argue that the set should be weighted less, as they contribute non-independent sources of information to the global estimate.

Date Requirements

The data required to implement a minimal management-objective based approach are low compared to the other approaches described here. In this perspective, a reference data set is not strictly necessary as test sites are evaluated with respect to objectives external to integration and assessment framework. This increases the importance of setting correct thresholds of concern, as these are given *apriori* rather than determined empirically. The number of thresholds that need to be established is at most, the product of the number of indicators, the number of significant temporal strata (e.g. season), and the number of spatially homogenous regions. This is because managers will need to set individual objectives tailored for each different sub-region, each indicator (Pantus and Dennison, 2005) and possibly also for multiple seasonal time-bands. Since this approach assumes that natural variability in indicator scores has been accounted for, the spatial partitioning and the thresholds of concern (treated as inputs to this integration method) must be of very high quality.

Inference, validity and uncertainty

The inference permitted by a ’compliance with objective’ based assessment is relatively weak. The report card may be interpreted as to the degree to which the indicators are meeting set objectives. It does not permit ‘strong’ interpretation of a report card in terms of representing some underlying ecological state. An objective-based approach will result in output strongly sensitive to variation in thresholds and spatial partitioning. Since there exists no formal statistical model of the relationship between indicators and human impact (modulated by natural factors) there is no way to quantify the uncertainty associated with the inference on human impact given indicators. The thresholds of concern, in particular, are subject to unquantified error.

The objective-compliance method transforms raw indicator data into a compliant / non-compliant code. Thus, it does not take into account the degree to which an indicator measurement does not meet a threshold. A more general criticism is that it assumes a binary (compliant / non-compliant) relationship between the indicator and the underlying construct (whether water quality or ecological health) rather than assuming a continuous relationship (whether linear or non-linear). As discussed above, simple objective-compliance integrates using averaging. A practical consequence of this is that that averages tend not to produce very high or very low scores, possibly leading to report-card grades that are all very similar.

We have noted that compliance based assessment is extremely sensitive to what objectives are chosen, and the spatial partitioning method used. Unless there is convincing justification that the objective thresholds and regionalisation are appropriate for that indicator, there is a significant risk that this sensitivity will compromise the resulting report card grades. Thus, whilst the demands of an objective-based approach on empirical data are relatively light, the quality of inference (on underlying impact / disturbance) is low, and the uncertainty associated with the output is unquantified. Compliance with objective based assessment treats individual indicators in isolation, rather than adopting a multivariate perspective. Thus, redundancy amongst indicators is not taken into account.

With the application of Monte Carlo or bootstrapping methods, the reliability and sensitivity associated with objective-based assessment can be assessed to some degree. An advantage of the simple compliance with objective approach is transparency: report-card grades can be readily interpreted as the global level of compliance, averaged over regions and indicators.

General compliance

Naturally, there are more sophisticated ways in which to assess indicators with respect to objectives, and in performing the integration of compliance scores. We present a modification of one reasonable approach (CCME, 2001) that has been documented and tested to some extent (Gartner Lee Ltd, 2006).

In any assessment and integration scheme, heterogeneous measurements (tests) must be integrated with respect to objective levels to produce a global score. A direct and sensitive approach is to expand the Canadian Council of Ministers for the Environment (CCME, 2001)'s water quality index to incorporate four factors ***F*** of indicator exceedence with respect to thresholds:

1. *scope* (the proportion of distinct indicators that exceed threshold);
2. *frequency* (proportion of failed tests);
3. *amplitude* (degree to which failed tests exceed threshold);
4. *extent* + *duration* (the geographical and temporal extent of failed tests).

This approach has the advantage of incorporating multiple aspects of the pattern of measurements exceeding objective condition, rather than only the proportion of failed tests.

For a given reporting region, the global score is given by:

$$I = \frac{\sqrt{F_1^2 + F_2^2 + F_3^2 + F_4^2}}{2}$$

which ranges between 0 and 100, high scores indicating high impact (a low level of compliance) and low scores indicating low impact (a high level of compliance). The first two factors are treated as proportions, with simple calculation with respect to indicators and with respect to test sites:

$$F_1 = 100 \times \frac{\text{\# failed indicators}}{\text{total \# indicators}},$$

$$F_2 = 100 \times \frac{\text{\# failed tests}}{\text{total \# tests}}.$$

The third factor takes into account the degree to which failed tests exceed threshold and is calculated in the following way: exceedence *e* is represented as

$$e_i = \begin{cases} \dfrac{t_i}{o_j} \text{-1 if the criteria is } t_i < o_j \\ \dfrac{o_j}{t_i} \text{-1 if the criteria is } \mathrm{o_j} < t_i \end{cases}$$

with respect to the indicator test score *t* and the objective *o*.

The average exceedence

$$\bar{e} = \frac{1}{I}\sum_{i}^{I} e_i$$

is transformed so as to approach 100 asymptotically as the ratio of exceedence to objective tends to infinity. Figure 10 below illustrates the behaviour of the exceedence factor graphically:

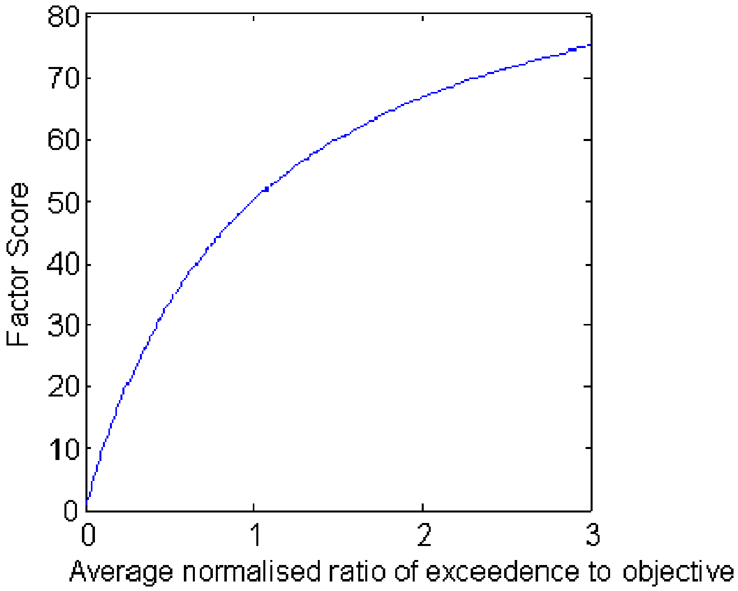


Figure 10: Change in exceedence factor score as the ratio of test score to objective increases.

The final factor estimates the extent and duration during which a component was exceeding the objective.

$$F_4 = \frac{\text{Volume of failed region}}{\text{total Volume}}$$

Numerical calculation of factor 4 may be done using interpolation or nearest-neighbour techniques. Specific methods for doing so depend on the form of the specific indicator. Suppose the spatial extent of the indicator is one-dimensional (e.g. a river course), and random measurements are taken over time and space. Figure 11 provides a graphical representation of the calculation of factor 4 in this case.

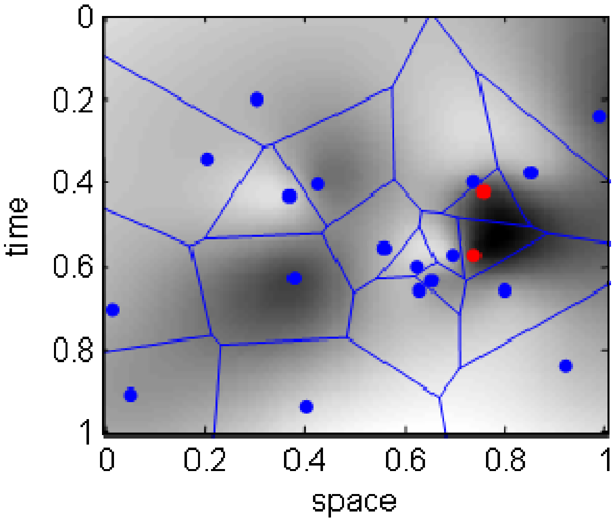


Figure 11: Transformation of indicator point-measurements (blue = non exceedance, red = exceedence) so as to represent exceedence over a time-space volume.

The multiple factors incorporated into the expanded CCME assessment and integration method provides multiple 'perspectives' of a group of indicators, with tests distributed over space and time, with respect to the given thresholds of concern. Thus, it may represent a more sensitive approach than simple compliance assessment. It has some interesting properties, in considering multiple indicators simultaneously, incorporating the degree to which test scores exceed objective, and in considering the temporal and spatial extent of non-compliance. One problem is that in calculation of 'degree of exceedence' is that it is scale-dependent: adding a constant to test scores and thresholds changes the factor score. Thus, it would assume that a, exceedence ratio of, say, 1:1.5 is 'equally bad' for all indicators, which is unrealistic. This problem could be addressed by estimating a scaling factor to be associated with each threshold.

In general, the data requirements, and the degree of inference and validity associated with the general compliance method are similar to those discussed for the simple compliance method. The difference is that some increase in the level of sensitivity and (conceptual) validity is gained at the expense of some transparency in interpretation.

Model based assessment and integration

In the compliance-based assessment methods described above, the assessment and integration methods is implemented subsequent and separately to assessing establishing indicator validity and efficacy (as discussed in the previous section). Hopefully, information pertaining to the validity and efficacy of indicators will be incorporated in some way into the spatial partitioning, thresholding, indicator organisation and weighting schemes. However, since they are treated as separate and distinct processes, the process of incorporating validation information is necessarily informal and/or ad hoc.

A model based approach treats indicator validation indicator assessment and integration in a single integrated process. Essentially, models are given indicator data and structuring environmental co-variates as input, and used to estimate the underlying ecological construct (some measure of disturbance). The *model estimates* at test sites (which are simply a weighted combination of the indicators) are treated as the global (indicator combined) score, and the integration technique is simply the model itself.

Most conventional modelling procedures combine multivariate information and produce estimates of the target variable with an associated estimate of uncertainty. The modelling procedure intrinsically optimises weights so as to account for indicator error and covariance. The covariates provide information as to natural variability, and this may eliminate the need for a separate process of spatial partitioning.

In summary, model-based integration provides a principled and statistically valid method for empirically transforming and integrating indicators into a global estimate of ecological health / human impact. Modern statistical modelling methods intrinsically deal with otherwise very hard-to-resolve problems related to uncertainty and covariance. Perhaps the most attractive feature of a model based approach is that validation, accommodating natural variability, indicator weighting and integration can all be accomplished within a single (and potentially quite simple) modelling framework. An example of using this approach for freshwater fish indicators is provide by Kennard et al. (2006)

Model based methods require access to the baseline data set (as specified in section 3.2.1 above) and this data set must be large enough to permit robust baseline modelling in order to implement this approach. However, most of this data should already have been collected in the previous indicator evaluation phase. Integration in a model-based approach takes place spatially according to reporting regions (since integration across indicators occurs within the modelling procedure itself).

3.4 Conclusions

Regardless of the indicator evaluation, assessment, and integration methods chosen, substantial work is required to understand and account for natural variability, measurement error, co-variability and validation with the reference distribution. The alternatives presented differ in terms of where the responsibility is placed for this development work. Compliance-based approaches assume that this work is done prior to assessment and integration, and therefore do not require a baseline dataset. Model-based approaches tackle these issues empirically with reference to a baseline data set. Whichever method is chosen, care must be taken to ensure that it satisfactorily addresses the issues and challenges detailed here.

We repeat here the most critical criteria for indicator evaluation:

- *Each indicator should have a reasonably high responsiveness to human disturbance*.
- *Minimise natural variability and measurement error*. Many candidate indicators will probably fail this test of having acceptable error bounds for their estimates to warrant inclusion in a report card.

There are three cardinal *do nots* for indicator assessment. Do not assess indicators for which:

- Responsiveness and validity are poor or unknown.
- There is excessive or unexplained measurement error or natural variability.
- There are no theoretical or empirical grounds to set thresholds of concern.

4. Mature indicators for a Great Barrier Reef Integrated Report Card

4.1 Pressure-Vector-Response Framework for indicator selection

To help organise the indicators within the IRC framework, a Pressure-Vector-Response model was employed (Figure 12). Traditional Pressure-Vector-Response (PVR) models outlined in Friend & Rapport (1979), form the basis of much of the State of Environment Reporting (SOE) that occurs around Australia. Recently a Pressure-Vector-Response (PVR) framework has been adopted by Queensland Agencies for the implementation of the Stream and Estuary Assessment Program (SEAP) (e.g. Negus & Marsh, 2006). A PVR framework was endorsed by the Reef Water Quality Partnership and therefore in conjunction with the 15 guidelines outlined in Section 3.1 are used to guide the indicator selection process for the GBR in this report.

The GBR IRC is specifically aimed at looking at the response of the marine system to changes in land management from adjacent catchments. However, as the streams and rivers are a conduit of sediments, nutrients and pesticides, as well as a home for many marine species during part of their lifecycle, it is not possible to evaluate marine health without including freshwater condition. Therefore within the PVR context the ‘Pressure’ represents land use (type and condition), the ‘Vector’ is represented by the freshwater condition and the ‘Response’ is the marine ecology and water quality.

There are two important reasons for including catchment condition in the GBR IRC. Firstly, due to the extensive temporal lags within natural systems between land use change and ecological response, it is important to be able to demonstrate improvements in land management before there is an improvement in ecological health. That way land managers may be “rewarded” for their efforts potentially decades before the ecosystems will respond. Secondly, information on land use condition is needed to test indicator responsiveness to disturbance, and therefore if the data is going to be collected for this purpose, then it would be worthwhile to utilise the data for potential socio-economic assessments that may/will be included within the GBR IRC in future years.

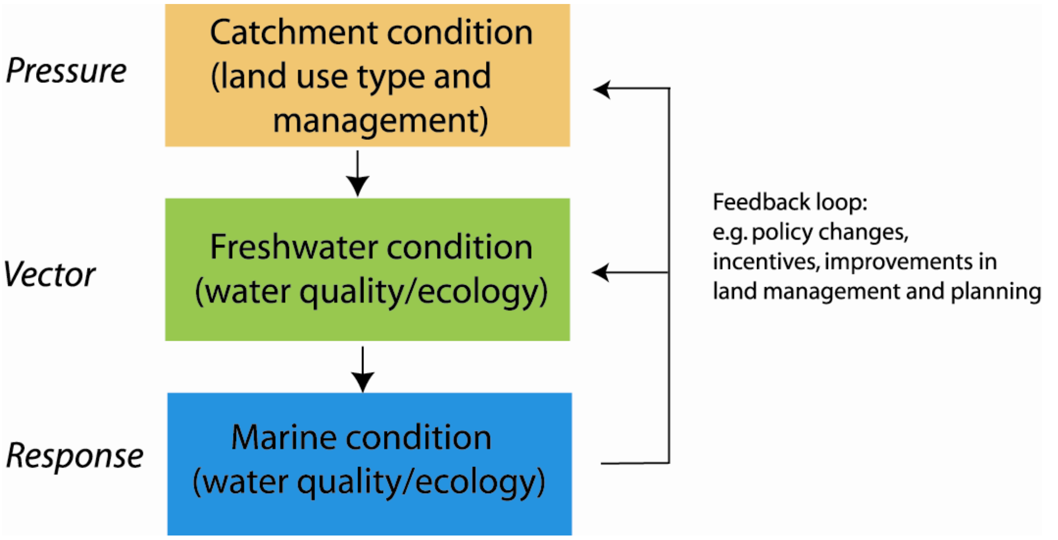


Figure 12: Pressure-Vector-Response Framework being used for indicator selection.

Based on reviews of literature on catchment, freshwater and marine research in the GBR catchments (e.g. Arthington and Pearson, 2007; Bartley et al., 2007; Coles et al., 2005; Connolly et al., 2007; Fabricius et al., 2007; Fabricius, 2005; Fabricius and De'ath, 2001; Fabricius et al., 2005; Honchin et al., Draft; Humphrey et al., In Press; Mackay et al., 2007; Mitchell et al., 2007; Pearson et al., 2007; Pusey et al., 2007; Steven et al., In Review; Sweatman, 2007; Sweatman et al., 2004), as well as discussions with key research scientists in the GBR region, an initial set of indicator groups were established. A full list of potential indicators for the catchment, freshwater and marine components are given in Appendix 1 (Table 18, Table 19, Table 20). From the initial group of 82 potential indicators, a list of 53 ‘mature’ indicators were determined. In summary, a mature indicator in this report is defined as having one or more of the following characteristics:

- A reasonably well established link (documented in scientific reports or papers) between the indicator and a known impact on, or response to, some pressure linked to land use condition or management;
- Is considered as a critical component of the ecosystem, and although the relationship between the indicator and land use management is not yet well established, testing is underway;
- There are good data sets currently available for these indicators or there are plans for these data to be collected in the future.

A proposed list of mature indictors for the catchment, freshwater and marine sections are given in Table 2, Table 3 and Table 4. It is important to note that the characteristics of the mature indicators do not necessarily adhere to the 15 guidelines for indicator selection outlined in Section 3.1. Considerable work is currently underway in a number of the MTSRF research programs to help determine the disturbance gradient impacts, thresholds of concern, associated management objectives and associated cost of monitoring these indicators in space and time. It is envisaged that this initial list of 53 mature indicators will be reduced further as more results from the research programs become available. There may also be additional indicators added or changes in the cluster categories over time. ***The next stage is now for the ‘domain’ research groups within the MTSRF program to evaluate each of the mature indicators according to the 15 guidelines presented in Section 3.1 of this report*** (Table 2, Table 3 and Table 4). This may take several years and it will be an iterative process. This step will need to be done in conjunction with all of the organisations responsible for monitoring catchment, freshwater and marine condition and trend in the GBR catchments (see Table 1). The outcome of this process will be a list of indicators that is suitable for use in a GBR report card.

Table 1: Guide to the source of data that may be used for the GBR IRC (note this list is likely to evolve and is not necessarily complete).

Type of data	Organisation responsible	Contact
General GIS data (topography, soils, geology, land use)	Geoscience Australia, Bureau of Rural Sciences, NLWRA, QNRW	Various
Seagrass data	QDPI	Rob Coles
Wetland data	QEPA	Jonathon Hodge
River flow hydrological data	QNRW	Chief Hydrographer
Load monitoring data (catchments)	QNRW	Simon Catzikiris
Freshwater ambient water quality	QEPA	Andrew Moss
Freshwater bugs	QNRW	Satish Choy
Freshwater fish		QNRW
Water Quality monitoring data (Marine)	AIMS/GBRMPA	Miles Furnas/ David Haynes
Marine coral and fish	AIMS/GBRMPA	Hugh Sweatman
Pesticide and Fertiliser use in cane lands	Cane Growers	Tim Wrigley

Table 2: Table of potential mature pressure indicators for catchment condition. ✓ should be used to indicate that specific guideline has been considered.

Indicator Number	Catchment indicators	Consideration of 15 Guidelines for Selecting Relevant Indicators (Jackson et al. 2000)															Relevant to	Data sites in GBR (custodian/program)	Reference or threshold value or WQ objective
	Land management	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15			
1	Fertiliser use: Application rates nitrogen																freshwater and marine ecology	see peak industry bodies (e.g. canegrowers, MLA)	not relevant
2	Fertiliser use: Application rates phosphate									✓							freshwater and marine ecology	see peak industry bodies (e.g. canegrowers, MLA)	not relevant
3	Pesticides: application rates				***The next stage is now for the 'domain' research groups within the MTSRF program to evaluate each of the mature indicators according to the 15 guidelines presented in Section 3.1 of this report***												freshwater and marine ecology	see peak industry bodies (e.g. canegrowers, MLA)	not relevant
4	Groundwater extraction rates																environmental flows	QNRW or QDPI	not relevant
5	No. of properties with farm management plans																land management change	see peak industry bodies (e.g. canegrowers, MLA)	not relevant
	Land type and condition																		
6	land use type (% agricultural land)											✓	✓				expected ecological impact	Geoscience Australia, Bureau of Rural Sciences, NLWRA, QNRW	not relevant
7	Outputs from point sources																freshwater and marine ecology	QEPA	not relevant
8	Ground cover (e.g. pasture condition, % C class)																land management change	QDPI, Bureau of Rural Sciences, NLWRA, QNRW	not relevant
9	riparian % cover (both banks)																light, erosion, habitat quality	Bureau of Rural Sciences, NLWRA, QNRW	not relevant
10	Vegetation Management																terrestrial ecology	Bureau of Rural Sciences, NLWRA, QNRW	not relevant

Table 3: Table of potential mature pressure indicators for freshwater condition. ✓ should be used to indicate that specific guideline has been considered.

Indicator Number	**Freshwater Indicators**	**Consideration of 15 Guidelines for Selecting Relevant Indicators (Jackson et al. 2000)**															**Indicator of**	**Data sites in GBR (custodian/program)**	**Reference or threshold value or WQ objective**
	Physical/Chemical/Hydrology	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15			
11	DIN (Nitrate, Ammonia) and DIP																Fertiliser use	see Henry et al (2007)	Under development
12	Particulate N and P				***The next stage is now for the 'domain' research groups within the MTSRF program to evaluate each of the mature indicators according to the 15 guidelines presented in Section 3.1 of this report***												Land use management	see Henry et al (2007)	Under development
13	Total suspended solids																Land use management	see Henry et al (2007)	Under development
14	Turbidity/clarity																Land use management	see Henry et al (2007)	Under development
15	Chlorophyll a																phytoplankton productivity	see Henry et al (2007)	Under development
16	Diuron																pesticides	see Henry et al (2007)	Under development
17	pH - Diel (24hr) range																	see Henry et al (2007)	Under development
18	Conductivity - Diel (24hr) range																	see Henry et al (2007)	Under development
19	Temperature - Diel (24hr) range																	see Henry et al (2007)	Under development
20	Disolved oxygen - Diel (24hr) range																	see Henry et al (2007)	Under development
21	Discharge (Mean Annual Flow)																Relative discharge	see Henry et al (2007)	Under development
22	Measure of flow variability																Dams and water diversions	see Henry et al (2007)	Under development
23	Sediment and nutrient loads																Total amounts of sediment and nutrients	see Henry et al (2007)	Under development
24	Coarse sediment load																Habitat impact	see Henry et al (2007)	Under development

Indicator Number	**Freshwater Indicators**	**Consideration of 15 Guidelines for Selecting Relevant Indicators (Jackson et al. 2000)**															**Indicator of**	**Data sites in GBR (custodian/program)**	**Reference or threshold value or WQ objective**
	Ecology (bugs and fish)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15			
25	Macroinvertebrates: species richness																light, sediments, nutrients, habitat quality	see Henry et al (2007) report	Under development
26	Macroinvertebrates: Family richness																light, sediments, nutrients, habitat quality	see Henry et al (2007)	Under development
27	Fish: O/E species richness and assemblage																light, sediments, nutrients, habitat quality	see Henry et al (2007)	Under development
28	Fish: No. of alien species																light, sediments, nutrients, habitat quality	see Henry et al (2007) report	Under development
29	Fish: % of alien species																light, sediments, nutrients, habitat quality	see Henry et al (2007)	Under development
	Habitat condition (incl. macrophytes)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15			
30	% riparian vegetation (both banks)																riparian management	State of Rivers Reporting	Under development
31	Macrophytes: % macrophyte cover edge habitat																light, sediments, nutrients, habitat quality	State of Rivers Reporting	Under development
32	Macrophytes: % alien taxa																light, sediments, nutrients, habitat quality	State of Rivers Reporting	Under development
33	Macrophytes: % poaceae																light, sediments, nutrients, habitat quality	State of Rivers Reporting	Under development

Table 4: List of potential mature response indicators for the Marine environment. ✓ should be used to indicate that specific guideline has been considered.

Indicator Number	Marine Indicators	Consideration of 15 Guidelines for Selecting Relevant Indicators (Jackson et al. 2000)															Indicator of	Data sites in GBR (custodian/program)	Reference, threshold value or water quality objective
	Nearshore and Reef ecology	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15			
34	Hard coral cover																?	Yes (AIMS LTMP)	under development
35	Hard coral richness				***The next stage is now for the 'domain' research groups within the MTSRF program to evaluate each of the mature indicators according to the 15 guidelines presented in Section 3.1 of this report***												Disturbance	Yes (AIMS LTMP)	under development
36	Hard coral recruits																sedimentation	Yes (AIMS LTMP)	under development
37	Soft coral richness																Turbidity	Yes - Fabricius	under development
38	Macro-bioeroders in Porites																POM	Yes (AIMS LTMP)	under development
39	Herbivorous fish density, biomass, spp composition																macroalgal biomass	Yes (AIMS LTMP)	under development
40	cover																DIN/DIP	Yes (AIMS LTMP)	under development
41	Foraminifera																light, nutrients, sedimenation	AIMS -Sven Uthicke	under development
42	% seagrass cover																light, sedimentation	Yes - Rob Coles	under development
43	Seagrass species composition, biomass																light, sedimentation	Yes - Rob Coles	under development
	Marine Water Quality***																		
44	Chlorophyll a																phytoplankton productivity	Furnas or MODIS	River mouth 1.0 m < 10 m isobath na > 10 m isobath na
45	Secchi Disk (light)																sediment	AIMS	River mouth 2.0 ug/l < 10 m isobath 0.5 ug/l > 10 m isobath 0.3 ug/l

Indicator Number	Marine Indicators	Consideration of 15 Guidelines for Selecting Relevant Indicators (Jackson et al. 2000)															Indicator of	Data sites in GBR (custodian/program)	Reference, threshold value or water quality objective
46	Nitrate																Fertiliser	AIMS	River mouth 15.0 ug/l < 10 m isobath 2.0 ug/l > 10 m isobath 2.0 ug/l
47	Particulate N																land use?	AIMS	under development
48	Particulate P																land use?	AIMS	under development
49	CDOM																land use?	AIMS	under development
50	Turbidity																land use?	AIMS	River mouth 10.0 ug/l < 10 m isobath 1.0 ug/l > 10 m isobath 1.0 ug/l
51	Suspended sediment																		under development
52	Sediment composition, grain size																land use?	AIMS	under development
53	Pesticide/Herbicide																pesticide use	Mueller, UQ	available for a variety of pesticides (see Honchin et al., 2006)

** Note: For most water quality indicators, the ecological system relies on some input as part of natural processing, however, the threshold level between natural and impacted is not yet well established.

4.2 Data integration and visualisation options

In an IRC context, results are not considered useful if the information cannot be effectively communicated to community stakeholders, managers, the media and policy makers effectively (Johnson, 2006). The report card must provide information at several levels of detail to meet the needs and expectations of the entire set of stakeholders (Aurricht, 2004; McKane, 2002).

A discussion of the various approaches for data integration and visualisation was given in Auricht (2004) and briefly in Browne et al., (2007). Instead of providing a range of options, our favoured integrated and visualisation approach, the 'data wheel' approach, is described. The data wheel approach as described in McKane (2002) provides an elegant and potentially consistent mechanism for visually integrating a range of different data sets. Here we refer to the data wheel as the 'GBR water quality data wheel'.

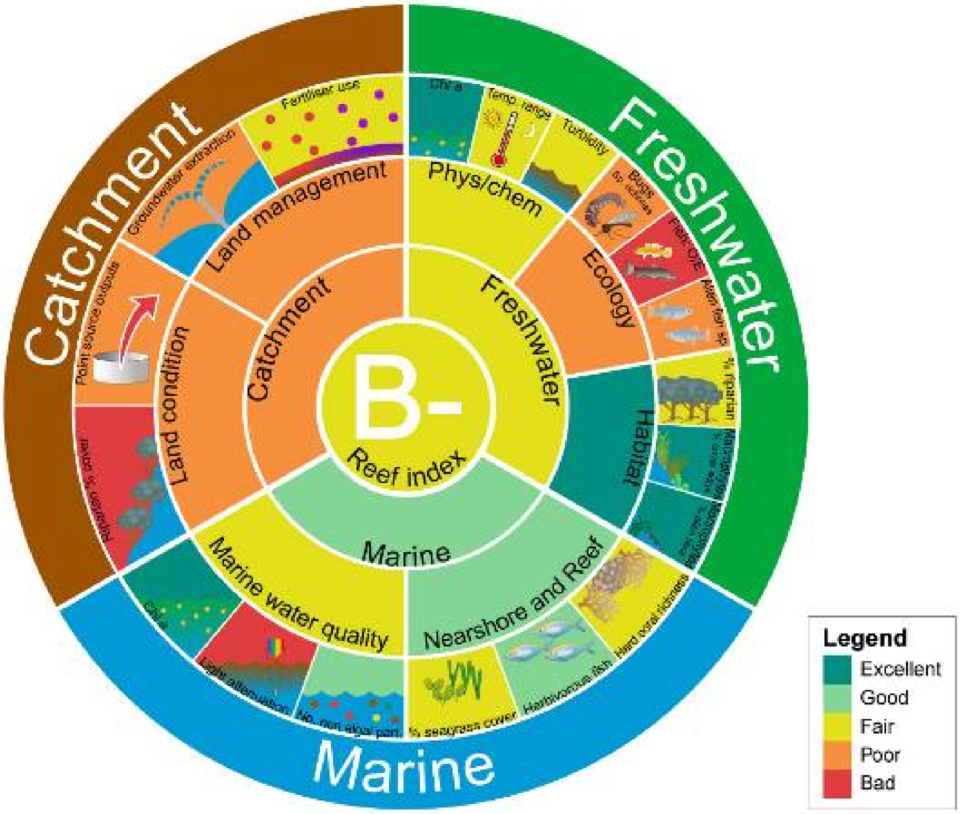


Figure 13: Example of GBR water quality data wheel approach that would be suitable for use in an integrated GBR report card

Within the data wheel framework, there are 4 levels of data integration as shown in Figure **14**:

1. Indicators
2. Clusters (average of indicators)
3. Groups (average of clusters)
4. Index (average of groups)

Within the GBR report card framework, the three ‘Groups’ follow the PVR framework (as described previously) and represent: Catchment, Freshwater and Marine zones. Within each of the groups are a set of ‘Cluster’ categories. The pressure or catchment indicator cluster categories are: (1) land use management and (2) land use type and condition. For the vector freshwater indicators, the cluster categories are: (1) physical, chemical and hydrological indicators; (2) ecology (bugs and fish) and (3) habitat condition. For the response or marine indicators, the cluster categories are: (1) nearshore and reef ecology, (2) marine water quality (see Figure 14). These cluster groups were derived during the selection of mature indicators in section 4.1.

Table 5 uses a sub-set of the indicators (for demonstration purposes only) to show how the individual indicator data can be reported at a range of levels (indicator, cluster or group) as well as be integrated into an overall ‘index’ of GBR catchment health.

The data wheel approach described in McKane (2002) is considered a suitable approach for the GBR IRC for a number of reasons:

- The approach is flexible in that new indicators, or indicator groups (e.g. socio-economic indicators) can be adapted at any stage;
- This type of visualisation option removes the need for different visualisation approaches for the scientific versus community report cards;
- Different layers of the data wheel can be removed to reduce the complexity of visualisation where required (See Figure 15);
- Variations in the colour scheme can be used to demonstrate data quality and availability and potentially highlight where more data are needed for certain indicators (See Figure 16);
- This approach promotes consistency between the catchment, freshwater and marine approaches, rather than having different visualisation options as done for other report cards;
- The data wheel could be used to present data for a single catchment, or for a region such as the ‘wet tropics’ (See Figure 17); and
- This approach allows for different reporting time frames, as it may be more appropriate for the catchment condition information to be reported on a 5 yearly cycle, whereas other data may be presenting annually.

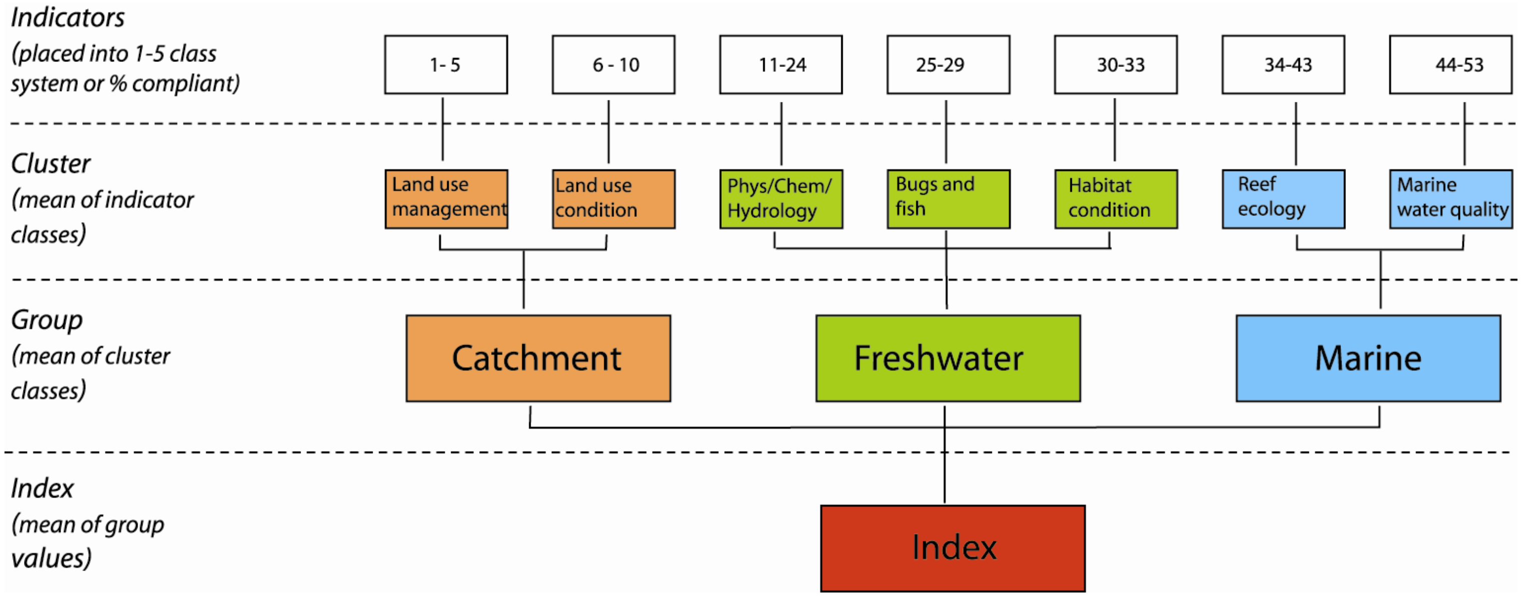
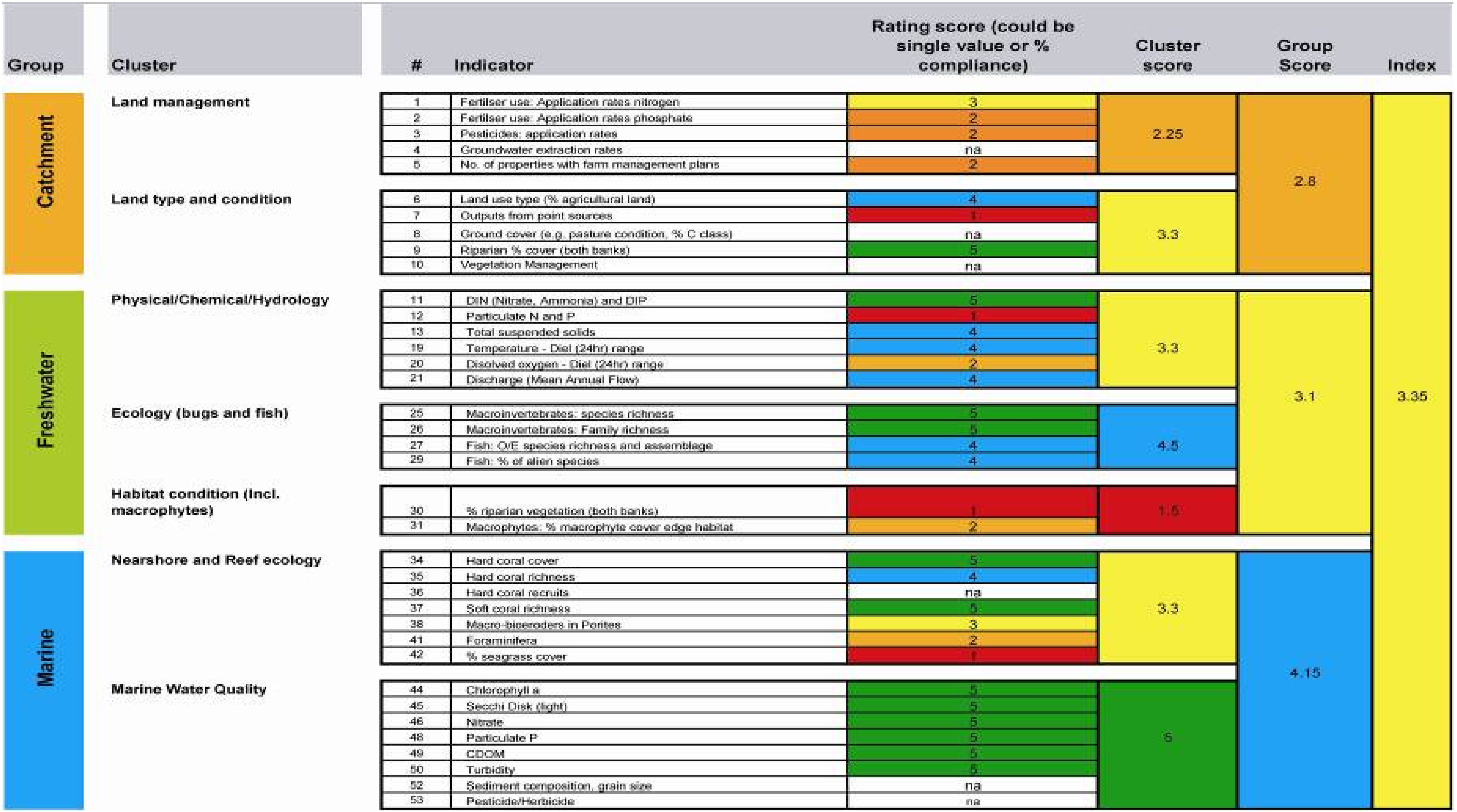


Figure 14: Example of the relationship used to build the index presented in the GBR water quality data wheel in Figure 13.

Table 5: Example of how the individual indicator measurement scores (either as a score from 1-5, % compliance or other measure) are averaged up to cluster scores, groups scores and finally an Index. Note only example indicators have been used and all of the values are synthetic. Colours match the legend in Figure 13.



Group	Cluster	#	Indicator	Rating score (could be single value or % compliance)	Cluster score	Group Score	Index
Catchment	Land management	1	Fertilser use: Application rates nitrogen	3	2.25	2.8	3.35
		2	Fertilser use: Application rates phosphate	2			
		3	Pesticides: application rates	2			
		4	Groundwater extraction rates	na			
		5	No. of properties with farm management plans	2			
	Land type and condition	6	Land use type (% agricultural land)	4	3.3		
		7	Outputs from point sources	1			
		8	Ground cover (e.g. pasture condition, % C class)	na			
		9	Riparian % cover (both banks)	5			
		10	Vegetation Management	na			
Freshwater	Physical/Chemical/Hydrology	11	DIN (Nitrate, Ammonia) and DIP	5	3.3	3.1	
		12	Particulate N and P	1			
		13	Total suspended solids	4			
		19	Temperature - Diel (24hr) range	4			
		20	Disolved oxygen - Diel (24hr) range	2			
		21	Discharge (Mean Annual Flow)	4			
	Ecology (bugs and fish)	25	Macroinvertebrates: species richness	5	4.5		
		26	Macroinvertebrates: Family richness	5			
		27	Fish: O/E species richness and assemblage	4			
		29	Fish: % of alien species	4			
	Habitat condition (Incl. macrophytes)	30	% riparian vegetation (both banks)	1	1.5		
		31	Macrophytes: % macrophyte cover edge habitat	2			
Marine	Nearshore and Reef ecology	34	Hard coral cover	5	3.3	4.15	
		35	Hard coral richness	4			
		36	Hard coral recruits	na			
		37	Soft coral richness	5			
		38	Macro-bioeroders in Porites	3			
		41	Foraminifera	2			
		42	% seagrass cover	1			
	Marine Water Quality	44	Chlorophyll a	5	5		
		45	Secchi Disk (light)	5			
		46	Nitrate	5			
		48	Particulate P	5			
		49	CDOM	5			
		50	Turbidity	5			
		52	Sediment composition, grain size	na			
		53	Pesticide/Herbicide	na			

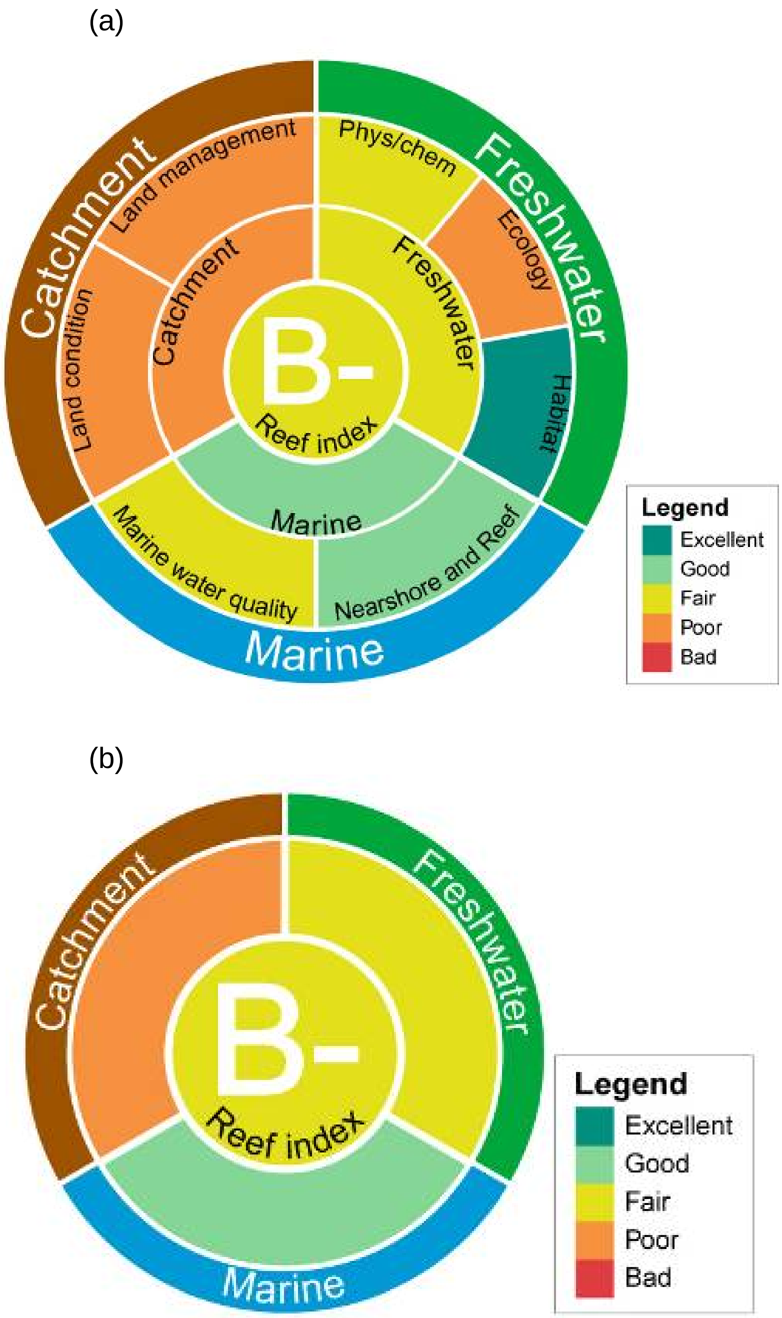


Figure 15: Examples of the data wheel with (A) the outer indicator layers removed and (B) showing only the Index layer. This may be useful depending on the audience.

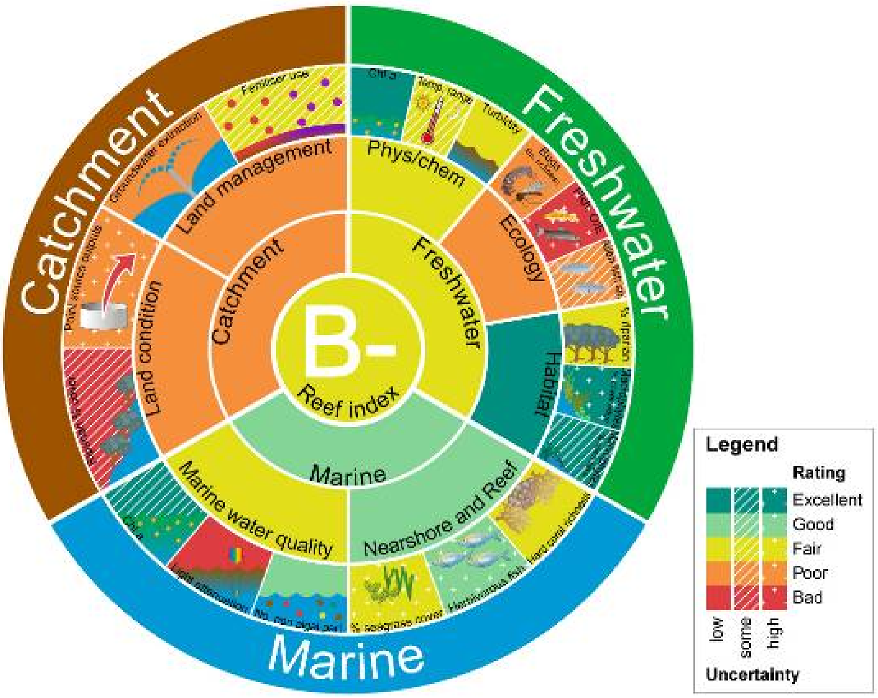


Figure 16: Example showing how different shaded colours could be used to demonstrate variation in data quality or uncertainty

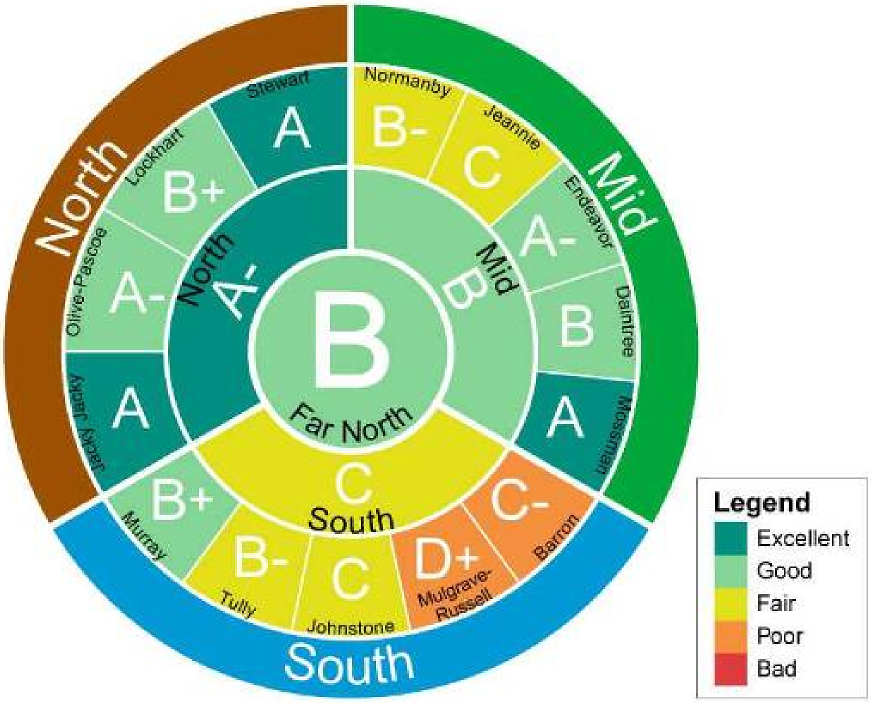


Figure 17: The data wheel could also be used to present the results different catchments or regions within a catchment management boundary such as the Wet Tropics.

5. Approaches to spatial partitioning for assessing compliance against targets and thresholds of concern

Having considered the conceptual framework for developing an integrated report card (Section 2), provided a process for the development and evaluation of indicators (Section 3), as well as providing a list of mature indicators within a pressure-state-response model (Section 4), the focus now moves toward statistical approaches for “spatial partitioning” that enable appropriate spatially focused reporting against targets or thresholds of concern.

Spatial partitioning, in the context of developing an integrated report card, is critical for two aspects. The first involves establishing reporting zones, a requirement in legislation, where the focus is on reporting environmental health. There are a number of approaches to “spatial partitioning” where the focus is often on providing a method for defining ecologically homogenous zones.

The other critical aspect focuses on assessing compliance against targets and thresholds of concern that ensure the spatial context in which the assessment is made is appropriate. For example, comparing the health of a waterway in an estuarine area using headwater stream reference condition data will provide a biased assessment of health. The focus of this project is on the latter aspect of “spatial partitioning” for developing an integrated report card.

5.1 Introduction

A robust and meaningful statistical approaches that enables spatially focused reporting against targets or threshold of concern is vital to any environmental monitoring program. Such approaches ensures that an indicator’s response to anthropogenic impacts will be interpreted correctly. “Spatial partitioning” is used to reflect the process of subdividing a geographic space into one or more homogeneous sub-regions or groups based on data that reflects differences in environmental and/or biological characteristics, where the main objective is to provide the spatial zones against which indicator evaluation, assessment and integration is required.

In the literature, spatial partitioning has been applied within three broad contexts:

1. Derivation of the sampling design (Marshall et al., 2006; Smith et al., 2001);
2. Identification of reference condition (Marshall et al., 2006a; Negus and Marsh, 2006b; Simpson and Norris, 2000; Smith et al., 2001); and
3. Reporting (Accad et al., 2005; Pullar et al., 2005; Rochester et al., 2004).

These three spatial partitioning contexts are discussed in detail in Section 3.

The quality of a spatial partitioning depends on the environmental variables used to create it. Some researchers advocate using a bottom-up approach, where spatial partitioning is based on biological data, such as macroinvertebrates or fish (Marshall et al., 2006a; Wells et al., 2002). Other groups promote the use of a top-down spatial partitioning, which is based on broad-scale environmental characteristics, such as geology type or climate, thought to influence or constrain biological community composition (Bailey, 1983; Bailey et al., 1995; Omernik, 1987b; Pantus and Dennison, 2005; Smith et al., 2001). Both approaches have support in the literature, but the bottom-up approach may be more difficult to implement (Barton and Metzeling, 2004). For example, the bottom-up approach is dependent upon reference sites, which may not exist or may not have been surveyed. Furthermore, bottom-up approaches may work well at regional scales, but fail to represent finer-scale variability that occurs between neighbouring catchments (Barton and Metzeling, 2004).

There is some evidence that top-down approaches based on broad-scale factors, such as climate or topography are related to biotic variation at sites, but this finding is inconsistent (Hawkins et al., 2000). Interestingly, the ability to explain variability in biotic communities appears to increase as the scale of the spatial partitioning decreases (Hawkins and Vinson, 2000; Metzeling et al., 2006; Waite et al., 2000). The top-down approach has the advantage of being practical to implement because it is not based on site-scale data. Geographic information system (GIS) datasets of land use, topography, or climate, are currently available for most parts of Australia, which facilitates the implementation of a top-down approach at the regional scale.

Many spatial partitioning methodologies currently in use are qualitative in the sense that they rely entirely on expert opinion to draw boundaries (Omernik, 1987b; Snelder and Biggs, 2002). Despite the popularity of this approach, it is difficult to perceive differences and delineate homogenous regions when a large number of variables are being considered in the elicitation process (Kaufman and Rousseeuw, 2005). It is also difficult to produce an objective and reproducible qualitative spatial partitioning because two people can be expected to evaluate region membership differently as a result of their prior biases, theories, and background (Underwood et al., 2000). An example of this is the approach taken by Shears et al. (2004) who compared quantitatively and qualitatively defined habitat types on shallow temperate subtidal reefs in New Zealand using biological data to determine whether the two methods produced biologically meaningful results. Their main finding was that although both methods identified some previously defined habitat types, the quantitative approach was able to identify additional classes that were found to be biologically distinct.

In the past, researchers may have relied on qualitative methods because quantitative methods were unavailable or it was not clear what combination of statistical methods could be used to achieve an ecologically plausible result. However, the computational capacity of computing systems has increased enormously and many GIS datasets are easily acquired at the regional scale. Furthermore, the science of quantitative classification has also rapidly developed over the last 30 years and more recent approaches have focussed on using pattern analysis techniques such as clustering and ordination (Digby and Kempton, 1987; Gauch, 1982; Mardia et al., 1979). As pointed out in Section 3, these approaches can also be considered subjective because they depend on the type of distance metric and clustering algorithm chosen. Furthermore, the approach results in a “hard” style of clustering, where sites are allocated to clusters rather than providing a probability of membership.

The “soft” style of clustering is an interesting alternative. This results in an assignment of a probability of membership to sites using a mixture model (Fraley and Raftery, 2002; Kuhnert, 2003; Rasmussen, 2000). An extension of this approach uses expert opinion to define the regional boundaries through priors that describe the shape, location and size of clusters (Accad et al., 2005; Pullar et al., 2005; Rochester et al., 2004). Given recent advances in quantitative classification, it seems prudent to incorporate expert opinion into a quantitative framework in order to account for any lack of understanding about a complex natural environment.

In this section, we consider two top-down approaches based on quantitative methods. The first is a freshwater spatial partitioning methodology that accounts for hydrologic connectivity and nested catchments by incorporating spatial partitioning attributes at three spatial scales: the reach, the reach contributing area (RCA), and the catchment. The second is a methodology targeted at marine zones, where grid or pixel based measurements are recorded over a specific marine spatial extent. Although we advocate using expert opinion in defining each spatial partitioning, we only use it to validate the regions identified through the relevant clustering techniques in this report. Investigation of its use in a more formal setting appears in the discussion.

5.2 Methods

5.2.1 Freshwater Spatial Partitioning

Background

It has long been recognised that the surrounding landscape influences stream condition (Hynes, 1975). Catchment attributes have been shown to be correlated with chemical (Davies et al., 2000; Nelson et al., 1992; Richards et al., 1996) and physical habitat (Davies et al., 2000; Jeffers, 1998; Maddock, 1999; Osborne and Wiley, 1988), as well as, biological communities (Dauer et al., 2000; Snyder et al., 2005). Freshwater ecosystems provide a challenge for spatial partitioning because streams form dendritic networks, catchments are nested within one another, and the ecosystem is connected by flow (Dunne, 1978). Nevertheless, these are fundamental characteristics of a freshwater ecosystem, which affect every aspect of stream condition (Fagan, 2002; Olden et al., 2001; Vannote et al., 1980). Therefore, we contend that it is vital to represent these processes in a freshwater spatial partitioning methodology.

The majority of freshwater spatial partitioning currently in use fail to represent the unique characteristics found in freshwater ecosystems (Bailey, 1983; Barton and Metzeling, 2004; Chiang et al., 2002; Myers et al., 2006; Omernik, 1987a; Rao and Srinivas, 2006; Rochester et al., 2004; Wardrop et al., 2005; Wolock et al., 2004). They produce contiguous regions that are not constrained by catchment boundaries (Detenbeck, 2000), which may not be appropriate given the dendritic nature of stream networks and the influence that network position has on the ecosystem (Fagan, 2002; Schlosser, 1991). For example, the biological community found that two first order (Strahler, 1957), or headwater streams, may be more similar than those found in first and third order streams, regardless of their proximity (Waite et al., 2000). For that reason, it may not be useful to force dendritic stream networks into spatially contiguous regions.

Catchments provide a natural hierarchical framework in which to structure a spatial partitioning (Fausch et al., 2002; Frissell et al., 1986; Poff, 1997) and provide an appropriate spatial unit for management and assessment (Omernik and Bailey, 1997; Wardrop et al., 2005). However, upstream catchments are hierarchically nested within larger downstream catchments, which in the past has made them time-consuming to calculate and difficult to represent visually across a large area. Consequently, it is not uncommon to see spatial partitionings performed on a set of spatially disjoint catchments (Chiang et al., 2002; Myers et al., 2006; Rao and Srinivas, 2006) or on interbasin areas (Barton and Metzeling, 2004; Wardrop et al., 2005; Wolock et al., 2004). An interbasin, or hydrologic unit, differs from a true catchment because it does not always include the entire area that flows to an outlet point (Griffith et al., 1999). Yet, we believe that there are serious disadvantages to both of these approaches.

Using disjoint catchments results in a spatial partitioning with gaps that must be given a “no region” assignment (Chiang et al., 2002; Myers et al., 2006) or “filled” in based on the neighbouring region assignments (Rao and Srinivas, 2006). However, a Euclidean neighbourhood may not be optimal if we are interested in hydrologic relationships (Benda et al., 2004; Ganio et al., 2005; Olden et al., 2001). A tessellation of interbasins obviates the need to fill gaps, but it fails to represent flow connectivity, which is one of the fundamental aspects of freshwater ecosystems (Allan, 1995). Catchment nesting and hydrologic connectivity should be an integral spatial component in a freshwater spatial partitioning, rather than removing the effects for computational convenience.

The spatial framework developed by Snelder and Biggs (2002) appears to be the most environmentally relevant framework for spatial partitioning in freshwater streams. The River Environment Classification (REC) is a hierarchical top-down approach used to classify individual stream segments. It is comprised of a tessellation of interbasins, with each stream segment explicitly linked to a single interbasin. The interbasins are also linked to one another using the structure of the stream network, which allows upstream interbasins to be identified and catchment attributes calculated for each segment in the stream network. This method is especially useful since there has been evidence that an approach based on a finer-scale spatial unit, such as stream reaches or small catchments, is more appropriate than broad contiguous spatial partitionings for freshwater ecosystems (Hawkins and Vinson, 2000; Metzeling et al., 2006; Waite et al., 2000). The ability to use the stream segment, which represents both local and upstream conditions, as a spatial unit for spatial partitioning is, in our opinion, an enormous improvement over previous methods. It enables differences attributable to network position and hydrologic connectivity to be represented and it accomplishes this without removing the effect of catchment nesting.

Although Snelder and Biggs (2002) have created an ecologically relevant spatial framework for spatial partitioning , their methodology is not statistically rigorous. Their approach is qualitative and entirely dependent upon expert opinion. Catchment and local characteristics are categorised at each of the six classification levels based on a list of “category membership criteria” (Snelder et al., 2004). Information is lost when continuous variables, such as temperature or percent land use, are categorised and the categories may be somewhat arbitrary. For example, if a catchment is 51 percent forested and 49 percent bare is it really a forested catchment? This is an extreme example, but it demonstrates how artificial categories based on expert opinion may not be appropriate for spatial partitioning.

Rochester et al. (2004) attempt to combine expert opinion about regional boundaries for South-East Queensland with statistical methods that attempt to structure clusters around environmental variables. However, their approach operates at the grid scale and does not incorporate hydrological relationships important in freshwater systems.

Recently developed tools for ArcGIS v9 have made it much easier to implement the spatial framework developed by Snelder and others (2002) over large areas (Theobald et al., 2005). These tools allowed us to use a modified version of the spatial framework developed by Snelder et al. (2002) in concert with clustering methodologies to come up with a more ecologically and statistically valid approach to freshwater spatial partitioning.

GIS Methodology

Landscape Network

The Landscape Network (LSN) is the spatial framework for the spatial partitioning and it is created using the Functional Linkage of Watersheds and Streams (FLoWS) toolset (Theobald et al., 2005) developed for ArcGIS v9. These tools allow the user to extract catchment characteristics derived from ancillary data sources, such as geology or elevation, at multiple scales for each stream reach in the network. The LSN provided a means of generating spatial partitioning attributes at three spatial scales: the reach (Figure 18a), the reach contributing area (RCA) (Figure 18b), and the catchment (Figure 18c).

Hydrologic Network.

The hydrologic network is a topologically corrected stream network made up of reaches. For our purposes, we define a reach as a single polyline segment in the GIS dataset (Figure 18). We used the FLoWS toolset to ensure that there were no topological errors in the streams dataset and that the reaches were digitized in the downstream direction (Theobald et al., 2005). These edits were necessary to facilitate the modelling of flow through the network.

Reach

Reach attributes (Figure 18a) were calculated by converting the hydrologic network to a raster, with each cell containing a reach identifier. If the cell did not represent a stream reach it was given a “no data” value. Reach attributes were calculated and stored in the hydrologic network attribute table.

Reach Contributing Areas (RCA)

An RCA is a detailed tessellation of non-overlapping, edge-matching interbasins (Figure 18b). The RCA boundaries are based on topography and include nearby areas that would theoretically contribute overland flow, where it occurs to a given reach (Theobald et al., 2005). There is a one-to-one relationship between reaches and RCAs, which allows the surrounding landscape to be explicitly linked to the stream reach. We used the FLoWS toolset to delineate the RCAs based on the hydrologic network and the local topography. The RCA attributes were calculated and stored in the hydrologic network attribute table.

Catchments

Catchment attributes were calculated for each stream using the RCA attributes and the topology of the hydrologic network (Figure 18c). The FLoWS toolset provides a function which allows the user to “accumulate” reach attributes downstream (Theobald et al., 2005). This function provided a way to calculate a suite of catchment attributes for the downstream node of each stream segment. This information was stored in the attribute table of the hydrologic network.

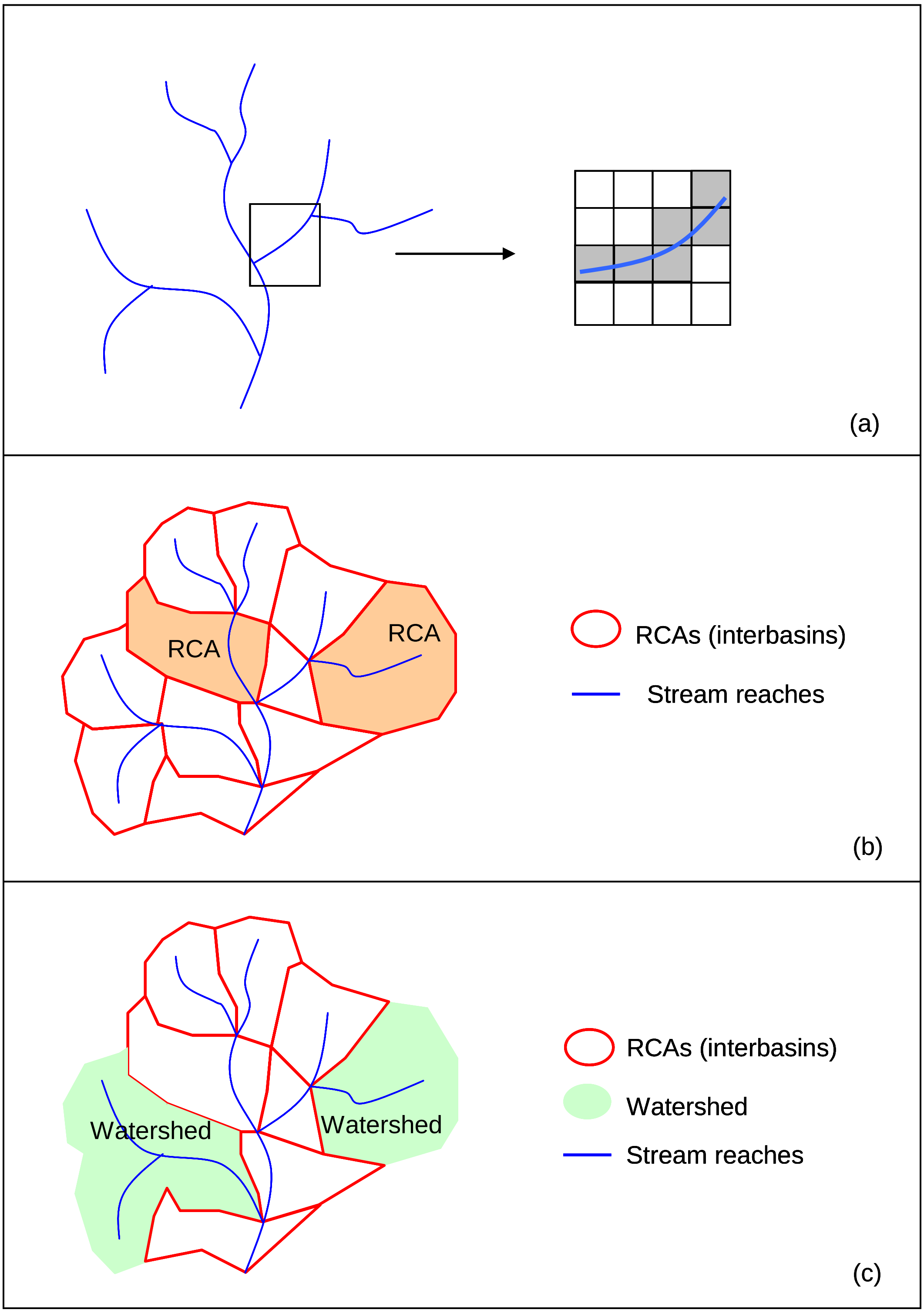


Figure 18: The three scales of freshwater spatial partitioning attributes: the reach (1a), the RCA (1b), and the catchment (1c). Each stream reach has three levels of attributes, which are stored in the hydrologic network attribute table.

5.2.2 Marine Spatial Partitioning

Background

Unlike the freshwater spatial partitioning, the approaches applied to marine zones tend to operate at one scale. For example, the data may consist of remote sensing information captured on a grid where each pixel may represent a 1km by 1km area. Alternatively, we may have site data collected from a carefully planned experimental design of the marine area under investigation.

In the past, many marine spatial partitioning approaches have relied solely on expert opinion to define regions. This typically involves a panel of experts, which come together to discuss methods for subdividing the region into ecologically meaningfully pieces. Although data such as benthic and spatial characteristics, or biota (if a biospatial partitioning is considered), are taken into account in the discussions, they are not formally integrated into any statistical analysis. Examples of these types of approaches include the work of Butler et al. (2001), who relied on expert opinion and statistical summaries of the data and model outputs to define a biologically relevant spatial partitioning for the south-east marine region of Australia.

Other approaches have used statistical methods to identify appropriate regions in collaboration with experts. Here, the experts are either used to:

- identify appropriate variables that are incorporated into the cluster analysis, and/or
- assess the validity of the regions that are suggested by the statistical approach.

An example of this approach includes the work of Snelder et al. (2007) who develop a very detailed statistical methodology for selecting environmental variables that are used to cluster marine areas of New Zealand. Finally, although less statistically rigorous, there are a number of applications that have purely relied on statistical summaries and visual displays of the data through GIS to define suitable boundaries (DEH, 2005; Heap et al., 2005).

Rochester et al. (2004) demonstrated an approach whereby they formally incorporated expert information into a Bayesian mixture model through priors. Their approach, although applied to a freshwater system was applied to data captured by GIS layers and therefore could be implemented in a marine context.

We use statistical clustering techniques that are outlined below to develop a spatial partitioning for marine zones in the GBR (See Section 6). The regions identified are ideally discussed with experts to determine their validity with respect to the study area.

5.2.3 Statistical Methods for Spatial Partitioning

The statistical methods used for the freshwater and marine spatial partitioning are very similar and will be summarised in this section for brevity. Section 6, presents an application of these methodologies to a GBR pilot area, the Tully-Murray catchment and marine zones.

Exploratory Analysis

The spatial partitioning attributes were standardised to avoid dependence on the measurement unit (Kaufman and Rousseeuw, 2005). We used Spearman's rank correlation statistic and Hoeffding's D statistic (Hoeffding, 1948) to identify and remove effectively redundant variables. A certain degree of collinearity was expected and tolerated since RCA and catchment attributes will always be identical for headwater streams (Figure 18b & Figure 18c). Principal Component Analysis (PCA) (Jolliffe, 1986; Mardia et al., 1979; Townend, 2003) allowed us to identify linear combinations of the standardised variables that economically account for the as much variability as possible in the data. This technique was used as an exploratory tool to identify and remove additional variables that did not contribute any considerable proportion of the total variability.

Cluster Analysis

We had no prior knowledge about the appropriate number of clusters to choose. Consequently we tested a variety of algorithms, metrics, and methods. Hierarchical clustering algorithms provide useful information about the cluster structure because the number of clusters is driven by the data rather than the user (Kaufman and Rousseeuw, 2005). We explored three hierarchical clustering algorithms: two agglomerative methods (AGNES (Kaufman and Rousseeuw, 2005) and hclust (Everitt, 1974; Hartigan, 1975) and one divisive method (DIANA) (Kaufman and Rousseeuw, 2005), which provided information about the number and definition of the clusters.

Distance metrics coupled with a clustering algorithm produce a variety of cluster shapes such as elongated, spherical, or circular and this characteristic can affect the cluster membership (Kaufman and Rousseeuw, 2005). We tested a suite of metrics and clustering algorithms based on the standardised data to help identify common clusters. The distance metrics investigated included the Euclidean distance, McQuitty (McQuitty, 1966), Gower (Gower, 1971) and the Bray-Curtis metrics (Bray and Curtis, 1957). Clustering algorithms explored included complete-linkage, average and Wards method (Kaufman and Rousseeuw, 2005).

Choosing the Number of Clusters

Determining the number of clusters can be highly subjective. As a result we used a combination of qualitative and quantitative methods to inform the decision process.

We started by visually examining the hierarchical cluster structure to determine whether a common pattern emerged. We used two partitioning algorithms: a partitioning around medoids (PAM) and a fuzzy analysis (FANNY) algorithm (Kaufman and Rousseeuw, 1990), for a range of cluster sizes using different distance metrics, namely the Euclidean distance, Gower, and Bray-Curtis metrics.

We examined plots showing the average silhouette width (a measure of dissimilarity) for each cluster number to determine which algorithm, metric, and cluster size provided the best separation into cluster groups (that is, maximized the separation between clusters and minimized the separation within clusters). Silhouette widths close to one are an indication that sites are well clustered, while those close to zero or negative, indicate a poor separation of sites into groups. Kaufmann and Rousseeuw (1990) offer informal decision rules to assist in the choice of the number of clusters.

We also utilized the Gap statistic (Tibshirani et al., 2001a), a technique that was developed to examine a range of clustering solutions for any metric and clustering algorithm applied to the data with the aim of selecting the optimal cluster size. This is achieved by calculating the cluster dispersion and comparing it to what is expected under a reference or null distribution.

Variable Importance and Dimension Reduction

We felt that it was important to examine the influence of each variable on the structure of the clustering solution that we arrived at. Once the regions were constructed, we examined the variables input into the cluster analysis to determine which had the strongest influence on cluster definition. We achieved this using two exploratory approaches.

The first used a stepwise greedy approach which involved performing a cluster analysis (such as PAM) and using one variable at a time, computing the average silhouette width which corresponded to a cluster analysis for the required number of groups. At each iteration, the variable that gave the best separation into the desired number of groups was chosen. This process was then repeated until all variables entered into the analysis.

In general, we found the average silhouette width decreased as more variables were included in the analysis. As a result, we were able to omit some variables (without changing the cluster structure) due to their performance in separating out the data into similar groupings.

We should emphasise that this approach is exploratory and represents a means for examining a variable's influence in partitioning the data. We used this approach to identify key variables since there was no real expert information to inform us about ecologically important variables for spatial partitioning . We also acknowledge that by removing variables from the analysis, although we may be producing clusters that are well separated (and less fuzzy), they may have less ecological relevance due to the removal of what is considered important variables. To avoid this, we recommend performing an analysis with and without the variables to examine differences in the cluster structure. For the marine example presented in this report, we found that the removal of variables did not alter the number, size and shape of clusters produced.

The second approach involved fitting a random forest of classification trees (Breiman, 2001) to explore the relationship between the cluster membership and the environmental variables representing the covariates. At first glance this approach may appear to be re-using the environmental data in a possibly questionable way. However the random forest merely supplies an alternative way of giving a rule-based specification of the groups already defined, with two additional properties that the original specification lacks, namely:

1. It allows us to classify additional points into the groups defined by the original analysis without re-doing the analysis; and more importantly
2. It allows us to assess the importance of the contributing variables to the clustering itself.

At the time of writing this report, it has come to our attention that random forests can be used to conduct unsupervised learning (Liaw and Wiener, 2002) and return a proximity matrix which can be used as a similarity measure and input into a clustering or multi-dimensional scaling framework. This presents an alternative approach for defining a metric that makes use of all the data and highlights important variables. Although not implemented here, the approach will be examined more closely in future work.

5.3 Discussion

A robust and meaningful spatial partitioning is vital to any environmental monitoring program because it helps to ensure that an indicator's response to anthropogenic impacts will be interpreted correctly. To do this, the spatial partitioning must be both ecologically and statistically valid. We believe that the methodologies presented here achieve these two goals.

For the freshwater spatial partitioning, we modified Snelder and Biggs (2002) spatial framework to calculate continuous, rather than categorical variables. This is important because catchments, RCAs, and even reaches may be extremely heterogeneous units. As such, it may not simply be the dominant characteristic that is driving the system. A model based on categorical variables, such as dominant type, oversimplifies a complex system and may not be sensitive enough to describe the ecological relationships that we are attempting to model. Therefore, incorporating multi-scale continuous variables allows us to represent a complex mixture of characteristics, which are likely to affect the ecosystem.

Categorical variables make it much easier to qualitatively define regions when a large number of variables are used. This same technique would be difficult for continuous variables because differences become difficult to discern in more than three dimensions (Kaufman and Rousseeuw, 1990). Quantitatively defining the clusters provides evidence that regions based on continuous variables are truly distinct, in the sense that information redundancy is minimal. However, regions can be distinct without being ecologically meaningful. Therefore, the variables that are included in a spatial partitioning must be chosen based on expert knowledge concerning the dominant ecological processes of the area. No simple data-based rule for inclusion can achieve the desired ecological result.

In the case of the marine spatial partitioning, we proposed using standard clustering techniques in collaboration with experts to define suitable regions that explain differences in benthic, spatial and biotic characteristics (if applicable).

5.3.1 Application to Other Catchments in the GBR

Large-scale spatial partitioning are data intensive and require a significant amount of data storage space, processing time, and computational power. For example, the spatial framework used in the freshwater spatial partitioning would have been difficult to implement only a few years ago. However, the increased availability of regional GIS datasets, new software, and the general increase in computing power have enabled us to implement these methods in a relatively large area in a short amount of time. The two methodologies described here could be applied to catchments or regions in the GBR. In fact, we apply these methodologies to one GBR catchment, the Tully-Murray (Section 6) and investigate the performance of a variety of indicators used in reporting.

The spatial framework proposed for the freshwater spatial partitioning could be scaled to the entire GBR at a coarse spatial resolution. There are a number of regional datasets available for this purpose: the Australian Soil Resource Information System (ASRIS) (McKenzie et al., 2005), SILO climate data (Jeffrey et al., 2001a), and geology type (Department of Natural Resources Mines and Energy, 2004). Geoscience Australia currently provides a 9 second DEM (approximately 250 metre spatial resolution), which they suggest using for modelling surface shape and drainage structure (Geoscience Australia, 2007). Streams generated from DEMs would make a more appropriate regional streams dataset because they contain a relatively minimal number of topographical errors and therefore. In addition, the GBR is composed of multiple catchments, which would allow the regional streams dataset to be generated in pieces.

5.3.2 Extensions and Future Work

We used a very straightforward clustering approach for both the marine and terrestrial spatial partitioning approaches which involved using standard clustering techniques, redundancy methods to determine which variables were important in finding cluster structure and expert opinion (where relevant) as a final check to determine whether the cluster structure found makes ecological sense.

Extensions to this approach, which we could not implement due to time restrictions include:

- Incorporating expert opinion as a prior in a Bayesian clustering algorithm. We are currently investigating Rasmussen (2000) as a suitable approach for the clustering of sites and its feasibility and flexibility of incorporating expert information.

- Incorporating measurement error into the clustering approach. In the freshwater spatial partitioning the challenge would be to define a method to propagate the error at the reach scale right up to the catchment, rather than using data at three separate spatial scales.
- Incorporating marine data at different spatial scales (grid based and site measurements).
- Examining a random forest approach by Liaw and Wiener, (2002) to defining a “proximity” metric that can be input into a clustering or multidimensional scaling algorithm.

6. Case Study: the Tully/Murray catchment

6.1 Background

The Tully-Murray catchment is part of the Wet Tropics of North Queensland. The catchment includes the Tully, Murray and Hull rivers and their tributaries and comprises an area of approximately 2,910 square kilometres (Figure 19). The primary river is the Tully River which is approximately 130 kilometres in length and extends from the Cardwell Range to Rockingham Bay, which is the main outlet to the sea. Dominant land uses are grazing, horticulture, plantation forestry, sugarcane and nature conservation. Areas in the upper parts of the catchment which have higher elevation are predominantly taken up by rainforest (Bohnet et al., 2007).

The primary reason for using the Tully catchment as a case study catchment for the IRC was due to the major investment in catchment and water research and monitoring being undertaken as part of the Department of Environment and Water Resources Coastal Catchments Initiative (CCI) program, as well as within the CSIRO Water for a Healthy Country Flagship program (see Table 6) and the MTSRF indicator development and evaluation projects for 2006/2007. The programs within the Tully Catchment are at various stages in their lifecycle and many of the projects are still in developmental phases, and therefore access to data was not as easy as first thought. A lot of the water quality data that has been collected in the Tully was done using a range of different objectives, using a range of field and lab techniques, and the custodians of the data are currently not prepared to share this data for fear of inappropriate interpretation. It is also important to point out that many sub-catchment and paddock scale data collection has been carried out at sites due to the willingness of the landholders to be involved, yet these sites may not be appropriate for a statistically robust monitoring design due to spatial location within the catchment (i.e. not necessarily downstream of important/appropriate monitoring points). Despite the initial difficulties with access to data, we were able to obtain sufficient data to demonstrate a range of statistical approaches for “spatial partitioning”, indicator assessment and integration.

Table 6: Key research projects being conducted in the Tully as part of the WQIP and WfHC projects (source: Frederieke Kroon, CSIRO).

Project	Project leader	Organisation
Interim Marine WQ guidelines for the GBR Marine Park	L. Gray	GBRMPA
GBR Marine monitoring program	J. Prange, B. Schaffelke	GBRMPA, AIMS
Receiving water model	J. Brodie	JCU
AIMS and BSES long term data analysis	J. Brodie	JCU
Sub-catchment land use monitoring	J. Brodie	JCU
Tully-Murray point sources	B. Roberts	Consultant
Sediment finger printing	P. Nelson	JCU
Overbank flow i.e. Mike 21 modelling	J. Wallace	CSIRO
SedNet/ANNEX modelling	J. Armour	QNRW
WQ effects of management practices in cane and bananas	J. Faithful	JCU
Nitrogen fertilizer management in sugar	T. Webster	CSIRO
Phosphorus fertiliser management of bananas	S. Lindsay	
Sediment run-off and constructed wetlands	D. Green	
Pesticides in cane review	D. Calcino	
WQ and Tully river rock armouring	R. Lait	
Wetland filter function review	D. McJannet	CSIRO
Wetland audit and prioritisation	D. Sydes	
Tully BMP benchmarking	P. Roebeling	CSIRO
BMP review – Effectiveness in WQ improvement	P. Roebeling	CSIRO
BMP review – Plot level financial-economic assessment	P. Roebeling	CSIRO
BMP review – Drivers promoting the adoption of BMPs	P. Roebeling	CSIRO
BMP review – Spatial prioritisation	P. Roebeling	CSIRO
BMP review – Efficient water quality improvement targets	P. Roebeling	CSIRO
BMP review – Effect of future growth and climate change	P. Roebeling	CSIRO
SMART partnerships	C. Robinson, B. Taylor	CSIRO

6.1.1 Catchment and Freshwater Data

We obtained GIS data for the catchment, which included information about streams (Geosciences Australia, 2004), geology (DNRME, 2004), soils (McKenzie et al., 2005; Wilkinson et al., 2004), climate (Jeffrey et al., 2001b), and topography (Smith and Brough, 2006) (Appendix 2). These datasets were pre-processed to ensure that they had a common projection and spatial scale.

Stream reach attributes, such as percent geology type or mean elevation, were calculated at three spatial scales: the reach, the reach contributing area (RCA), and the watershed as described in Section 5 (see Figure 18).

In total, eighty-five variables were calculated based on these datasets and these were included in the freshwater spatial partitioning (Appendix II). We deliberately chose variables, such as topography or climate, which are not considered to be anthropogenically affected at short or medium time-scales. This was important because our goal was to create a spatial partitioning that was sensitive to changes at freshwater sites caused by anthropogenic impacts.

Site specific information about freshwater condition was also available for analysis and assessment. The State of Rivers reporting (Johnson, 1998) describes the ecological and physical condition of the Tully and Murray rivers. We examined two indicators from the State of Rivers reporting process, namely riparian vegetation and macrophyte percent bed cover. Modelled output from the catchment is available in the form of total suspended sediment (TSS) yield as captured by the SedNet model (Hately et al., 2006). This is reported in tonnes per year and can be evaluated at a number of sites in the Tully-Murray catchment area. We also used the modelled TSS as a third indicator in our demonstration of an assessment for the Tully-Murray catchment and freshwater region.

Figure 19: Map of the Tully-Murray river catchments (indicated by the dotted black and yellow lines) and marine zone (indicated by the solid orange line) (Source: Bohnet et al., 2007).

6.1.2 Marine Data

The marine extent of the Tully-Murray catchment region was determined by the Tully-Murray Water Quality Improvement Plan (WQIP) (Bohnet et al., 2007) and is highlighted by the solid orange line in Figure 19. Marine data for the Tully-Murray consists of three types:

- Pixel based measurements (1km x 1km) consisting of bathymetry, aspect, slope, seabed stress and spatial variables representing the distance across the coast and the distance out to the furthermost reef (Pitcher et al., 2007). This data was used in the spatial partitioning process to identify regions with similar physical characteristics (not subject to anthropogenic pressures). The data is summarised in Figure 24.
- MODIS satellite imagery of near-surface concentrations of chlorophyll-a, suspended solids and vertical attenuation of light coefficients (Schaffelke et al., 2006). We were given a median twelve month summary of chlorophyll-a, vertical light attenuation and suspended solids. Median values have also been provided for both wet (November – April) and dry (May – October) periods (not used in this report). This data capture was

based on measurements taken from the 1st of November 2004 and the 30th of October 2005 and is visualised in Figure 25.

- Monitoring data collected for a variety of indicators were captured in the Tully-Murray marine region. The data represents medians taken from two samples (one wet and one dry) at three sites within the marine extent shown in Figure 19. Information captured consists of water quality, coral, pesticides and seagrass (GBRMPA data, 2007). Although made available at the time of constructing this report, this dataset was not used because of the small number of samples taken. Table 7 summarises the two main datasets used in the Tully-Murray marine analysis.

Table 7: Summary of variables available for analysis in the Tully-Murray marine region.

Spatial Partitioning		
Variable	**Range of Values**	
	Min	**Max**
Topography		
bathymetry	-35.682	-0.234
aspect	12.82	339.412
slope	0.007	2.553
Seabed Stress	0.002	0.462
Spatial		
distance along the coast	0.000	0.465
distance out to reef	0.000	0.377

Indicators		
Variable	**Range of Values**	
	Min	**Max**
Chlorophyll-a		
12 month	0.051	4.220
dry (May-Oct)	0.041	4.857
wet (Nov-Apr)	0.000	4.991
SS		
12 month	0.233	20.838
dry (May-Oct)	0.550	0.936
wet (Nov-Apr)	0.000	12.834
Light Attenuation		
12 month	0.534	0.875
dry (May-Oct)	0.550	0.936
wet (Nov-Apr)	0.000	0.822

[1] Suspended sediment (SS) in this context refers to the number of non-algal particles.

6.2 Spatial partitioning

6.2.1 Freshwater

The Spatial Framework

We developed a landscape network to calculate spatial partitioning attributes at three spatial scales: the reach, the reach contributing area (RCA), and the watershed (see Section 5). The hydrologic network was developed using a 1:100,000 streams dataset containing 2906 stream reaches (Appendix 2). Reach attributes were calculated by converting the hydrologic network to a 25 metre raster, with each cell containing a reach identifier. We calculated zonal statistics, such as mean slope or total area, for each variable and stored the value in the hydrologic network attribute table.

We used the Functional Linkage of Watersheds and Streams (FLoWS) toolset (Theobald et al., 2005) to delineate the RCAs using the hydrologic network and a 25 metre DEM as input (Appendix 2). There were 61 stream reaches that were not associated with an RCA and this occurred for a number of reasons. In a few cases the stream reach was less than 25 metres and the resulting RCA was too small to delineate using a 25 metre DEM. The RCA boundaries were not always delineated for every stream reach when a group of reaches were located in a waterbody, such as a lake or reservoir. Instead, the entire waterbody tended to be grouped within a single RCA and the area assigned to one stream reach. The

RCA boundaries for seven additional stream segments located along the coast were not delineated properly. The error occurred because the spatial extent of the streams dataset was smaller than the DEM, which extended further to the east. This caused the RCAs for the seven reaches to be unrealistically large, which did not represent the true contributing area. The RCA attributes for the remaining 2845 reaches were calculated and stored in the hydrologic network attribute table. Watershed attributes were also calculated for each stream reach using the RCA attributes and the topology of the hydrologic network (Peterson and Kuhnert, 2007).

The spatial framework was relatively fine given the size of the Tully-Murray watershed. The majority of the streams in the 1:100,000 Tully-Murray dataset represented headwater streams; with 1457 of the 2845 reaches (51 %) being 1st order streams. Most stream reaches were also less than 1.5 kilometres in length (Table 8). The RCA areas were generally less than 1 square kilometre in size, with the exception of a few large RCAs found in the central portion of the watershed. Watershed size ranged between 0.03 and 1466 square kilometres.

Table 8. Mean length and area measurements for the freshwater spatial partitioning spatial framework.

Variable	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
Reach length (km)	0.4136	0.7019	1.1520	1.3590	1.7350	10.75
Reach area (km^2)	0.0006	0.0219	0.0363	0.0429	0.0544	0.3406
RCA area (km^2)	0.0013	0.2994	0.5631	0.9156	1.0460	48.85
Catchment area (km^2)	0.0331	0.6406	1.5760	29.0600	5.8100	1466.00

Exploratory Data Analysis

We calculated 85 variables for the freshwater spatial partitioning , which included a variety of climate, geology, soils, and topography attributes (Appendix 2). The variables were standardised to have mean zero and unit variance to avoid dependence on the measurement unit (Kaufman and Rousseeuw, 2005). The data for the 61 stream reaches where RCAs were not delineated properly were removed from further analyses.

Twenty-seven of the original 85 variables displayed evidence of collinearity based on Hoeffding's D statistic (Hoeffding, 1948) and Spearman's Rank Correlation coefficient values ($\rho > 0.8, \Pr(\rho = 0) < 0.05$) and were removed from further analysis (Appendix 2). An examination of the PCA loadings (Principal Component Analysis) enabled us to identify a subset of variables that did not explain a considerable proportion of the total variance (Appendix 2). Variables which explained a very small proportion of the variance (less than 0.15 say) were removed from further analysis. This corresponded to 26 variables. The remaining 32 variables represented soil characteristics, climate, topography, geology, and erosion potential (

Table **9**). All of the climate variables, with the exception of mean maximum temperature in the watershed, were removed during exploratory data analysis. This is most likely due to the coarse spatial scale of the climate data (5 km) and the correlation that is commonly found between climate and elevation.

Cluster Analysis

We implemented three hierarchical clustering algorithms: two agglomerative methods (AGNES (Kaufman and Rousseeuw, 2005) and hclust (Murtagh, 1985)) and one divisive method (DIANA (Kaufman and Rousseeuw, 2005)), which provided information about the number and definition of the data clusters. The algorithms and metrics that we tested were outlined in Section 5.

Determining the number of clusters can be subjective and so we used a combination of qualitative and quantitative methods to inform the decision. We limited the maximum number of clusters to 10 because our goal was to identify fairly general regions within the watershed. A visual analysis of plots produced by the hierarchical clustering algorithms indicated that the optimal number of clusters was somewhere between 3 and 6. The cluster structure was consistently well-defined with the agglomerative coefficients greater than 0.91 and divisive coefficients greater than 0.94 for all hierarchical algorithms and metrics.

We used two partitioning algorithms: a partitioning around medoids (PAM) and a fuzzy analysis (FANNY) algorithm (Kaufman and Rousseeuw, 2005), to provide further quantitative evidence about the cluster number. Plots of average silhouette width produced by the FANNY algorithm using the Gower and Euclidean distance metrics indicated that four clusters produced the largest average silhouette width values of 0.45 and 0.30, respectively. The plot of the Bray-Curtis metric indicated that 3 clusters were optimal, but the average silhouette width was only 0.40. The results produced by the PAM algorithm were nearly identical to those of FANNY. The fuzzy four cluster solution based on Gower’s metric was deemed optimal (Figure 20) and all further results presented here are in reference to this cluster solution.

Table 9. Thirty-two variables used as input into clustering algorithms.

Scale	Variable	Description
Reach	perLoamRch	Percent loam soil type (%)
	perFelsicRch	Percent felsic rock type (%)
	perFelsputRch	Percent felsic plutonic rock type (%)
	perUnconsRch	Percent unconsolidated rock type (%)
	rchPhMn	Mean soil pH
	rchrklsMn	Erosion potential produced from the SedNet model
	rchSlopeMn	Mean slope at the reach (degrees)
RCA	perFelsicRca	Percent felsic rock type (%)
	perFelsputRca	Percent felsic plutonic rock type (%)
	perUnconsRca	Percent unconsolidated rock type (%)
	perLmsndRca	Percent loam sand soil type (%)
	perCllmRca	Percent clay loam soil type (%)
	rcaCecMn	Mean cation exchange capacity (cmol(+)/kg)
	rcaExbMn	Mean sum of exhangeable bases in the top soil layer (cmol(+)/kg)
	rcaOcMn	Mean organic carbon (%)
	rcaPhMn	Mean soil pH
	rcaClayMn	Mean percent clay (%)
	rcaRklsMn	Erosion potential produced from the SedNet model
	rcaElevMn	Mean elevation (m)
	rcaSlopeMn	Mean slope (degree)
Catchment	h2ORklsMn	Mean erosion potential produced from the SedNet model
	h2OClayMn	Mean percent clay in the soil (%)
	h2OTmaxMn	Mean annual maximum temperature (°C)
	h2OPhMn	Mean soil pH
	h2OOcMn	Mean organic carbon (%)
	h2OExbMn	Mean sum of exhangeable bases in the top soil layer (cmol(+)/kg)
	h2OCecMn	Mean cation exchange capacity (cmol(+)/kg)
	h2OElevMn	Mean elevation (m)
	perFelsicH2O	Percent felsic rock type (%)
	perFelsputH2O	Percent felsic plutonic rock type (%)
	perUnconsH2O	Percent unconsolidated rock type (%)
	perLmsndH2O	Percent loam sand soil type (%)

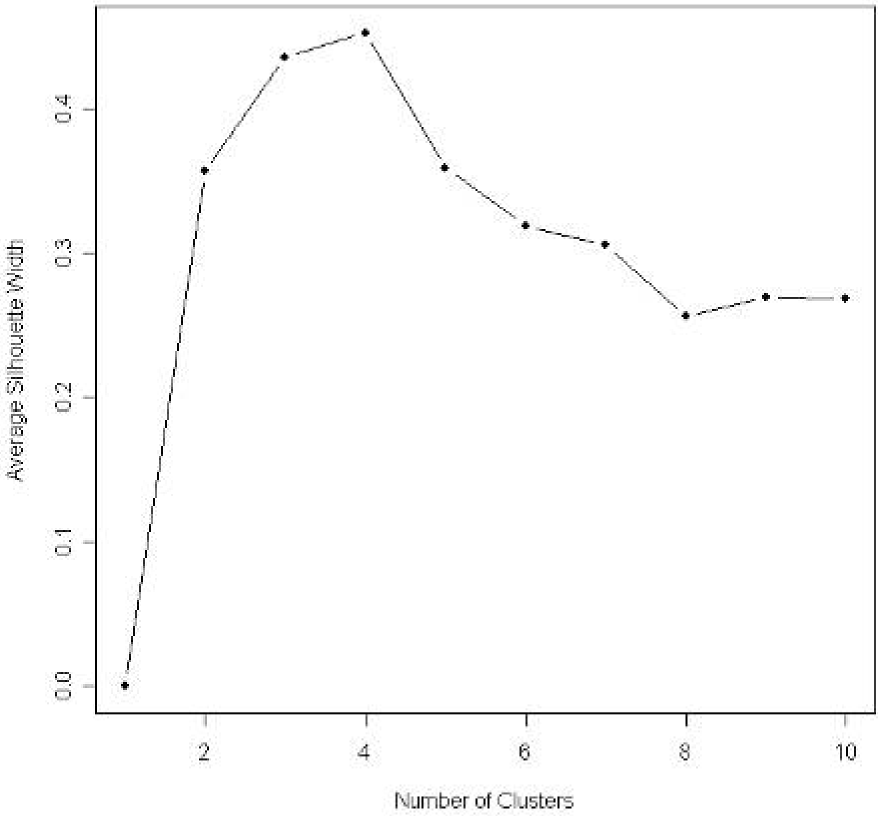


Figure 20. A comparison of average silhouette widths produced by the FANNY algorithm using a Gower dissimilarity metric.

We reclassified the output from the fuzzy four cluster model to artificially create a “hard” or single cluster membership for each stream reach based on the largest inclusion probability. The “hard” cluster membership was assigned to the RCAs, which are explicitly linked to individual stream reaches, to produce a map showing the regions that were developed for the Tully-Murray (Figure 21). The four regions were unequal in area and display a range of spatial patterns. Regions 1 and 4 are fairly contiguous and do not appear to follow drainage patterns closely. In contrast, regions 2 and 3 follow drainage patterns more closely with region 3 prevalent in lowland areas (mean watershed area = 72.76 km^2) and region 2 occupying upland areas (mean watershed area = 4.79 km^2).

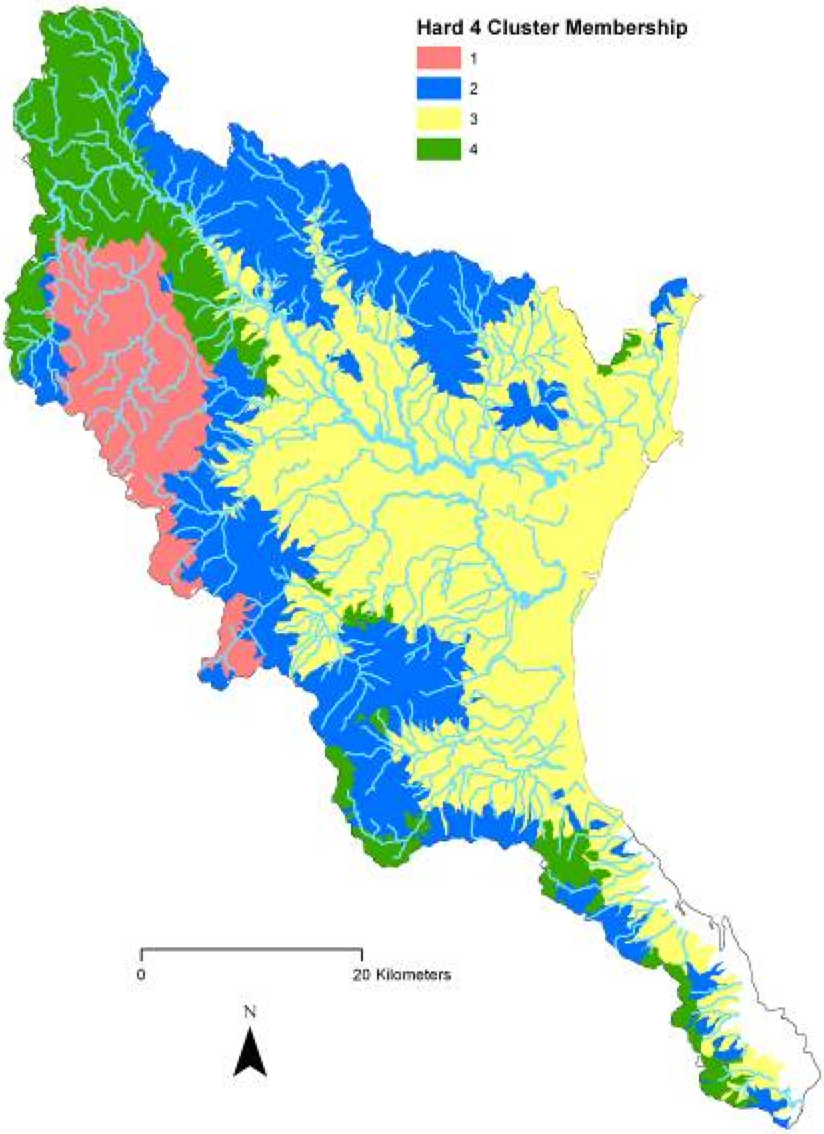


Figure 21. Hard 4 cluster regions for the Tully-Murray watershed based on fuzzy clustering using Gower's dissimilarity metric.

A random forest of 100,000 regression trees (Breiman, 2001) was fit using the hard cluster membership as the response and the 32 environmental variables as the covariates with the aim of exploring the importance of variables with regard to the cluster structure. All 32 variables were considered at each split. The importance values produced by the random forest regression tree showed that the most influential variables were found at the RCA and watershed scale, but that a small number of reach scale attributes were also important (Figure 22). The most influential included percent clay in the watershed, mean RCA slope, mean annual maximum temperature in the watershed, percent clay in the RCA, and mean slope in the reach (Figure 22).

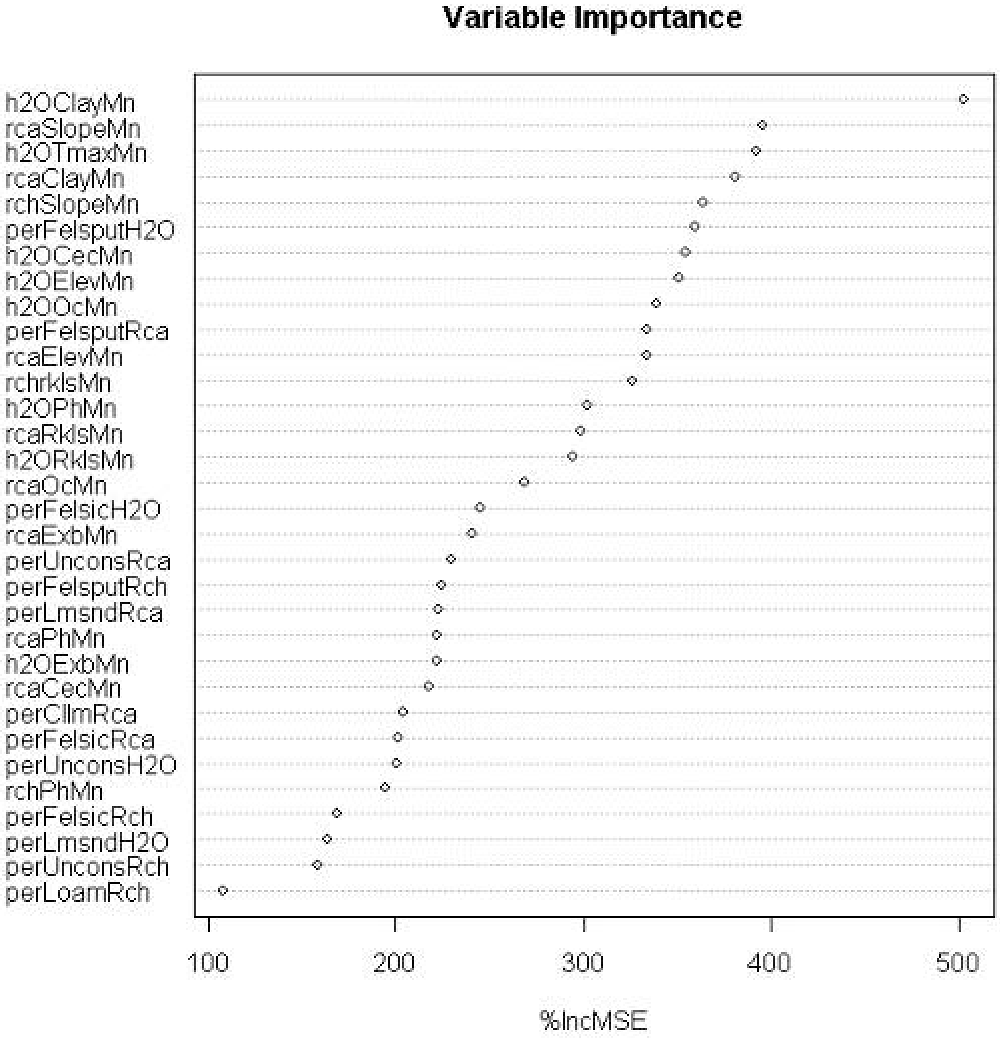


Figure 22. The percent increase in mean square error (MSE) for each variable if it were omitted from the analysis. Variables displayed appear in Appendix 2.

Hard cluster assignments make it is easy to forget that there is a level of uncertainty associated with each reach or RCA assignment (Figure 21). The average silhouette width for the fuzzy four cluster solution was less than 0.5, which indicated that there was in fact some overlap in the clusters. We used the inclusion probabilities produced by the fuzzy clustering algorithm to identify reaches with a high level of uncertainty (Figure 23). When we set the minimum acceptable inclusion probability to 0.7, 9.1% of the reaches did not meet the criteria. The majority of these reaches lie at the boundary of two or more regions and probably represent transitional areas between regions.

It is useful to be able to visualise and identify reaches with a high level of uncertainty associated with their region assignment. Simply being able to identify the areas allows the uncertainty to be taken into consideration when the indicators are assessed. Maps can also be used to elicit expert opinion concerning the appropriate region assignment and consequently refining the spatial partitioning used for indicator assessment.

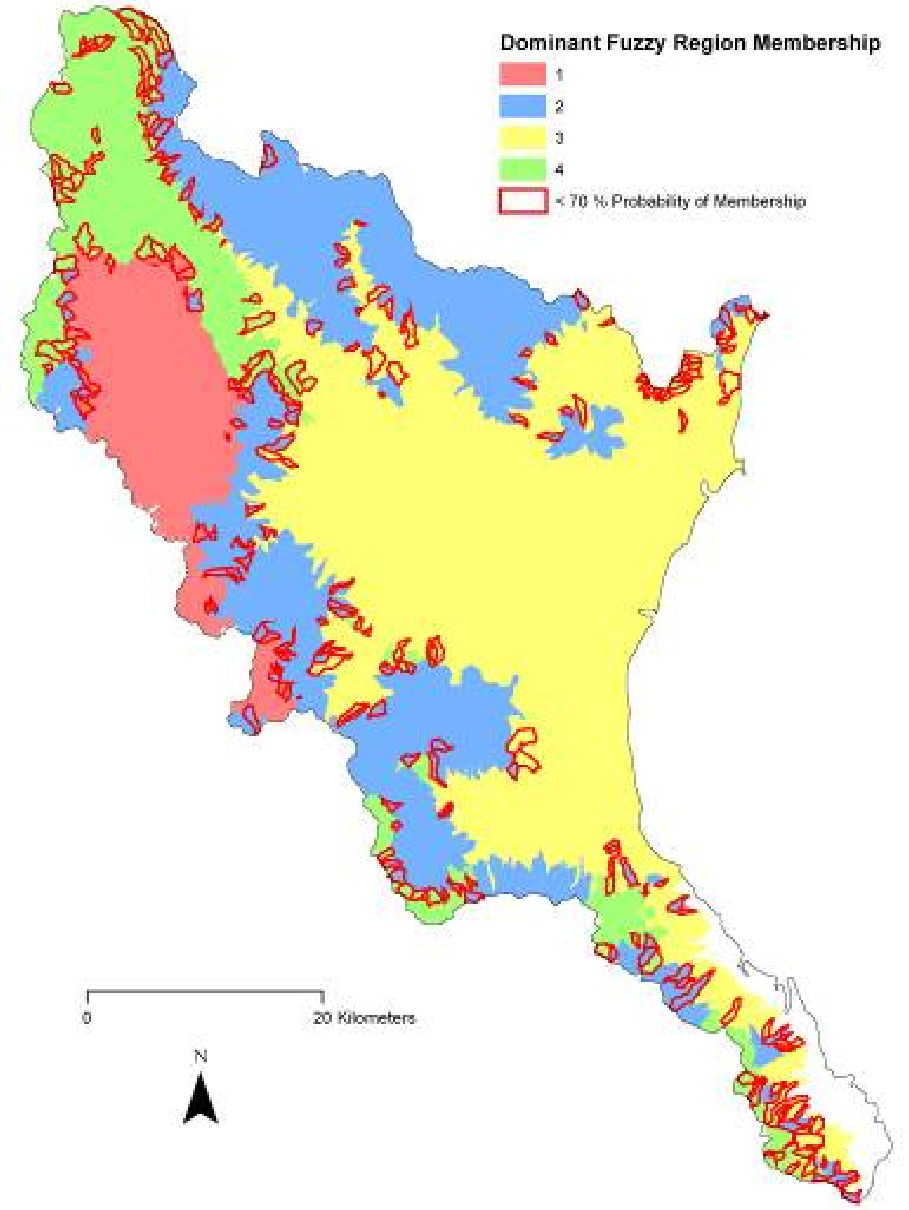


Figure 23. Uncertainty in region assignments found in the fuzzy four cluster solution.

6.2.2 Marine

We adopted a pragmatic approach and conducted an initial exploratory data analysis using a range of multivariate methods, including variable clustering via spearman rank correlations (which we refer to as a “correlation analysis”). These methods were described in Section 5.

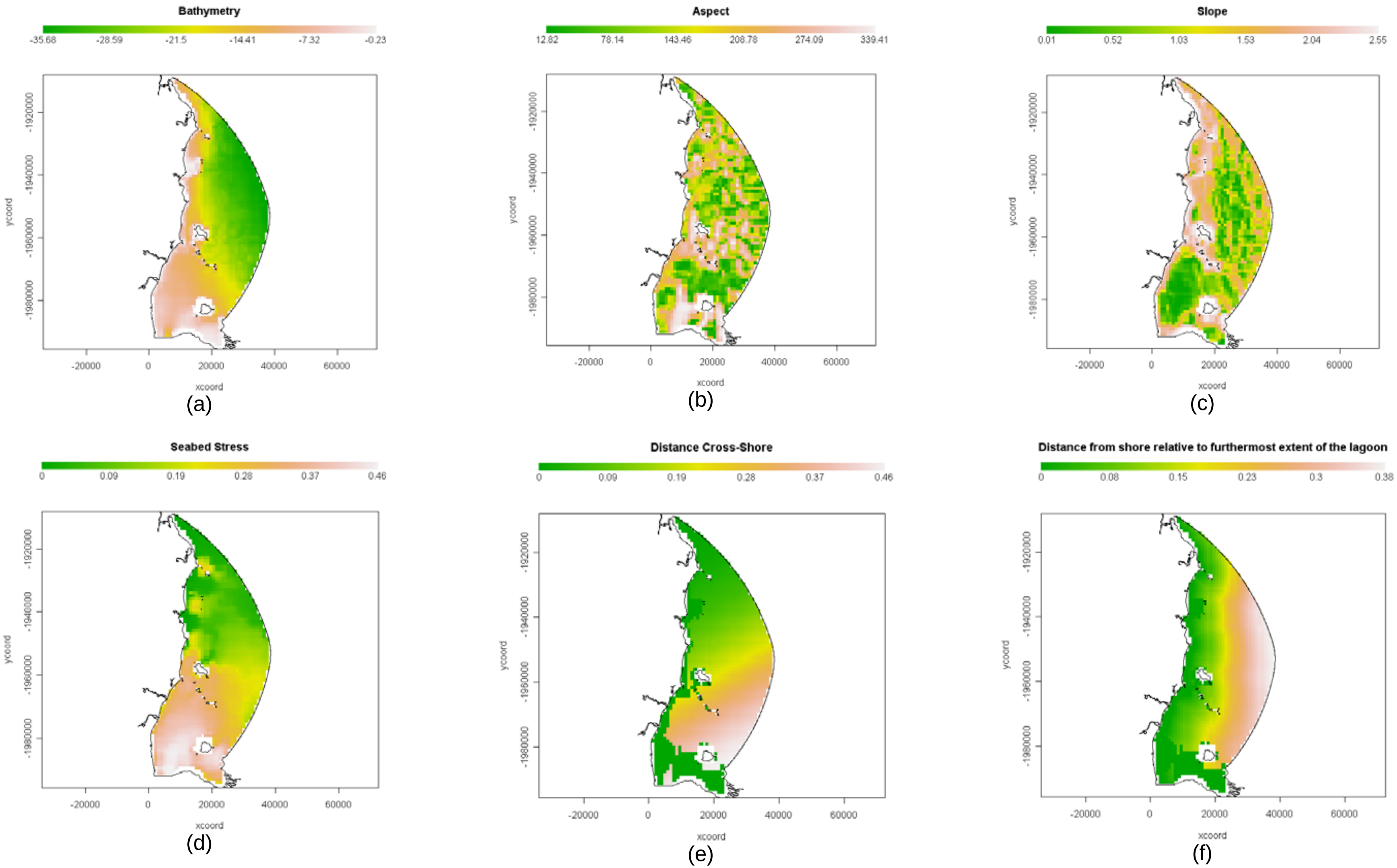
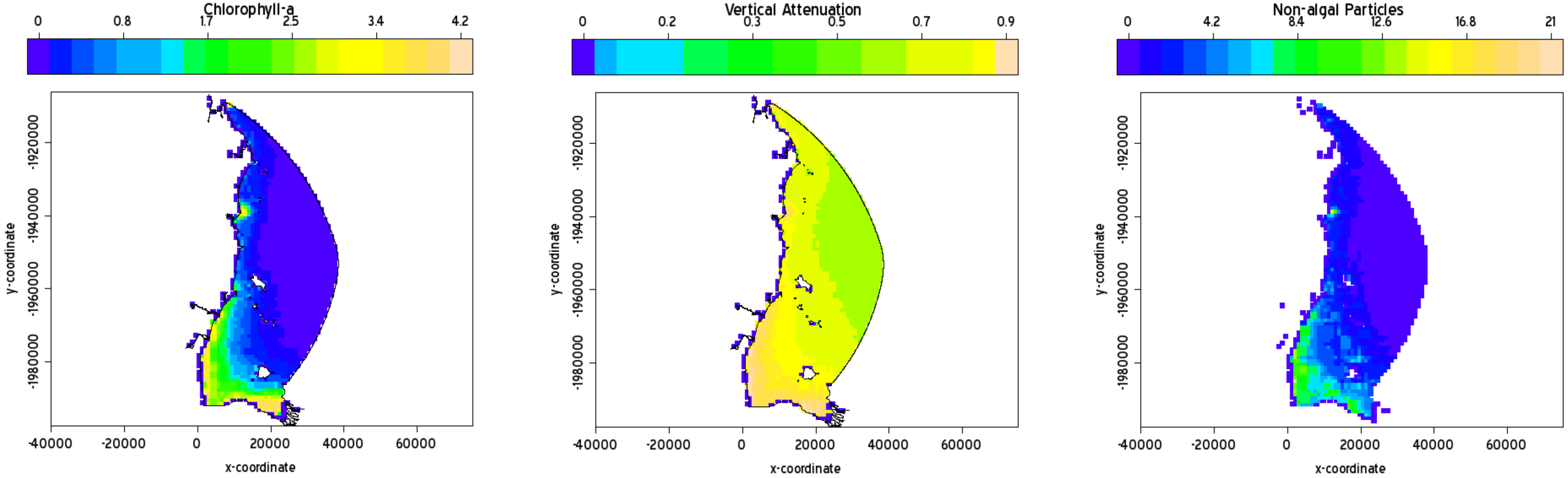


Figure 24: Variables input into the spatial partitioning for the Tully-Murray marine extent. Variables shown are (a) bathymetry, (b) aspect, (c) slope, (d) seabed stress, (e) distance cross-shore and (f) distance from shore relative to the furthermost extent of the lagoon.



(a) (b) (c)

Figure 25: Indicator variables used in the assessment for the Tully-Murray marine extent. Indicators shown are (a) chlorophyll-a, (b) vertical light attenuation and (c) number of non-algal particles, a surrogate for total suspended sediment.

We performed the analysis on 1885 pixels in the Tully-Murray region using bathymetry, aspect, slope, seabed stress and two distance measures which represent the distance along the coast and the distance out to the further most reef. Although we acknowledge that there may be other variables available for this purpose, these variables were the only ones made available for this project. Furthermore, these variables were chosen subject to having minimal anthropogenic impact and were therefore considered appropriate for the spatial partitioning process. Figure 24 shows the distribution of these variables across the Tully-Murray region. We used a correlation analysis to investigate:

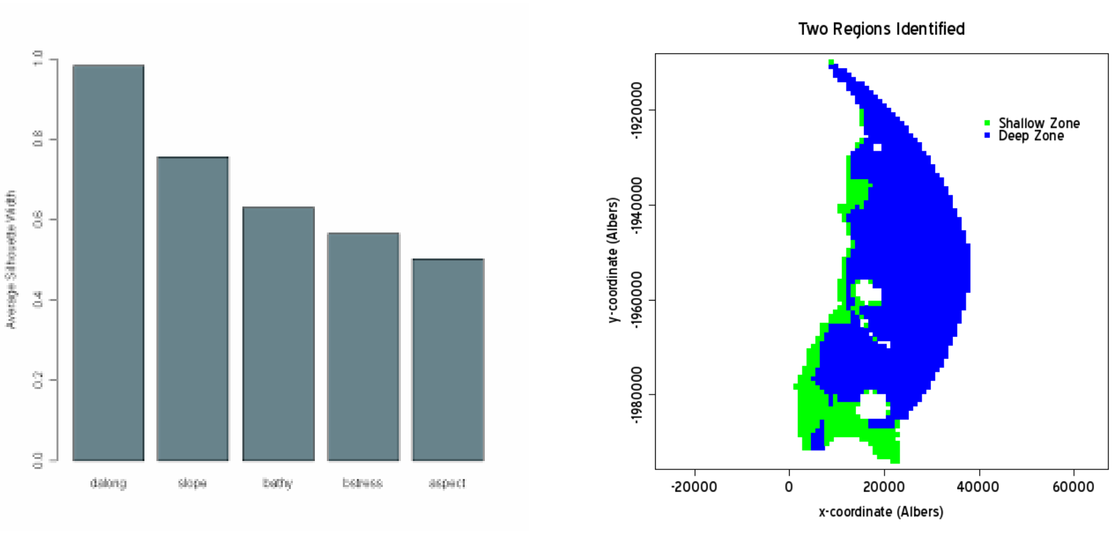
- Important variables in identifying cluster structure; and
- Potential variables to omit from the analysis (which were highly correlated and therefore would provide redundant information to the cluster analysis.)

This analysis highlighted that both bathymetry and the distance cross shore were highly negatively correlated ($\rho = -0.854, \Pr(\rho = 0) < 0.0001$). As a result, we omitted distance cross shore from further analyses.

We examined a number of cluster analysis techniques discussed in Section 5, ranging from agglomerative clustering approaches to divisive measures. We examined the silhouette width resulting from each cluster analysis and also calculated the Gap statistic (Tibshirani et al., 2001b) to investigate the appropriate number of clusters. Both analyses identified two distinct clusters representing shallow, inshore areas (including sites north of Hinchinbrook Island) and deeper regions in the Tully-Murray marine extent. This is illustrated in Figure 26(b). Here, we used Partitioning Around Medoids (PAM) (Kaufman and Rousseeuw, 1990) to partition a distance matrix formed from the dataset using the Gower metric for a range of cluster solutions between 2 and 10. The two cluster solution yielded the highest average silhouette width from this analysis.

We also explored the importance of variables in defining the cluster structure using the “greedy approach” described in Peterson and Kuhnert (2007). Here, we used PAM to cluster using each variable separately to identify the variable which yielded the highest average silhouette width for that clustering solution. Once identified, a search through the remaining variables was conducted to find the variable that together with the first yielded the highest average silhouette width and therefore was able to separate out the clusters effectively. This process continued until all remaining variables entered into the analysis. The results are shown in Figure 26(a) and illustrate that the most important variables for this analysis were distance across shore, bathymetry and slope. It also showed that as we included more variables into the analysis, the silhouette width decreased gradually indicating higher overlap in the clusters formed. Based on this analysis, we only included bathymetry, slope and distance across shore in further analyses of the data. As pointed out in Peterson and Kuhnert (2007), we need to be careful regarding the removal of variables from the analysis. We compared the clustering results from using the entire dataset and the reduced dataset consisting of three variables. We found negligible differences between the clusters formed for both.

Clustering these variables using PAM and the Gower metric resulted in the two cluster solution shown in Figure 26(b). Ideally, this clustering solution would be discussed with stakeholders to determine its validity and whether changes to the “hard” boundaries are required. For example, the expert may provide an opinion about where they think the boundaries lie. This information could be incorporated at the clustering stage at the very beginning as discussed in Section 5. For the purpose of this report however, we assume that the regions identified in Figure 26(b) are appropriate.



(a) (b)

Figure 26: Figures showing the (a) average silhouette width for each variable included in the cluster analysis and (b) the marine spatial partitioning with two distinct clusters representing a shallow and a deep zone In Figure (a) the higher the average silhouette width the higher the importance placed on the variable.

6.3 Indicator Assessment and Integration

6.3.1 Methods for Assessment

There are many ways to assess indicators and integrate these assessments to form an overall evaluation of ecosystem health as described in Section 3. The approaches range from methods that are data driven and attempt to use science to determine relevant thresholds of concern, to thresholds that are based solely on management objectives and are largely qualitative and defined through an expert driven process (Pantus and Dennison, 2005). The latter is the focus of this report as we have not been given suitable data to evaluate indicators nor identify appropriate thresholds of concern. What follows therefore, is a summary of two approaches for assessment when we have thresholds of concern predefined for a particular region.

Simplistic Assessment

For the Tully-Murray river catchment and marine regions we consider two approaches for assessment. The first approach applies the methods of Pantus and Dennison (2005) and only considers the raw data, x_i assuming no uncertainty in the values recorded for each indicator. It uses a threshold of concern defined for a particular management objective defined within each region, r (via the spatial partitioning process) to determine the proportion of sites or proportion of the area that does not exceed the threshold, t_r. This is what we refer to as “compliance”. As an example of this process, stakeholders may define a management objective consisting of the following: “*improving water quality in marine zones*”.

One of the indicators chosen to assess compliance with respect to this objective may be chlorophyll-a for example. Based on extensive consultation with relevant stakeholders we may find that the threshold of concern in deeper regions should be different to shallow areas of the marine zone to reflect the water quality within each region. This may result in thresholds of 1 µg/l and 2 µg/l after some extensive consultation which we can use to make an assessment in each region.

The appropriate thresholds of concern are then used to test for compliance at each of the sites and provide a proportion of the region that complies with the nominated threshold. This evaluation is expressed through the formula shown in Equation 1. Here $I(x_i \le t_r)$ represents the indicator function and returns a one when the expression inside the brackets is satisfied otherwise a zero is returned. Note, in Equation 1, it is assumed that compliance is measured as an indicator having a value lower than the nominated threshold. This is not always the case as can be seen in the example below for riparian vegetation where we are interested in sites that have a percentage of riparian vegetation greater than a set amount.

$$\hat{p}_1 = \frac{\sum_{i|i \in r} I(x_i \le t_r)}{n_r} \qquad (0.1)$$

Equation 1

The second approach we propose uses bootstrap techniques (Efron and Tibshirani, 1993) to examine the variability around the indicator. More importantly, this approach allows bootstrap confidence intervals around the statistic of interest to be calculated at each site. Here we aim to estimate a 95% bootstrap confidence interval for the 80th percentile of values recorded at neighbouring sites.

The bootstrap itself is applied to a neighbourhood surrounding each site. We define the spatial neighbourhood as a circular region of radius, k. Empirical semivariograms were generated based on the exponential model and used to visually estimate the range of spatial autocorrelation, which is the approximate distance over which sites are considered to be correlated (Cressie, 1993). The effective range of the exponential model includes 95% of the variance among uncorrelated data (Goovaerts, 1997). We wanted to restrict the neighbourhood to areas with a relatively strong correlation to the site and so we set k to approximately ½ the effective range.

Ninety-five percent confidence intervals were calculated using the percentile approach which calculates the quantiles of the bootstrap distribution at values of 2.5 and 97.5 (Efron and Tibshirani, 1993). We based all estimates on 5000 bootstrap samples.

Using the bootstrap confidence interval, we can then determine whether a site is:

- Compliant (*bootstrap confidence interval is below the threshold*) or
- Non-compliant (*bootstrap confidence interval is above the threshold*) or whether the assessment is
- Uncertain (*bootstrap confidence interval includes the threshold*) at the site under investigation.

Algorithm 1 illustrates the process.

Algorithm 1: Bootstrap algorithm for computing the 95% confidence interval for a statistic $s(\cdot)$ and assessing against compliance.

1. Calculate the empirical semivariogram for the indicator of interest and determine an appropriate value for k which represents the radius of a spherical neighbourhood.

2. Assuming a spherical neighbourhood with radius k, identify the nearest neighbours at site i.

3. Select B independent bootstrap samples $x^{*1}, x^{*2}, \dots, x^{*B}$, each consisting of samples with replacement from the defined neighbourhood.

4. Evaluate the relevant statistic for each bootstrap sample,
$$\hat{\theta}^{*}(b) = s\left(x^{*b}\right) \quad b = 1, 2, \dots, B.$$

5. Calculate the 95% bootstrap confidence intervals for the statistic of interest,
$$CI : \left(\hat{\theta}_{lo}, \hat{\theta}_{up}\right) = \left(\hat{\theta}^{*(\alpha_1)}, \hat{\theta}^{*(\alpha_2)}\right)$$
where α_1 and α_2 represent quantiles corresponding to the lower and upper bounds of the 95% confidence interval.bounds.

6. Test against the defined threshold, t:
$$\hat{p}_2 = \begin{cases} \sum_{i|i \in r} I\left(\hat{\theta}_{up} \le t_r\right) / n_r & \text{compliant} \\ \sum_{i|i \in r} I\left(\hat{\theta}_{lo} > t_r\right) / n_r & \text{non-compliant} \\ \sum_{i|i \in r} I\left(t_r \in \left[\hat{\theta}_{lo}, \hat{\theta}_{up}\right]\right) / n_r & \text{uncertain} \end{cases}$$

Freshwater (Simplistic) Assessment

Indicators collected as part of the State of the Rivers report (Johnson, 1998) were examined and tested against a suite of hypothetical thresholds defined for each of the four regions constructed in the spatial partitioning process. Two indicators were deemed appropriate for analysis: riparian vegetation and percent macrophyte bed cover. We also used output from the SedNet model in the form of total suspended sediment, otherwise known as TSS. Table 10 displays the thresholds of concern defined for each indicator in each of the four regions. In some instances, the thresholds of concern are the same irrespective of the region. We emphasise that the thresholds of concern shown in this table are purely for demonstration purposes only.

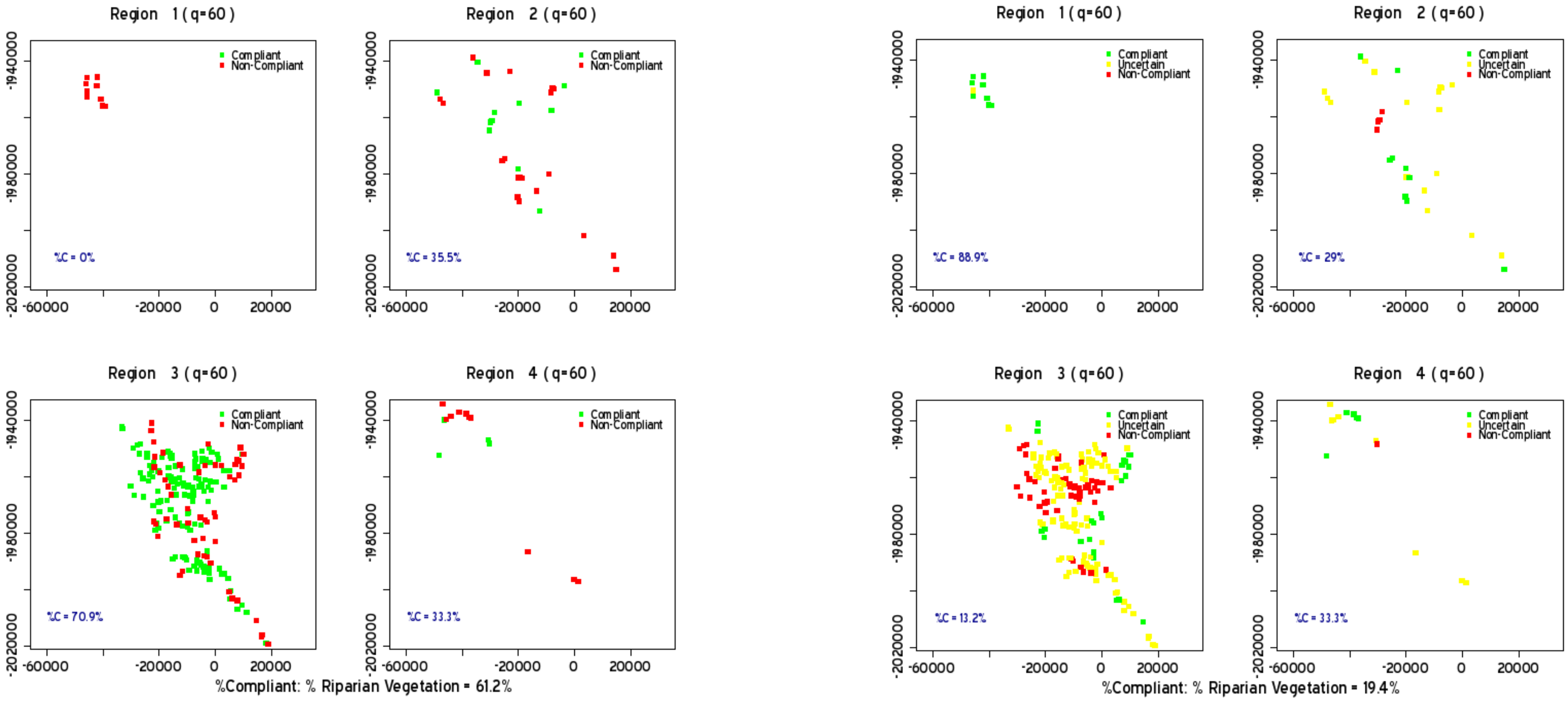
We applied the methodology defined above to assess compliance against the thresholds defined in Table 10. We defined a 3500 metre radius for the neighbourhood for riparian vegetation and percent macrophyte cover and a 7000 metre radius for the TSS neighbourhood. These neighbourhood structures were based on exploratory analysis of the data.

The results are displayed in Figure 27 to Figure 29 for the raw and bootstrap assessments made for riparian vegetation, TSS and percent macrophyte cover, respectively. In each figure, two maps are shown. The first map shown on the left is the result corresponding to the raw assessment. The second map is based on bootstrapping the neighbouring sites of each grid cell. The raw data results show total lack of compliance in region 1 of the catchment and poor compliance in regions 2 and 4. Approximately 70.9% of sites in region 3 of the catchment were compliant, that is, sites with riparian vegetation cover greater than 60%.

The bootstrapped assessment highlights some uncertainty with the assessment, particularly in the third region. Many of the sites that were rated as being compliant are now regarded as uncertain. This results in a much lower compliance rating when compared to the raw assessment.

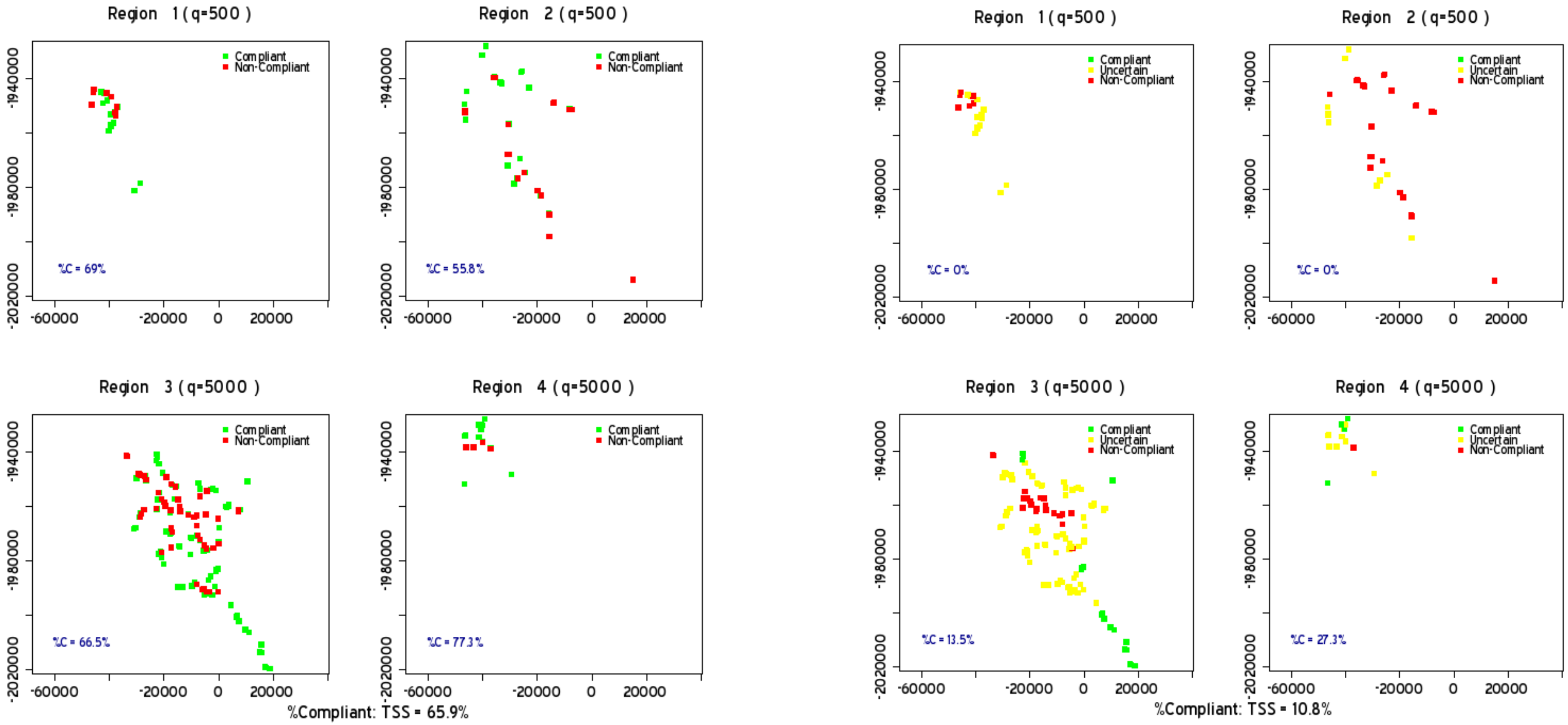
Table 10: Thresholds of concern chosen for freshwater indicators within each of the defined regions shown in Figure 21. Note the sign indicates how the compliance is assessed. A greater than sign (>) indicates that compliance is met when the indicator falls above the nominated threshold. A less than sign (<) indicates that compliance is met when an indicator falls below the nominated threshold.

Threshold of Concern	Riparian Vegetation	TSS (tonnes/year)	% Macrophyte Bed Cover
Region 1: pink	>60%	<500	<5
Region 2: blue	>60%	<500	<5
Region 3: yellow	>60%	<5000	<5
Region 4: green	>60%	<5000	<5



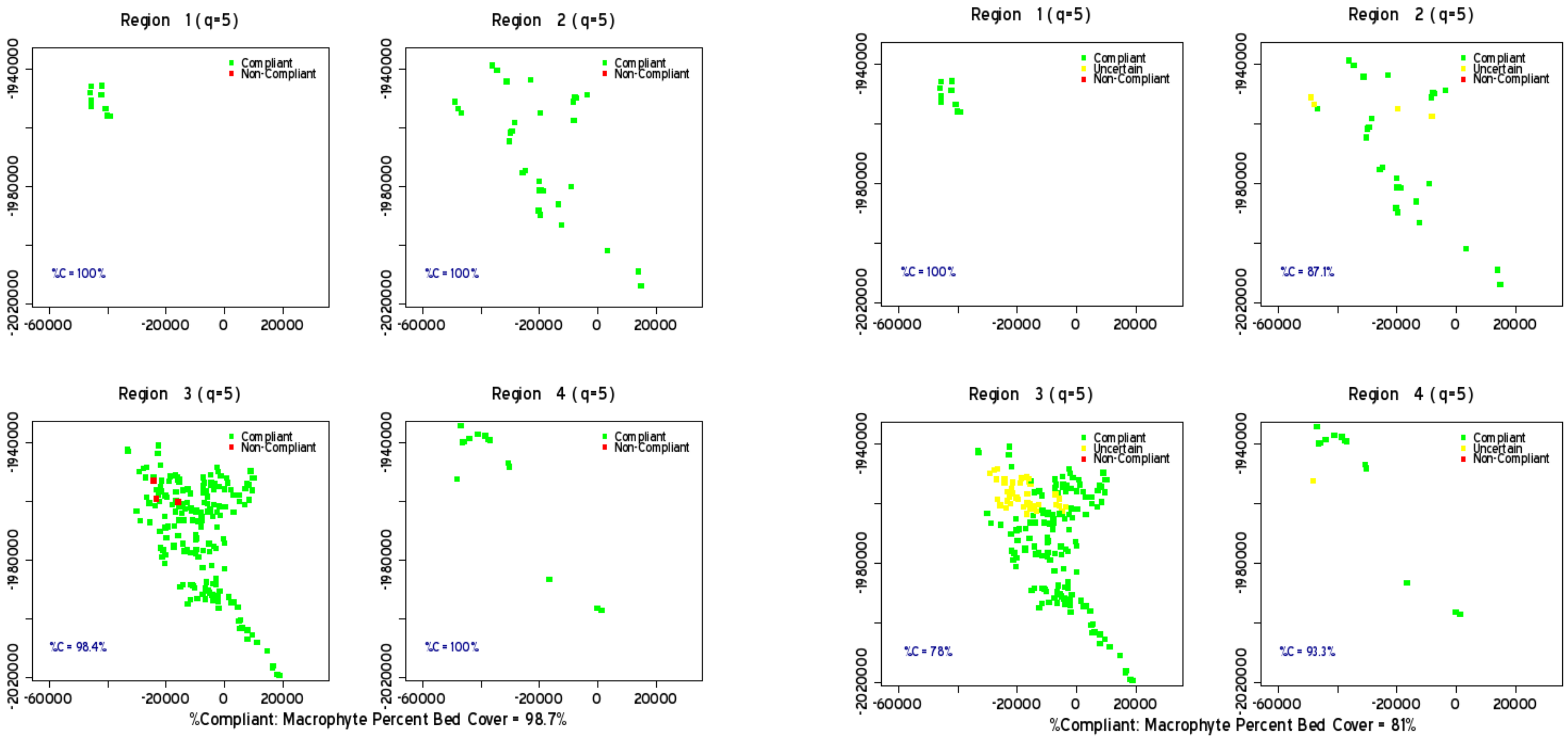
(a) (b)

Figure 27: Maps showing (a) the raw assessment and (b) the bootstrapped assessment for riparian vegetation for each region. An overall percentage of compliance is given beneath each figure. Green cells represent sites which were compliant, while red cells indicate sites there were not compliant. Yellow cells in Figure (b) indicate an “uncertain” assessment.



(a) (b)

Figure 28: Maps showing (a) the raw assessment and (b) the bootstrapped assessment for TSS for each region. An overall percentage of compliance is given beneath each figure. Green cells represent sites which were compliant, while red cells indicate sites there were not compliant. Yellow cells in Figure (b) indicate an “uncertain” assessment.



(a)

(b)

Figure 29: Maps showing (a) the raw assessment and (b) the bootstrapped assessment for percent macrophyte bed cover for each region. An overall percentage of compliance is given beneath each figure. Green cells represent sites which were compliant, while red cells indicate sites there were not compliant. Yellow cells in Figure (b) indicate an “uncertain” assessment.

Marine Assessment

As outlined in an earlier section, we have MODIS satellite imagery as well as site data from three monitoring locations. As the site data consists of very few observations, we concentrate only on the satellite data for the purpose of making an assessment for the Tully-Murray marine region.[1]

Table 11 displays the thresholds of concern chosen for each indicator within each region identified through the spatial partitioning process. Once again, we stress that these values are purely for demonstration purposes as other MTSRF programs are currently trying to determine appropriate thresholds of concern for a large range of indicators. Using these thresholds, we examined each site within each region for compliance using the two approaches: raw calculation of compliance and a bootstrap compliance estimate. For the bootstrap approach, we fit a set of anisotropic (directional) semivariograms to each of the indicators and examined the range. In the marine environment, all of the indicators had a range value of approximately 15 kilometres in all four directions. To avoid oversmoothing, we used 7.5 kilometres as the radius of the spherical neighbourhood with which to take bootstrap samples from.

Table 11: Thresholds of concern chosen for marine indicators within each of the defined regions. Note the sign indicates how the compliance is assessed. A greater than sign (>) indicates that compliance is met when the indicator falls above the nominated threshold. A less than sign (<) indicates that compliance is met when an indicator falls below the nominated threshold.

Threshold of Concern	Chlorophyll-a	Vertical Light Attenuation	No. of Non-algal Particles
Region 1: Shallow	<2μg/l	<0.75	<10
Region 2: Deep	<1μg/l	<0.5	<5

The results are displayed in Figure 30 to Figure 32 for chlorophyll-a, vertical light attenuation and the number of non-algal particles respectively. In each figure, two maps are shown. The first summarises the raw assessment and shows those sites that were compliant with regards to the threshold shown in Table 11. The percentage of sites compliant for the region is shown in the bottom right hand corner of the plot and a weighted compliance measure, representing the overall percentage of compliance for that indicator is shown at the base of the two maps.

The results show some obvious differences between the raw compliance calculation and the bootstrap calculation which takes into account variability due to neighbouring regions. In all three examples, the bootstrap measure of compliance is lower than the raw calculation.

6.3.2 Methods for Integration

Integration methods were discussed in Section 3.2 for a variety of applications. For the methods described and implemented above for the Tully-Murray region, integration by way of averaging proportions across sites to obtain regional and global assessments and furthermore, averaging assessments across indicators can occur irrespective of the approach used for the assessment.

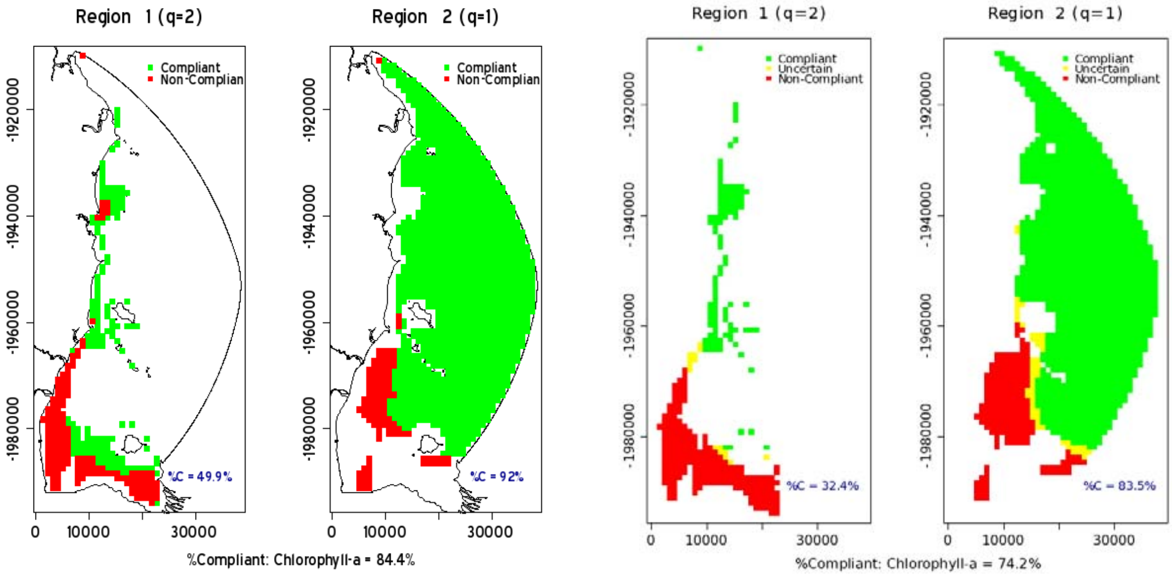
---

[1] Other than computing an assessment straight from the raw data, we really cannot offer a sensible, statistically robust assessment and therefore suggest avoiding making an assessment using the monitoring sites until further data has been collected at these and neighbouring sites.

Using either the raw calculation or the bootstrap estimate of the proportion of compliance for each indicator, we can:

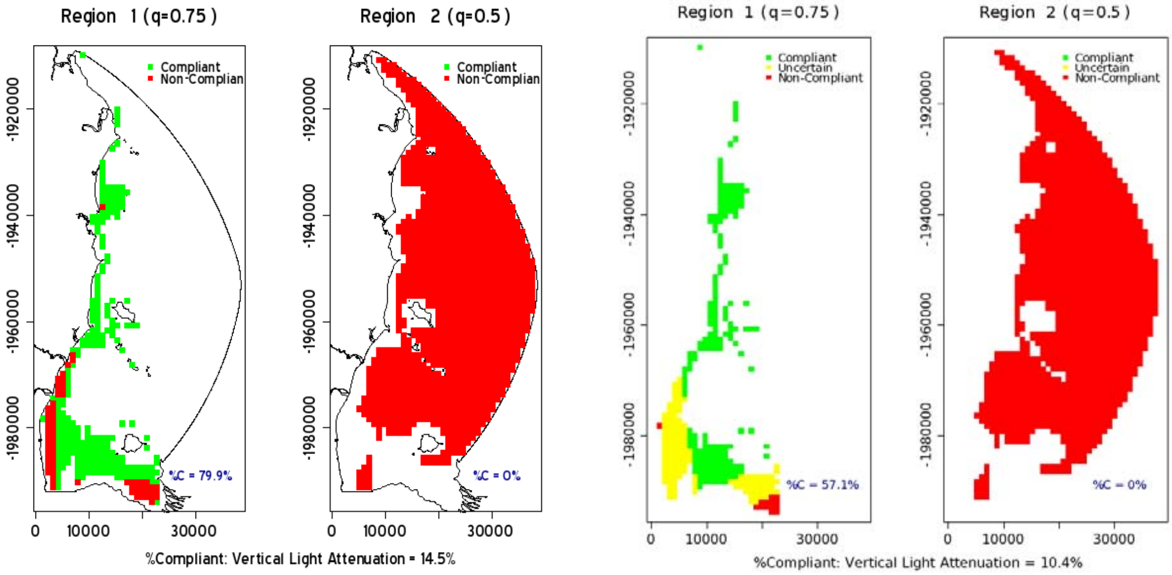
- Average within regions to obtain a **regional assessment**.
- Perform a weighted average across regions to obtain an **indicator assessment.** Here the weighting corresponds to the sample size of points in each region.
- Average across indicators (irrespective of region) to obtain a **global assessment** of compliance.

We can also weight indicators by their sample size to reflect the amount of data that actually went into the assessment.



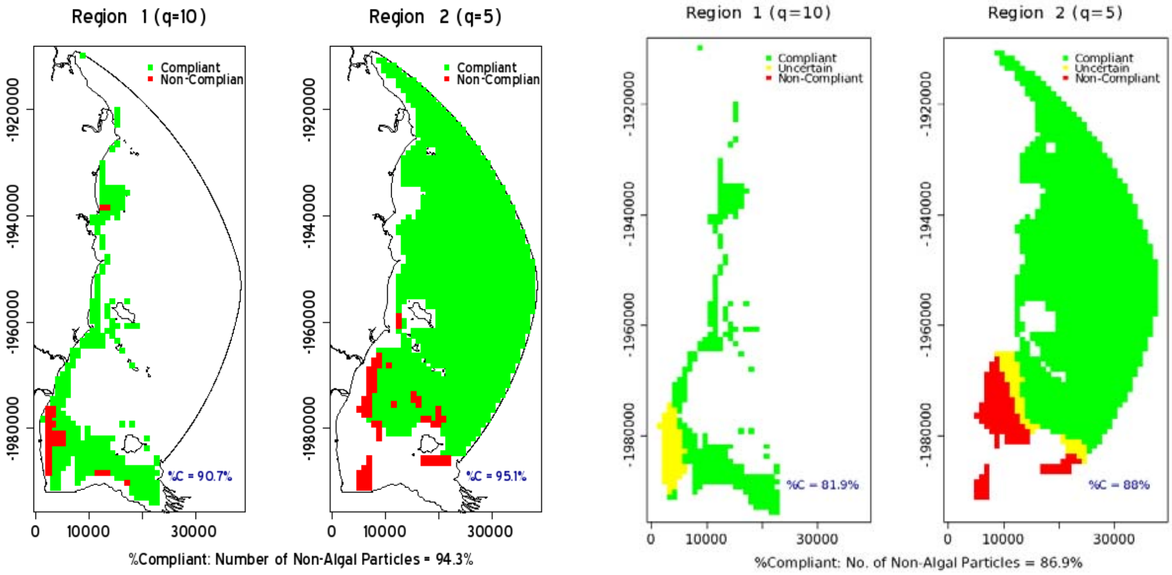
(a) (b)

Figure 30: Maps showing (a) the raw assessment and (b) the bootstrapped assessment for chlorophyll-a for each region. An overall percentage of compliance is given beneath each figure. Green cells represent sites which were compliant, while red cells indicate sites there were not compliant. Yellow cells in Figure (b) indicate an “uncertain” assessment.



(a) (b)

Figure 31: Maps showing (a) the raw assessment and (b) the bootstrapped assessment for vertical light attenuation for each region. An overall percentage of compliance is given beneath each figure. Green cells represent sites which were compliant, while red cells indicate sites there were not compliant. Yellow cells in Figure (b) indicate an “uncertain” assessment.



(a) (b)

Figure 32: Maps showing (a) the raw assessment and (b) the bootstrapped assessment for number of non-algal particles for each region. An overall percentage of compliance is given beneath each figure. Green cells represent sites which were compliant, while red cells indicate sites there were not compliant. Yellow cells in Figure (b) indicate an “uncertain” assessment.

Integration of Freshwater and Marine Assessments

Tables 12 to 15 show the results of integrating the water quality assessments produced from the raw and bootstrapped data for the freshwater and marine assessments respectively. The assessment produced for each indicator at each region has been transferred to the respective tables and coloured according to a very simple colour graduation scheme: blue (81-100% compliance), green (61-80% compliance), yellow (41-60% compliance), orange (21-40% compliance) and red (0-20% compliance). This colour scheme has been used purely for demonstration purposes.

In the case of the marine assessment, this particular colour scheme highlights poor compliance of vertical light attenuation measurements since the assessment across regions was calculated to be 14.5% and 10.4% respectively for the raw and bootstrapped assessments. On closer inspection, this poor assessment has come about because of poor compliance in the deeper regions of the Tully-Murray marine spatial extent, leading to a moderate global assessment of 64.4% and 57.2% respectively.

Table **14** and 15 show a summary of the raw and bootstrapped assessments for the freshwater region. These tables summarise the assessments into two reporting categories: habitat condition, which comprises riparian vegetation and percent macrophyte cover and land use management, which comprises TSS. The integration of the reporting categories appears in Table 16 and Table 17.

.

Table 12: Raw Assessments of water quality for the Tully-Murray marine zone.

Indicator	**Region**		**Indicator Assessment**
	Shallow	**Deep**	
Chlorophyll-a	49.9%	92%	84.4%
Vertical Light Attenuation	79.9%	0%	14.5%
No. of Non-Algal Particles	90.7%	95.1%	94.3%
Regional Assessment	73.5%	62.4%	64.4%

Table 13: Bootstrap Assessments of water quality for the Tully-Murray marine zone using a neighbourhood consisting of a 7.5km radius.

Indicator	**Region**		**Overall Indicator Assessment**
	Shallow	**Deep**	
Chlorophyll-a	32.4%	83.5%	74.2%
Vertical Light Attenuation	57.1%	0%	10.4%
No. of Non-Algal Particles	81.9%	88%	86.9%
Regional Assessment	57.1%	57.2%	57.2%

Table 14: Raw Assessments for the Tully-Murray freshwater region for (a) habitat condition and (b) land use management.

(a) Indicator	**Region**				**Indicator Assessment**
	1	**2**	**3**	**4**	
Riparian Vegetation	100%	67.7%	29.7%	66.7%	39.7%
% Macrophyte Cover	100%	100%	98.4%	100%	98.7%
Overall Assessment	100%	83.9%	64.1%	83.9%	69.2%

(b) Indicator	**Region**				**Indicator Assessment**
	1	**2**	**3**	**4**	
TSS	69%	55.8%	66.5%	77.3%	65.9%

Table 15: Bootstrap assessments for the Tully-Murray freshwater region for (a) habitat condition and (b) land use management.

(a) Indicator	**Region**				**wtd Indicator Assessment**
	1	**2**	**3**	**4**	
Riparian Vegetation	88.9%	29%	13.2%	33.3%	19.4%
% Macrophyte Cover	100%	87.1%	78%	93.3%	81%
Overall Assessment	94.5%	58.1%	45.6%	63.3%	50.2%

(b) Indicator	**Region**				**wtd Indicator Assessment**
	1	**2**	**3**	**4**	
TSS	0%	0%	13.5%	27.3%	10.8%

Using the thresholds of concern defined in Table 10, percentage macrophyte cover almost leads to 100% compliance in the raw assessment (Table 14) and slightly lower levels in the bootstrap assessment (Table 15). The bootstrap assessment shows low percentage of compliance for riparian vegetation in region 3, where a large proportion of sites lie. TSS received a “moderate” assessment (65.9%) using the raw data and a very poor assessment (10.8%) when the bootstrap approach was implemented.

Table 16 and Table 17 integrate the assessments made for each indicator in terms of two reporting categories: habitat condition and suspended sediment generation using the raw and bootstrapped approaches respectively. For the raw assessment, habitat condition and land use management across the four regions were rated as “moderate” (69.2% and 65.9% respectively). This produced a global average assessment of 67.5%.

In the bootstrapped assessment, however, the results were more pessimistic, with habitat condition rated as average (50.2%) and suspended sediment generation rated as very poor (10.8%). This lead to a global score of 30.5%.

Table 16: Integrated raw assessments of habitat condition (riparian vegetation and percent macrophyte bed cover) and land use management (TSS).

Indicator	**Region**				**Overall Indicator Assessment**
	1	**2**	**3**	**4**	
Habitat Condition	100%	83.85%	64.05%	83.85%	69.2%
Suspended sediment generation	69%	55.8%	66.5%	77.3%	65.9%
Regional Assessment	84%	69.8%	65.3%	80.6%	67.5%

Table 17: Integrated bootstrapped assessments of habitat condition (riparian vegetation and percent macrophyte bed cover) and land use management (TSS).

Indicator	**Region**				**Overall Indicator Assessment**
	1	**2**	**3**	**4**	
Habitat Condition	94.5%	58.1%	45.6%	63.3%	50.2%
Suspended sediment generation	0%	0%	13.5%	27.3%	10.8%
Regional Assessment	47.3%	29.1%	29.6%	45.3%	30.5%

6.4 Examples of data visualisation for the Tully catchment

This section demonstrates how the recommended visualisation approach, the data wheel (Section 4.2), can utilise the overall indicator assessment outlined in the previous section. In this section we apply the data wheel to synthetic data from the Tully/Murray catchment (as there was not enough data to complete the wheel). It is important to note that all of the images in this section are for demonstration only. Figure 33 provides an example of how the results from the data wheel can be used to reflect the index scores for catchment, freshwater and marine condition. Alternatively, Figure 34 shows how the data wheel can be used to reflect the index scores for the various sub-catchments or regions in the Tully/Murray, and Figure 35 shows how the data wheel can be used to look at data for individual indicators or for the whole catchment. Finally, Figure 36 is an example of how the data wheel may be used in a community level report card for the Tully catchment.

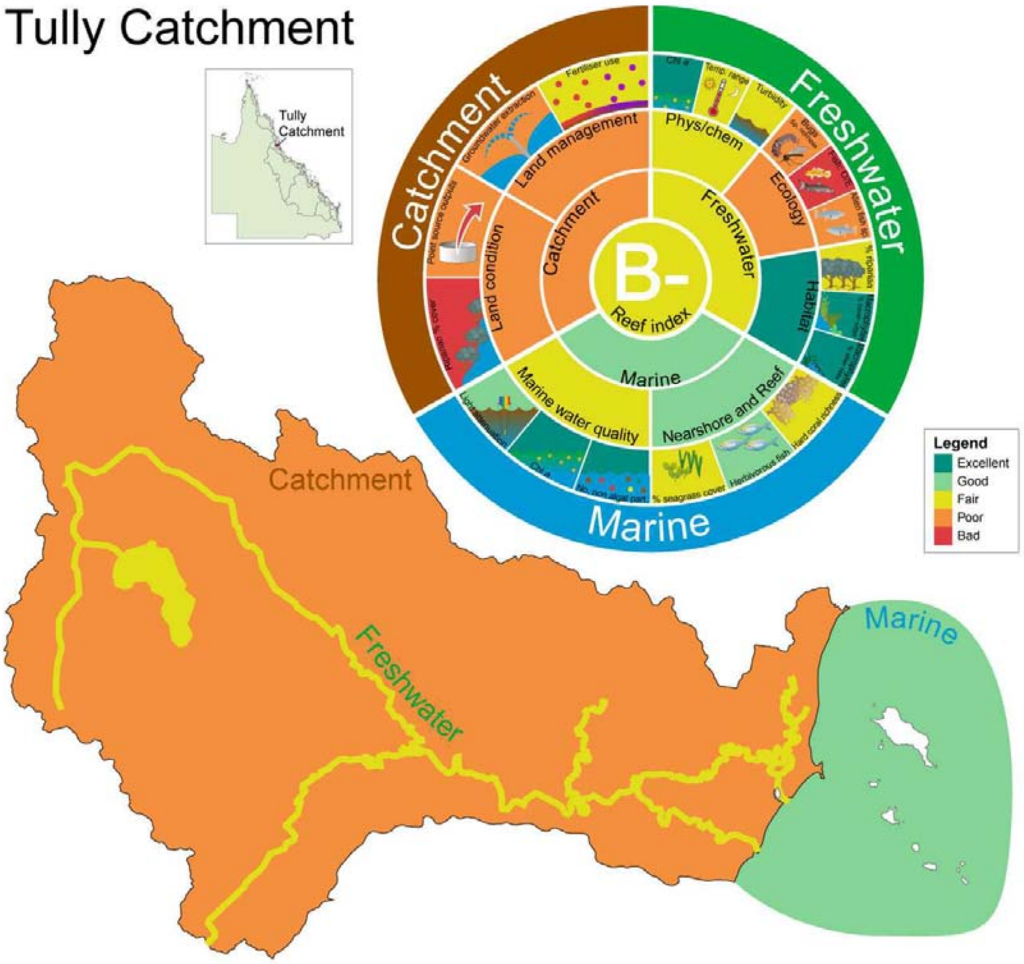


Figure 33: Example of how the data wheel approach can be linked to geographical maps of the catchment showing the different ratings for the catchment, freshwater and marine zones.

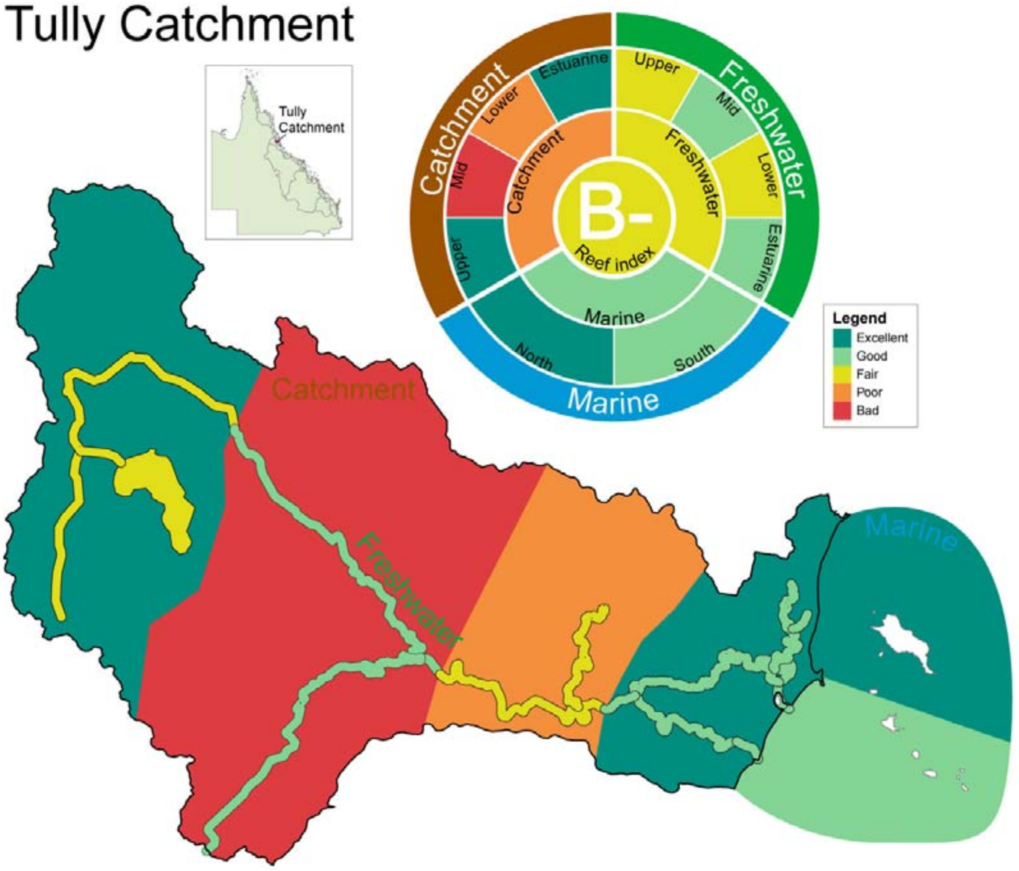


Figure 34: An alternative to the figure above showing how the data wheel can be used to show the score in different ‘regions’ or sub-catchments of a single catchment

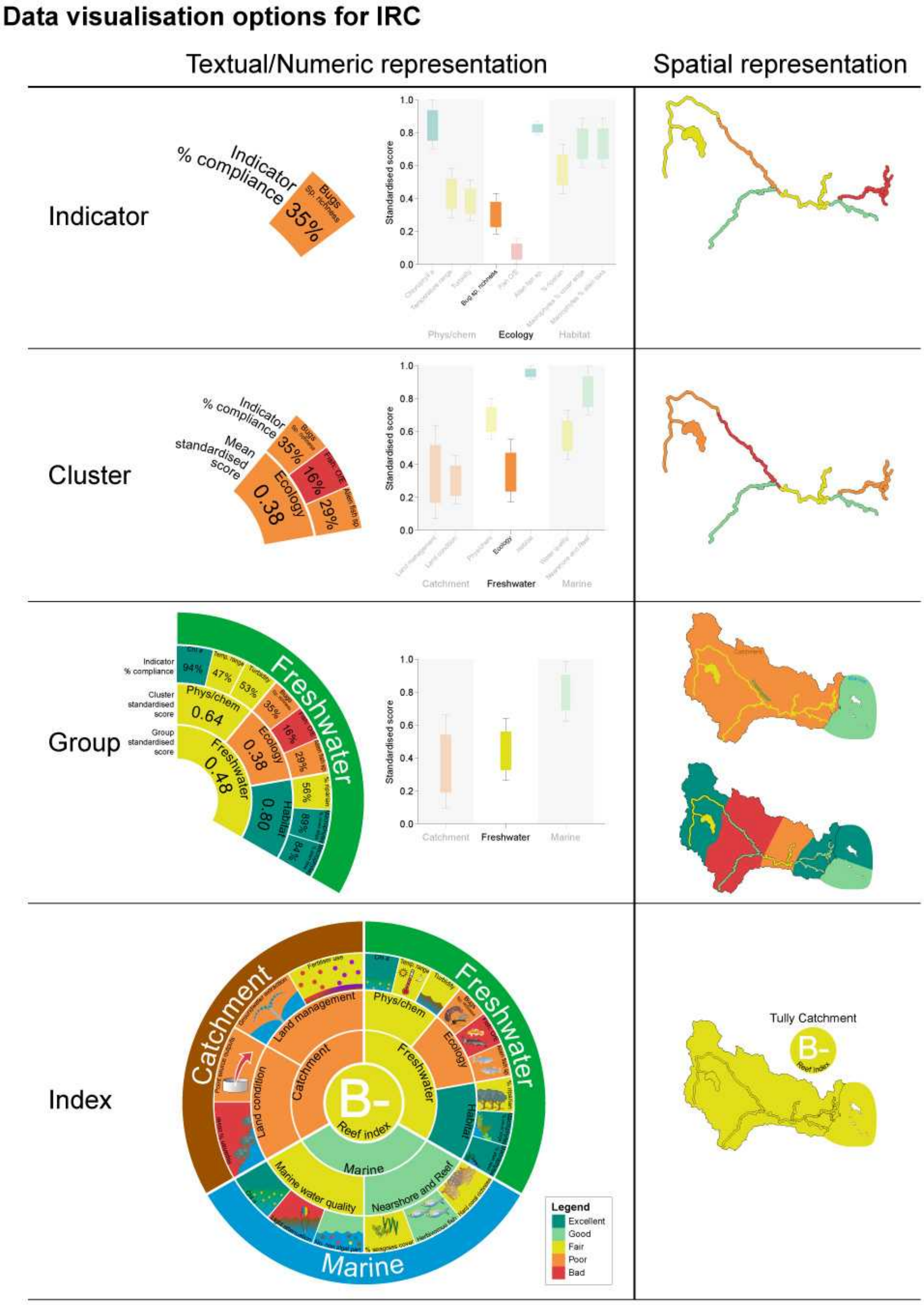
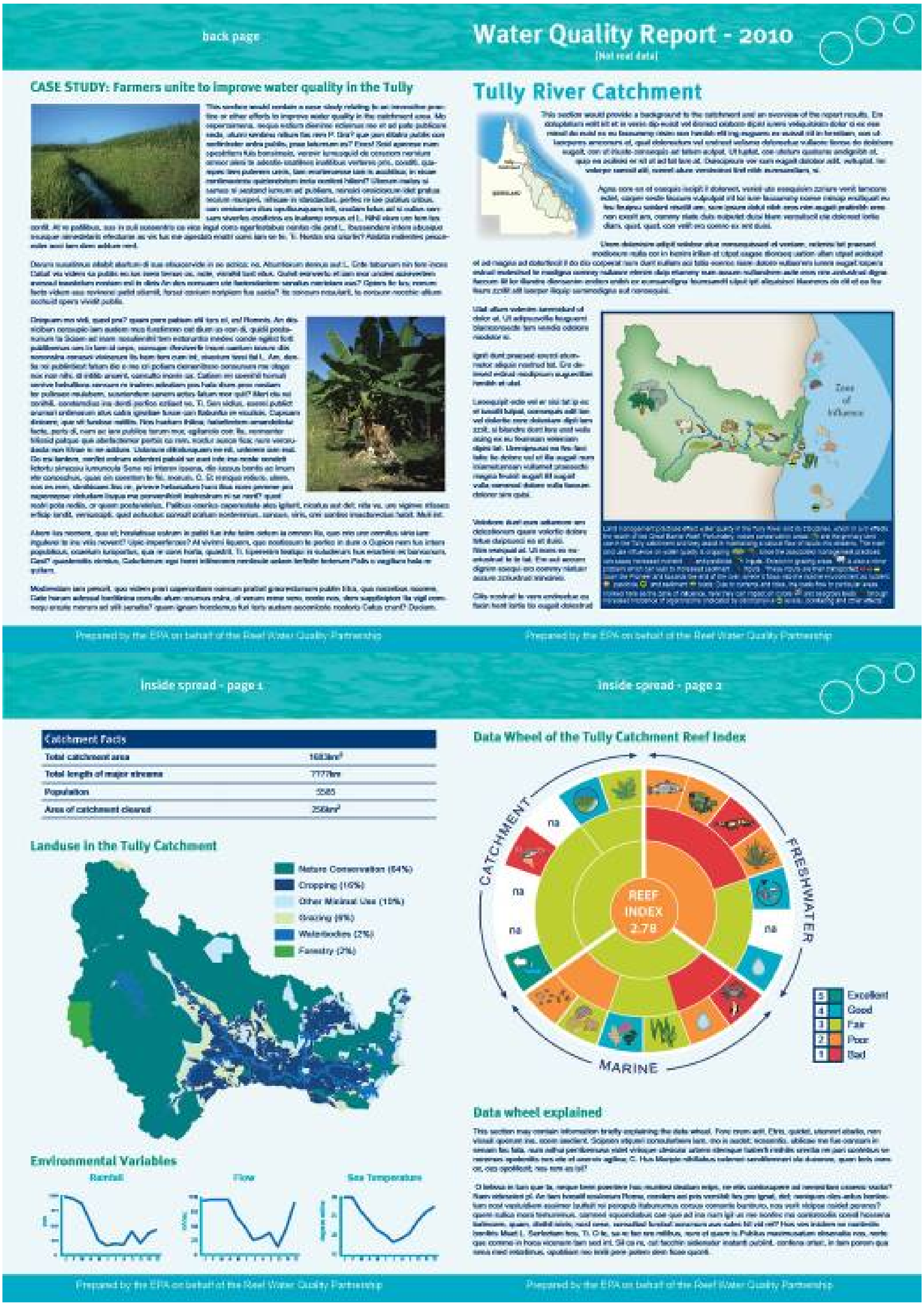


Figure 35: Example of how data can be visualised at different levels within the data wheel depending on the audience



back page

Water Quality Report - 2010

(Not real data)

CASE STUDY: Farmers unite to improve water quality in the Tully

Tully River Catchment

Prepared by the EPA on behalf of the Reef Water Quality Partnership

Prepared by the EPA on behalf of the Reef Water Quality Partnership

Inside spread - page 1

Inside spread - page 2

Catchment Facts	
Total catchment area	1683km²
Total length of major streams	7777km
Population	3585
Area of catchment cleared	256km²

Landuse in the Tully Catchment

Environmental Variables

Data Wheel of the Tully Catchment Reef Index

Data wheel explained

Prepared by the EPA on behalf of the Reef Water Quality Partnership

Prepared by the EPA on behalf of the Reef Water Quality Partnership

Figure 36: Example layout of a report card for the Tully catchment (developed by Tory Grice and Maria Vandergragt of QEPA)

6.5 Discussion

The spatial partitioning, indicator assessment and integration methods presented here are tools that can be used as part of a robust integrated reporting framework. Our goal was to create statistically defensible methods that were deeply rooted in an ecologically valid framework.

Defining regions for compliance assessment against targets and thresholds of concern can be a subjective process. We used a suite of quantitative methods to help ensure that the regions were truly distinct. We attempted to represent ecological processes, such as flow connectivity in freshwater systems and near shore effects in marine environments, using quantitatively defined spatial neighbourhoods. These neighbourhoods allowed us to account for the condition and effect of nearby locations rather than treating each unit as spatially independent from one another.

The ability to accurately represent uncertainty in the spatial partitioning and assessment was critical. Uncertainty is a fundamental aspect of environmental monitoring and the power to locate, quantify, communicate, and incorporate it into the decision making process instils confidence in the ecological assessment and the management decisions that are based upon it. We proposed a bootstrap approach that examined the variability around neighbouring sites to produce a confidence interval and a corresponding assessment of compliance at each site.

Data quality and spatial scale strongly affect the results of the spatial partitioning. If the scale of GIS and remotely sensed data used in the top-down spatial partitioning are too coarse it becomes difficult to account for the conditions found at multiple scales, such as the reach, RCA, and watershed. Data from coarse scales can be re-sampled to a finer spatial resolution, but the quality of the data does not change. As a result, the quality of the finer-scale representation may be less than that of the coarser scales and may not appear to contribute to the spatial partitioning structure. However, this does not mean that the finest scale data available is always the most appropriate. As the spatial scale of the data decreases, the computational requirements of the analysis increase. Therefore, the scale of the data should be chosen to be appropriate for the scale of assessment.

The quality of the indicator data is perhaps more important than that of the spatial partitioning data because it is actually used to assess the condition of the environment. We have made use of data collected from other studies, summarised data (e.g. medians from 12 monthly images) and modelled data to demonstrate our methods here. However, it is inappropriate to use these quality of data to create an integrated report card. Other studies may have been based on biased sample designs, used assorted sampling methods, sampled at various times of the year or even different years, or had different or even competing objectives than those of the monitoring program. Modelled data, such as the total suspended solid loads produced by SedNet (Hately et al., 2006) or remotely derived maps of chlorophyll-a (Schaffelke et al., 2006; Steven et al., In Review), may have a large amount of uncertainty associated with model predictions and making use of median twelve monthly summaries is clearly not taking into account the temporal nature of the data nor is it taking into account extreme fluctuations in a measured indicator. Instead, the monitoring program must invest in its own data collection and make sure that it is based on a statistically robust sampling design. This is a critical part of the assessment process because it ensures that the data 1) reflect the spatial partitioning and the reference conditions, 2) represent an unbiased estimate of the condition, and 3) were collected to specifically meet the objectives of the monitoring program.

We used two methods for assessing compliance that could feed into a reporting framework and deal with the nature of the data that we were given. The first used the raw data to test against a defined threshold. In this report the thresholds chosen were somewhat arbitrary and used as an illustration of the approach. The second approach used a bootstrap technique to estimate a 95% confidence interval for the 80th percentile around a defined neighbourhood of points. This confidence interval was then compared with the threshold to determine compliance.

The purpose of the bootstrap was two fold. The first allowed us to calculate a confidence interval around the statistic of interest (the 80th percentile) and therefore allowed us to examine the variability in the neighbourhood of points surrounding each data point and assess how this affects the compliance rating. We chose an 80th percentile rather than a median so that we could arrive at an assessment that focussed on extreme points in the data. Unfortunately, we were provided with median summaries to begin with, which made the assessment difficult to implement and interpret.

The second reason for proposing a bootstrap approach was that it allowed us to identify points within each region where compliance was not immediately obvious. This occurred when the bootstrap confidence interval included the threshold.

In the application presented in this report, we found that the assessments could differ substantially when we took into account the variability in the neighbourhood of points surrounding each site using the bootstrap. This highlights the importance of incorporating uncertainty into the assessment instead of using a single raw value, which is often used in practice.

We have demonstrated statistically valid and robust methods for spatial partitioning and compliance. However, the quality of the assessment is strongly related to the quality of the data. We have shown that incorporating uncertainty into the assessment is an important part of the process even in data rich situations. Furthermore, we highlighted that the assessment can change dramatically when variability around the value of an indicator at a site is taken into account.

Ideally there should be dedicated monitoring programs in place for the purpose of making valid and scientifically sound assessments within a region that incorporate uncertainty. Unfortunately this will not be the case for many catchments in the Great Barrier Reef region. If data from disparate sources are used, any assessments conducted using these techniques will need to be interpreted with caution.

Finally, this section provided examples of how the statistical methods presented can be packaged into an overall Reef Water Quality Index and can be visually represented using a range of data wheel approaches.

7. Summary and recommendations from this report

Developing a scientifically robust IRC for the GBR catchments and marine region will take a minimum of 4 years. This first year, and this report in particular, has provided some important foundations for the continued development of the GBR IRC. This report has met the milestone activities as follows:

- Outline of initial framework and report card shell (Sections 2 and 4.2)
- Present the recommended regionalisation approach (Section 5 and 6)
- Recommend options for visualisation of the reported condition and trend (Sections 4.2 and 6.4)
- List of mature indicators and developmental indicators (Section 4.1)
- Outline spatial and temporal monitoring strategies for indicators (Section 3)
- Issues relating to data integration (Section 3.3 and 6.3)
- Recommended framework for development and implementation in Years 2 to 4 (see below)

The remainder of this section provides some specific recommendations which should help move the IRC development forward in subsequent years.

Recommendation 1: Indicator development and Incorporation into an IRC

Dedicated working groups within each of the IRC disciplinary classes (catchment, freshwater and marine) need to be set up to facilitate the further development and finalisation of the indicators to be used in the report card. The working groups will need to consist of the research scientists testing the indicators, the government bodies undertaking the routine monitoring, as well the statisticians that will be involved in the final analysis and integration of the data. Where appropriate regional body stakeholders may also be involved to provide community feedback on indicator development. Each of the working groups needs to complete Table 2, Table 3 and Table 4 using the 15 points outlined in Section 3.1.

The outcome of the working group process will hopefully lead to a final list of indicators that are suitable for use in an IRC for the GBR, as well as match the reporting requirements of each of the agencies involved. It is important to note that while we are suggesting at least 3 working groups, it is vital that the different components (catchment, freshwater and marine) remain connected and continue to focus on the linkages at the catchment to reef scale. The connectivity between the groups could be facilitated via the Reef Partnership.

Recommendation 2: Set up pilot monitoring programs

As a follow on to Recommendation 1, there is the need to set up pilot monitoring programs to assist with indicator testing and selection. This will be crucial for developing the thresholds of concern which are expected to be very different between wet and dry catchments (and associated marine monitoring zones) and between upland and estuarine conditions. The pilot programs may build on existing monitoring programs (e.g. AIMS Long term Marine Monitoring Program, or QNRW's Stream and Estuary Assessment Program), however, the programs may need to be re-focused to address specific indicators. It is acknowledged that monitoring across the entire GBR is not necessarily practical, therefore a preliminary spatial partitioning exercise, similar to the one conducted in this report for the Tully catchment, would be useful for the entire GBR region so that areas with different biophysical conditions can be identified and targeted for sampling and establishing targets and/or thresholds of concern. The pilot programs could also be set up alongside other research programs such as

CSIRO's National Research Flagship on Water for a Healthy Country so that existing research can help inform the indicator development and testing.

Recommendation 3: Data management

There is a need for a central inventory of data sets that are available on the indicators suggested in this report for the whole of the GBR region. This will include both freshwater and marine systems, as well as the catchment based data (e.g. fertiliser use rates). The freshwater water quality aspect of this is being undertaken by staff at QNRW, and will be crucial for understanding data availability, quality, geographic extent and future needs. Similarly, staff at GBRMPA and AIMS have access to most of the current marine data. Information such as fertiliser and pesticide application rates in cane systems are available for ~25% of the Queensland cane region (pers. comm. Tim Wrigley), however, funding is required to generate the data into a suitable format. The equivalent should be done for other catchment indicators and agricultural systems such as grazing and horticulture.

Having a list of available data is only the first of a series of data management steps that need to be undertaken. There is an enormous amount of data available in report format for all aspects of the GBR system, however, very little of the data described in these reports are in a suitable format for use in a report card (i.e. in spreadsheet or GIS format). A web based data capture, storage and distribution system that can be used by all parties working on issues to do with GBR water quality (that may be accessed via a Reef Partnership web page for example) would be extremely useful.

The data inventory and storage processes will require data sharing agreements allowing data to be shared between organisations. The lack of data sharing agreements was a major impediment to accessing data for use in this report. Sensitivity may be required when attempting to obtain data on land use management practice, and specific agreements need to made between industry and government on the way in which this data will be used within a report card framework.

Recommendation 4: Using modelled data within a report card framework

Due to the size of the GBR catchments, a number of indicator variables will need to be modelled in some form, rather than directly measured. This may require the use of sediment transportation models, such as SedNet and associated modelling platforms such as E2, as well as approaches such as remote sensing. It is important to communicate this intention with model developers so that more effort can go into improving the accuracy and uncertainty estimates on specific parameters. This will facilitate the use of these types of data within a report card framework.

Recommendation 5: Responsibility for the GBR integrated report card

Within the natural resource management realm of the GBR, there are a large number of different reporting requirements for State, Federal and Regional bodies (e.g. State of Environment and State of Catchment reports). The organisation finally responsible for the GBR IRC needs to be very clear about how the GBR IRC is different (or similar) to other existing reporting frameworks used in the GBR to avoid duplication as well to help form synergies between different government departments.

Recommendation 6: Ownership and leadership of an integrated report card

For the GBR IRC to become a reality requires ownership and leadership by a single organisation. The Reef Water Quality Partnership is the best placed organisation to do this, however, it needs to be equipped with the appropriate resources and it will also need to expand its focus away from just reef 'water quality' to encapsulate other issues in the GBR (e.g. Rainforest, Torres Strait, etc).

Recommendation 7: Communication strategy

The communication of a final report card product is an enormous task, and a communication strategy should be developed early for the successful transfer of the final report card to the wider community.

8. References

Accad, A., Low-Choy, S.J., Pullar, D. and Rochester, W.A., 2005. Bioregion classification using model-based clustering: a case study in north eastern Queensland, Modelling and Simulation Conference, MODSIM05, Melbourne, Australia.

Allan, J.D., 1995. Stream Ecology: Structure and function of running waters. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Arthington, A.H. and Pearson, R.G., 2007. Biological Indictors of Ecosystem Health in Wet Tropics Streams. Final Report Task 3, Catchment to Reef Research Program, Townsville.

Auricht, C.M., 2004. Report Card Review and concept Development, National Land and Water Resources Audit.

Bailey, R.G., 1983. Delineation of ecosystem regions. Environmental Management, 7: 365-373.

Bailey, T.C., Gatrell, A.C., Bailey, T.C. and Gatrell, A.C., 1995. Interactive spatial data analysis, Interactive spatial data analysis. Pearson Education Limited, Essex, England.

Bartley, R., Hawdon, A., Post, D.A. and Roth, C.H., 2007. A sediment budget in a grazed semi-arid catchment in the Burdekin basin, Australia. Geomorphology, 87: 302-321.

Barton, J.L. and Metzeling, L., 2004. The Development of Biological Objectives for Streams in a Single Catchment: A Case Study on the Catchment of Western Port Bay, Victoria, Australia. Environmental Monitoring and Assessment, 95(1-3): 239-256.

Bayne, B.L., Addison, R.F., Capuzzo, J.M., Clarke, K.R., Gray, J.S., Moore, M.N. and Warwick, R.M., 1988. An overview of the GEEP Workshop. Marine Ecology Progress Series, 46(235-243).

Benda, L., Poff, N.L., Miller, D., Dunne, T., Reeves, G., Pess, G. and Pollock, M., 2004. The Network Dynamics Hypothesis: How Channel Networks Structure Riverine Habitats. BioScience, 54(5): 413-427.

Bohnet, I., Kinjun, C., Haug, K., Kroon, F., Sydes, D., Pert, P. and Roberts, B., 2007. Community uses and values of waters in the Tully-Murray catchment. Final Report for FNQ NRM Ltd., CSIRO Sustainable Ecosystems, Atherton.

Bowman, M.F. and Somers, K.M., 2006. Evaluating a novel test site analysus (TSA) bioassessment approach. Journal of the North American Benthological Society, 25(3): 712-727.

Bray, J.R. and Curtis, J.T., 1957. An ordination of the upland forest communities of southern Wisconsin. Ecological Monographs, 27(4): 325-349.

Breiman, L., 2001. Random Forests. Machine Learning, 45(1): 5-32.

Browne, M., 2006. A geometric approach to density estimation. Pattern Recognition, 40: 134-140.

Browne, M., Kuhnert, P., Peterson, E., Bartley, R., Steven, A. and Harch, B., 2007. Review of existing approaches used to develop integrated report card frameworks (IRCF) and their relevance to catchments draining to the Great Barrier Reef, CSIRO, Brisbane.

Bunn, S.E. and Smith, M.J., 2007. Design and implementation of an ecosystem health monitoring program for streams and rivers in Southeast Queensland, Australia: an overview, DIBM3 technical report for Healthy Waterways Australia.

Butler, A., Harris, P., Lyne, V., Heap, A., Passlow, V. and Porter-Smith, R., 2001. An interim bioregionalisation for the continental slope and deeper waters of the South-East marine region of Australia, Final Report to the National Oceans Office, CSIRO Marine Research and Geoscience Australia, Hobart.

C4ES, 2004. Indendepent Environmental Audit of the Sugar Industry in Queensland, New South Wales & Western Australia, Canegrowers, Canberra.

CCME, 2001. Canadian water quality guidlines for the protection of aquatic life, Candian Councl of Ministers for the Environment (CCME), Winnipeg, http://www.ccme.ca/assets/pdf/wqi_techrprtfctsht_e.pdf.

Chiang, S., Tsay, T. and Nix, S., 2002. Hydrologic regionalization of watersheds. 1: Methodology development. Journal of Water Resources Planning and Management, January/February: 3-11.

Coles, R., Delean, S., Furnas, M., McKenzie, L., Mortimer, M., Muller, J., Negri, A.S., H. and Thompson, A., 2005. Report on status and trends of water quality and ecosystem health in the Great Barrier Reef World Heritage Area, Reef Cooperative Research Centre, Townsville.

Connolly, N., Pearson, R.G. and Pearson, B.A., 2007. Macroinvertebrates as Indicators of Ecosystem Health in Wet Tropics Streams. Final Report Task 3, Catchment to Reef Research Program, Townsville.

Cressie, N., 1993. Statistics for Spatial Data, Revised Edition. John Wiley and Sons, New York, pp. 900 p.

Dauer, D.M., Ranasinghe, J.A. and Weisberg, S.B., 2000. Relationships Between Benthic Community Condition, Water Quality, Sediment Quality, Nutrient Loads, and Land Use Patterns in Chesapeake Bay. Estuaries, 23(1): 80-96.

Davies, N.M., Norris, R.H. and Thoms, M.C., 2000. Prediction and assessment of local stream habitat features using large-scale catchment characteristics. Freshwater Biology, 45: 343-369.

De'ath, G., 2007. The spatial, temporal and structural composition of Water Quality of the Great Barrier Reef, and indicators of Water Quality and Mapping Risk. MTSRF milestone report, Australian Institute of Marine Science, Townsville.

DEH, 2005. National Marine Bioregionalisation of Australia, Department of Environment and Heritage, Canberra.

Department of Natural Resources Mines and Energy, 2004. Geological Mapping Data, Geoscience Data, Queensland Mapping Data (GEOLDATA).

Detenbeck, N.E., 2000. A test of watershed classification systems for ecological risk assessment. Environmental Toxicology and Chemistry, 19(4): 1174-1181.

Digby, P.G.N. and Kempton, R.A., 1987. Multivariate analysis of ecological communities. Chapman and Hall, London.

DNRME, 2004. Geological Mapping Data. In: D.o.N.R.a.M.a. Energy (Editor), Geoscience Data, Queensland Mapping Data (GEOLDATA).

Dunne, T.L.L.B., 1978. Water in Environmental Planning. W.H. Freeman and Company, United States of America.

Efron, B. and Tibshirani, R.J., 1993. An Introduction to the bootstrap. Chapman and Hall, New York.

Everitt, B., 1974. Cluster Analysis. Heinemann Education Books, London.

Fabricius, K., Uthicke, S., Cooper, T., Humphrey, C., De'ath, G. and Mellors, J., 2007. Candidate bioindicator measures to monitor exposure to changing water quality on the Great Barrier Reef. Catchment to Reef Research Program, Project 7.2 in Program C, Australian Institute of Marine Science, Townsville.

Fabricius, K.E., 2005. Effects of terrestrial runoff on the ecology of corals and coral reefs: review and synthesis. Marine Pollution Bulletin, 50: 125-146.

Fabricius, K.E. and De'ath, G., 2001. Biodiversity on the Great Barrier Reef: Large-scale Patterns and Turbidity-Related Local Loss of Soft Coral Taxa. In: E. Wolanski (Editor), Oceanographic Processes of Coral Reefs. CRC Press, Florida, pp. 127-144.

Fabricius, K.E., De'ath, G., McCook, L., Turak, E. and Williams, D.M., 2005. Changes in algal, coral and fish assesmblages along water quality gradients on the inshore Great Barrier Reef. Marine Pollution Bulletin, 51: 384-398.

Fagan, W.F., 2002. Connectivity, Fragmentation, and extinction Risk in Dendritic Metapopulations. Ecology, 83(12): 3243-3249.

Fausch, K.D., Torgersen, C.E., Baxter, C.V. and Li, H.W., 2002. Landscapes to Riverscapes: Bridging the Gap between Research and Conservation of Stream Reaches. BioScience, 52(6): 483-498.

Fellows, C.S., Clapcott, J.E., Udy, J.W., Bunn, S.E., Harch, B.D., Smith, M.J. and Davies, P.M., 2006. Benthic Metabolism as an Indicator of Stream Ecosystem Health. Hydrobiologia, 572(1).

Fraley, C. and Raftery, A.E., 2002. Model-based clustering, discriminant analysis and density estimation. Journal of the American Statistical Association, 97: 611-631.

Frissell, C.A., Liss, W.J., Warren, C.E. and Hurley, M.D., 1986. A Hierarchical Framework for Stream Habitat Classification: Viewing Streams in a Watershed Context. Environmental Management, 10(2): 199-214.

Ganio, L.M., Torgersen, C.E. and Gresswell, R.E., 2005. A geostatistical approach for describing spatial pattern in stream networks. Frontiers in Ecology and the Environment, 3(3): 138-144.

Gartner Lee Ltd, 2006. A sensitivity of the Canadian Water Quality Index.

Gauch, H., 1982. Multivariate analysis in community ecology. Cambridge University Press, Cambridge.

Geoscience Australia, 2007. GEODATA 9 Second Digital Elevation Model (DEM-9S) Version 2. . Geoscience Australia.

Geosciences Australia, 2004. NATMAP 1:100.000 scale topographic maps of Australia. Geosciences Australia.

Goovaerts, P., 1997. Geostatistics for Natural Resources, Geostatistics for Natural Resources. Oxford University Press, New York, New York.

Gower, J.C., 1971. A general coefficient of similarity and some of its properties. Biometrics, 27: 623-637.

Griffith, G.E., Omerniik, J.M. and Woods, A.J., 1999. Ecoregions, watersheds, basins, and HUCs: How state and federal agencies frame water quality. Journal of soil and water conservation, 54(4): 666-667.

Hartigan, J.A., 1975. Clustering Algorithms. Wiley, New York.

Harwell, M.A., Myers, V., Young, Y., Bartuska, A., Gassman, N., Gentile, J.H., Harwell, C.C., Appelbaum, S., Barko, J., Causey, B., Johnson, C., McLeon, A., Smola, R., Templet, P. and Tosini, S., 1999. A framework for an ecosystem integrity report card. Bioscience 49(7): 543-556.

Hately, L., Armour, J.D., Pitt, G., Sherman, B.S., Read, A., Chen, Y., Brodie, J. and Cogle, A.L., 2006. Sediment and nutrient modelling in the Far North Queensland NRM region, Department of Natural Resources.

Hawkins, C.P., Norris, R.H., Gerritsen, J., Hughs, R.M., Jackson, S.K., Johnson, R.K. and Stevenson, R.J., 2000. Evaluation of the use of landscape classifications of the prediction of freshwater biota: synthesis and recommendations. Journal of the North American Benthological Society, 19: 541-556.

Hawkins, C.P. and Vinson, M.R., 2000. Weak correspondence between landscape classifications and stream invertebrate assemblages: implications for bioassessment. Journal of the North American Benthological Society, 19(3): 501-517.

Haynes, D., Waterhouse, J., Innes, J., Vella, K., Furnas, M. and Schaffelke, B., 2005. Great Barrier Reef Water Quality Protection Plan (Reef Plan), Great Barrier Reef Marine Park Authority, Townsville.

Heap, A.D., Harris, P.T., Hinde, A. and Woods, M., 2005. Benthic marine bioregionalisation of Australia's Exclusive Economic Zone. Report to the National Oceans Office on the development of a national benthic marine bioregionalisation in support of region marine planning., Australian Government, Canberra.

Hoeffding, W., 1948. A non-parametric test of independence. The Annals of Mathmatical Statistics, 19: 546-557.

Honchin, C., Gray, L., Brodie, J., Haynes, D. and Moss, A., Draft. Interim Marine Water Quality Guidelines for the Great Barrier Reef Marine Park, Great Barrier Reef Marine Park Authority.

Humphrey, C., Codi-King, S. and Klumpp, D.W., In Press. A multi-biomarker approach in barramundi (Lates calcarifer) to measure exposure to contaminants is estuaries of tropical North Queensland. Marine Pollution Bulletin.

Hynes, H.B.N., 1975. The stream and its valley. Verhandlugen der Internationalen Vereinigung fur Theoretische und Angewandte Limnologie, 19: 1-15.

Jeffers, J.N.R., 1998. Characterization of river habitats and prediction of habitat features using ordination techniques. Aquatic Conservation: Marine and Freshwater Ecosystems, 8: 529-540.

Jeffrey, S.J., Carter, J.O., Moodie, K.B. and Beswick, A.R., 2001a. Using spatial interpolation to construct a comprehensive archive of Australian climate data. Environmental Modelling and Software with Environment Data News, 16(4): 309-330.

Jeffrey, S.J., Carter, J.O., Moodie, K.B. and Beswick, A.R., 2001b. Using spatial interpolation to construct a comprehensive archive of Australian climate data. Environmental Modelling and Software, 16: 309-330.

Johnson, D.P., 1998. An ecological and physical assessment of the condition of streams ain the drainage basins of the Tully and Murray rivers. ISBN 0 7248 8010 3, Queensland Department of Natural Resources, Brisbane.

Jolliffe, I.T., 1986. Principal Component Analysis. Springer-Verlag, New York.

Kaufman, L. and Rousseeuw, P.J., 1990. Finding Groups in Data. Wiley, New York.

Kaufman, L. and Rousseeuw, P.J., 2005. Finding groups in data: An introduction to cluster analysis. Wiley Series in Probability and Statistics. John Wiley and Sons, Inc., Hoboken, New Jersey, USA.

Kennard, M.J., Arthington, A.H., Pusey, B.J. and Harch, B., 2005. Are alien fish a reliable indicator of river health. Freshwater Biology, 50(1): 174-193.

Kennard, M.J., Harch, B.D., Pusey, B.J. and Arthington, A.H., 2006a. Accurately Defining the Reference Condition for Summary Biotic Metrics: A Comparison of Four Approaches. Hydrobiologia, 572(1): 151-170.

Kennard, M.J., Pusey, B.J., Arthington, A.H., Harch, B.D. and Mackay, S.J., 2006b. Development and Application of a Predictive Model of Freshwater Fish Assemblage Composition to Evaluate River Health in Eastern Australia. Hydrobiologia, 572(1): 33-57.

Kuhnert, P., 2003. New methodology and comparisons for the analysis of binary data using Bayesian and tree based methods, Queensland University of Technology, Brisbane.

Legendre, L. and Legendre, P., 1983. Numerical Ecology. Developments in Environmental Monitoring 3. Elsevier, Amsterdam.

Liaw, A. and Wiener, M., 2002. Classification and Regression by random Forest. R News, 2/3: 18-22.

Mackay, S.J., James, C.S. and Arthington, A.H., 2007. Aquatic Macrophytes as Indictors of Catchment Land-use and Water Quality in Wet Tropics Streams. Final Report Task 3, Catchment to Reef Research Program, Townsville.

Maddock, I., 1999. The importance of physical habitat assessment for evaluating river health. Freshwater Biology, 41: 373-391.

Mardia, K.V., Kent, J.T. and Bibby, J.M., 1979. Multivariate Analysis. Academic Press, London.

Marshall, J., Prior, A., Steward, A. and McGregor, G., 2006a. Freshwater Bioregionalisation of Queensland's riverine ecosystems, Development of interim freshwater biogeographic provinces, Queensland Natural Resources, Mines and Water, Brisbane.

Marshall, J.C., Steward, A.L. and Harch, B.D., 2006b. Taxonomic Resolution and Quantification of Freshwater Macroinvertebrate Samples from an Australian Dryland River: The Benefits and Costs of Using Species Abundance Data. . Hydrobiologia 572(1): 171-194.

McCulloch, M., Fallon, S., Wyndham, T., Hendy, E., Lough, J. and Barnes, D., 2003. Coral record of increased sediment flux to the inner Great Barrier Reef since European Settlement. Nature, 421: 727-730.

McKane, D., 2002. State of the catchment 2002; Performance Assessment Reporting System (PARS) for the South Australian Murray-Darling Basin, South Australian Government, Department of Water, Land and Biodiversity Conservation, Adelaide.

McKenzie, N.J., Jacquier, D.W., Maschmedt, D.J., Griffin, E.A. and Brough, D.M., 2005. The Australian Soil Resource Information System: technical specifications. . The National Committee on Soil and Terrain Information, Australian Collaborative Land Evaluation Program, Canberra.

McQuitty, L.L., 1966. Similarity analysis by reciprocal pairs for discrete and continuous data. Educational and Psychological Measurement, 26: 825-831.

Metzeling, L., Tiller, D., Newall, P., Wells, F. and Reed, J., 2006. Biological Objectives for the Protection of Rivers and Streams in Victoria, Australia. Hydrobiologia, 572(1): 287-299.

Mitchell, A., Regenzani, J., Furnas, M., De'ath, G., Brodie, J. and Lewis, S., 2007. Nutrients and suspended sediments in the Tully River: spatial and temporal trends. Report Number 06/06, ACTRF, Townsville.

Murtagh, F., 1985. Multidimensional clustering algorithms. In: J.M. Chambers, J. Gordesch, A. Klas, L. Lebart and P.P. Sint (Editors), COMPSTAT lectures, 4 (lectures in computational statistics). Physica-Verlag., Vienaa, Wurzburg, pp. 131.

Myers, W.L., McKenney-Easterling, M., Hychka, K., Griscom, B., Bishop, J.A., Bayard, A., Rocco, G.L., Brooks, R.P., Constantz, G., Patil, G.P. and Taillie, C., 2006. Contextual clustering for configuring collaborative conservation of watersheds in the Mid-Atlantic Highlands. Environmental and Ecological Statistics, 13(4): 391-407.

Negus, P. and Marsh, N., 2006a. Water Quality Condition Assessment Framework. National Action Plan - Water Quality Report No. Qld NAP WQ02-ASS01/1, Queensland NR&M, Brisbane.

Negus, P. and Marsh, N., 2006b. Water quality condition assessment framework, National Action Plan for Salinity and Water Quality, Water Quality State-level Investment Project, Queensland Department of Natural Resources and Mines, Indooroopilly, Queensland, Australia.

Neil, D.T., Orpin, A.R., Ridd, P.V. and Yu, B., 2002. Sediment yield and impacts from river catchments to the Great Barrier Reef Lagoon. Marine and Freshwater Research, 53: 733-752.

Nelson, R.L., Platts, W.S., Larsen, D.P. and Jenson, S.E., 1992. Trout distributions and habitat in relation to geology and geomorphology in the North Fork Humbolt River Drainage, Northeastern Nevada. Transactions of the American Fisheries Society, 121(4): 405-426.

NWC, 2006. Discovery Phase, Australian Water Resources, Australian Government National Water Commisssion, Canberra, http://www.nwc.gov.au/nwi/docs/AppendixD_River_Health.pdf.

Olden, J.D., Jackson, D.A. and Peres-Neto, P.R., 2001. Spatial isolation and fish communities in drainage lakes. Oecologia, 127: 572-585.

Omernik, J.M., 1987a. Ecoregions of the conterminous United States. Annals of the Association of American Geographers, 77(1): 118-125.

Omernik, J.M., 1987b. Map Supplement: Ecoregions of the Conterminous United States. Annals of the Association of American Geographers, 77(1): 118-125.

Omernik, J.M. and Bailey, R.G., 1997. Distinguishing between watersheds and ecoregions. Journal of the American Water Resources Association, 33(5): 935-949.

Osborne, L.L. and Wiley, M.J., 1988. Empirical Relationships Between Land Use/Cover and Stream Water Quality in an Agricultural Watershed. Journal of Environmental Management, 26: 9-27.

Pantus, F.J. and Dennison, W.C., 2005. Quantifying and Evaluating Ecosystem Health: A Case Study from Moreton Bay, Australia. Environmental Management, 36(5): 757-771.

Pearson, R.G., Arthington, A.H., Connolly, N., Mackay, S.J. and Pusey, B.J., 2007. Summary and synthesis: integrated protocols for monitoring the ecosystem health of Australian Wet Tropics streams. Final Report Task 3, Catchment to Reef Research Program, Townsville.

Peterson, E.E. and Kuhnert, P.M., 2007. A recommended regionalisation for the Tully-Murray catchment and marine zones with implications for the Great Barrier Reef reporting region, CSIRO Mathematical and Information Sciences, Brisbane.

Pitcher, R., Doherty, P., Arnold, P., Hooper, J. and Gribble, N., 2007. Seabed Biodiversity on the Continental Shelf of the Great Barrier Reef World Heritage Area. Final Task Report to CRC Reef Research, FRDC and National Oceans Office.

Poff, N.L., 1997. Landscape filters and species traits: towards mechanistic understanding and prediction in stream ecology. Journal of the North American Benthological Society, 16(2): 391-409.

Pullar, D., Low-Choy, S.J. and Rochester, W.A., 2005. Ecoregion classification using a Bayesian approach and model-based clustering, Modelling and Simulation Conference, MODSIM05, Melbourne, Australia.

Pusey, B.J., Kennard, M.J. and Arthington, A.H., 2007. Freshwater fish as Indictors of Ecosystem Health in Wet Tropics Streams. Final Report Task 3, Catchment to Reef Research Program, Townsville.

Queensland Government, 2003. Reef water quality protection plan. www.thepremier.qld.gov.au/reefwater.

Rao, A.R. and Srinivas, V.V., 2006. Regionalization of watersheds by hybrid-cluster analysis. Journal of Hydrology, 318(1-4): 37-56.

Rasmussen, C., 2000. The infinite Gaussian mixture model, Advances in Neural Information Processing Systems MIT Press, Cambridge, MA.

Reynoldson, T.B., Norris, R.H., Resh, V.H., Day, K.E. and Rosenberg, D.M., 1997. The reference condition: a comparison of multimetric and multivariate approaches to assess water-quality impairment using benthic macroinvertebrates. Journal of the North American Benthological Society, 16(4): 833-852.

Richards, C., Johnson, L.B. and Host, G.E., 1996. Landscape-scale influences on stream habitats and biota. Canadian Journal of Fisheries and Aquatic Science, 53(Suppl. 1): 295-311.

Rochester, W.A., Accad, A., Low-Choy, S.J., Neldner, V.J., Pullar, D. and Williams, K., 2004. UQ-EPA ARC/SPIRT subregion classification project, Final Report, University of Queensland, Brisbane.

Schaffelke, B., Carleton, J., Zagorskis, I., Furnas, M.J., Skuza, M., Wright, M., Dekker, A.G., Blondeau-Patissier, D. and Brando, V.E., 2006. Nearshore marine water quality monitoring. , CRC Reef Consortium. Water Quality and Ecosystem Monitoring Programs–Reef Water Quality Protection Plan. Final Report. August 2006 (revision November 2006). CRC Reef Research, Townsville.

Schlosser, I.J., 1991. Stream Fish Ecology: A Landscape Perspective. BioScience, 41(10): 704-712.

Shears, N.T., Babcock, R.C., Duffy, C.A.J. and Walker, J.W., 2004. Validation of qualitative habitat descriptors commonly used to classify subtidal reef assemblages in north-eastern New Zealand. New Zealand Journal of Marine and Freshwater Research, 38(4): 743-752.

Simpson, J.C. and Norris, R.H., 2000. Biological assessment of river quality: development of AUSRIVAS models and outputs. In: J.F. Wright, D.W. Sutcliffe and M.T. Furse (Editors), Assessing the biological quality of fresh waters: RIVPACS and other techniques. Proceedings of an International Workshop held in Oxford, UK, 16-18 September 1997, Freshwater Biological Association, Cumbria, UK.

Smith, D.M. and Brough, D.M., 2006. Modelling the land surface: A report on the creation of digital elevation models for high priority regions of Queensland, The Queensland Department of Natural Resources, Mines and Water.

Smith, M.J., Bunn, S.E., Storey, A.W., Harch, B.D. and Redfern, F.M., 2001. Process undertaken to devlop an EHMP for rivers and streams in south east Queensland. In: A.W. Storey et al. (Editors), Task DIBM3: Design and implementation of baseline monitoring 3, Final Report. Moreton Bay Waterways andCatchments, , Brisbane, Australia.

Snelder, T., Biggs, B.J. and Weatherhead, M., 2004. New Zealand River Environment Classification User Guide, New Zealand Ministry for the Environment, Wellington, New Zealand.

Snelder, T.H. and Biggs, B.J.F., 2002. Multiscale River Environment Classification for water resources management Journal of the American Water Resources Association, 38(5): 1225-1239.

Snelder, T.H., Leathwick, J.R., Dey, K.L., Rowden, A.A., Weatherhead, M.A., Fenwick, G.D., Francis, M.P., Gorman, R.M., Grieve, J.M., Hadfield, M.G., Hewitt, J.E., Richardson, K.M., Uddstrom, M.J. and Zeldis, J.R., 2007. Development of an ecologic marine classification in the New Zealand region. Environmental Management, 39(1): 12-29.

Snyder, M.N., Goetz, S.J. and and Wright, R.K., 2005. Stream health rankings predicted by staellite derived land cover metrics. Journal of the American Water Resources Association, 41(3): 659-677.

Stebbing, A.R.D. and Dethlesfsen, V., 1992. Introduction to Bremerhaven workshop on biological effects of contaminants. Marine Ecology Progress Series, 91: 1-8.

Steven, A., Brando, V.E., Dekker, A.G., Fathring, B., Alston, C. and Mengersen, K., In Review. Applying remote sensing to monitoring water quality trends and compliance in the Great Barrier Reef Lagoon. Journal of Remote Sensing.

Stoddard, J.L., Larsen, D.P., Hawkins, C.P., Johnson, R.K. and Norris, R.H., 2005. Setting expectations for the ecological condition of streams: the concept of reference condition. Ecological Applications, 20: 181-197.

Strahler, A.N., 1957. Quantitative analysis of watershed geomorphology. Trans. Am. Geophys. Union, 21: 913-920.

Sweatman, H., 2007. Coral reef health indicators and thresholds of concern. MTSRF Milestone Report Project 1.1.2, Australian Institute of Marine Science, Townsville.

Sweatman, H., Thompson, A., Delean, S., Davidson, J. and Neale, S., 2004. Status of near-shore reefs of the Great Barrier Reef 2004. CRC Reef Research Centre Technical Report, Australian Institute of Marine Science, Townsville.

Theobald, D.M., Norman, J., Peterson, E. and Ferraz, S., 2005. Functional Linkage of Watersheds and Streams (FLoWs): Network-based ArcGIS tools to analyze freshwater ecosystems. Environmental Systems Research Institute, Inc.

Tibshirani, R., Walther, G. and Hastie, T., 2001a. Estimating the number of clusters in a data set via the gap statistic. Journal of the Royal Statistical Society Series B-Statistical Methodology, 63: 411-423.

Tibshirani, R., Walther, G. and Hastie, T., 2001b. Estimating the number of clusters in a data set via the gap statistic. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 63(2): 411-423.

Townend, 2003. Principle Component Analysis, Practical statistics for environmental and biological scientists. John Wiley & Sons Ltd., West Sussex, England, pp. 205-218.

Udy, J.W., Fellows, C.S., Bartkow, M.E., Bunn, S.E., Clapcott, J.E. and Harch, B.D., 2006. Measures of Nutrient Processes as Indicators of Stream Ecosystem Health. Hydrobiologia, 572(1): 89-102.

Underwood, A.J., Chapman, M.G. and Connell, S.D., 2000. Observations in ecology: you can't make progress on processes without understanding the patterns. Journal of Experimental Marine Biology and Ecology, 250(1-2): 97-115.

Vannote, R.L., Minshall, G.W., Cummins, K.W., Sedell, J.R. and Cushing, C.E., 1980. The River Continuum Concept. Canadian Journal of Fisheries and Aquatic Sciences, 37: 130-137.

Waite, I.R., Herlihy, A.T., Larsen, D.P. and Klemm, D.J., 2000. Comparing strengths of geographic and nongeographic classifications of stream benthic macroinvertebrates in the Mid-Atlantic Highlands, USA. Journal of the North American Benthological Society, 19(3): 429-441.

Wardrop, D.H., Bishop, J.A., Easterling, M., Hychka, K., Myers, W., Patil, G.P. and Taillie, C., 2005. Use of landscape and land use parameters for classification and characterization of watersheds in the mid-Atlantic across five physiographic provinces. Environmental and Ecological Statistics, 12(2): 209-223.

Wells, F., Metzeling, L. and Newall, P., 2002. Macroinvertebrate Regionalisation for use in the Management of Aquatic Ecosystems in Victoria, Australia. Environmental Monitoring and Assessment, 74(3): 271-294.

Wilkinson, S., Henderson, A. and Chen, Y., 2004. SedNet User Guide, CSIRO Land and Water, Canberra.

Wolock, D.M., Winter, T.C. and McMahon, G., 2004. Delineation and Evaluation of Hydrologic-Landscape Regions in the United States Using Geographic Information System Tools and Multivariate Statistical Analyses. Environmental Management, 34: S71-S88.

Wrigley, T., 2005. Canegrowers Public Environment Report, Canegrowers, Brisbane.

Appendix 1: Summary tables of draft indicators

Table 18: List of potential catchment indicators for the IRC.

Indicator Number	Catchment (*Pressure*) indicators	Considered useful/suitable by domain experts
1	Fertilser use	
a	Application rates nitrogen	Yes
b	Application rates phosphate	Yes
c	No. of properties with farm management plans e.g. COMPASS	Maybe
2	Land use/vegetation type (ground cover)	
a	land use type (% agricultural land)	Yes
b	point sources	Yes
c	land use condition (e.g. C class pasture)	Yes
3	Gullies and/or Drains	
a	Density	Yes
b	Cover	Yes
4	Riparian Zones	
a	% cover (both banks)	Yes
b	Width	Yes
c	Connectivity	Developmental
5	Pesticide use	
a	Proportion of catchment using pesticide	Yes
b	Application rates of Diuron, Atrazine, Simazin	Yes
6	Hydrology (Irrigation, Environmental flows)	
a	Environmental flows	spell analysis, total flows
b	Ground water extraction rates	Developmental - influence on wetlands

Table 19: List of potential freshwater indicators for the IRC.

Indicator Number	**Freshwater (*State*) Indicators**	**Considered useful/suitable by domain experts**
1	Physical/Chemical + hydrology	
a	DIN (Nitrate, Ammonia) and DIP	Yes
b	Particulate N and P	Yes
c	Total suspended solids	Yes
d	Turbidity/clarity	Yes
e	Chlorophyll a	Yes
f	Diuron	Yes
g	pH - Diel (24hr) range	Yes
h	Conductivity - Diel (24hr) range	Yes
I	Temperature - Diel (24hr) range	Yes
j	Disolved oxygen - Diel (24hr) range	Yes
k	Discharge (MAF)	Yes
l	Sediment and nutrient loads	Yes
2	Aquatic Macroinvertebrates	
a	species richness	Yes
b	Family richness	Yes
c	Signal	No
d	PET	No
3	Fish	
a	O/E species richness and assemblage	Yes
b	No. of alien species	Yes
c	% of alien species	Yes
d	trophic guild	No
e	mean density fish species	No
f	K abundance	No
4	Macrophytes	
a	% macrophyte cover edge habitat	Yes
b	species richness	No
c	% submerged taxa	No
d	% emergent taxa	No
e	% native taxa	No
f	% alien taxa	Yes
g	% poaceae	Yes
5	Habitat condition	
a	% riparian vegetation (both banks)	Yes
b	Bank integrity	developmental
c	habitat integrity (complexity)	developmental
d	habitat connectivity	developmental
e	median grain size O/E	developmental

Indicator Number	**Freshwater (*State*) Indicators**	**Considered useful/suitable by domain experts**
6	Ecosystem Processes	
a	Stream metabolism	?
b	Food web ecology	?

Table 20: List of potential marine indicators for the IRC.

Indicator Number	Marine Indicators	Considered useful/suitable
1	Corals and Reef Ecology	
b	Hard coral cover	yes
d	Hard coral richness	yes
c	Hard coral recruits	yes
e	Soft coral richness	yes
f	RNA : DNA ratio	developmental
g	Coral colour	developmental
h	Tissue thickness in massive coral	developmental
i	Macro-bioeroders in Porites	yes
j	Deepest depth of reef development	developmental
k	Herbivorous fish density, biomass, spp composiiton	yes
2	Macroalgae	
a	Cover	yes
b	community composition	maybe
3	Seagrasses	
a	% seagrass cover	yes
b	Seagrass species composition, biomass	yes
c	tissue nutrient content	developmental
4	Biofilms	
a	Foraminifera	yes
b	Benthic diatoms	developmental
c	benthic bacteria	developmental
5	Water Quality	
a	Chlorophyll a	yes
b	Secchi Disk (light)	yes
c	Nitrate	maybe
d	Particulate N	yes
e	Particulate P	yes
f	CDOM	yes
g	Turbidity	yes
h	Suspended sediment	yes
i	Sediment composition, grain size	maybe
j	Pesticide/Herbicide (which one/s???)	yes
6	Barramundi / Crabs / Fish	
a	EROD	developmental
b	DNA damage	developmental
c	RNA/DNA	developmental

Indicator Number	Marine Indicators	Considered useful/suitable
d	AChE	developmental
e	Fluorescent aromatic compounds	developmental
f	Condition factor	developmental
g	S.serrata (mud crub) bioaccumulation	possible although limited data

Appendix 2: Metadata for freshwater regionalisation data sets

Table 21: Source, spatial scale, and pre-processing methods.

Dataset	**Variables**	**Scale**	**Source**
SILO	mean annual daily rainfall (1961-1990) * 365.25	5 km	Jeffrey et al., 2001
	mean annual minimum temperature during the coldest period (1961-1990)		
	mean annual maximum temperature during the warmest period (1961-1990)		
	mean annual maximum temperature during the warmest period (1961-1990) - mean annual minimum temperature during the coldest period (1961-1990)		
	mean annual solar radiation (1961-1990)		
	mean annual vapour pressure (1961-1990)		
	mean annual evapotranspiration (1961-1990)		
	potential evapotranspiration/rainfall ratio		
The Australian Soil Resource Information System (ASRIS)	Sand (see texture reclass table for details)	1:50000, 1:100000, & 1:250000	McKenzie et al., 2005
	Loamy sand (see texture reclass table for details)		
	Loam (see texture reclass table for details)		
	Clay loam (see texture reclass table for details)		
	Light clay (see texture reclass table for details)		
	Peat (see texture reclass table for details)		
	pH		
	Sum of exchangeable bases		
	Percent clay		
	Cation exchange capacity		
	Electrical conductivity		
	Organic content		

Dataset	Variables	Scale	Source
Geoscience Geologic Mapping Data	Mafic rock type (see geology reclass table) Felsic rock type (see geology reclass table) Felsic plutonic rock type (see geology reclass table) Sedimentary rock type (see geology reclass table) Unconsolidated rock type (see geology reclass table)	1:100,000	Natural Resources Mines and Energy, 2004
Digital elevation model	Elevation - also used to calculate slope	25 meter	Smith and Brough, 2006
AUSLIG 1:100,000 Topographic Maps	Streams were edited by Queensland Natural Resources and Water. Additional reaches were digitised based on ASLIG repromats and all line segments were flow corrected.	1:100,000	Geoscience Australia, 2004
Erosion potential	Rkls (r*k*l*s) produced from the SedNet model	1 km	Wilkinson et al., 2004

Table 22: Geologic mapping data reclass table.

geol_reclass	DOMINANT_R
MAFIC	"BASALT"
SEDIM	"ARENITE-MUDROCK"
SEDIM	"MUDROCK"
FELSIC	"FELSITES (LAVAS, CLASTICS & HIGH-LEVEL INTRUSIVES)"
FELSPLUT	"GRANITOID"
UNCONS	"ALLUVIUM"
UNCONS	"SAND"
UNCONS	"MISCELLANEOUS UNCONSOLIDATED SEDIMENTS"

Table 23: ASRIS soil texture reclass table.

Texture	Descript	Group	Grp_nam	Northcote	Northc_descr
CL	Clay loam	CL	Clay loam	4	Clay Loam
FSCL	Fine sandy clay loam	CL	Clay loam	4	Clay Loam
MSCL	Medium sandy clay loam	CL	Clay loam	4	Clay Loam
KSCL	Coarse sandy clay loam	CL	Clay loam	4	Clay Loam
CLA	Sapric clay loam	CL	Clay loam	4	Clay Loam
CLI	Fibric clay loam	CL	Clay loam	4	Clay Loam
CLS	Clay loam, sandy	CLS	Clay loam, sandy	4	Clay Loam
CLFS	Clay loam, fine sandy	CLS	Clay loam, sandy	4	Clay Loam
CLMS	Clay loam, medium sandy	CLS	Clay loam, sandy	4	Clay Loam
CLKS	Clay loam, coarse sandy	CLS	Clay loam, sandy	4	Clay Loam
CLSA	Sapric clay loam, sandy	CLS	Clay loam, sandy	4	Clay Loam
CLSI	Fibric clay loam, sandy	CLS	Clay loam, sandy	4	Clay Loam
ZCL	Silty clay loam	ZCL	Silty clay loam	4	Clay Loam
ZCLA	Sapric silty clay loam	ZCL	Silty clay loam	4	Clay Loam
ZCLI	Fibric silty clay loam	ZCL	Silty clay loam	4	Clay Loam
LC	Light clay	LC	Light clay	5	Light Clay
SLC	Sandy light clay	LC	Light clay	5	Light Clay
FSLC	Fine sandy light clay	LC	Light clay	5	Light Clay
MSLC	Medium sandy light clay	LC	Light clay	5	Light Clay
KSLC	Coarse sandy light clay	LC	Light clay	5	Light Clay
ZLC	Silty light clay	LC	Light clay	5	Light Clay
LCA	Sapric light clay	LC	Light clay	5	Light Clay
LCI	Fibric light clay	LC	Light clay	5	Light Clay
LMC	Light medium clay	LMC	Light medium clay	5	Light Clay
ZLMC	Silty light medium clay	LMC	Light medium clay	5	Light Clay
SLMC	Sandy light medium clay	LMC	Light medium clay	5	Light Clay

Texture	Descript	Group	Grp_nam	Northcote	Northc_descr
FSLMC	Fine sandy light medium clay	LMC	Light medium clay	5	Light Clay
MSLMC	Medium sandy light medium clay	LMC	Light medium clay	5	Light Clay
KSLMC	Coarse sandy light medium clay	LMC	Light medium clay	5	Light Clay
LMCA	Sapric light medium clay	LMC	Light medium clay	5	Light Clay
LMCI	Fibric light medium clay	LMC	Light medium clay	5	Light Clay
FSC	non-standard - equates to FSLC	LC	Light clay	5	Light Clay
SC	non-standard - equates to SLC	LC	Light clay	5	Light Clay
ZC	non-standard - equates to ZLC	LC	Light clay	5	Light Clay
L	Loam	L	Loam	3	Loam
LA	Sapric loam	L	Loam	3	Loam
LI	Fibric loam	L	Loam	3	Loam
ZL	Silty loam	ZL	Silty loam	3	Loam
ZLA	Sapric silty loam	ZL	Silty loam	3	Loam
ZLI	Fibric silty loam	ZL	Silty loam	3	Loam
SCL	Sandy clay loam	SCL	Sandy clay loam	3	Loam
SCLFS	Sandy clay loam, fine sand	SCL	Sandy clay loam	3	Loam
SCLA	Sapric sandy clay loam	SCL	Sandy clay loam	3	Loam
SCLI	Fibric sandy clay loam	SCL	Sandy clay loam	3	Loam
SL	Sandy loam	SL	Sandy loam	2	Loamy Sand
FSL	Fine sandy loam 5	SL	Sandy loam	2	Loamy Sand
MSL	Medium sandy loam	SL	Sandy loam	2	Loamy Sand
KSL	Coarse sandy loam	SL	Sandy loam	2	Loamy Sand
SLA	Sapric sandy loam	SL	Sandy loam	2	Loamy Sand
SLI	Fibric sandy loam	SL	Sandy loam	2	Loamy Sand
LFSY	non-standard - equates to FSL	SL	Sandy loam	2	Loamy Sand

Texture	Descript	Group	Grp_nam	Northcote	Northc_descr
AP	Sapric peat	AP	Sapric peat	0	Peat
SP	Sandy peat	SP	Sandy peat	0	Peat
LP	Loamy peat	LP	Loamy peat	0	Peat
CP	Clayey peat	CP	Clayey peat	0	Peat
GP	Granular peat	GP	Granular peat	0	Peat
HP	Hemic peat	HP	Hemic peat	0	Peat
IP	Fibric peat	IP	Fibric peat	0	Peat
S	Sand	S	Sand	1	Sand
FS	Fine sand	S	Sand	1	Sand
MS	Medium sand	S	Sand	1	Sand
KS	Coarse sand	S	Sand	1	Sand
SA	Sapric sand	S	Sand	1	Sand
SI	Fibric sand	S	Sand	1	Sand
LS	Loamy sand	LS	Loamy sand	1	Sand
LFS	Loamy fine sand	LS	Loamy sand	1	Sand
LMS	Loamy medium sand	LS	Loamy sand	1	Sand
LKS	Loamy coarse sand	LS	Loamy sand	1	Sand
LSA	Sapric loamy sand	LS	Loamy sand	1	Sand
LSI	Fibric loamy sand	LS	Loamy sand	1	Sand
CS	Clayey sand	CS	Clayey sand	1	Sand
CFS	Clayey fine sand	CS	Clayey sand	1	Sand
CMS	Clayey medium sand	CS	Clayey sand	1	Sand
CKS	Clayey coarse sand	CS	Clayey sand	1	Sand
CSA	Sapric clayey sand	CS	Clayey sand	1	Sand
CSI	Fibric clayey sand	CS	Clayey sand	1	Sand

Table 24: Reach, RCA (reach contributing area), and watershed variables used in the freshwater regionalisation.

Scale	Variable	Description
Reach	Shape_Length *	Length of the stream reach (m)
	rchAreakm2 †	Reach area based on a raster representation (km2)
	perPeatRch †	Percent peat soil type (%)
	perLtclRch †	Percent light clay (%)
	perLoamRch	Percent loam soil type (%)
	perLmsndRch *	Percent loamy sand soil type (%)
	perCllmRch †	Percent clay loam soil type (%)
	perFelsicRch	Percent felsic rock type (%)
	perFelsputRch	Percent felsic plutonic rock type (%)
	perMaficRch †	Percent mafic rock type (%)
	perSedimRch †	Percent sedimentary rock type (%)
	perUnconsRch	Percent unconsolidated rock type (%)
	rchElevMn *	Mean elevation (m)
	rchTrangeMn *	Mean annual temperature range (°C)
	rchTminMn *	Mean annual minimum temperature (°C)
	rchTmaxMn *	Mean annual maximum temperature (°C)
	rchRadMn *	Mean annual radiation (mj/day)
	rchPetMn *	Ratio of mean annual rainfall to mean annual potential evapotranspiration
	rchMarMn *	Mean annual rainfall (mm/year)
	rchEvapMn *	Mean daily potential evapotranspiration (mm/day)
	rchVpMn *	Mean vapour pressure (millibars)
	rchCecMn *	Mean cation exchange capacity (cmol(+)/kg)
	rchClayMn *	Percent clay in the reach area (%)
	rchEcMn †	Mean electrical conductivity (dS/m)
	rchExbMn *	Mean sum of exhangeable bases in the top soil layer (cmol(+)/kg)
	rchOcMn *	Mean organic carbon (%)
	rchPhMn	Mean soil pH
	rchrklsMn	Erosion potential produced from the SedNet model
	rchSlopeMn	Mean slope at the reach (degrees)

Scale	Variable	Description
RCA	rca_Area_k †	RCA area (km2)
	perMaficRca †	Percent mafic rock type (%)
	perFelsicRca	Percent felsic rock type (%)
	perFelsputRca	Percent felsic plutonic rock type (%)
	perSedimRca †	Percent sedimentary rock type (%)
	perUnconsRca	Percent unconsolidated rock type (%)
	perPeatRca †	Percent peat soil type (%)
	perLtclRca †	Percent light clay soil type (%)
	perLoamRca †	Percent loam soil type (%)
	perLmsndRca	Percent loam sand soil type (%)
	perCllmRca	Percent clay loam soil type (%)
	rcaCecMn	Mean cation exchange capacity (cmol(+)/kg)
	rcaEcMn †	Mean electrical conductivity (dS/m)
	rcaExbMn	Mean sum of exhangeable bases in the top soil layer (cmol(+)/kg)
	rcaOcMn	Mean organic carbon (%)
	rcaPhMn	Mean soil pH
	rcaEvapMn *	Mean daily potential evapotranspiration (mm/day)
	rcaMarMn *	Mean annual rainfall (mm/year)
	rcaPetMn *	Ratio of mean annual rainfall to mean annual potential evapotranspiration
	rcaRadMn *	Mean annual radiation (mj/day)

* Variables that were removed due to collinearity

† Variables removed after the PCA analysis

Scale	Variable	Description
RCA	rcaTmaxMn †	Mean annual maximum temperature (°C)
	rcaTminMn *	Mean annual minimum temperature (°C)
	rcaTrangeMn *	Mean annual temperature range (°C)
	rcaVpMn *	Mean vapour pressure (millibars)
	rcaClayMn	Percent clay in the reach area (%)
	rcaRklsMn	Erosion potential produced from the SedNet model
	rcaElevMn	Mean elevation (m)
	rcaSlopeMn	Mean slope (degree)
Watershed	h2OAreakm2[†]	Watershed area (km^2)
	h2ORklsMn	Mean erosion potential produced from the SedNet model
	h2OClayMn	Percent clay in the reach area (%)
	h2OVpMn *	Mean vapour pressure (millibars)
	h2OTrangeMn[†]	Mean annual temperature range (°C)
	h2OTminMn *	Mean annual minimum temperature (°C)
	h2OTmaxMn	Mean annual maximum temperature (°C)
	h2ORadMn *	Mean annual radiation (mj/day)
	h2OPetMn *	Ratio of mean annual rainfall to mean annual potential evapotranspiration
	h2OMarMn[†]	Mean annual rainfall (mm/year)
	h2OEvapMn *	Mean daily potential evapotranspiration (mm/day)
	h2OPhMn	Mean soil pH
	h2OOcMn	Mean organic carbon (%)
	h2OExbMn	Mean sum of exhangeable bases in the top soil layer (cmol(+)/kg)
	h2OEcMn[†]	Mean electrical conductivity (dS/m)
	h2OCecMn	Mean cation exchange capacity (cmol(+)/kg)
	h2OElevMn	Mean elevation (m)
	perMaficH2O[†]	Percent mafic rock type (%)
	perFelsicH2O	Percent felsic rock type (%)
	perFelsputH2O	Percent felsic plutonic rock type (%)
	perSedimH2O[†]	Percent sedimentary rock type (%)
	perUnconsH2O	Percent unconsolidated rock type (%)
	perPeatH2O †	Percent peat soil type (%)
	perLtclH2O †	Percent light clay soil type (%)
	perLoamH2O †	Percent loam soil type (%)
	perLmsndH2O	Percent loam sand soil type (%)
	perCllmH2O †	Percent clay loam soil type (%)
	h2OSlopeMn †	Mean slope in the watershed (degrees)

* Variables that were removed due to collinearity

[†] Variables removed after the PCA analysis