

Reef Rescue Research & Development

Technical Report

Developing integrated assessment metrics for reporting
of water quality in the Great Barrier Reef lagoon

RRRD016



Vittorio E Brando, Michelle Devlin, Melissa Dobbie,
Aaron MacNeil, Britta Schaffelke and Thomas Schroeder



Developing integrated assessment metrics for reporting of water quality in the Great Barrier Reef lagoon

(Project RRRD016)

Vittorio E Brando¹, Michelle Devlin², Melissa Dobbie¹, Aaron MacNeil³, Britta Schaffelke³ and Thomas Schroeder¹

¹CSIRO

²TropWATER James Cook University

³Australian Institute of Marine Science



Australian Government

Department of the Environment

Supported by the Australian Government's
Reef Rescue Research and Development Program



© CSIRO 2014. To the extent permitted by law, all rights are reserved & no part of this publication covered by copyright may be reproduced or copied in any form or by any means except with written permission of CSIRO.

National Library of Australia Cataloguing-in-Publication entry:

ISBN: 978-1-925088-43-4

This report should be cited as:

Brando, V.E., Devlin, M.J., Dobbie, M., MacNeil, A., Schaffelke, B. and Schroeder, T. (2014). Developing integrated assessment metrics for reporting of water quality in the Great Barrier Reef lagoon. Report to the Reef Rescue Water Quality Research & Development Program. Reef and Rainforest Research Centre Limited, Cairns (53pp.). ISBN: 978-1-925088-43-4.

The Reef Rescue Research and Development Program (Reef Rescue R&D) was funded under the Australian Government's Caring for our Country program, and consisted of 18 projects. The Reef Rescue R&D is designed to improve our understanding of the link between land management practices and environmental impacts, and improve water quality across the Great Barrier Reef by informing on-ground projects to reduce the amounts of nutrients, pesticides and sediments reaching the Reef from agricultural lands.

This publication is copyright. The Copyright Act 1968 permits fair dealing for study, research, information or educational purposes subject to inclusion of a sufficient acknowledgement of the source.

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. While reasonable effort has 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.

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.

Cover photographs: Top left: Sampling the Tully plume; Top right: Plume waters from a Wet Tropics river; Bottom left: River plume discharging into the GBR; Bottom right: Full plume extent visualised from MODIS imagery (Courtesy NOAA).

This report is available from the Reef Rescue R&D website: <http://www.reefrescueresearch.com.au>

Contents

List of Figures.....	iii
List of Tables.....	iv
Acronyms.....	vi
Abbreviations	vi
Acknowledgements	vi
Executive Summary	1
1. Introduction.....	4
1.1. Background.....	4
1.2. Project Objectives.....	5
1.3. Summary of review of national and international water quality assessment and reporting approaches	7
2. Methods	9
2.1. Study area.....	9
2.1.1. <i>Ambient water quality sites</i>	9
2.1.2. <i>Flood plume sites</i>	9
2.2. Data attributes and availability	10
2.3. A proposed spatial scale for analysis and reporting	16
2.4. A proposed temporal scale for analysis and reporting	19
2.5. Proposed data integration approach	21
2.5.1. <i>Overview</i>	21
2.5.2. <i>Step 1: collating/checking/cleaning data for each indicator</i>	22
2.5.3. <i>Step 2: form a modeled response surface for each indicator</i>	22
2.5.4. <i>Step 2a: merging data sources for an indicator</i>	22
2.5.5. <i>Step 2b: a proposed model framework</i>	23
2.5.6. <i>Step 3: Calculate a condition “score” surface for each indicator</i>	26
2.5.7. <i>Step 4: Assessment of condition relative to regional, spatial and seasonal guidelines</i>	27
2.5.8. <i>Step 4a: calculate a condition “score” surface – combined indicators</i>	28
2.5.9. <i>Step 4b: aggregate condition scores – separate indicators</i>	28
2.5.10. <i>Step 5: combine aggregated scores across indicators</i>	28
3. Results	29
3.1. Proof of concept of water quality reporting framework	29
3.2. Calculation of assessment scores for TSS.....	30

3.2.1.	<i>Calculate seasonal averages (per pixel)for TSS concentrations using statistical model.....</i>	30
3.2.2.	<i>Assess the magnitude of non-compliance of TSS with the water quality guidelines</i>	32
3.3.	Calculation of assessment scores for DIN and PN.....	35
3.3.1.	<i>Calculate seasonal averages (per grid) for DIN and PN concentrations using in-situ data only.....</i>	35
3.4.	Integration of regional assessment scores between water quality variables	40
4.	Discussion	43
4.1.	<i>Statistical model.....</i>	43
4.2.	<i>Proof of concept – integrated assessment.....</i>	43
4.3.	<i>Future research</i>	45
References		46

List of Figures

Figure 1-1: The section of first Annual Report Card showing the Land Practice, Catchment and Marine baselines, based on 2008-09 data.	5
Figure 1-2: Conceptual process of eutrophication responses contrasted in temperate and tropical waters.....	7
Figure 2-1: Location of flood plume sites sampled during the 2010-2011 wet season for the Tully region only. The edge of the plume is illustrated as high frequency of plumes recorded for at least 13 out of 22 weeks in wet season (red line), medium frequency recorded at 9 out of 22 weeks (yellow line) and low frequency recorded for at least 4 weeks (green line).	10
Figure 2-2: Spatial representation of data availability for the four data sources in the Wet Tropics region in front of the Tully and Murray river mouths. The Logger is located at Dunk Island.	15
Figure 2-3: Temporal representation of data availability for the four data sources in the Wet Tropics region in front of the Tully and Murray river mouths. The Logger is located at Dunk Island.	15
Figure 2-4: Spatial extent of river influenced waters as defined by Coloured dissolved organic matter (CDOM) as a proxy for reduced salinity (Schroeder et al., 2012). The box represents the reporting area as defined by the cross shelf extent of the river influenced waters (red to green).	17
Figure 2-5: Spatial extent of river influenced waters as defined by ocean colour as a measure of distribution and frequency of plume influenced water (Devlin et al., 2012; Alvarez-Romero et al., 2013). The boxes represents the reporting area as defined by the cross shelf extent of the river influenced waters (red to green).....	18
Figure 2-6: Spatial extent of river influenced waters as defined by a cumulative exposure index based on a GBR hydrodynamic model (from Furnas et al., 2013). The colour bar indicates the calculated cumulative exposure ("Conc.Days", concentration x days) above 1% of the incoming concentration (see Furnas et al. 2011 for details). Contours show 0.1, 1.0 and 10.0 Conc.Days exposure levels and the reporting boundary is defined by the 10 Conc.Days contour line.	18
Figure 2-7: Spatial extent of river influenced waters as defined by GBR bathymetry	19
Figure 2-8: The location of three locations within the Tully reporting area.....	20
Figure 2-9: Seasonal averages for satellite-derived TSS for three locations in the Wet Tropics NRM region that are located in the vicinity of the Tully and Murray River mouths in the 2010-11 water reporting year.months.	20
Figure 2-10: Overview of proposed GBR water quality reporting framework	21
Figure 3-1:Flow chart outlining the process of assessment using four water quality parameters, including TSS, chl-a, DIN and PN. Note that in this report, chl-a is not included.	30
Figure 3-2: Modeled average TSS concentration (mg L^{-1}) for four reporting quarters (Q1 to Q4) during the 2010-11 Water Year. a) Q1= October 2010 to December 2010, b) Q2= January 2011 to March 2011, c) Q3= April 2011 to June 2011, d) Q4= July 2011 to September 2011.31	

Figure 3-3: Compliance of average TSS concentration with the seasonally adjusted GBR water quality guideline trigger values (GL;GBRMPA 2010) in four reporting quarters (Q1 to Q4) during the 2010-11 Water Year. Green pixels have a quarterly average < GL and are allocated a compliance score C of 0, red pixels are >GL and have a C of 1.	32
Figure 3-4: TSS assessment scores in four reporting quarters during the 2010-11 Water Year, based on assessments of the magnitude of non-compliance of average TSS concentrations. See text for further explanation of the method.	34
Figure 3-5: Example reporting area for the Tully marine area. Area highlighted in (a) represents the area aligned to the 91 km ² sub-region of the Wet Tropics NRM region reported in Figure 2.3. (b) Reporting is expressed for DIN and PN per 10km ² grid. The frequency of data points are illustrated in (b) by the colour, with mean average DIN reported in each reporting grid. Grid squares are identified in (c)	36
Figure 3-6: The mean DIN and PN values for 2010-2011 as compared against the WQ guideline and the ECS presented for two grid sub-reporting areas including (a) 18S12) and (b) 18S22. Water quality guidelines for DIN is presented as an annual mean (Moss et al., 2005) and seasonally adjusted for PN (see GBRMPA, 2010). The quarterly ECS has been calculated from the long term historical data (1988 – 2010) within each sub-reporting area.....	38
Figure 3-7: Conceptual diagram of overall water quality assessment with the aggregation of each water quality assessment scores.	42

List of Tables

Table 2-1: Summary of available data for making a GBR water quality condition assessment	13
Table 2-2: Description of spatial data sources that can be used to describe a sub-region reporting area.....	16
Table 2-3: Spatial and temporal footprints for point observations. Note that Ambient-Niskin monitoring was only carried out typically three times per year: nominally February, June and September/October, whereas Event-Niskin monitoring was carried out only when “events” occurred. A temporal footprint encompasses the day of the observation so represents condition before and after the monitoring has occurred. Dry season = July, August, September. Wet season = January, February, March. ‘In-between’-seasons are the months between wet and dry seasons.	23
Table 3-1: The values calculated for TSS for the reporting area (See Figure xx), including the average TSS value predicted from surface model, the average magnitude on non-compliance against the WQ guidelines, the average magnitude o per quarter average MNC (magnitude of non-compliant pixels) and overall compliance score for TSS reported for each quarter within the selected reporting area for the Tully region.....	33
Table 3-2: Mean DIN and PN value calculated for each quarter in the 2010-2011 year for each grid code within the Tully marine reporting area.	37

Table 3-3: Average values for DIN and PN for each grid area, reported as compliance scores (c), ECS values and MNC (magnitude of non-compliance) are shown for each grid contained in the Tully marine reporting area. Orange shading denotes grids that do not have sufficient data points (>10) to calculate an ECS or no data was collected in 2010-11. Cells with *** denotes grids where MNC could not be calculated as the ECS was not available or it was lower than the guideline values..... 39

Table 3-4: Average DIN values for the Tully region, reported for each quarter. The percent value of non-compliant DIN is calculated from the number of grid squares with value greater than WQ guideline. The DIN assessment score is calculated from the average MNC score per grid square..... 40

Table 3-5: Average PN values for the Tully region, reported for each quarter. The percent value of non-compliant DIN is calculated from the number of grid squares with value greater than WQ guideline. The DIN assessment score is calculated from the average MNC score per grid square..... 40

Table 3-6: Aggregated scores taken across all water quality indicators in the Tully region (see Figure 2-9). The aggregated score is calculated from average of the assessment scores for the three water quality parameters used in this report. The cells are colour coded according to the resulting report card rating. 41

Acronyms

RRRC	Reef and Rainforest Research Centre Limited
RRRD	Reef Rescue Research & Development
CSIRO	Commonwealth Scientific Industrial Research Organisation
GBR	Great Barrier Reef
GBRMPA	Great Barrier Reef Marine Park Authority
JCU	James Cook University
NRM	Natural Resource Management
P2R	Paddock to Reef Integrated Monitoring, Modelling and Reporting
RRRC	Reef and Rainforest Research Centre Limited
RRRD	Reef Rescue Research & Development
WfHC	Water for a Healthy Country

Abbreviations

- Chl-a: Chlorophyll-a
- MODIS: Moderate Resolution Imaging Spectroradiometer
- MODIS products are organized into different processing levels:
- Level 0 (L0) product: raw radiance counts from all bands;
- Level 1B (L1B) product: calibrated and geolocated radiances;
- Level 2 (L2) product: geophysical product for each pixel (after application of atmospheric correction and bio-optical algorithms).
- GBR: Great Barrier Reef
- CDOM: coloured dissolved organic matter
- Case 1 waters: waters in which phytoplankton (with their accompanying and covarying retinue of material of biological origin) are the principal agents responsible for variations in optical properties of the water
- Case 2 waters: waters influenced not just by phytoplankton and related particles, but also by other substances, that vary independently of phytoplankton, notably inorganic particles in suspension and yellow substances

Acknowledgements

The authors would like to acknowledge the support and funding from the Australian Government Reef Rescue R&D Program and CSIRO through the Wealth from Oceans Flagship who made the development of this modelling framework possible. In-kind support from CSIRO, AIMS and JCU is acknowledged in the completion of this project. We also acknowledge that this work has been undertaken in collaboration with projects in the Reef Rescue Marine Monitoring Program.

Executive Summary

This project aimed to develop integrated assessment metrics for the reporting of water quality in the Great Barrier Reef (GBR) lagoon. The Reef Water Quality Protection Plan, including the Australian Government funded Reef Programme (formerly Reef Rescue) have the express purpose of improving the condition of water quality and ecosystem health in the Great Barrier Reef Marine Park. Long-term monitoring of spatial and temporal trends in broad-scale water quality within the GBR Marine Park is central to assessment of Reef Plan performance. This monitoring has been carried out since 2005 with funding from the Australian Government. The MMP is part of the Paddock to Reef (P2R) Integrated Monitoring Modelling and Reporting Program (PRIMMR), which is a key action under Reef Plan to evaluate the efficiency of the implementation of Reef Plan and the progress towards its goal and targets.

The MMP water quality monitoring uses a combination of three complementary approaches to collect data at various spatial (site, location, region, and whole GBR lagoon) and temporal (snapshot, daily, 10-minute) scales, using traditional direct water sampling from research vessels, in situ data loggers at a small number of selected inshore reef locations and remote sensing techniques. Two indicator variables, suspended solids (an indicator of water clarity) and the plant pigment chlorophyll (an indicator of phytoplankton biomass and a proxy for nutrient availability), are measured by all three techniques. In addition, direct water sampling provides information on a wider suite of water quality variables, such as nutrients. While data loggers provide detailed information on the local variability in water quality parameters, remote sensing observations provide extensive spatial coverage at 1 km resolution, nominally on a daily basis (except e.g. overcast days).

Given the spatial and temporal variability of the data, an integrated framework that enables a comprehensive assessment of condition and trends of GBR water quality is required to provide environmental managers and policy makers with the information needed to make decisions for the management of land runoff. The integrated assessment method developed as part of this project combines a selection of key indicators to enable a reasonable evaluation of the overall status of coastal and marine waters. The assessment framework was developed to include the analyses of spatial and temporal variability of each data source, the development of metrics/indices that combine and scale up the three approaches used to collect water quality data (direct sampling, in situ data loggers and remote sensing), the estimation of uncertainty of combined indices, as well as the development of statistical methods to assess performance with regard to the existing GBR Water Quality Guidelines (GBRMPA, 2010).

The development of the water quality assessment and reporting framework was closely aligned with other nationally and international approaches for ecological monitoring, assessment and reporting. A review was carried out to assess the suitability for application in the GBR of national and international marine water quality assessment frameworks and their approaches and methods. This review found that the monitoring and reporting was often carried out at high temporal frequency and in small reporting areas where it should be reasonable to expect measurable change in response to variations of the environmental pressures that affect marine water quality.

In light of the findings of the review, we defined the criteria to determine the appropriate spatial and temporal scale for analysis and reporting for the assessment framework:

- In the GBR, the direct influence of land run-off occurs at a smaller scale than at the scale of entire marine NRM regions, as currently used for P2R reporting. Hence we propose to define smaller regions that are directly influenced by land run-off and consistent with local oceanography (e.g., residence times, bathymetry, hydrodynamics). In these smaller reporting areas it is reasonable to expect measurable change over the next decade in response to improvement in land management practices on the GBR catchment.
- To allow consideration of the variable influences during the dry and wet season, we propose to assess the condition of water relative to the seasonal guidelines (four seasons per year). This should provide an improved capacity to untangle the effects of land management practices from those directly due to the climate events, overcoming some of the limitations of the current P2R reporting of annual water quality summary data.

To aggregate multiple sources of data across various spatial and temporal scales in the reporting framework, a spatio-temporal statistical process model was developed. This allowed modelling of the average “response surfaces” for each indicator as well as the associated uncertainties. These seasonal “average response surfaces” for each indicator can then be used to assess compliance to the guidelines as well as provide a measure of the year-to-year variability and, ultimately, long-term trends.

For the proof of concept of the reporting framework we selected an area with the most data, in both space and time, from direct water sampling. A small reporting region around the Tully River mouth/Rockingham Bay was defined, using the criteria outlined above, which encompassed most of these data points. The reporting year 2010-11 was selected for this study, as in this year extreme floods occurred along the whole of the GBR coast. By selecting the most data-rich region and the most extreme year we aimed to maximise the capacity of the proposed reporting framework to show seasonal and spatial differences.

For the proof of concept of the reporting framework, two parameters were selected: (i) Total Suspended Solids measured by remote sensing, direct water sampling both during floods and ambient conditions, and water quality instruments on coral reefs, and (ii) Dissolved Inorganic Nitrogen and Particulate Nitrogen values measured only by direct water sampling both during floods and ambient conditions.

We demonstrated that the approach would be fully compatible with the current Paddock to Reef reporting and builds on the current model of tiered reporting across regional areas. The implementation of the proposed framework would need to be considered in the context of potential future changes to the Paddock to Reef monitoring and reporting structure.

For this case study, we have focused on a small reporting area (Tully marine area) as it the most data rich in the GBR. Implementation on a regular basis for whole Great Barrier Reef World Heritage Area (GBR WHA) is a complex computational exercise. To be able to perform the data integration across temporal and spatial scales at a GBR scale, the spatio-temporal statistical process model developed for the proof of concept would need to be implemented in parallel computing environments.

The proof of concept highlighted several research and development needs related to the implementation of the reporting framework for the inshore waters across the whole GBRWHA:

- The proposed framework included the definition of new smaller, localised coastal reporting regions to provide a better reporting base for water quality assessments based on the influence of the river flow and water quality associated with that water flow. These smaller reporting regions need to be defined, based on the decision-making criteria proposed in the assessment framework.
- In the proposed framework, all variables were normalised and no weighting was applied to the combination of the water quality variables. Future applications of this approach needs to consider the weighting of the different water quality variables based on understanding of relevant ecological thresholds for GBR ecosystems such as corals and seagrasses. Weighting and/or thresholds need to be considered in both seasonal and geographical context
- Further development is required of the water quality guidelines for the GBR to define trigger values that are specific for regional characteristics and/or for coral and seagrass environments.
- The future integration of other water quality indicators, such as pesticides, will require an investigation of the appropriateness of the proposed framework for spatial and temporal data aggregations.

1. Introduction

1.1. Background

Coastal areas around the world are under increasing pressure from human population growth, intensifying land use and urban and industrial development. As a result, increased loads of suspended sediment, nutrients and pollutants, such as pesticides and other chemicals, invariably enter coastal waters and lead to a decline in estuarine and coastal marine water quality. This increase in sediment, nutrients and other pollutants results in eutrophication and increased turbidity. Many tropical coastal regions are considered to be at great risk because of strong economic and population growth paired with limited environmental management. However, after decades of decline, some areas along the coasts of wealthier countries, generally in the temperate northern hemisphere, are showing signs of water quality improvements due to significant regulatory and policy intervention over the last two decades (Cloern, 2001; Nixon, 2009).

It is well documented in the scientific literature that sediment and nutrient loads carried by land runoff into the coastal and inshore zones of the Great Barrier Reef (GBR) have increased since European settlement (Kroon et al. 2012). This increase has been implicated in the decline of some coral reefs and seagrass meadows in these zones (reviewed in Brodie et al. 2011, 2012). Concern about these negative effects of land runoff triggered the formulation of the Reef Water Quality Protection Plan (Reef Plan) for catchments adjacent to the GBR World Heritage Area by the Australian and Queensland governments in 2003 (Anon. 2003). The Reef Plan was revised and updated in 2009 (Anon. 2009a) and has two primary goals:

- immediate goal - to halt and reverse the decline in quality of water entering the Reef by 2013;
- Long-term goal - to ensure that by 2020 the quality of water entering the Reef from adjacent catchments has no detrimental impact on the health and resilience of the Great Barrier Reef.

Reef Plan actions also include the establishment of water quality monitoring programs extending from the paddock to the Reef (Anon. 2009b), to assess the effectiveness of the Reef Plan's implementation. The Reef Rescue Marine Monitoring Program (MMP), formerly known as Reef Water Quality Protection Plan Marine Monitoring Programme (Reef Plan MMP), was designed and developed by the Great Barrier Reef Marine Park Authority (GBRMPA) and routine monitoring of water quality in the coastal and inshore waters of the GBR lagoon commenced in 2005. A summary of the MMP's overall goals and objectives and a description of the sub-programs is available at: <http://e-atlas.org.au/content/rrmmp> and <http://www.gbrmpa.gov.au/about-the-reef/how-the-reefs-managed/science-and-research/our-monitoring-and-assessment-programs/reef-rescue-marine-monitoring-program>.

The MMP forms an integral part of the *Paddock to Reef Integrated Monitoring, Modelling and Reporting Program*, which is a key action of Reef Plan 2009 and is designed to evaluate the efficiency and effectiveness of implementation and report on progress towards the Reef Plan and Reef Rescue goals and targets. A key output of the Paddock to Reef Program is an annual report card, including an assessment of Reef water quality and ecosystem condition to which the MMP contributes assessments and information. To date, three Annual Report

Cards were released (available at www.reefplan.qld.gov.au). The first Annual Report Card, based on the 2008-09 data, which will serve as a baseline for future assessments, was released in August 2011. The Annual Report Cards for 2009-10 and 2010-11 were released in April and July 2013. In these three report cards the Paddock-to Reef marine water quality metric was based on the assessment of compliance to the Water Quality Guidelines (GBRMPA 2010; hereafter called 'Guidelines') using only remote sensing data (Figure 1-1).

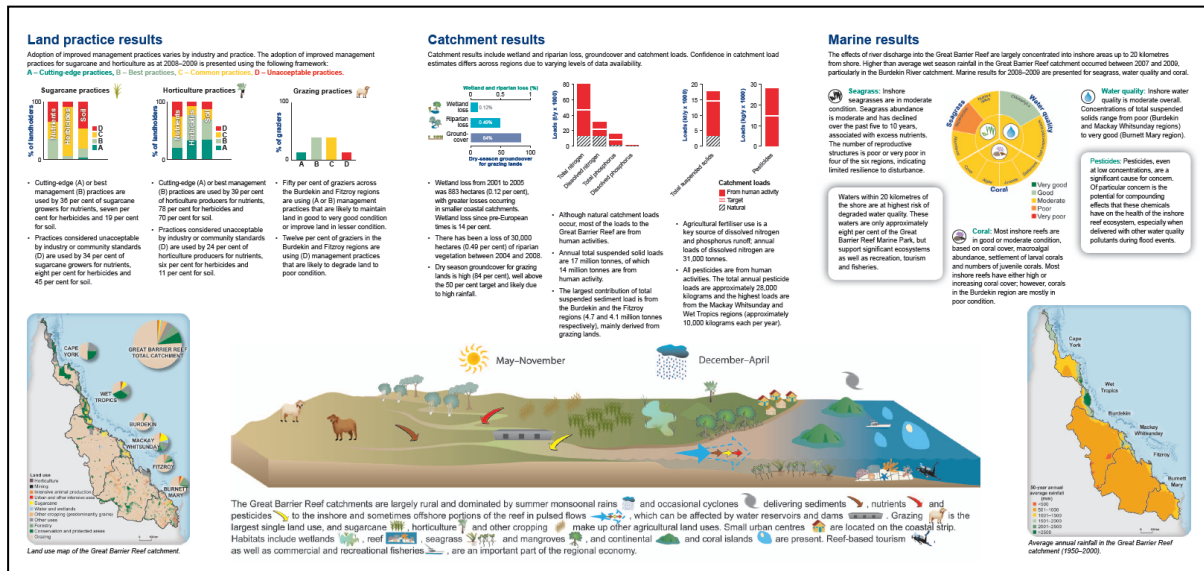


Figure 1-1: The section of first Annual Report Card showing the Land Practice, Catchment and Marine baselines, based on 2008-09 data.

1.2. Project Objectives

The MMP water quality monitoring uses a combination of three complementary approaches to collect data at various spatial and temporal (snapshot, daily, 10-minutely) scales: traditional direct water sampling from research vessels, *in situ* data loggers at a small number of selected inshore and midshelf reef locations and remote sensing techniques. Two water quality indicator variables, suspended solids (an indicator of water clarity) and the plant pigment chlorophyll (an indicator of phytoplankton biomass and a proxy for nutrient availability), are measured (or derived) by all three techniques. Direct water sampling provides additional information on a wider suite of water quality variables, such as nutrient concentrations. While data loggers provide detailed information on the local variability in water quality parameters, remote sensing observations provide extensive spatial coverage at 1 km² resolution, nominally on a daily basis (except e.g. overcast days). Given the multiple spatial and temporal sources of water quality data and the numerous indicators of water quality that have been recorded, an integrated assessment method, e.g. a water quality index, will allow us to use all relevant information to provide a comprehensive assessment of GBR water quality. This will provide environmental managers and policy makers with the information needed to make decisions for the management of land runoff.

The integrated assessment method developed as part of this project combines a selection of key indicators to enable a reasonable evaluation of the overall condition of coastal and marine waters. In developing such methods we need to address the spatial and temporal variability of measurements from each data stream, develop metrics/indices that combine and scale up these three approaches used to collect water quality data, address how to estimate uncertainty in the combined indices, as well as develop statistical methods to assess condition relative to the existing GBR Water Quality Guidelines (GBRMPA 2010).

The development of water quality assessment metrics for the GBR was closely aligned with other water quality condition assessments and reporting frameworks that have been developed nationally and internationally

The ultimate outcome of this project is a method for the ongoing reporting of a marine water quality index, which would be able to be fully integrated into future PRIMMR report cards. It is not possible, nor feasible to monitor or assess all parts of a marine system. As such, the development of an integrated assessment method, which makes use of a variety of data types, is critical to the assessment of inshore GBR water quality, particularly with respect to nutrient enrichment/eutrophication.

The original objectives of this project were:

1. Review international approaches on the assessment of WQ status/eutrophication;
2. Develop methodologies to integrate the different WQ data types into an overall WQ assessment. Subsequently test the reporting framework prototype using actual Reef Rescue MMP data for one NRM region.
3. Delivery of WQ reporting framework for a GBR-wide scale for input into the PRIMMR Water Quality 2013 WQ report card.

In consultation with key stakeholders, the second objective evolved towards the development of a proof-of-concept of the proposed reporting framework for a particular sub-region in a particular reporting year. Due to the time required for the development of the statistical tools, a full-scale implementation at a GBR-wide scale was not deemed feasible within the resources of the projects.

This report is structured as follows:

- The method section of this report summarises the findings of the review and presents the rationale for the development of the proposed reporting framework. It also provides the details of the spatio-temporal statistical model that enables the integration of different water quality variables across different spatial and temporal scales. The model delivers a predicted average response surface for each indicator and the predicted associated uncertainties.
- The results section describes the proof-of-concept of the reporting framework to a sub-region that was carefully selected to have the most data, in both space and time, from direct water sampling.
- The discussion section outlines how the proposed integrating reporting framework should be implemented within the Paddock-to-Reef (P2R) reporting structure.

1.3. Summary of review of national and international water quality assessment and reporting approaches

Eutrophication in tropical waters is driven by high nutrient values; however the delivery of these nutrients tends to be during high episodic pulses experienced during wet monsoonal summer periods. Tropical waters, particularly coral reefs, are dominated by picoplankton and can experience shifts to larger taxa with on-going consequences for the food web (Furnas et al., 2005; Brodie et al., 2011; Devlin et al., 2012b). Macroalgae can also increase and shift dominance from coral and seagrass systems to macroalgal dominated systems. A more complete analysis of the complex effects of nutrient loading in coral reef ecosystems has been reviewed by Fabricius (2005) and includes other confounding terrestrial pollutants such as suspended sediments. Overall it is clear that nutrient enrichment and subsequent eutrophication present quite different symptoms and consequences in tropical coral reef ecosystems compared with temperate marine ecosystems (Rabalais 2002, Figure 1-2).

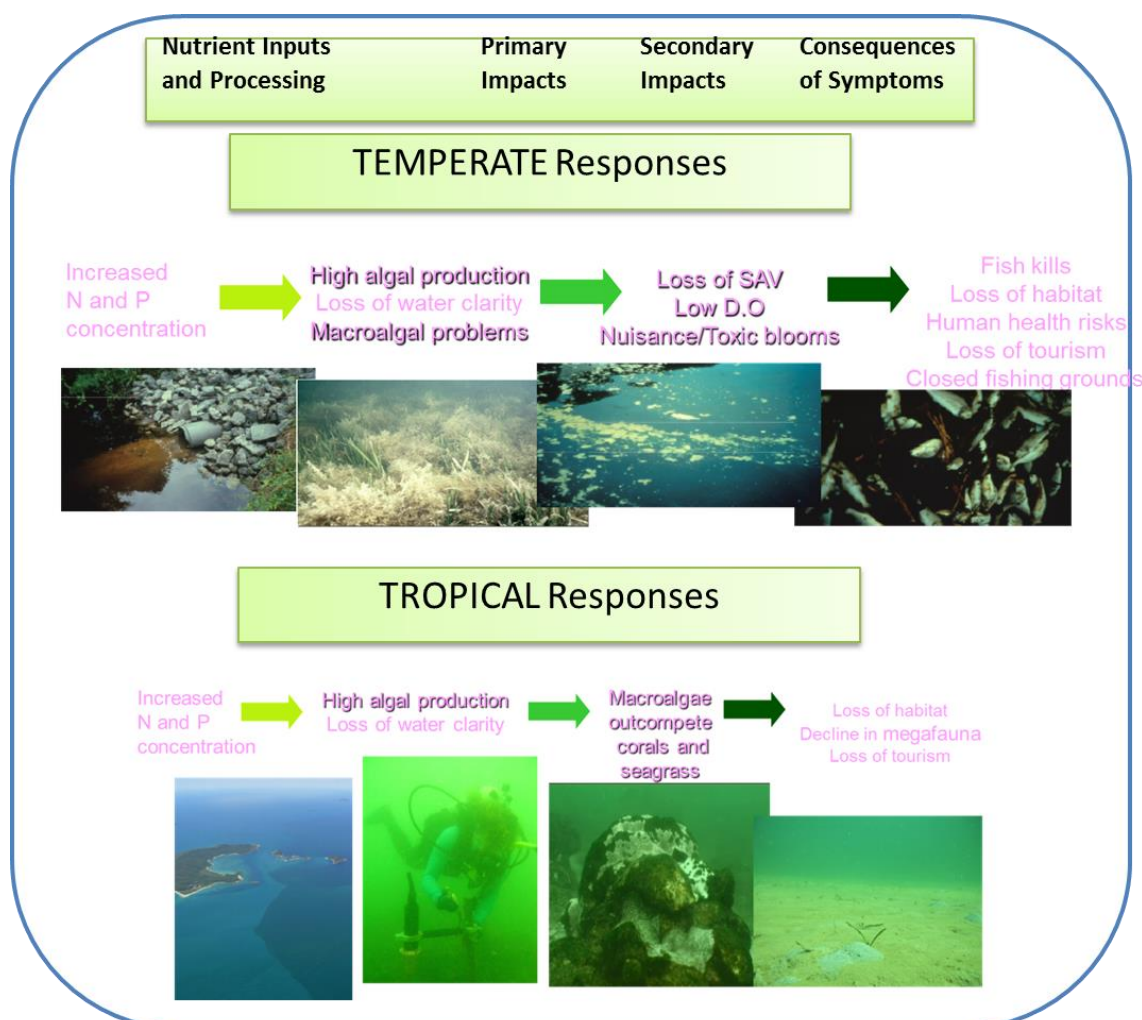


Figure 1-2: Conceptual process of eutrophication responses contrasted in temperate and tropical waters.

This has to be considered when looking for ‘traditional’ (i.e. temperate) indicators of eutrophication e.g. dissolved oxygen sags, in a tropical coral reef system versus the known indicators of coral reef response.

The last scientific consensus statement (Brodie et al., 2008a, 2008b) and recent reviews on eutrophication in coral reefs (Fabricius, 2005), and within the Great Barrier Reef (Brodie *et al.*, 2011; 2012a) do encapsulate that parameters that can indicate eutrophication in tropical systems can be both generic to other eutrophication assessment programs (i.e. increase in chlorophyll biomass) and unique to the Great Barrier Reef (regionally variable responses to nutrient enrichment across large latitudinal areas). This review outlined criteria which have been successfully applied in other approaches and could be applied in a GBR assessment of water quality status.

The national and international approaches reviewed in this report (Appendix A1) are both comparable in selection of indicators; however they vary across temporal and spatial resolution and in the integration of data into an overall assessment of water quality.

Programs differ in scope, definitions, spatial extent, and how the different factors combine for the final assessment. Slight differences in approaches, reference conditions and statistical tools can have significant impacts on the outcomes of any eutrophication assessment.

The commonality between all programs is related to selection of indicators and metrics to assess eutrophication.

- Measure of dissolved oxygen, chlorophyll and nutrient concentrations in most programs
- Waterbodies/assessment areas – typically one catchment –river – marine area
- In situ sampling programs still supplying most of the monitoring data
- High spatial and temporal data collection is required (monthly sampling over many sites per waterbody)
- Restructuring of monitoring programs for improved confidence in assessment tools

Major differences between programs relate to the integration of metric outcomes, the thresholds for “good” and “bad” variable between programs and a variable reporting scale. This review identified the most appropriate indicators are similar across all program and this is also reflective in the MMP indicators. The review provides recommendations on eutrophication criteria and assessment process for the Great Barrier Reef are made. Difficulties in applying other assessment processes relate to the variable cross shelf and latitudinal differences and the sheer size of the assessment area.

Improved assessment and integration method for the MMP water quality data should be based on principles similar to these existing programs to ensure continued stakeholder uptake. An improved assessment and integration method for the MMP water quality data should include a spatial model to integrate the available data in defined reporting zones or regions before a compliance assessment in comparison with the GBR Water Quality Guidelines. The review also recommended the analysis of how the event (flood plume) data, ambient water quality and logger data can be included in a statistical model.

2. Methods

2.1. Study area

The study area is located in a discrete, regional cross-shelf transect in front of the Tully and Murray river mouths within the Wet Tropics region (16° - 18°S). In this region water quality data is collected within the MMP using a combination of three complementary approaches: traditional direct water sampling from research vessels, *in situ* data loggers selected reef locations and remote sensing techniques.

2.1.1. Ambient water quality sites

In the wet tropic region, at Dunk Island is located one the 14 fixed water quality sampling locations, collocated with a 'core' sites of the inshore coral reef monitoring component of the MMP. The 14 sites were selected within each region along a gradient of exposure to runoff, largely determined as increasing distance from a river mouth in a northerly direction to reflect the predominantly northward flow of surface water forced by the prevailing south-easterly winds (Larcombe et al. 1995, Brinkman et al. 2011). At these sites, detailed manual and instrumental water sampling was undertaken (see below) as well as annual surveys of reef status (see Thompson et al. 2011).

2.1.2. Flood plume sites

The study area is located in a discrete, regional cross-shelf transect within the Wet Tropics (16° - 18°S). Flood plume waters are discharged into inshore Great Barrier Reef waters during high river flow events resulting from cyclonic and monsoonal rains (Devlin et al. 2001). Therefore, within each transect, multiple sites have been sampled during wet season conditions (ca. December to April), from 2006 to 2013 and the location of the flooding rivers. The intensity of sampling varied between regions in relation to the logistics of sampling and the frequency of high flow periods. The design of the flood monitoring program under the current Marine Monitoring Program is detailed in Devlin et al. (2011). Sampling is over depth and surface and encompasses many additional parameters than reported in this paper. This paper focused on the measurement of Chlorophyll-a through surface in-situ sampling of chlorophyll biomass. Sites measured during the wet season under the flood plume monitoring program are shown in Figure 2-1 with edge of plume highlighted over the study area for the 2010-2011 wet season.

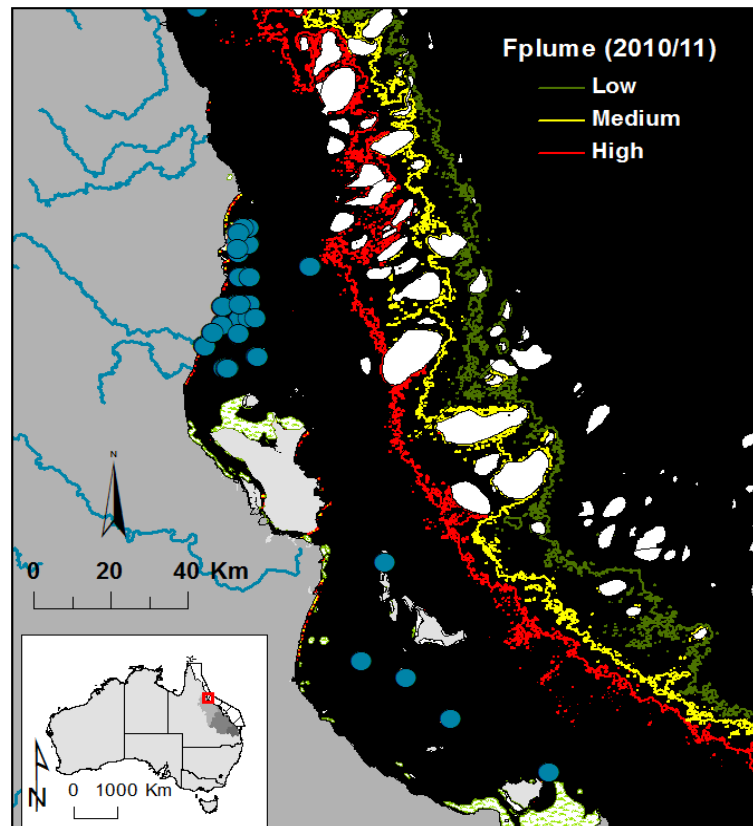


Figure 2-1: Location of flood plume sites sampled during the 2010-2011 wet season for the Tully region only. The edge of the plume is illustrated as high frequency of plumes recorded for at least 13 out of 22 weeks in wet season (red line), medium frequency recorded at 9 out of 22 weeks (yellow line) and low frequency recorded for at least 4 weeks (green line).

2.2. Data attributes and availability

Three complementary approaches are used to collect data at two spatial (point and areal) and three temporal (snapshot, daily, 10-minutely) scales: traditional direct water sampling from research vessels, *in situ* data loggers located at a small number of selected inshore reef locations, and remote sensing instrumentation.

In the MMP, a variety of indicators are measured across the three approaches, with only one indicator –Chlorophyll *a* (a surrogate nutrient measure)- common to all approaches (see www.gbrmpa.gov.au for comprehensive details of the available data for each method). Total Suspended Solids (TSS) was recorded through all methods excluding the data logger. The data logger recorded turbidity, which was converted into TSS units, using an equation based on *in situ* parallel sampling of the two measures (see e.g., Schaffelke et al. 2012). Colour-dissolved Organic Matter (CDOM; a measure of freshwater extent) was measured via the satellite and the event-based *in situ* sampling but not through the data logger or regular *in situ* sampling. A range of other physical and chemical water quality indicators was measured through regular and event-based *in situ* sampling only (Table 2-1). Full details about the available data, sampling and analysis methods are provided in the MMP reports (Schaffelke et al., 2013; Devlin et al., 2013; Brando et al., 2013; www.gbrmpa.gov.au)

The differing spatial and temporal scales of the MMP are summarised more explicitly below:

Spatial (encompassing large-scale and small-scale spatial concepts):

- i. Three regional marine boundaries (coastal zones): near/inshore, midshelf, offshore (roughly W-E differentiation)
- ii. Six regional NRM boundaries (coastal regions): Cape York, Wet Tropics, Burdekin, Mackay-Whitsunday, Fitzroy, Burnett Mary (roughly forming a N-S differentiation)
- iii. Two spatial units (also known as the spatial footprint or spatial support for a single measurement): **areal** (1 km² pixel resolution from remote sensing) and **point** (direct water sampling and data logger)
- iv. Two depth representations: 5 m depth (data logger) and surface (direct sampling and remote sensing)

Temporal (encompassing frequency of measurement and length of data record available):

- i. Direct ambient water sampling – about 3 times/year (typically Feb, June, Sep) since Aug 2005 (referred to as Ambient-Niskin hereafter)
- ii. Direct event-based water sampling –collected between 1994 – 2006 by multiple organisations (AIMS, JCU, GBRMPA). Intensive sampling under the MMP has been undertaken since 2007 by JCU. Multiple sites within the study area have been sampled during wet season conditions (ca. December to April), from 1994 to 2013.) (referred to as Event-Niskin hereafter)
- iii. Data logger – 10-minutely since Oct 2007
- iv. Satellite – daily images since Nov 2002)

Figure 2-2 illustrates the sparseness of the direct water sampling and data logger locations in relation to the remote sensing coverage (1 km pixels for a 91 km² sub-region of the Wet Tropics NRM region). There are very few point source locations (7 unique sites for regular and 17 unique sites for event sampling) although this particular sub-region is fairly “data-rich” for point source locations compared to other sub-regions. The event-based direct water sampling locations are generally clustered in the inshore zone around river mouths, in order to capture river outputs, and are more numerous than the ambient direct water sampling locations. We also note that reefs and islands are masked from remote sensing (equates to an absence of grey dots in Figure 2) so direct water sampling in those locations is important to provide more comprehensive spatial coverage.

In Figure 2-3, TSS is plotted (on 4th root scale due to data skewness) against time for the 2010-11 water reporting year to illustrate the temporal intensity of the data recorded in the same 91 km² sub-region of the Wet Tropics NRM region. The TSS satellite data in this sub-region range in concentration from ~0.1 mg/L (recorded offshore) to >50 mg/L (recorded close to shore). There are large gaps in the TSS remotely-sensed data due to cloud cover that are evident in the wet season (November-March). In this sub-region, the logger is located at Dunk Island and the range of the values derived from the logger is approximately one-third of the range of the remotely-sensed TSS data range, reflecting the location of the logger with respect to distance in the cross shelf gradient. The periodical nature of the data logger

measurement response is likely due to tidally-driven resuspension close to the sea bed. This process is not captured in the remotely-sensed data as the logger provides a measure of TSS integrated in the water column. The direct ambient water sampling data span the lower range of the TSS satellite data, again representing the cross-shelf gradient of water quality and the fact that these data are generally not capturing flood events but sample during three seasons. In contrast, the direct event-based water sampling data that were collected in the wet season cover the higher range of the satellite data.

Figure 2-3 also highlights the temporal sparseness of the direct water sampling in relation to the remote sensing coverage TSS: no direct water sampling data was collected in April, May, July, August, and there were 10 days overlap between direct water sampling dates and remote sensing dates. For 2010-11 water reporting year, direct water sampling data makes up only 0.07% of all available observations (across space and time).

Table 2-1: Summary of available data for making a GBR water quality condition assessment

Custodian	Data Acquisition Description	Temporal extent	Spatial footprint of each observation	Spatial extent	Parameters	Comments
CSIRO	MODIS Aqua Satellite	Daily images ; Nov 2002 – current	1 km ² pixel	All zones and regions	CHL TSS CDOM	Pixels which comprise reef or other obstructions (in water or sky) are masked and are therefore recorded as “missing” values
AIMS	Regular <i>in situ</i> sampling					
	1. Data loggers	10-minutely records; Oct 2007 to Oct 2011 (with some gaps due to instrument failures)	Local	14 fixed sites on reef slope at 5m depth (LAT), in “open coastal” or “midshelf” water bodies, 4 NRM regions; WT, BUR, MW, FTZ.	- Chlorophyll fluorescence ($\mu\text{g L}^{-1}$) - Turbidity (NTU)	Most useful are likely to be daily means per site. Data in Oracle database
	2. Direct water sampling	3 times/year (generally Feb, June, Sep) since Aug 2005. To date 17 sampling occasions	Local	6 fixed open water sites in Cairns region 14 fixed reef sites, here 2 locations sampled: at data logger and depth profile just off the reef (few 100 m away)	- Chlorophyll a ($\mu\text{g L}^{-1}$) - Total suspended solids (TSS, mg L^{-1}) - Secchi depth (m) - Dissolved nutrients(all in μM): - ammonium, - nitrate, - nitrite, - phosphate, - total dissolved N, - total dissolved P - Dissolved organic carbon (μM) - Particulate nitrogen (μM) - Particulate phosphorus (μM) - Particulate carbon (μM) - Salinity	Most useful are likely to be depth-weighted means per site and sampling occasion Data in Oracle database

Custodian	Data Acquisition Description	Temporal extent	Spatial footprint of each observation	Spatial extent	Parameters	Comments
JCU	Grab sampling during flood events	Only during flood/extreme rainfall events ; since 2007	Local	Inshore and mid-shore zones; WT, BUR, FTZ, MW regions?	- Chlorophyll a ($\mu\text{g L}^{-1}$) - Total suspended solids (TSS, mg L^{-1}) - Dissolved nutrients(all in μM): - ammonium, - nitrate, - nitrite, - phosphate, - total dissolved N, - total dissolved P - Dissolved organic carbon (μM) - Particulate nitrogen (μM) - Particulate phosphorus (μM) - Dissolved Oxygen (mg/L) - Salinity - Coloured Dissolved Organic Matter (CDOM) - Phytoplankton (cells/L) - Pesticide (PSII) concentrations	Salinity, Light, Temperature, DO, are depth profile. All the others are mainly sampled at the water surface. Data in Access database. Data available from www.jcu.edu.au. Plume maps sourced from www.e-Atlas.gov.au
	RS Imagery from NASA	Daily MODIS Aqua (level-0) images from Jun 2002 to current	From 1 km^2/pixel up to 250 m^2/pixel	All zones and regions	Chl-a, CDOM, TSS and true colour image	Level 0 and Level 2 products.
Other	Extraneous					
	River discharges from Department of Environment and Resource Management (DERM –Queensland Government)	Daily river discharge for 39 specified rivers		All zones and regions	Maximum, minimum and mean daily discharge in Mega-liters	Data is collected at a regular basis after three months of its generation for allowing data validation by DERM. Data in Access database

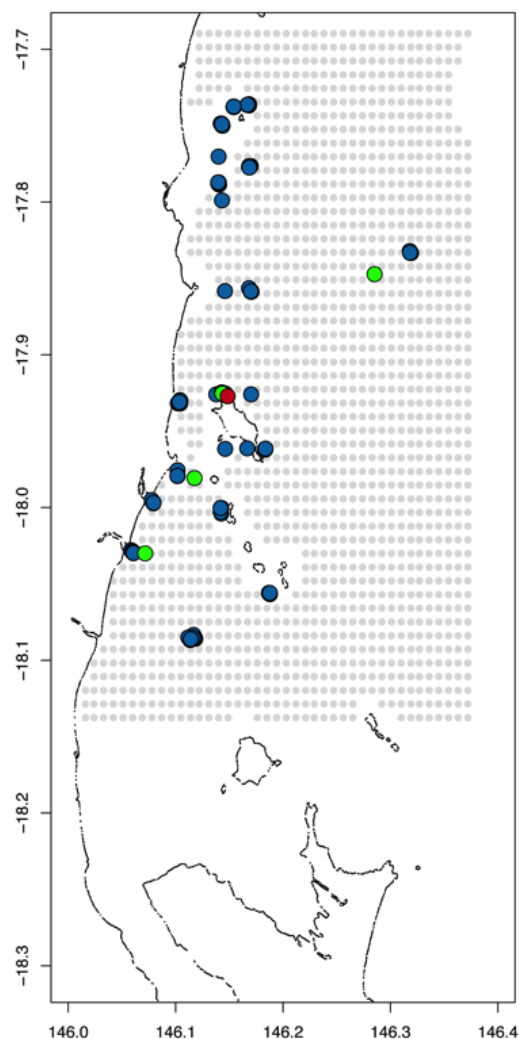


Figure 2-2: Spatial representation of data availability for the four data sources in the Wet Tropics region in front of the Tully and Murray river mouths. The Logger is located at Dunk Island.

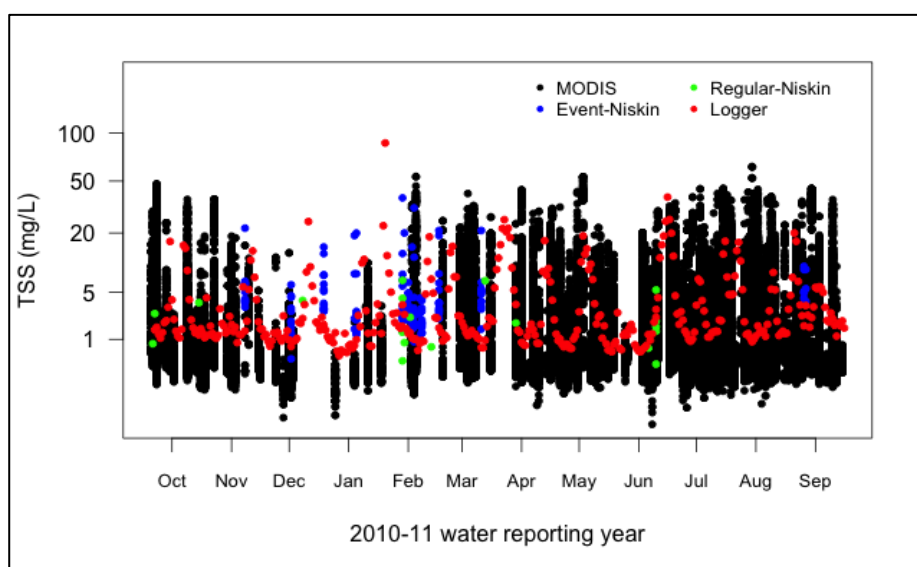


Figure 2-3: Temporal representation of data availability for the four data sources in the Wet Tropics region in front of the Tully and Murray river mouths. The Logger is located at Dunk Island.

2.3. A proposed spatial scale for analysis and reporting

Our objective is to develop methodologies for calculating an overall WQ assessment with testing at the regional level. However, it became evident that NRM regions are not meaningful reporting units because the direct influence of land run-off is at a smaller scale. Consequently we needed to define smaller regions that are:

- directly influenced by land run-off
- consistent with local oceanography (e.g residence times, bathymetry, hydrodynamics) and,
- where it is reasonable to expect **measurable change** over the next decade in response to improvement in land management practices.

Multiple sources of spatial information can be used to delineate a region in order to objectively determine a sub-region that meets the specified requirements. Examples of relevant sources of information include coloured dissolved organic matter (CDOM) measurements as a proxy for salinity, the extent and frequency of flood plumes, hydrodynamic modeling and bathymetry. Further details about these data sources are described in Table 2-2.

Table 2-2: Description of spatial data sources that can be used to describe a sub-region reporting area.

Spatial information	Source	Description	Resolution
Coloured dissolved organic matter (CDOM) (as a proxy for salinity)	Remote sensing – L2 products. MODIS imagery, CSIRO algorithm <i>(Schroeder et al., 2012)</i>	CDOM concentrations at 1km ² pixel. CDOM value converted to salinity. Plume edge defined by salinity thresholds (30 +/- 4). Annual measurements of exceedance (1) or non exceedance (0)	Spatial – 1km ² resolution Temporal -2001 – 2012. Annual maps converted to shape files (ArcView)
Extent of flood plume (river influenced) waters	Remote sensing – ocean colour, MODIS imagery <i>(Devlin et al., 2012; Alvarez-Romero et al., 2013)</i>	Ocean colour converted to 6 classes representing flood plume conditions. Full extent measured per year through wet season (Nov to April).	Spatial – 1km ² resolution Temporal - 2005 – 2012. Annual maps converted to shape files (ArcView)
GBR hydrodynamic modelling	AIMS/CSIRO (Fitch et al; 2012)	Cumulative exposure - 1% of the sources concentration.	
GBR bathymetry (depth)			

In order to demonstrate proof-of-concept, a study area was selected according to the desired criteria. This sub-region is located in a discrete, regional cross-shelf transect within the Wet Tropics NRM region (16° - 18° S). Our proof-of-concept focused on the analysis of surface TSS measurements. The Tully-Murray subregion is illustrated for the selection criteria using the spatial distribution of CDOM (Figure 2-4, flood plume extent (Figure 2-5), an exposure index from a GBR hydrodynamic model (Figure 2-6) and bathymetric data (Figure 2-7). The area of reporting was defined by the extent of the river plume as defined by both plume imagery and the hydrodynamic model. Bathymetric data (20m depth) and the extent of the hydrodynamic model have been used to describe the latitudinal extent of the Tully reporting area with the cross shelf extent describe by the extent of the flood plume waters as described by CDOM and ocean colour.

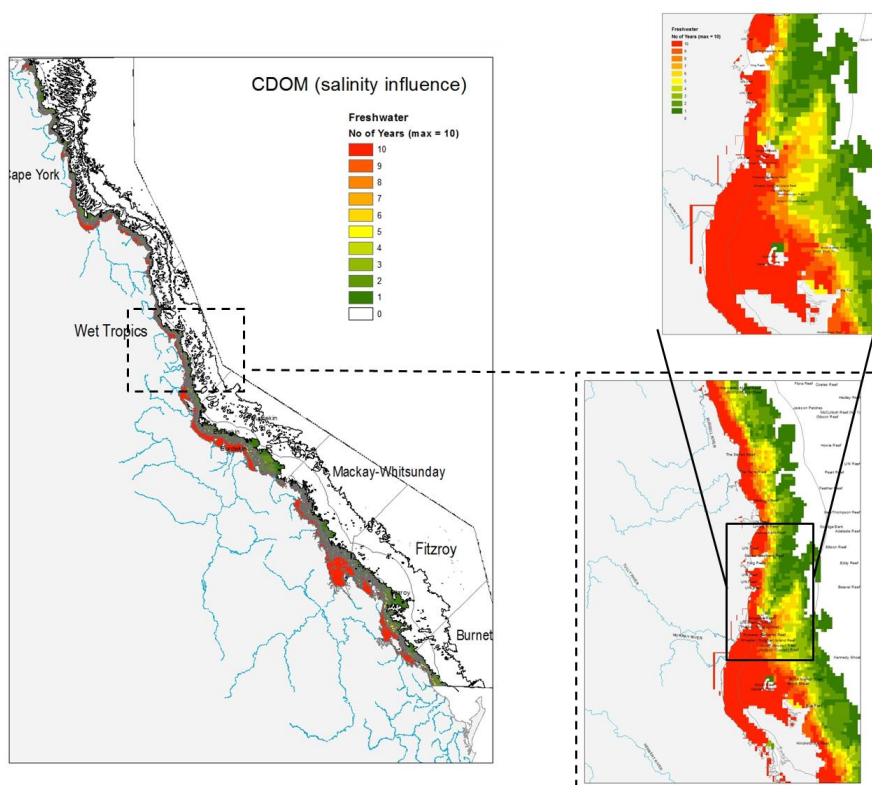


Figure 2-4: Spatial extent of river influenced waters as defined by Coloured dissolved organic matter (CDOM) as a proxy for reduced salinity (Schroeder et al., 2012). The box represents the reporting area as defined by the cross shelf extent of the river influenced waters (red to green).

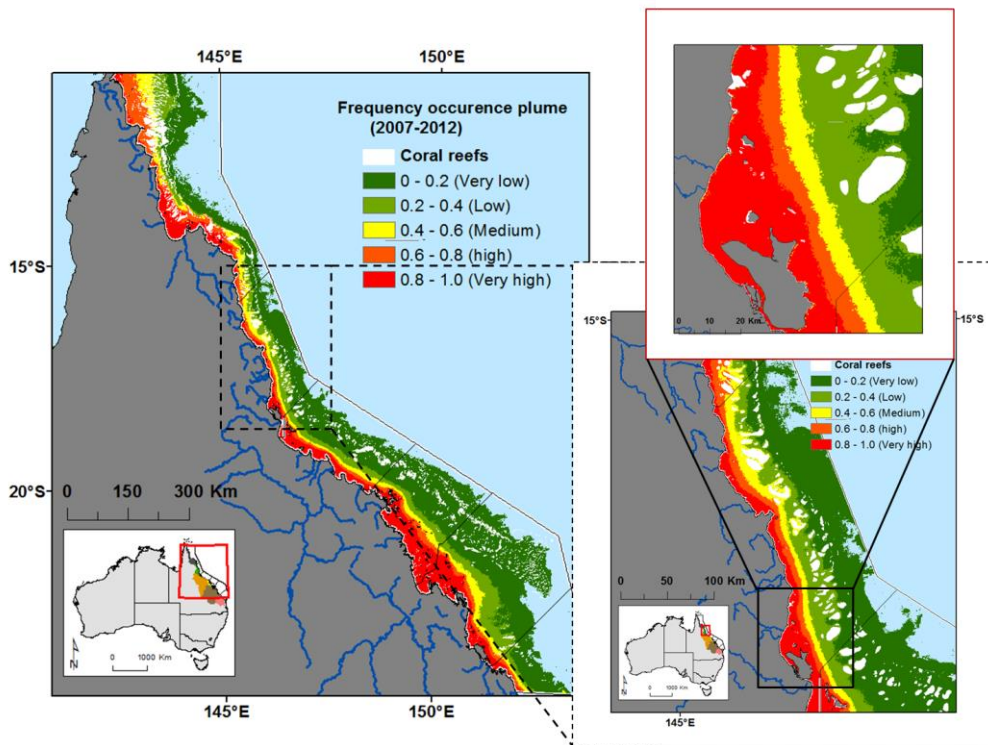


Figure 2-5: Spatial extent of river influenced waters as defined by ocean colour as a measure of distribution and frequency of plume influenced water (Devlin et al., 2012; Alvarez-Romero et al., 2013). The boxes represents the reporting area as defined by the cross shelf extent of the river influenced waters (red to green).

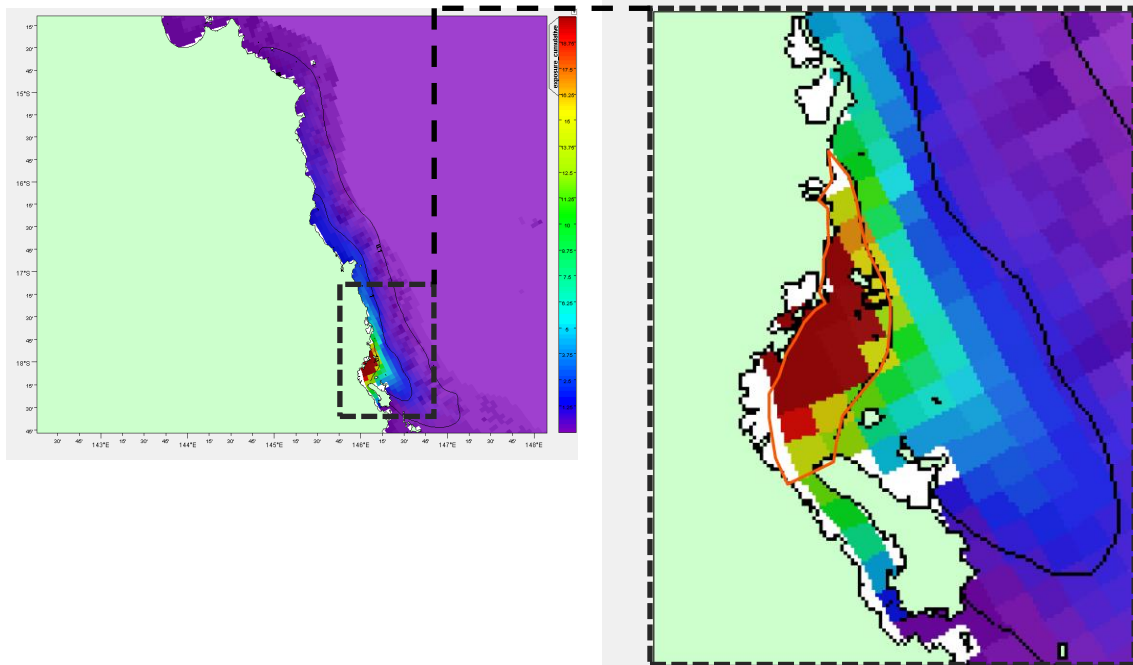


Figure 2-6: Spatial extent of river influenced waters as defined by a cumulative exposure index based on a GBR hydrodynamic model (from Furnas et al., 2013). The colour bar indicates the calculated cumulative exposure ("Conc.Days", concentration x days) above 1% of the incoming concentration (see Furnas et al. 2011 for details). Contours show 0.1, 1.0 and 10.0 Conc.Days exposure levels and the reporting boundary is defined by the 10 Conc.Days contour line.

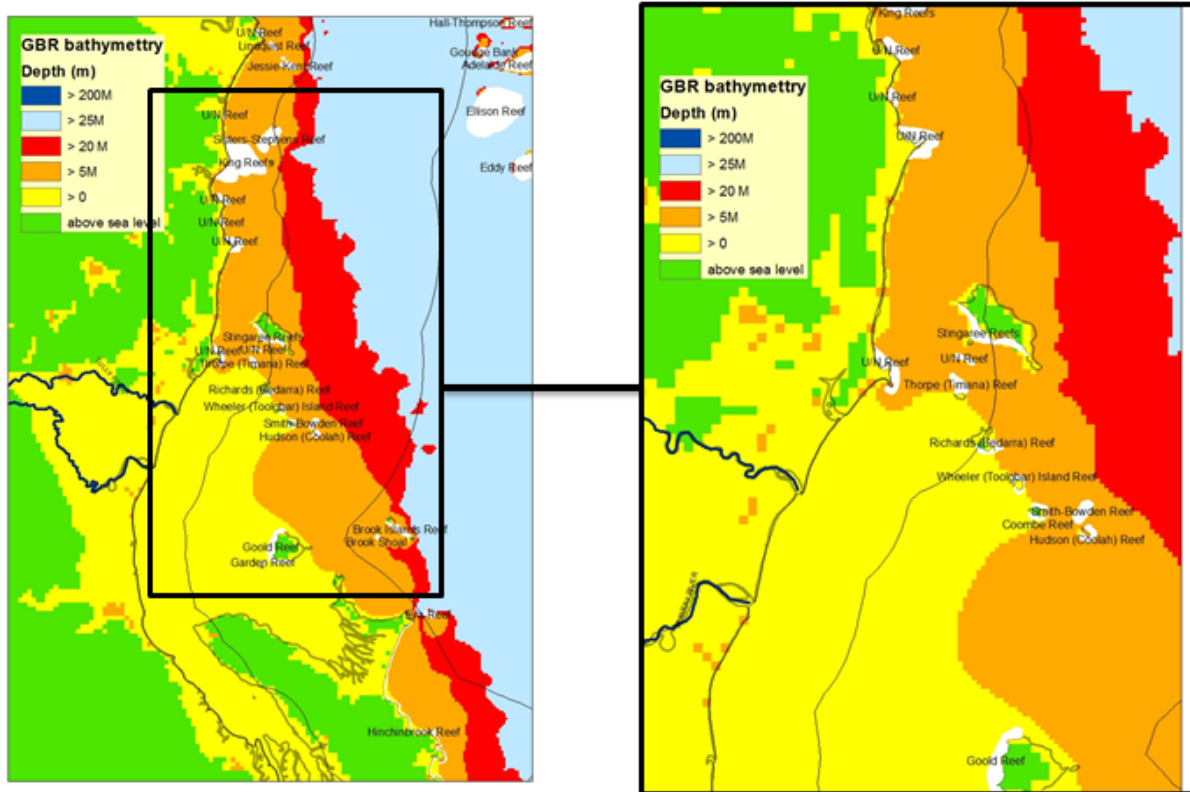


Figure 2-7: Spatial extent of river influenced waters as defined by GBR bathymetry

2.4. A proposed temporal scale for analysis and reporting

The temporal scale suggested for the reporting is the whole of year including:

- The assessment of the condition of water relative to the seasonal guidelines (4 seasons per year).
- Improved capacity to interpret data with respect to land-runoff events.

Sites along a water quality gradient can be representative of a larger reporting area, however variation within sites needs to be considered in the overall reporting process. Sites within the Tully reporting areas (Figure 2-8) provides an example of how seasonal aggregations can differ (Figure 2-9) with a time series of TSS satellite data for the three data points and a comparison of seasonal averages against the water quality guidelines. Using information from all sites, over the whole year, allows an assessment of the intensity of wet season and if the reduced water quality persisted or improved in the dry season.

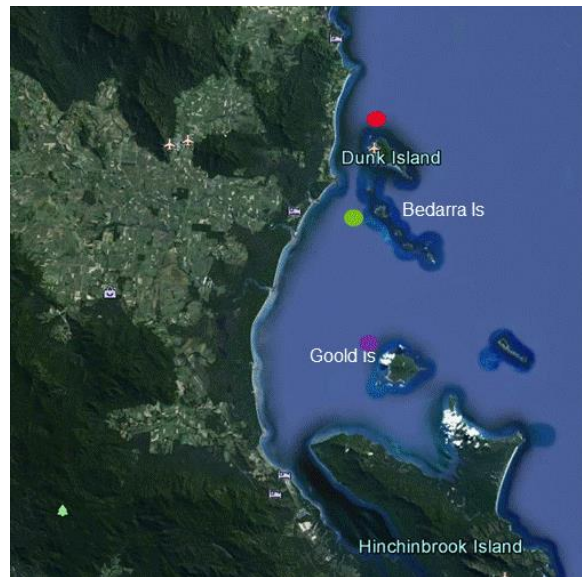


Figure 2-8: The location of three locations within the Tully reporting area

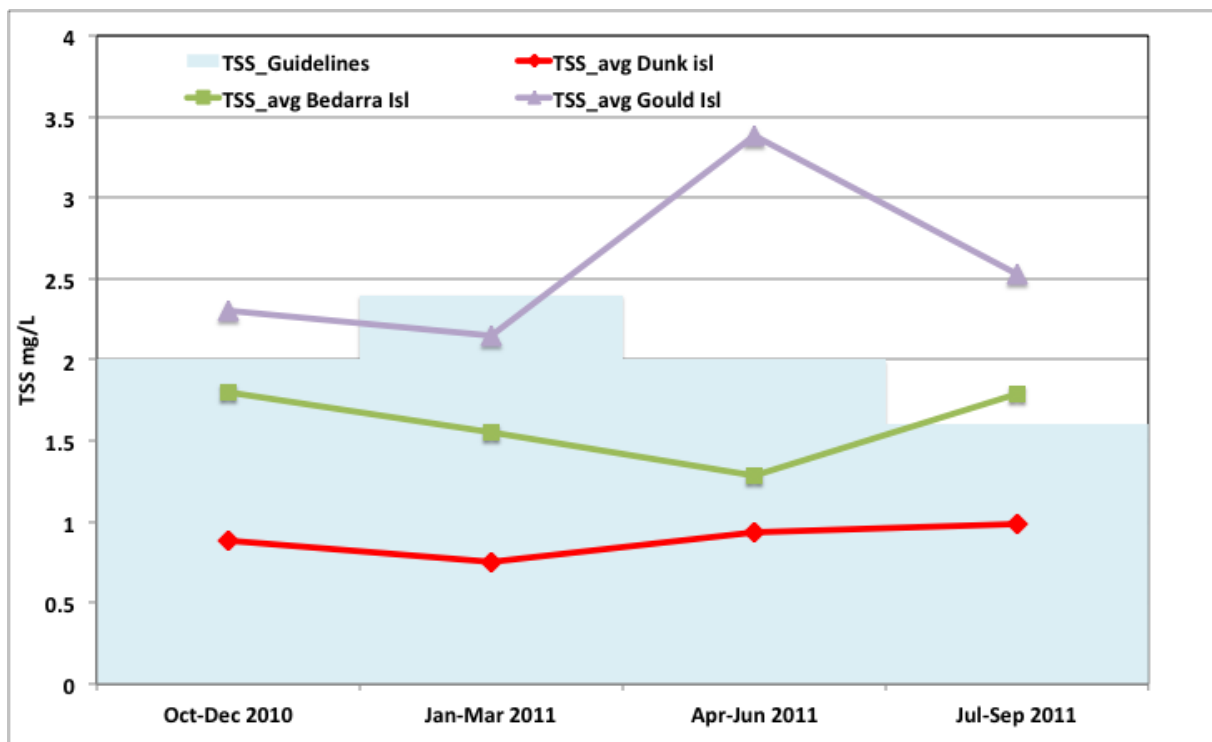


Figure 2-9: Seasonal averages for satellite-derived TSS for three locations in the Wet Tropics NRM region that are located in the vicinity of the Tully and Murray River mouths in the 2010-11 water reporting year.

2.5. Proposed data integration approach

2.5.1. Overview

There are typically many steps involved in developing a report card for the quality of an aquatic resource, with the number of steps positively related to the complexity of the data contributing to the report card. For the GBR report card for water quality, there are multiple spatial and temporal scales as well as multiple variables that need to be combined to give the overall report card score or qualitative grade. The order in which the data are combined is flexible but typically follow the general steps of temporal and/or spatial aggregation followed by indicator integration. Figure 2-10 illustrates the proposed general framework for calculating a report card.

The aim of this project is to develop a more comprehensive and informative approach for assessing water quality in the GBR, which

- utilizes all available and relevant data
- is transparent in its development
- is statistically defensible (i.e. uncertainty is quantified and inferences are meaningful)

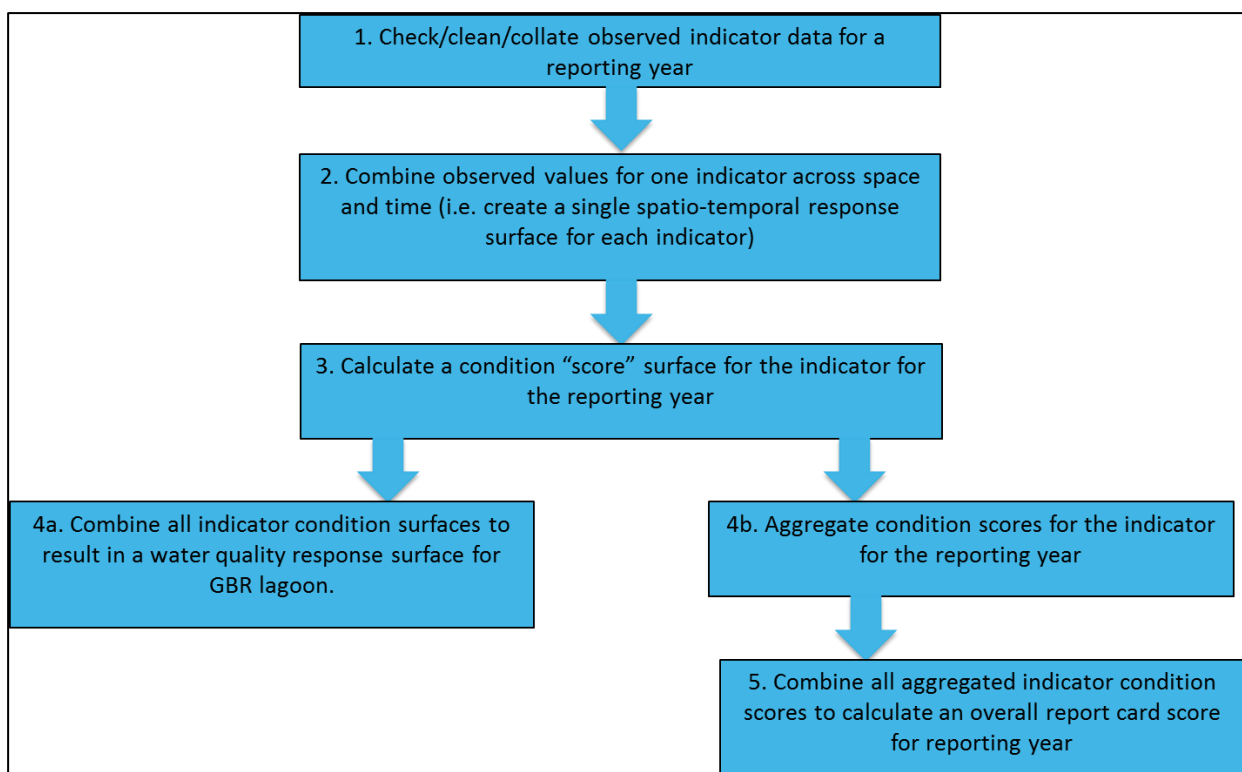


Figure 2-10: Overview of proposed GBR water quality reporting framework

2.5.2. Step 1: collating/checking/cleaning data for each indicator

In this step, the different data sets should be collated, checked, cleaned accordingly, and organised for one water reporting year (October – September) and for each indicator separately.

2.5.3. Step 2: form a modeled response surface for each indicator

This is a very complex yet pivotal step in the report card development, encompassing data merging, modeling the underlying (latent and unobservable) process and quantifying uncertainty of model predictions. Complexity arises because spatio-temporal data are inherently auto-correlated (i.e. the water quality indicator value for a pixel on a particular day is not only typically correlated with values from neighbouring pixels and from preceding times, but also from neighbouring pixels for preceding times.) so any model that describes the underlying process will need to account for this.

2.5.4. Step 2a: merging data sources for an indicator

Remotely-sensed data will make up the majority of the information contributing to a report card for each spatial region given its high spatial and relatively high temporal frequency. The point data that have been recorded through direct water sampling contribute less than 0.5% of all observations (across space and time). Critically however, point data may be available at times when, and places where, there are no remotely-sensed data values due to masking effects or instrumentation failure. So a valid use for the point data is to "fill in" or in a statistical expression, impute, the missing remotely-sensed data. In the small number of cases where the remotely-sensed data are coincident with point data, we suggest integrating the values by averaging them, although other more sophisticated approaches could be used (e.g. decreasing weight/contribution of point data value to the remotely-sensed value with increasing distance of the point data value from pixel centre).

Mathematically, we write this step as

$$y^*(s_y, t) = \begin{cases} x(s_x, t) & \text{if } y(s_y, t) = NA \\ \frac{1}{2}(y(s_y, t) + x(s_x, t)) & \text{if } s_x \subseteq s_y \\ y(s_y, t) & \text{otherwise} \end{cases}$$

Where direct water sampling data, $x(s_x, t)$, is used to either impute missing remote sensing data (in space and time, denoted by $y(s_y, t)$) or average coincident values.

The Ambient-Niskin data are considered the benchmark data as they are collected and analysed by accepted standard procedures and are particularly valuable to incorporate.

However, there are few records of this data source having only been recorded from few sites approximately three times each year (typically February, June and September/October). The Event-Niskin data were typically recorded during the summer months when flood events are most frequent and have greatest magnitude and are concentrated spatially where flood plumes occur in the Great Barrier Reef lagoon.

Each point observation has a spatial and temporal "footprint" (also known as a region of influence or support region) and the footprint differs according to the time of year, spatial location, and purpose of monitoring (ambient vs event). Table 2-3 provides details of the suggested spatial and temporal footprints for the point data.

Table 2-3: Spatial and temporal footprints for point observations. Note that Ambient-Niskin monitoring was only carried out typically three times per year: nominally February, June and September/October, whereas Event-Niskin monitoring was carried out only when "events" occurred. A temporal footprint encompasses the day of the observation so represents condition before and after the monitoring has occurred. Dry season = July, August, September. Wet season = January, February, March. 'In-between'-seasons are the months between wet and dry seasons.

Season	Ambient-Niskin and Event-Niskin	
	Spatial footprint	Temporal footprint
Wet	5 km radius	3-day period
In-between	10 km radius	5-day period
Dry	15 km radius	7-day period

2.5.5. Step 2b: a proposed model framework

This step concerns taking the merged daily response surfaces (with any data gaps) for each indicator and developing an appropriate model that accounts for any spatio-temporal correlation and enables us to make predictions of the response at a nominal spatial scale as well as how confident we are with those predictions. Hierarchical models are one approach used for combining incompatible data sets, for example, data sets generated under different spatial and/or temporal resolutions. Henderson (2003) presents a brief review of the use of hierarchical models for combining incompatible spatial/temporal data, whilst Gelfand et al (2004) give them a more thorough treatment.

We explore the use of hierarchical models for combining incompatible spatial data before considering how they can be used for combining both incompatible spatial and temporal data.

The problem of combining incompatible spatial data is not new and has been encountered in several fields of study. Gotway and Young (2002) present an overview of the statistical issues associated with combining incompatible spatial data for modeling and inference. Gotway and Young (2007) and Young and Gotway (2007) consider and demonstrate a general method that can be used for aggregation, disaggregation, side-scaling and intensity estimation.

Due to its origins and developments made in multiple fields of study, there are numerous terms for describing one or more facets of the incompatible spatial data problem as well as

the various solutions to it (Gotway and Young, 2002). These include the ecological inference problem, modifiable areal unit problem (MAUP), spatial data transformations, areal interpolation, inference with spatially-misaligned data, and multi-scale and multi-resolution modeling, among others.

In spatial statistics, combining incompatible spatial data is known as the change of support problem (COSP). The term 'support' relates not only to the geometrical size or volume of each data value but also the shape and spatial orientation. The support of the data is scale dependent and it also depends on the measurement process. Changing the support of a variable creates a new variable, which is related to the original one but has different statistical and spatial properties. More specifically, the COSP is the problem of how the spatial variation in one variable associated with a given support relates to that of the other variable with a different support. COSP involves two inferential issues: the *scale* or *aggregation* effect and the *zoning* or *grouping* effect. The former issue arises because the distribution of a variable that has been averaged over spatial units is different from that of the original variable. It tends to be more bell-shaped and has smaller variance. The situation is slightly more complicated in spatial statistics because the averaging also affects spatial autocorrelation properties. The latter inferential issue concerns the variability in results and inferences that are due to alternative formations of spatial units.

The GBR water quality data have been recorded on different spatial units that represent different spatial scales. The direct water sampling methods generate data that are recorded at point locations within the GBR lagoon whilst the satellite images generate data on an areal (1 km² pixel) scale. We needed to combine these data sets.

If we use one square-kilometre pixels as the common spatial unit for analyses, we need to upscale the direct water sampling data from the point locations to the (areal) pixel units. Block kriging was developed in data mining to provide predictions of areal averages from point data and Gotway and Young (2007) generalized this method for use in general change of support problems i.e. aggregation, disaggregation, side-scaling and intensity estimation.

Where sufficient data are available, i.e. with a large fraction of potential gridded observations actually made in space and time, the most appropriate models for spatio-temporal inference are often Bayesian hierarchical linear models. This large and flexible class of models allows coherent inferences to be made concerning the most likely values of a given indicator, with dependencies in space and time accounted for through appropriate covariance structures. The five basic components of these models include: a deterministic mean function (i.e. a linear model) that includes a intercept term and any potential covariates of interest; a mean-zero random effect for location-specific variability that is common at all times; a mean-zero random effect for time-specific variability that is common at all locations; a mean-zero random effect for the interaction of space and time; and a mean-zero random effect for observation-scale spatio-temporal variability that is modelled as 'white noise' (Cressie and Wikle, 2011).

Because these kinds of Bayesian hierarchical models can be specified completely, in the sense that every potential observation or pixel has a part in the model structure, they are particularly effective for analyzing spatio-temporal datasets that contain missing data. However, while missing data can be accommodated readily, the precision of the inferences made declines rapidly with increasing extent of data gaps. Unfortunately in the case of GBR

water quality data, the proportion of missing pixel values over the entire rainy season can be near 90%, making reliable inference difficult.

Further complicating the analysis of these water quality metrics is the high dimensionality of the problem. For each reporting region in each quarter, there are approximately 10^7 potential observations, leading to thousands of potential model parameters. This high dimensionality makes implementing these models impossible on personal computers and requires instead specialized distributed or super-computing approaches that were unavailable for this project. However, a specific type of hierarchical model, a Gaussian Predictive Process (GP) model (Banerjee et al. 2004; Banerjee et al. 2008), provides a framework for meeting the modelling aims and that is capable of handling the exceptionally large data set involved. GP models provide substantial dimension reduction compared to full Bayesian hierarchical model approaches but provide similar inferences in terms of spatial and temporal interdependence. Gaussian processes are essentially distributions of functions – in many cases a distribution of normal distributions – that allow efficient estimation when the underlying functional relationship is unknown. While increasingly used by statisticians, the implementation of customized GP models, particularly in the spatio-temporal realm, remains highly technical.

Ideally, the GP model developed to model GBR water quality annually should be fully spatio-temporal in nature, including most critically a spatio-temporal covariance function that handles spatial and temporal variance and covariance simultaneously (Stein 2005). However implementation of such a model requires specialized statistical methods at the cutting edge of spatial statistics. Although we were not able to implement the spatio-temporal covariance function in time, we developed a GP approximation that allowed us to illustrate the GP approach and retain the most important features of the full spatio-temporal model, namely the propagation of uncertainty through all stages of the modeling process.

The first stage in our approximation was to develop posterior mean and standard deviation surfaces for each day in the Tully region for which we had observations. The GP model for each surface variable (d) in each 1 km^2 pixel (i) is given by:

$$\begin{aligned} d_i &\sim N(f(x_i, y_i), V) \\ f &\sim GP(M, C) \\ M &= GP_{mean}(\mu) \\ p(\mu) &\propto 1 \\ C &= ME(x_i, y_i, \alpha, \theta, \kappa) \\ \alpha &\sim \text{Exp}(7e^{-5}) \\ \theta &\sim \text{Exp}(4e^{-3}) \\ \kappa &\sim U(0.01, 10) \\ V &\sim \text{Exp}(5e^{-9}) \end{aligned}$$

where x and y are the latitude and longitude coordinates for pixel i ; GP defines a Gaussian process; GP_{mean} is the Gaussian process mean proportional to overall mean μ ; and ME is the Matérn euclidian covariance function for spatial relatedness (Banerjee et al. 2004). The

hyper priors include the covariance amplitude α , scale θ , and difference degree κ , that control the smoothness of the random field, spatial scale, and the third-order Bessel function that is part of the Matérn function. Priors were chosen to be weakly informative as we had little prior information as to the exact spatial relatedness of the pixel values. For each daily GP surface we obtained 500 posterior realizations from the posterior surface, producing a representative sample of values at 5 km^2 pixels that approximate the posterior values at this scale. From these, we obtained means and standard deviations for each 5 km^2 cell on each day that at least 10 cells observed, completing our spatial approximation.

Given the daily means and standard deviations from the spatial GP model, we assumed them known, and modelled these values through using a simple Bayesian hierarchical model that assumed individual pixels where correlated through time:

$$\begin{aligned}\hat{d}_{i,t} &\sim N(\mu_{i,t}, \hat{\tau}_{i,t}) \\ \mu_{i,t} &\sim N(\beta_{0,i}, \tau_i) \\ \beta_{0,i} &\sim N(\gamma_0, \tau) \\ \gamma_0 &\sim N(0.0, 0.001) \\ \tau_{i,t}, \tau_i, \tau &\sim U(0, 100)^{-2}\end{aligned}$$

where, $\hat{d}_{i,t}$ and $\hat{\tau}_{i,t}$ were the assumed known daily GP means and precisions. Because subsequent days were rarely reported, we accommodated the temporal correlation for each pixel using a hierarchical structure, given a mean ($\beta_{0,i}$) and precision (τ_i) for each pixel in a given quarter. We also estimated a hierarchical overall mean (γ_0) value, however the $\beta_{0,i}$'s were used for our quarterly maps, showing the estimated average water quality values and their associated uncertainty. Both the spatial and temporal Bayesian model components were implemented in the PyMC package for the Python programming language.

2.5.6. Step 3: Calculate a condition “score” surface for each indicator

We used the pre-determined guidelines to calculate a condition “score” surface for each indicator separately. There are two parts to this step: calculate the condition (relative to guideline) and convert the relative condition into a score. To calculate the condition, it is typical to just calculate compliance/non-compliance with the guideline. However, it could be more informative if it incorporated how non-compliant an (aggregated) observed value was. The score should be standardized (by one of any number of ways) to make interpretation easier and results more directly comparable across indicators. A score of 1 suggests poor condition whilst a score of 0 indicates very good condition. There are two choices with how to proceed from this step: Step 4a or Step 4b.

2.5.7. Step 4: Assessment of condition relative to regional, spatial and seasonal guidelines

For each pixel in the region of interest, we want to assess the condition of water relative to the seasonal, regional and spatial GBRMPA water quality guidelines. These steps provide guidance on how to calculate the relative condition for our modeled response surface.

1. Firstly calculate an extreme case scenario (ECS) per pixel for each season using all available historical RS* data. In our case, we define the seasonal ECS to be the 95th percentile of all RS values in that season in that region in that pixel. The seasonal ECS will be a step function in time for a particular pixel comprising 4 different ECS values: one per season.
2. Next determine the seasonal guideline values for the region. If only wet and dry season guidelines are available, then we suggest calculating the in between season guidelines as the average of the wet and dry season guidelines. Alternatively, percentiles of the historical data could be used to empirically derive guideline values. Seasonal guideline values are a step function in time for a particular pixel. If spatially-explicit guideline values are available then these could be used to provide further input into establishing the most relevant guideline values for assessing compliance.
3. Then for each predicted mean value, $\hat{y}(s_y, t)$, in the modeled response surface, assess compliance, z , with the seasonal (and spatial?) water quality guideline, g , as well as how non-compliant a predicted mean value is:

$$z(s_y, t) = \begin{cases} 0 & \text{if } \hat{y}(s_y, t) < g_q \\ \max(1, (\hat{y}(s_y, t) - g_q) / (ECS(s_y, t) - g_q)) & \text{if } \hat{y}(s_y, t) \geq g_q \end{cases}$$

This will give a surface of values between 0 (compliant) and 1 (non-compliant) for each day.

4. Average the daily compliance surfaces to create an average seasonal surface

This average seasonal compliance surface is an informative way to communicate regional differences in compliance across space and time for a single indicator. However, we acknowledge that report cards are typically represented by a single number/score/grade, and thus suggest this is most simply obtained by averaging the average seasonal surface values. Alternatively, different summaries of the average seasonal surfaces could be calculated such as the proportion of pixels with average scores of greater than 0.5 or perhaps even the 80th percentile of values. If annual relative condition is of interest, then the four average seasonal maps could be combined by averaging the seasonal value.

The challenge remains as to how the uncertainty in these maps and/or single seasonal scores can be quantified to make our inferences about GBR water quality condition defensible and meaningful.

2.5.8. Step 4a: calculate a condition “score” surface – combined indicators

In this step the goal is to create a condition score surface which comprises contributions from all individual indicator score surfaces. Only Total Suspended Solids and Chlorophyll a will have predicted surfaces for the entire region of reporting interest. The other indicators of relevance to the report card calculation are for sampling points and thus they have a more limited distribution spatially. However there will be a subset of a reporting region for which multiple predicted indicator values will be available for a pixel in that subset area and combining those values will be of interest. Averaging the indicator values is one approach that has been used in other monitoring programs (references?), but we may have reason to want to weight a particular indicator(s) more heavily in the aggregation, thereby requiring unequal weights for contributions of individual score surfaces to the combined surface. This combined indicator surface would be the end goal in reporting the condition of the GBR lagoon waters.

2.5.9. Step 4b: aggregate condition scores – separate indicators

In this step, the response surface created in Step 3 is aggregated to form a single condition score across the entire GBR lagoon for each indicator separately. How is the response surface aggregated?

2.5.10. Step 5: combine aggregated scores across indicators

Following on from Step 4b, the individual indicator aggregated scores are combined to produce an overall aggregated score. Need to discuss equal weights vs unequal weights for contributions of individual aggregated scores to the overall aggregated score.

3. Results

3.1. Proof of concept of water quality reporting framework

For the proof of concept we selected an area with the most data, in both space and time, from direct water sampling. A small reporting region around the Tully mouth/Rockingham Bay was defined, using the criteria from the previous chapter, which encompassed most of these data points. The reporting year was the 2010-11 Water Year (October 2010 to September 2011), which was the year of the MMP sampling which saw the most extreme floods along the whole of the GBR coast. By selecting the most data-rich region and the most extreme year we aimed to maximise the ability of the proposed reporting framework to show seasonal and spatial differences.

The key parameter selected for the reporting framework proof of concept was TSS (measured by remote sensing, direct water sampling both during floods and ambient conditions) and TSS as measured by water quality instruments on a coral reef. Below, we outline the detailed steps how the assessment scores were derived for TSS.

To illustrate the integration between water quality variables, the TSS scores are then combined with the assessment scores for (ii) DIN and PN values as measured by direct water sampling both during floods and ambient conditions at various sites and time points in the in the reporting region during the 2010-11 Water Year. The process of combining the results from the TSS assessment using both remote sensed data and in-situ water quality with outputs from the DIN and PN assessment which use in-situ water quality only. Note that the TSS assessment calculates a TSS value over each 5km grid square in contrast to the DIN and PN assessment, which groups and averages values over a 10km² grid square. In addition, the flow chart identifies that future assessments would also include an assessment of chlorophyll-a using the same statistical model as presented for the TSS.

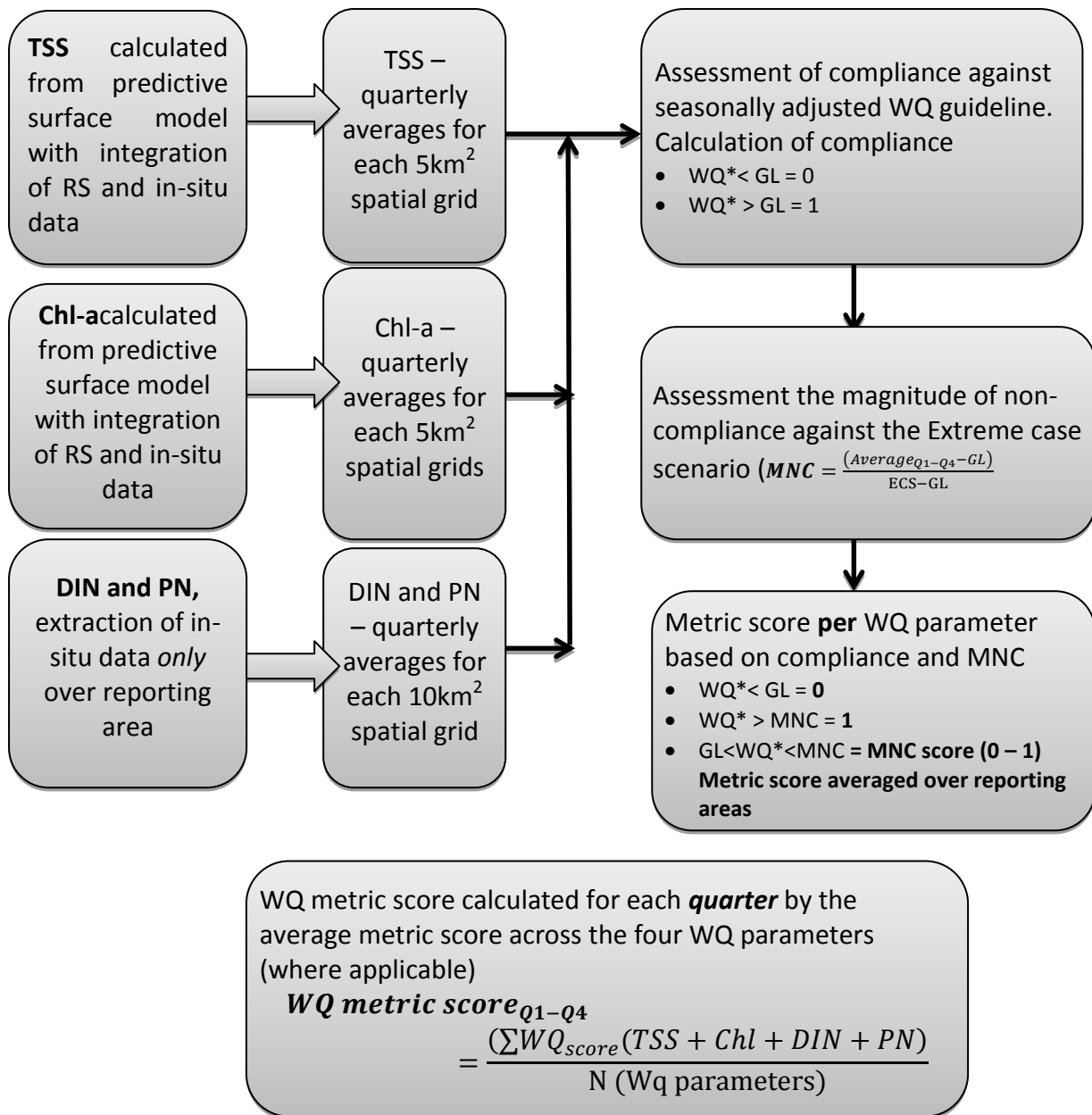


Figure 3-1: Flow chart outlining the process of assessment using four water quality parameters, including TSS, chl-a, DIN and PN. Note that in this report, chl-a is not included.

3.2. Calculation of assessment scores for TSS

3.2.1. Calculate seasonal averages (per pixel) for TSS concentrations using statistical model

As the first step in the assessment, a modeled data surface using remote sensing data, augmented with the available *in situ* data, from 2010-2011 only, was produced to give quarterly averages for each 5 km² spatial pixel (Figure 3-2). Total average TSS values were highest in Q4 (0.88 [0.65, 1.17] mg L⁻¹; Bayesian highest posterior density [95% uncertainty intervals]), followed by Q2 (0.62 [0.26, 1.43]) and Q3 (0.60 [0.44, 0.91]); Q1 had the lowest estimated TSS (0.49 [0.23, 0.81]).

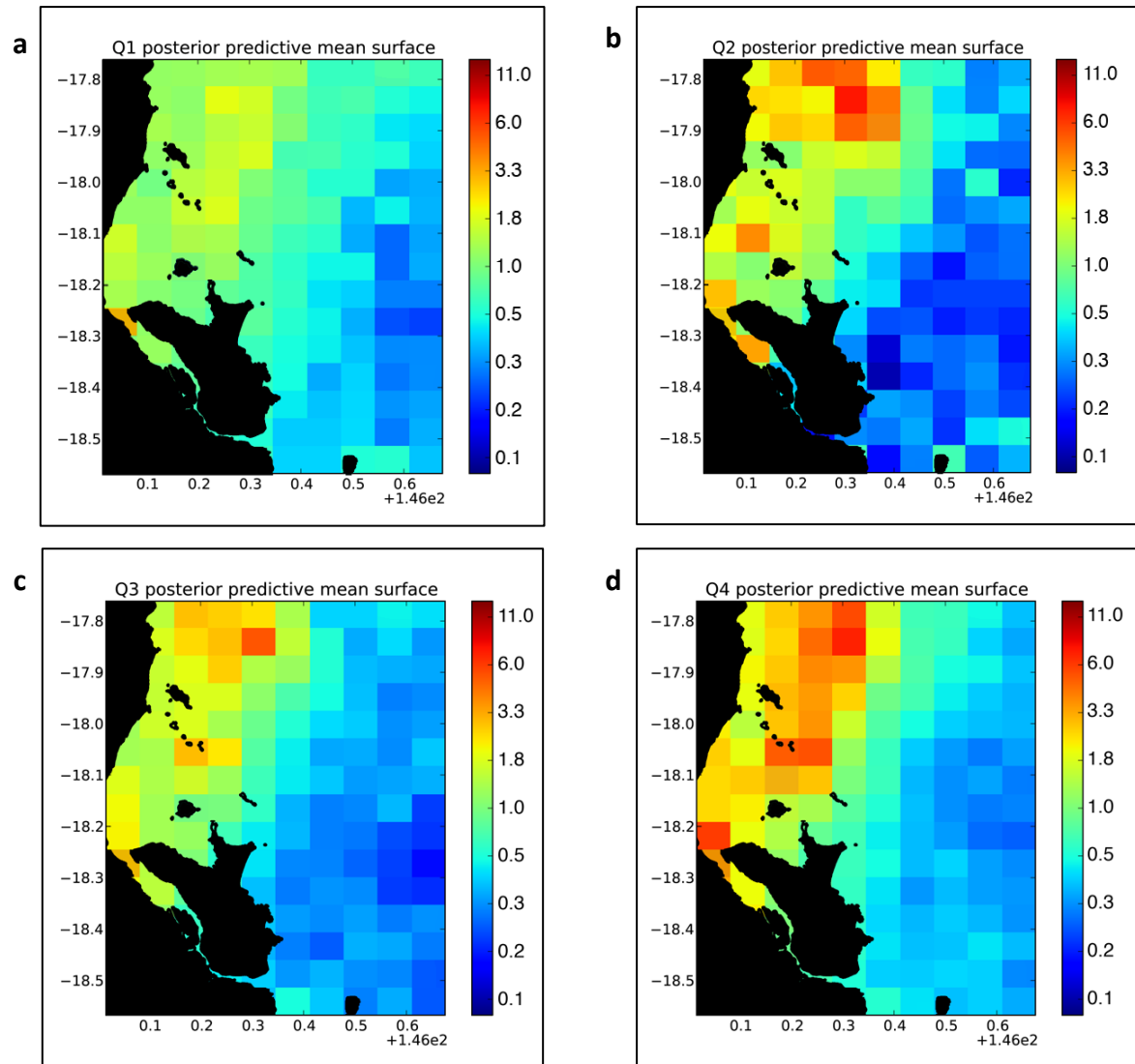


Figure 3-2: Modeled average TSS concentration (mg L^{-1}) for four reporting quarters (Q1 to Q4) during the 2010-11 Water Year. a) Q1= October 2010 to December 2010, b) Q2= January 2011 to March 2011, c) Q3= April 2011 to June 2011, d) Q4= July 2011 to September 2011.

In this step of the assessment, the average predicted TSS value per pixel per quarter is compared to the seasonally adjusted GBR water quality guideline trigger values (GL) for TSS (GBRMPA, 2010). For October to December (Q1) this value is 2.0 mg L^{-1} , for the main months of the wet season, January to March (Q2) the GL is 2.4 mg L^{-1} . For April to June (Q3) the GL is again 2.0 mg L^{-1} , and for the main months of the dry season, July to September the GL is 1.6 mg L^{-1} . Pixels that have a quarterly average below the GL were allocated a compliance score C of 0, while pixels with averages above the GL have a C of 1 (Figure 3-3). **Estimated percent non-compliance of pixels per quarter were 1% in Q1; 10% in Q2 and Q3; and 32% in Q4 (Figure 3-3).**

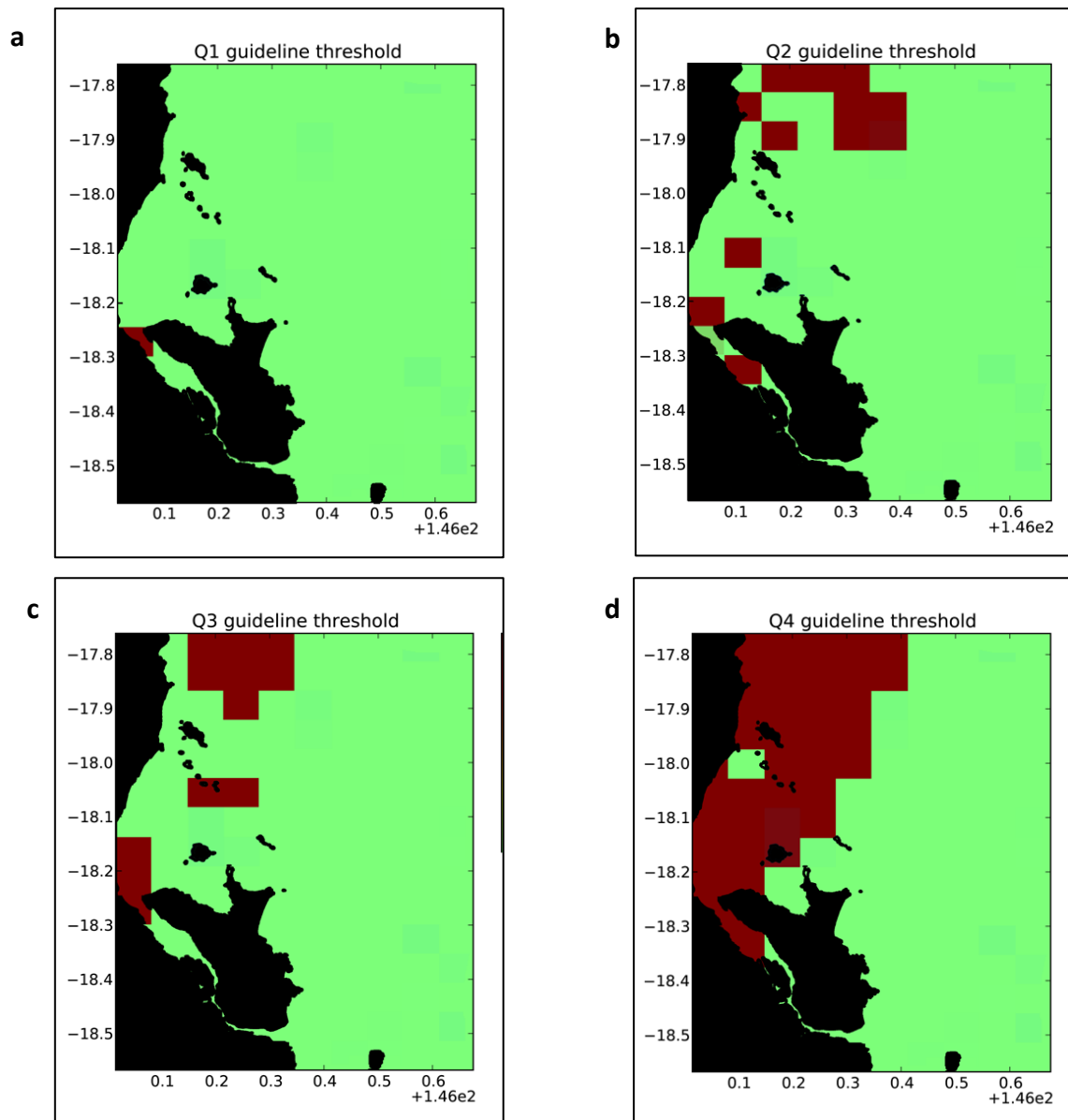


Figure 3-3: Compliance of average TSS concentration with the seasonally adjusted GBR water quality guideline trigger values (GL;GBRMPA 2010) in four reporting quarters (Q1 to Q4) during the 2010-11 Water Year. Green pixels have a quarterly average < GL and are allocated a compliance score C of 0, red pixels are >GL and have a C of 1.

3.2.2. *Assess the magnitude of non-compliance of TSS with the water quality guidelines*

In this step, the magnitude of non-compliance with the GBR Water Quality Guidelines (GBRMPA 2010) was quantified. For this assessment, the Extreme Case Scenario values (ECS) were calculated (see methods), which provide a context based on long-term data for each pixel and season. The magnitude of non-compliance (MNC) was then calculated using the equation:

$$MNC = \frac{(Average_{Q1-Q4} - GL)}{ECS - GL}$$

- where MNC (per non-compliant pixel) equals the magnitude of non-compliance,
- Average_{Q1..Q4} is the average TSS (or other WQ parameter) over the reporting quarters
- GL is the water quality guideline (seasonally adjusted)
- ECS is the Extreme Case Scenario values, calculated from the 95th percentile of all data.

In each quarter, **average** magnitudes of non-compliant pixels is thus the average of the maps in Figure 3-2, calculated only for the non-compliant pixels identified in Figure 3-3. The **average** magnitudes of non-compliant pixels were 0.44 in Q1; 0.76 in Q2; 0.46 in Q3; and 0.81 in Q4 (Table 3.1).

The values for the compliance score (**C**) from this step and for the MNC were then translated into TSS assessment scores for each quarterly pixel (Figure 3-4):

- Pixels with a quarterly average below the GL have a C=0 and a negative MNC value; these pixels were assigned a TSS score of 0.
- Pixels with a quarterly average above the ECS have an MNC of larger than 1 and are assigned a TSS score of 1.
- Pixels with quarterly values larger than the GL, have a C = 1, but lower than the ECS have an MNC between 0 and 1. ***In these cases, the assigned assessment score is the MNC value.***

Overall compliance scores, for TSS only, for each quarter throughout the Tully region were 0.0 for Q1; 0.07 for Q2; 0.08 for Q3; and 0.26 for Q4 (Table 3.1).

Table 3-1: The values calculated for TSS for the reporting area (See Figure xx), including the average TSS value predicted from surface model, the average magnitude on non-compliance against the WQ guidelines, the average magnitude o per quarter average MNC (magnitude of non-compliant pixels) and overall compliance score for TSS reported for each quarter within the selected reporting area for the Tully region.

Season	Average TSS values (calculated from posterior surface model) (mg L ⁻¹)	Non-compliant TSS against WQ guidelines, reported as percent of pixels per quarter	Average magnitude of non-compliance pixels	TSS compliance score (MNC)
Oct-Dec (Q1)	0.49	1%	0.44	0
Jan – Mar (Q2)	0.62	10%	0.76	0.07
Apr – Jun (Q3)	0.60	10%	0.46	0.08
Jul – Sep (Q4)	0.88	32%	0.81	0.26

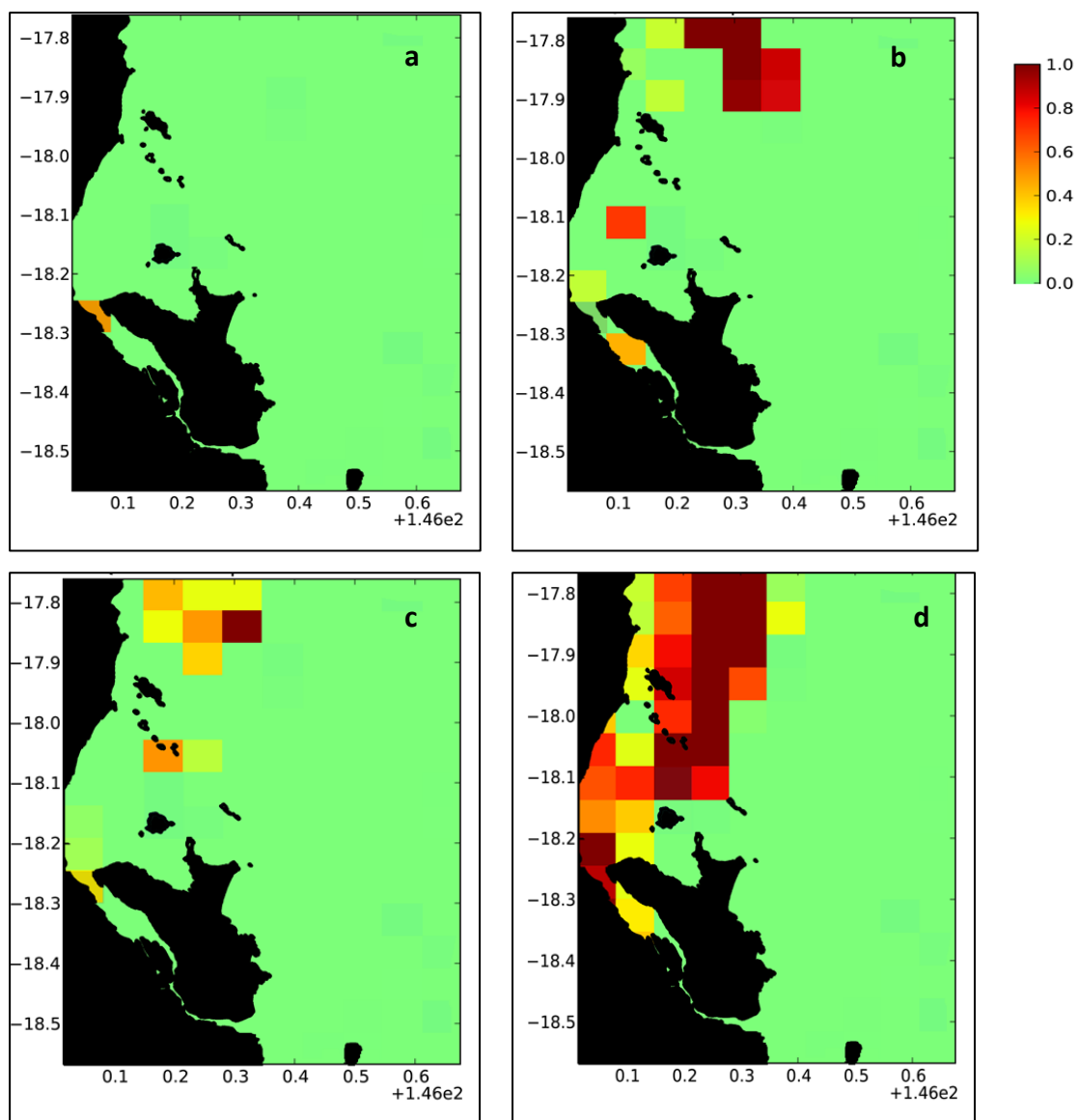


Figure 3-4: TSS assessment scores in four reporting quarters during the 2010-11 Water Year, based on assessments of the magnitude of non-compliance of average TSS concentrations. See text for further explanation of the method.

3.3. Calculation of assessment scores for DIN and PN

3.3.1. Calculate seasonal averages (per grid) for DIN and PN concentrations using in-situ data only.

Historical data has been collected through two programs, the MMP- flood plume program (Devlin et al., 2011; 2012), focusing on variability in the wet season and the MMP-AIMS water quality program focusing on long term data collection through the year at key sites along the inshore GBR (Schaffelke et al., 2011; 2012).

Figure 3-5 shows the full set of historical data for the whole GBR and the data available for the reporting area. The ECS value for both DIN and PN has been calculated from this long term data. Note that both the calculation of the 2010-2011 quarterly averages and the long term quarterly ECS values have been calculated from in-situ data only and have much higher data representation in the wet season due to high frequency wet season sampling. Dry season sampling has occurred from the AIMS MMP water quality monitoring, but is focused on a small sub-set of the reporting area.

In-situ data collected within the reporting area was analysed to give quarterly averages for 10km² spatial pixel (Figure 3-5). Note that the *in-situ* water quality data is aggregated into 10km² grid sections across the reporting area. The smaller grid aggregation employed by the RS modelling would not allow enough in-situ data points to calculate a long term ECS value and thus sub-reporting areas have been aggregated to a larger spatial resolution identified by a series of 10km² gridded boxes within the reporting area. Data frequency was highest within 40km of the river mouth, and thus reporting of the DIN and PN only occurred in 9 spatial pixels (Figure 3-5b) where frequency of in-situ data points was greater than 10 for the reporting year (2010-11). In-situ data is dominated by the seasonal measurements taken under the MMP – flood plume program and thus frequency is highest during the summer months (Oct to Dec, Q1 and Jan to Mar, Q2). Water quality measurements have also been taken by AIMS during the dry season, allowing some measure of seasonal average during Q3 and Q4. Total average DIN values were highest in Q2 (5.1µM [1.1, 12.7]), reflecting the influence of high river flow on the DIN concentration followed by Q1 (3.7 [0.7, 17.4]). Both Q3 and Q4 had low average values (1.3 [0.8, 1.9] and 1.1 [0.9, 1.2] reflecting the low river flow during this period and reduced sampling (Table 3-2).

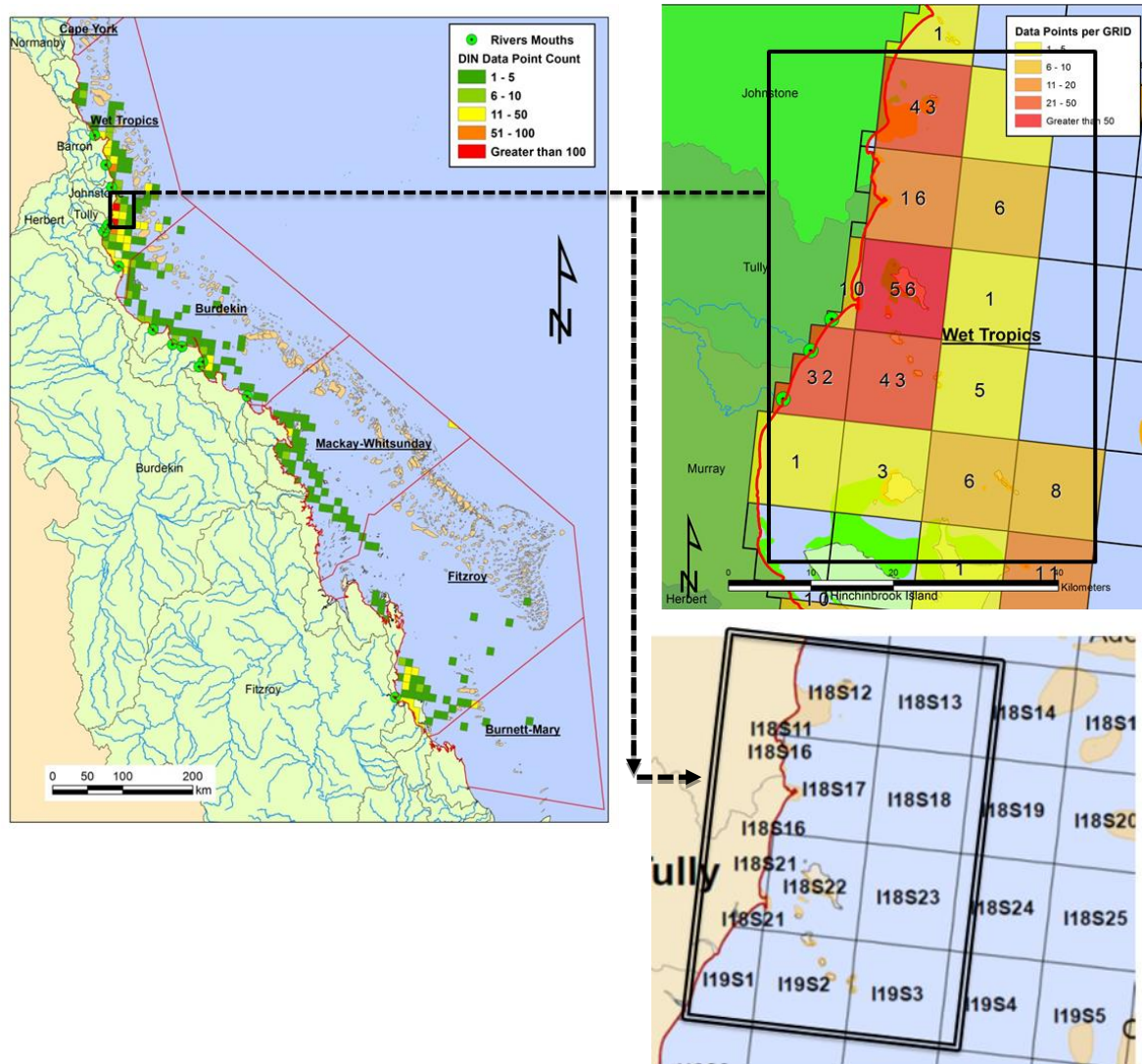


Figure 3-5: Example reporting area for the Tully marine area. Area highlighted in (a) represents the area aligned to the 91 km² sub-region of the Wet Tropics NRM region reported in Figure 2.3. (b) Reporting is expressed for DIN and PN per 10km² grid. The frequency of data points are illustrated in (b) by the colour, with mean average DIN reported in each reporting grid. Grid squares are identified in (c)

As with the TSS assessment, Step 2 of the process assessed the compliance of quarterly DIN and PN values collected in the 2010-2011 year against the water quality guidelines. In this step of the assessment, the average DIN and PN value per reporting area per quarter (see Table 3-2) are compared to the seasonally adjusted GBR water quality guideline trigger values (GL) for DIN and PN (GBRMPA 2010; Moss et al., 2005). The DIN and PN average values were calculated from the available *in situ* data, from 2010-2011 only, to give quarterly averages for each 10 km² spatial grid (Figure 3-5). Total average DIN values were highest in Q1 (2.1 [$\pm$ 2.1] μ M; mean [SE]), followed by Q2 (1.6 [$\pm$ 1.6]) and Q3 (1.4 [$\pm$ 0.6]); Q1 had the lowest estimated (1.1 [$\pm$ 0.2]). Total average PN values were highest in Q1 (4.0 [$\pm$ 0.8] μ M; mean [SE]), followed by Q2 (3.6 [$\pm$ 0.8]) and Q4 (1.5 [0.1]); Q3 had the lowest estimated PN (0.2 [$\pm$ 0.1]).

For DIN for October to August (Q1, Q2, Q3, Q4) this guideline value is 1.7µM. For PN for October to December (Q1), this value is 1.7 µM, for the main months of the wet season, January to March (Q2) the GL is 1.7 µM. For April to June (Q3) the GL is 1.4 µM, and for the main months of the dry season, July to September the GL is 1.4 µM. Reporting grids with a quarterly average below the GL were allocated a compliance score C of 0, while grids with averages above the GL have a C of 1. Estimated percent non-compliance of grid per quarter were 1% in Q1; 10% in Q2 and Q3; and 32% in Q4 (Figure 15).

Table 3-2: Mean DIN and PN value calculated for each quarter in the 2010-2011 year for each grid code within the Tully marine reporting area.

DIN - average 2010-11					PN - average 2010/11				
	Q2	Q3	Q4	Q1		Q2	Q3	Q4	Q1
I18S12	2.1			1	I18S12	2.6			2.2
I18S17	1.1		1.2	0.8	I18S17	2		1.4	2.9
I18S21	3.7			1.4	I18S21	2.5			3.6
I18S22	2.4	0.8	0.9	1	I18S22	2.1	0.1	1.6	1.8
I19S1	2.8	1.9		0.9	I19S1	7.6	0.2		4.9
I19S2	4			0.7	I19S2	3.2			1.9
I19S9									
L22S2	12.3			6.4	L22S2	2.5			6.9
L22S7	12.7			17.4	L22S7	6.3			7.9
Average	5.1	1.4	1.1	3.7	Average	3.6	0.2	1.5	4.0
SE	1.6	0.6	0.2	2.1	SE	0.8	0.1	0.1	0.8

In the final step, the magnitude of non-compliance with the GBR Water Quality Guidelines (GBRMPA 2010) was quantified. For this assessment, the Extreme Case Scenario values (ECS) were calculated (see methods), which provide a context based on long-term data for each grid and season.

The magnitude of non-compliance (MNC) was then calculated using the equation:

$$MNC = \frac{(Average_{Q1-Q4} - GL)}{ECS - GL}$$

- where MNC (per non-compliant pixel) equals the magnitude of non-compliance,
- Average_{Q1..Q4} is the average TSS (or other WQ parameter) over the reporting quarters
- GL is the water quality guideline (seasonally adjusted)
- ECS is the Extreme Case Scenario values, calculated from the 95th percentile of all data.

The magnitude of non-compliance (MNC) was then calculated using the equation: MNC = (avg-GL)/(ECS-GL). For DIN, in each quarter, average magnitudes of non-compliant pixels were 0.6 in Q1; 0.6 in Q2; 0.5 in Q3; and 0.7 in Q4 (Table x). For PN, in each quarter, average magnitudes of non-compliant pixels were 1.0 in Q1; 0.4 in Q2; 0.0 in Q3; and 0.3 in Q4 (Table 3-3).

The values for the compliance score C from assessment step 1 and for the MNC were then translated into DIN and PN assessment scores for each quarterly pixel (Figure 3-4):

- Pixels with a quarterly average below the GL have a C=0 and a negative MNC value; these pixels were assigned a DIN score of 0.
- Pixels with a quarterly average above the ECS have an MNC of larger than 1 and are assigned a DIN score of 1.
- Pixels with quarterly values larger than the GL but lower than the ECS have an MNC between 0 and 1. In these cases, the assigned assessment score is the MNC value.

Overall compliance scores for each quarter throughout the Tully region were 0.0 for Q1; 0.07 for Q2; 0.08 for Q3; and 0.26 for Q4.

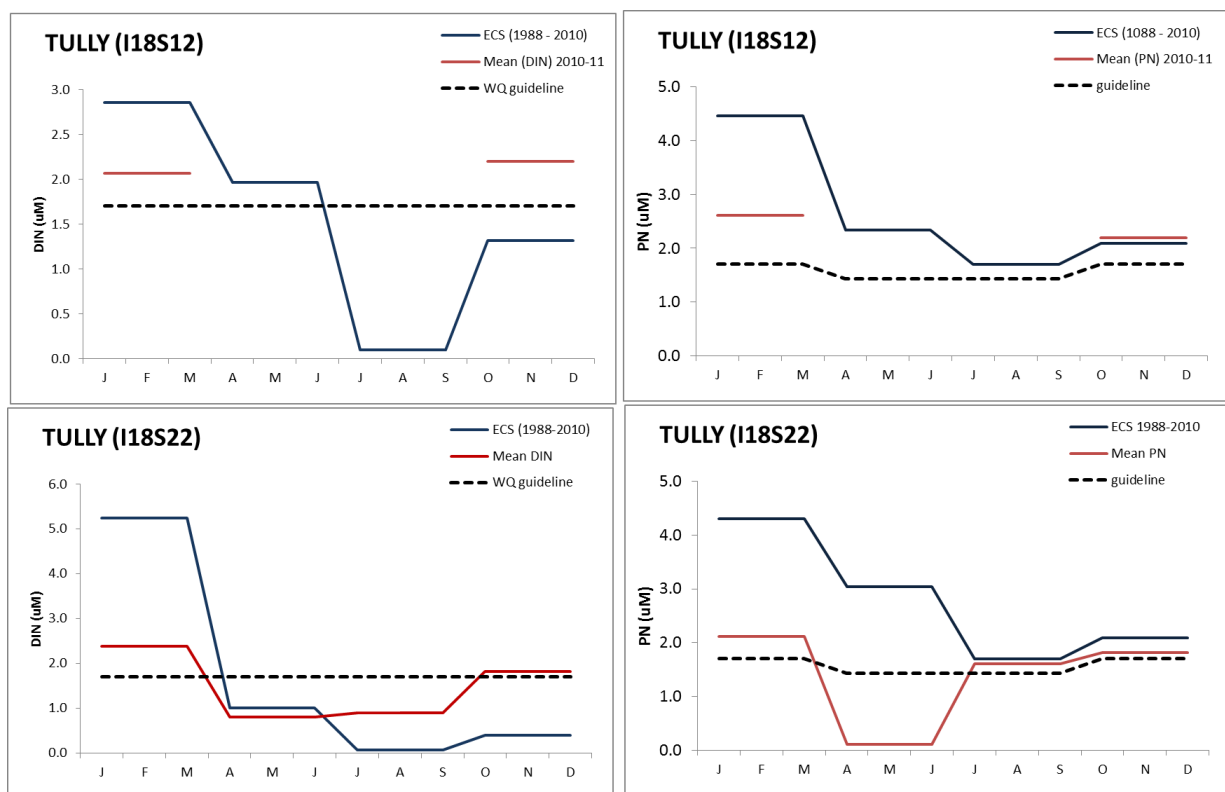


Figure 3-6: The mean DIN and PN values for 2010-2011 as compared against the WQ guideline and the ECS presented for two grid sub-reporting areas including (a) 18S12 and (b) 18S22. Water quality guidelines for DIN is presented as an annual mean (Moss et al., 2005) and seasonally adjusted for PN (see GBRMPA, 2010). The quarterly ECS has been calculated from the long term historical data (1988 – 2010) within each sub-reporting area.

Table 3-3: Average values for DIN and PN for each grid area, reported as compliance scores (c), ECS values and MNC (magnitude of non-compliance) are shown for each grid contained in the Tully marine reporting area. Orange shading denotes grids that do not have sufficient data points (>10) to calculate an ECS or no data was collected in 2010-11. Cells with * denotes grids where MNC could not be calculated as the ECS was not available or it was lower than the guideline values.**

DIN ₂₀₁₁ Q1						PN ₂₀₁₁ Q1					
	2010-11	GL	C	1988-2010			2010-11	GL	C	1988-2010	
	MEAN			ECS	MNC		MEAN			ECS	MNC
I18S12	1.01	1.7	0		0	I18S12	2.20	1.7	1	1.3	***
I18S17	0.79	1.7	0		0	I18S17	2.89	1.7	1		
I18S21	1.36	1.7	0		0	I18S21	3.57	1.7	1		
I18S22	0.99	1.7	0		0	I18S22	1.82	1.7	1	0.4	***
I19S1	0.86	1.7	0		0	I19S1	4.85	1.7	1	9.6	0.40
I19S2	0.65	1.7	0		0	I19S2	1.89	1.7	1		
I19S9						I19S9					
L22S2	6.35	1.7	1		***	L22S2	6.93	1.7	1		
L22S7	17.42	1.7	1		***	L22S7	7.85	1.7	1		
AVG	3.68				0	AVG	4.00				0.40
DIN ₂₀₁₁ Q2						PN ₂₀₁₁ Q2					
	2010-11	GL	C	1988-2010			2010-11	GL	C	1988-2010	
	MEAN			ECS	MNC		MEAN			ECS	MNC
I18S12	2.06	1.7	1	2.86	0.31	I18S12	2.61	1.7	1	4.46	0.33
I18S17	1.11	1.7	0	1.66	0	I18S17	2.04	1.7	1	3.06	0.25
I18S21	3.68	1.7	1	16.06	0.14	I18S21	2.46	1.7	1	4.92	0.24
I18S22	2.39	1.7	1	5.25	0.19	I18S22	2.12	1.7	1	4.30	0.16
I19S1	2.80	1.7	1	9.59	0.14	I19S1	7.64	1.7	1	6.35	1
I19S2	4.04	1.7	1	6.69	0.47	I19S2	3.17	1.7	1	4.77	0.48
I19S9				1.33		I19S9					
L22S2	12.34	1.7	1	11.22	1	L22S2	2.51	1.7	1	10.84	0.09
L22S7	12.67	1.7	1	21.64	1	L22S7	6.26	1.7	1	13.57	0.38
AVG	5.14				0.41	AVG	3.60				0.37
DIN ₂₀₁₁ Q3						PN ₂₀₁₁ Q3					
	2010-11	GL	C	1988-2010			2010-11	GL	C	1988-2010	
	MEAN			ECS	MNC		MEAN			ECS	MNC
I18S12				1.96		I18S12				2.34	
I18S17						I18S17					
I18S21						I18S21					
I18S22	0.81	1.7	0	1.00	0	I18S22	0.11	1.4	0	3.04	0
I19S1	1.89	1.7	0	11.26	0	I19S1	0.17	1.4	0	6.89	0
I19S2				1.21		I19S2				2.18	
I19S9						I19S9					
L22S2						L22S2					
L22S7						L22S7					
AVG	1.35				0	AVG	0.14				0
DIN ₂₀₁₁ Q4						PN ₂₀₁₁ Q4					
	2010-11	GL	C	1988-2010			2010-11	GL	C	1988-2010	
	MEAN			ECS	MNC		MEAN			ECS	MNC
I18S12				0.10		I18S12					
I18S17	1.21	1.7	0		0	I18S17	1.36	1.4	0	1.7	0
I18S21						I18S21					
I18S22	0.89	1.7	0	0.07	0	I18S22	1.61	1.4	1	1.7	0.69
I19S1						I19S1					
I19S2						I19S2					
I19S9						I19S9					
L22S2						L22S2					
L22S7						L22S7					
AVG	1.05				0	AVG	1.48				0.34

Table 3-4: Average DIN values for the Tully region, reported for each quarter. The percent value of non-compliant DIN is calculated from the number of grid squares with value greater than WQ guideline. The DIN assessment score is calculated from the average MNC score per grid square.

Season	Average DIN values (calculated from in-situ data) (mg L ⁻¹)	Non-compliant DIN against WQ guidelines, <i>reported as percent of reporting grids per quarter*</i>	DIN compliance score
Oct-Dec (Q1)	3.68	25%	0
Jan – Mar (Q2)	5.14	87.5%	0.41
Apr – Jun (Q3)	1.35	0%	0
Jul – Sep (Q4)	1.05	0%	0

Table 3-5: Average PN values for the Tully region, reported for each quarter. The percent value of non-compliant DIN is calculated from the number of grid squares with value greater than WQ guideline. The DIN assessment score is calculated from the average MNC score per grid square.

Season	Average PN values (calculated from in-situ data) (mg L ⁻¹)	Non-compliant PN against WQ guidelines, <i>reported as percent of reporting grids per quarter*</i>	PN compliance score
Oct-Dec (Q1)	4.00	100%	0.4
Jan – Mar (Q2)	3.60	100%	0.36
Apr – Jun (Q3)	0.14	0%	0
Jul – Sep (Q4)	1.48	50%	0.34

3.4. Integration of regional assessment scores between water quality variables

The final step in a water quality assessment is to combine all the aggregated indicator condition scores to calculate an overall report card score for reporting year (Table 3-2). For this report, we have calculated a seasonal aggregated assessment score across all pixel or grids within the Tully reporting area. To obtain one value per quarter, representative of all the water quality parameters within the assessment, by averaging the seasonal compliance scores of each water quality parameter have been averaged to provide one overall assessment score.

The last column of the table reports the aggregated assessment scores for each season, while the annual scores are reported in the last line. In the table all condition scores (i.e. the

compliance and aggregated assessment scores) are colour coded to represent a report card ranking for of five categories ranging from “very poor” to “very good”, represented by a colour scheme consistent with the Paddock to Reef reporting scheme (Figure 3.7).

The annual compliance scores for TSS and DIN were 0.10 representing a “very poor” condition while the PN score (0.28) represented a “poor” condition. The annual aggregated water quality assessment score (i.e. the value in the lower right corner of the table) of this case study was 0.16 representing a very poor water quality status for the Tully region for 2010/11.

Table 3-6: Aggregated scores taken across all water quality indicators in the Tully region (see Figure 2-9). The aggregated score is calculated from average of the assessment scores for the three water quality parameters used in this report. The cells are colour coded according to the resulting report card rating.

Season	TSS compliance score	DIN compliance score	PN compliance score	Aggregated assessment score
Oct-Dec (Q1)	0	0	0.40	0.13
Jan – Mar (Q2)	0.07	0.41	0.36	0.28
Apr – Jun (Q3)	0.08	0	0	0.03
Jul – Sep (Q4)	0.26	0	0.34	0.20
Oct 2010 – Sep 2011	0.10	0.10	0.28	0.16

Compliance score range	0 - 0.20	0.20 - 0.40	0.40 - 0.60	0.60 - 0.80	0.80 - 1
Report card rating	very poor	Poor	moderate	good	very good



Figure 3-7: Conceptual diagram of overall water quality assessment with the aggregation of each water quality assessment scores.

4. Discussion

4.1. Statistical model

The explosion of large spatiotemporal datasets arising from remote sensing technology has generated considerable interest in modelling geolocated data (Banerjee et al. 2008) and the water quality monitoring on the Great Barrier Reef is no exception. While the technology to collect exceptionally large volumes of data has expanded rapidly, the tools to analyze such data have not; although suitable methods have been developed, they cannot be readily implemented by the majority of statistical and computational researchers. As a result, the inability model these kinds of data is the greatest limitation to their achieving full, real-time integration into water-quality monitoring and assessment.

Implementation for large spatio-temporal water quality data requires three key skills:

1. Understanding spatio-temporal models
2. Computational power
3. Geostastical mapping

Despite the importance of space and time in all measurements of the natural world, there are remarkably few statisticians working in environmental science that specialize in spatio-temporal modelling. This is largely due to the complexities of understanding and implementing the various variance covariance components need to account for both spatial and temporal variation simultaneously; however it also includes issues around missing data and the biological interpretation of results. Only recently have these methods been complied comprehensively (Cressie and Wikle 2011), and the kinds of Gaussian process models appropriate for the report card are not even included (Patil et al. 2011).

The volume of water quality data available for the Great Barrier Reef annually is in the millions of observations. Handling and modelling such vast quantities of data in a single coherent model requires vast computational support for data wrangling, optimization, and parallel or super computing. While each of these elements is currently available nationally, dedicated persons are required to help implement the full water quality reporting framework across the entire GBR.

While building and implementing appropriate spatio-temporal models alone are difficult tasks, results must be presented visually, often as maps that require their own particular set of visualization skills. We see the Gaussian process-based maps currently implemented as part of the Malaria Atlas Project (<http://www.map.ox.ac.uk>) as being exactly the sort of visual mapping tools needed for water quality monitoring on the Great Barrier Reef. Here, a suite of tools has been developed to infer the global distribution of *Plasmodium falciparum* parasites annually, as well as finer-scale distributions and values for national or regional use. The sparse, opportunistic data collected for water quality are of exactly the same form and have the same intrinsic problems for inference.

4.2. Proof of concept – integrated assessment

The project aim of developing a comprehensive approach to reporting marine water quality on the GBR was achieved by producing a proof of concept.

The development of water quality assessment and reporting framework was closely aligned with other nationally and international approaches for ecological monitoring, assessment and reporting. A review of other assessment frameworks and their approaches and methods carried out in this project found that the reporting is often carried out in small reporting units where it should be reasonable to expect measurable change in response to variations to the environmental pressures.

In light of the findings of the review, we revisited the criteria to determine the appropriate spatial and temporal scale for analysis and reporting for the reporting framework:

- In the GBR, the direct influence of land run-off occurs at a smaller scale than the whole NRM regions currently used for P2R reporting. Hence there is a need to define smaller regions that are directly influenced by land run-off and consistent with local oceanography (e.g residence times, bathymetry, hydrodynamics). In these smaller reporting areas should be reasonable to expect measurable change over the next decade in response to improvement in land management practices.
- The current P2R reporting on the whole of year tends to have limited meaning, hence we propose to assess the condition of water relative to the seasonal guidelines (four seasons per year). This should provide an improved capacity to untangle the effects of land management practices from those directly due to land-runoff events.
- To include in the reporting framework a measure of the year to year variability, we introduced the concept of “Extreme Case Scenario” (ECS), as the upper limit of the “known” variability of the water quality parameter. ECS varies spatially and seasonally over the GBR and were computed from the ten years of remote sensing data, augmented with in situ observations

To aggregate multiple sources of data across various spatial and temporal scales in the reporting framework, a spatio-temporal statistical process model was developed. This allows to predict average “response surface” for each indicator as well as the associated uncertainties. In a spatial-temporal analysis, space and time are sometimes considered separately to provide some diagnostic capacity for the analysis. However, space and time were found to be inter-related dimensions in our context so were ultimately not separable; hence the combination of data from multiple sources was carried out across space and time simultaneously.

For the proof of concept of the reporting framework we selected an area with the most data, in both space and time, from direct water sampling. A small reporting region around the Tully mouth/Rockingham Bay was defined, using the criteria outlined above, which encompassed most of these data points. The reporting year was 2010-11, which was the year of the MMP sampling which saw the most extreme floods along the whole of the GBR coast. By selecting the most data-rich region and the most extreme year we aimed to maximise the ability of the proposed reporting framework to show seasonal and spatial differences.

The parameters selected for the proof of concept of the reporting framework concept were (i) TSS measured by remote sensing, direct water sampling both during floods and ambient conditions, and water quality instruments on coral reefs, and (ii) DIN and PN values measured only by direct water sampling both during floods and ambient conditions.

- Aspects of seasonality are included in this approach allowing consideration of the variable influences through the dry and wet season.

- For this proof of concept, several implementation issues need to be carefully considered, i.e. the definition of the reporting boundaries of the water bodies. It is agreed that definition of the new smaller, localised coastal reporting regions provide a better reporting base for WQ assessments based on the influence of the river flow and water quality associated with that water flow.
- Implementation on a regular basis for whole GBRWHA is a non-trivial computational exercise. To be able to do the data integration across temporal and spatial scales at GBR scale needs development of parallel computing proper operational software. For this case study, we have focused on a small reporting area (Tully marine area).
- Approach is fully compatible with the current Paddock to Reef reporting and builds on the current model of tiered reporting across regional areas.
- Implementation of the current approach would need to be considered in the context of potential future changes to the Paddock to Reef monitoring and reporting structure.

4.3. Future research

- Accounting for the variance and covariance of observations in space and time are critical if best-practice water quality maps are to be developed for the Great Barrier Reef. However implementing these methods requires specialized statistical, mapping, and computational skills that exceptionally few persons possess. We recommend that a dedicated spatio-temporal modeler familiar with Gaussian process models and mapping be retained, alongside a computer technician, to implement the full water quality framework.
- The proposed framework included the definition of the new smaller, localised coastal reporting regions to provide a better reporting base for WQ assessments based on the influence of the river flow and water quality associated with that water flow. These smaller reporting regions need to be defined, based on the decision-making criteria proposed in the assessment framework.
- In the proposed framework, all variables were normalised and no weighting was applied to the combination of the water quality variables. Future applications of this approach needs to consider the weighting of the different WQ variables based on ecological understanding of relevant thresholds for GBR ecosystems such as corals and seagrasses. Weighting and/or thresholds need to be considered in both seasonal and geographical context.
- The issue of the overlapping influence on the waterbodies will need detailed analysis of the influence of river flow and plume formation from individual rivers. .
- Further development of water quality guidelines for the GBR that is specific for regional characteristics and /or for coral and seagrass environments.
- In order to integrate the approach, other water quality indicators such as pesticides implementation issues such as the definition of ECS and spatial and temporal aggregations need to be investigated.
- For this case study, we have focused on a small reporting area (Tully marine area) as it the most data rich in the GBR. To be able to achieve a full GBR implementation of this approach it will be necessary to assess whether the MMP sampling design provides a sufficient spatial and temporal support for the statistical integration model.

References

- Anon, 2003. Reef Water Quality Protection Plan for catchments adjacent to the Great Barrier Reef World Heritage Area. The State of Queensland and Commonwealth of Australia. Queensland Department of Premier and Cabinet, Brisbane. p. 43.
- Anon, 2009a. Reef Water Quality Protection Plan 2009. For the Great Barrier Reef World Heritage Area and adjacent catchments. The State of Queensland and Commonwealth of Australia. Queensland Department of Premier and Cabinet, Brisbane. p. 30.
- Anon, 2009b. Paddock to Reef Program: Integrated Monitoring, Modelling and Reporting (Reef Water Quality Protection Plan). The State of Queensland (Department of the Premier and Cabinet), Brisbane, p. 8.
- Álvarez-Romero, J.G., Devlin, M., Teixeira da Silva, E., Petus, C., Ban, N.C., Pressey, R.L., Kool, J., Roberts, J.J., Cerdeira- Estrada, S., Wenger, A.S., Brodie, J., 2013. A novel approach to model exposure of coastal-marine ecosystems to riverine flood plumes based on remote sensing techniques. *Journal of Environmental Management* 119, 194-207.
- Banerjee, S., Carlin, B.P., and Gelfand, A.E. (2004). Hierarchical modeling and analysis for spatial data. Boca Raton: Chapman and Hall-CRC.
- Banerjee, S., Gelfand, A.E., Finley, A.O., and Sang, H. (2008). Gaussian predictive process models for large spatial data sets. *Journal of the Royal Statistical Society Series B* 70, 825-848.
- Brinkman R., Herzfeld M., Andrewartha A., Rizwi F., Steinberg C. and Spagnol S. (2011) Project 2.5i.1 Hydrodynamics at the Whole of GBR Scale. Final Project Report to the Marine and Tropical Sciences Research Facility. AIMS, Townsville.
- Brodie, J.E., Devlin, M.J., Haynes, D., Waterhouse, J., 2011. Assessment of the eutrophication status of the Great Barrier Reef lagoon (Australia). *Biogeochemistry* 106, 281-302. DOI 10.1007/s10533-010-9542-2.
- Brodie, J.E., F.J. Kroon, B. Schaffelke, E.C. Wolanski, S.E. Lewis, M.J. Devlin, I.C. Bohnet, Z.T. Bainbridge, J. Waterhouse, and A.M. Davis, Terrestrial pollutant runoff to the Great Barrier Reef: An update of issues, priorities and management responses. *Marine Pollution Bulletin*, 2012a. 65(4–9): p. 81-100.
- Cloern, J., 2001. Our evolving conceptual model of the coastaleutrophication problem. *Marine Ecology Progress Series* 210, 223–253.
- Department of the Premier and Cabinet (2011) Reef Water Quality Protection Plan: First Report 2009 Baseline. Available from www.reefplan.qld.gov.au
- Devlin, M., Waterhouse, J., Taylor, J., Brodie, J., 2001. Flood plumes in the Great Barrier Reef: spatial and temporal patterns in composition and distribution. GBRMPA Research Publication No. 68, Great Barrier Reef Marine Park Authority, Townsville, Australia.
- Devlin, M., Harkness, P., McKinna, L. and Waterhouse, J., 2011. Mapping the surface exposure of terrestrial pollutants in the Great Barrier Reef. Report to the Great Barrier Reef

Marine Park Authority, August 2010. Australian Centre for Tropical Freshwater Research. Report Number 10/12.

Devlin, M.J., L.W. McKinna, J.G. Álvarez-Romero, C. Petus, B. Abott, P. Harkness, and J. Brodie, Mapping the pollutants in surface riverine flood plume waters in the Great Barrier Reef, Australia. *Marine Pollution Bulletin*, 2012. 65(4–9): p. 224-235.

Car, N, Fitch, P, Lemon, D (2012). "Scoping Study: eReefs Work Package 2–Interoperable data and information systems." (2012).

Furnas, M., D. Alongi, D. McKinnon, L. Trott, and M. Skuza. "Regional-scale nitrogen and phosphorus budgets for the northern (14 S) and central (17 S) Great Barrier Reef shelf ecosystem." *Continental Shelf Research* 31, no. 19 (2011): 1967-1990.

Gotway, C.A. and Young, L.J. (2002) Combining incompatible spatial data. *Journal of the American Statistical Association* 97, 632-48.

Gotway, C.A. and Young, L.J. (2007) Ageostatistical approach to linking geographically aggregated data from different sources. *Journal of Computational and Graphical Statistics* 16, 115-35.

Great Barrier Reef Marine Park Authority (2010) Water quality guidelines for the Great Barrier Reef Marine Park, rev ed. Great Barrier Reef Marine Park Authority, Townsville

Henderson, B. (2003) On integrating data at different scales with hierarchical modeling. CMIS Technical Report Number 03/148.

Kroon, F. J., P.M. Kuhnert, B.L. Henderson, S.N. Wilkinson, A. Kinsey-Henderson, B. Abbott, J.E. Brodie, and R.D. Turner. "River loads of suspended solids, nitrogen, phosphorus and herbicides delivered to the Great Barrier Reef lagoon." *Marine Pollution Bulletin* 65, no. 4 (2012): 167-181.

Larcombe P, Ridd PV, Prytz A, Wilson B (1995) Factors controlling suspended sediment on inner-shelf coral reefs, Townsville, Australia. *Coral Reefs* 14:163–171

Nixon, S.W., 2009. Eutrophication and the macroscope. *Hydrobiologia* 629, 5-19

Schaffelke B, Carleton J, Costello P, Davidson J, Doyle J, Furnas M, Gunn K, Skuza M, Wright M, Zagorskis I (2012) Reef Rescue Marine Monitoring Program. Final Report of AIMS Activities 2010/11– Inshore Water Quality Monitoring. Report for the Great Barrier Reef Marine Park Authority. Australian Institute of Marine Science, Townsville (97 p.)

Schroeder, T., M.J. Devlin, V.E. Brando, A.G. Dekker, J.E. Brodie, L.A. Clementson, and L. McKinna, Inter-annual variability of wet season freshwater plume extent into the Great Barrier Reef lagoon based on satellite coastal ocean colour observations. *Marine Pollution Bulletin*, 2012. 65(4–9): p. 210-223.

Young, L.J. and Gotway, C.A. (2007) Linking spatial data from different sources: the effects of change of support. *Stochastic Environmental Research and Risk Assessment* 21, 589-600.