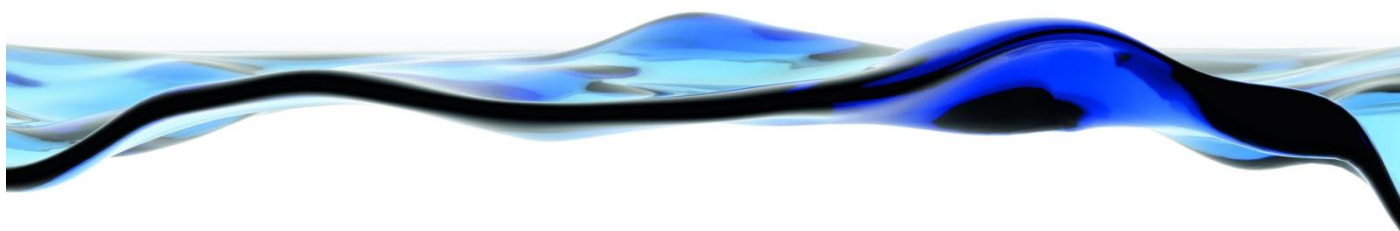


Integrated catchment water planning support for Adelaide Mount Lofty Ranges Water Allocation Planning (GWAP Project)

Task 5: Tiered Water Quality Risk Assessment

Ford, J.H., Ickowicz, A., Oliver, D., Hayes, K.R., and Kookana, R.



Goyder Institute for Water Research
Technical Report Series No. 15/4

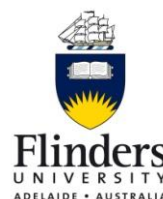


www.goyderinstitute.org



Goyder Institute for Water Research Technical Report Series ISSN: 1839-2725

The Goyder Institute for Water Research is a partnership between the South Australian Government through the Department of Environment, Water and Natural Resources, CSIRO, Flinders University, the University of Adelaide and the University of South Australia. The Institute will enhance the South Australian Government's capacity to develop and deliver science-based policy solutions in water management. It brings together the best scientists and researchers across Australia to provide expert and independent scientific advice to inform good government water policy and identify future threats and opportunities to water security.



The following Associate organisations contributed to this report:



Enquires should be addressed to: Goyder Institute for Water Research
Level 1, Torrens Building
220 Victoria Square, Adelaide, SA, 5000
tel: 08-8303 8952
e-mail: enquiries@goyderinstitute.org

Citation

Ford, J.H., Ickowicz, A., Oliver, D., Hayes, K.R., Kookana, R. 2015. GWAP Project , Task 5: Tiered Water Quality Risk Assessment, Goyder Institute for Water Research Technical Report Series No. 15/4, Adelaide, South Australia.

Copyright

© 2015 [CSIRO] To the extent permitted by law, all rights are reserved and no part of this publication covered by copyright may be reproduced or copied in any form or by any means except with the written permission of [CSIRO].

Disclaimer

The Participants advise that the information contained in this publication comprises general statements based on scientific research and does not warrant or represent the completeness of any information or material in this publication.

Contents

1	Introduction	1
2	Materials and Methods	5
2.1	Water quality data collation in the Mt Lofty Ranges	5
2.2	Tiered risk assessment process	5
2.2.1	Tier 1	6
2.2.2	Tier 2	7
2.2.3	Tier 3	8
2.2.4	Tier 4	8
3	Results	11
3.1	Screening level assessment of water quality in various catchments:	
	Tier 1	11
3.2	Relative risk ranking of catchments: Tier 2	16
3.3	Water quality maps of risk of exceedance hotspots: Tier 2	20
3.4	Temporal trends in water quality parameters: Tier 2	25
3.5	Effect of flow conditions on water quality	27
3.6	Correlation between water quality parameters: Tier 2	30
3.7	Effectiveness of sedimentation pond at Uraidla, Cox Creek	33
3.8	Catchment characteristics and attributes associated with water qual-	
	ity parameters: Tier 4	35
3.8.1	Key catchment characteristics, attributes and hotspots	35
3.8.2	Explanatory covariates: influence of landscape parameters	38
4	Discussion	42
4.1	Relative ranking of catchments (Tier 1 and Tier 2)	42
4.2	Temporal patterns in water quality	43
4.3	Relationship between flow and water quality	44
4.4	Correlation between water quality parameters	44
4.5	Effectiveness of sedimentation pond at Cox Creek Uraidla	44
4.6	Key catchments characteristics associated with water quality status	45
4.7	Limitations and recommendations	45
5	References	47

Appendices	50
Appendix A Data Sources	50
Appendix B Trends in Electrical Conductivity	52
Appendix C Tier 4 Spatio-temporal modelling: Methods and Results	53
C.1 Data	53
C.1.1 Water quality measurement frequency	53
C.1.2 Water quality measurement location	53
C.2 The Spatio-temporal model	56
C.2.1 Overview of the spatio-temporal approach	56
C.3 Land use information treatment	59
C.4 Results Nitrite + Nitrate as Nitrogen	61
C.5 Results Total Nitrogen	69
C.6 Results Total Phosphorus	75
C.7 Results Suspended Solids	83

List of Figures

1	Nitrate + Nitrite as Nitrogen (NO _x) for two sites with contrasting amount of data. Blue dashed line shows median, black solid and dashed lines show mean and standard error (respectively) and red line the ANZECC guideline.	12
2	Total Phosphorus for two sites with contrasting water quality exceedance risk. The blue line shows median, the black lines show mean (solid) and standard error (dashed) and red line the ANZECC guideline.	17
3	Visual representation of sites with risk of exceedance for Total Phosphorus and Reactive Phosphorus, increasing from low (blue) to highest (red) as determined in the Tier 2 assessment.	21
4	Visual representation of sites with risk of exceedance for Total Nitrogen and Nitrate + Nitrite as Nitrogen, increasing from low (blue) to highest (red) as determined in the Tier 2 assessment.	23
5	Visual representation of sites with risk of exceedance for Suspended Solids, increasing from low (blue) to highest (red) as determined in the Tier 2 assessment.	24
6	Box plots of monthly Total Nitrogen for two sites. Any values greater than 1.5 times the interquartile range are represented by a 'o'.	26
7	Monthly box plots and kernel density plots for log(flow) with log(NO _x) and log(Total Nitrogen) at two sites.	28
8	Monthly means (across all years) for log(flow) and log(NO _x) and log(Total Nitrogen) for all sites. Red '+' indicates means for flow for a site, and black 'o' indicates mean for NO _x and TN for a site.	29
9	Two Copulas modelling the joint correlation between two parameters for two different sites. The Gumbel copula is used to model strong upper tail dependence: the stronger the correlation contour, the stronger the tail dependence. The "narrowing" of the contours in the upper right quadrant indicates that the dependence between the two parameters becomes stronger as they become more extreme.	32

10	Data for Suspended Solids and Zinc for Cox Creek at Uraidla. Split lines indicate medians (blue dashed line), means and standard errors (black solid and dashed lines respectively) for data recorded before and after the installation of the sedimentation pond in 2006. The red dashed line indicates ANZECC guideline.	34
11	Prediction performance of the model for Total Nitrogen, at each site. Only one catchment (Onkaparinga River) has Total Nitrogen measurements. The vertical bar for each point is an error bar, stating the confidence interval for each prediction. Y-axis displays log(predictions) and x-axis log(observations). The legend ‘elev’ indicates elevation for each site.	36
12	Prediction performance of the model for Total Nitrogen at two Cox Creek sites. The blue dotted line represents the log of the recommended ANZECC guideline. The y-axis shows log(predictions) and the x-axis Year.	37
13	Land use and soil types PCA representation. The larger the circle, the bigger the weight. Green circles stand for a positive influence, and blue circles for a negative one.	39
14	Estimated values (and error bars) for the different parameters of the water quality model. Only significant influences are displayed. The x-axis indicates the order of magnitude for parameters in the model. This is a proxy for increasing risk, green indicates positive effect on concentration, so larger values in green indicate high positive effect on concentration, and conversely for blue. The parameters preceded by “alpha” are the parameters of the fixed covariates component of the model. The terms “sill” and “nugget” are used to describe the spatial correlation. “V1” is the first order temporal influence (a constant influence over time).	40
15	Land use and soil types composition of the monitored sites with sufficient data. The larger the circle, the bigger the weight. Green circles stand for a positive influence, and blue circles for a negative one. The y-axis for the two figures are four soil type and land use principal components respectively.	41

16	EC time series analysis for Mount Barker Ck US Bremer River junction. The top panel shows all data; the second panel season; the third panel overall trend; and the bottom panel the remaining variance. The x-axis (time) is years of data.	52
17	Mount Lofty Ranges map. Location of the WQ stations, and their agencies; and rainbow map of the different sub-catchments.	54
18	Mount Lofty Ranges map. Map of the sub-catchment with/without WQ stations; and elevation of the WQ stations.	55
19	Land use principal components. The zero-value components have been removed from the figure.	60

List of Tables

1	ANZECC guidelines for freshwater aquatic ecosystems (ANZECC/ARMCANZ, 2000). Guideline values for DOC were provided by SA Water.	6
2	Land use categories and abbreviations used in Tier 4.	10
3	Number of Sites per Catchment	11
4	Number of Sites with median above/below the associated ANZECC water quality guideline	13
5	Sites with number of water quality parameters (for all parameters, and for non-metals) above the associated ANZECC water quality guideline, as a ratio of parameters recorded at that site. EC is excluded from this table.	14
6	Sites with total number of water quality parameters in Tier 2 Risk assessment, average Tier 2 risk for all parameters with and without metals, and relative change in average Tier 2 risk when metals are excluded. The sites are listed from highest to lowest rank of average risk of exceedance excluding metals.	18
7	Sites with Tier 2 Risk (and rank within parameter) for 6 key parameters, and average risk across parameters. Sites are ranked by average risk.	19
8	Summary of copulas for sites with pairwise parameter combinations (split by '/') where Kendall tau was greater than 0.5 or less than -0.5. Abbreviations for parameters are used for clarity: Cu = Copper; NOx = Nitrate + Nitrite as Nitrogen; RP= Reactive Phosphorus; SS = Suspended Solids; Tb = Turbidity; TN = Total Nitrogen; TP = Total Phosphorus; Zn = Zinc.	31
9	Tier 2 Risk for Cox Creek, before and after installation (2006) of sedimentation pond. Note that pH and Turbidity were not recorded prior to 2006 so are not included in the table.	33
10	Summary of sources and contact details for Mt Lofty Ranges Water Quality Risk Assessment.	51
11	Discrepancy of the parameter measurements. The first column states the number of stations measuring the parameter. The last three columns relate to the number of measurements (in time) per station.	53
12	Listing of the secondary land use categories.	59

13	Principal component decomposition. Percentage of explained variance.	60
----	--	----

Acknowledgements

This study was supported by funding from the Goyder Institute for Water Research for the project ‘Mt Lofty Ranges Water Allocation and Planning’ led by Associate Professor Jim Cox. Thank you to Neil Power and Dr Michele Akeroyd for overseeing the Goyder project.

We would like to thank SA Water, SA EPA, DEWNR, AMLR NRMB and Dr Leon van der Linden for providing water quality data from the Mt Lofty Ranges for the risk assessment. Thank you also to staff from numerous South Australian Government Agencies who provided valuable feedback at presentations given about project outputs. We would also like to thank the reviewers for their constructive comments which have significantly improved the report.

Glossary of terms

TERM/ACRONYM	DEFINITION
AMLRNRMB	Adelaide and Mt Lofty Ranges Natural Resource Management Board
ANZECC	Australian and New Zealand Environmental Conservation Council
ARMCANZ	Agriculture Resource Management Council of Australia and New Zealand
Cu	Copper
DEWNR	Department of Environment, Water and Natural Resources (South Australian Government)
DOC	Dissolved organic carbon
EC	Electrical conductivity
EPA	Environment Protection Agency
Fe	Iron
GMP	Grazing modified pastures
IPH	Irrigated perennial horticulture
ISH	Irrigated seasonal horticulture
MLR	Mt Lofty Ranges
MRP	Managed reserve protection
NC	Nature conservation
NO ₃	Nitrate
NO ₂	Nitrite
NO _x	Nitrate and nitrite as nitrogen
Pb	Lead
PCA	Principle components analysis
pH	Measure of acidity or alkalinity
Res	Residential
RP	Reactive phosphorus
SA	South Australia
SAMDBNRMB	South Australian Murray Darling Basin Natural Resource Management Board
Ser	Services
SS	Suspended solids
Tb	Turbidity
TDS	Total dissolved solids
TKN	Total Kjeldahl nitrogen
TN	Total nitrogen
TP	Total phosphorus
WAP	Water Allocation Plan
Zn	Zinc

Executive Summary

The water resources of the Mt Lofty Ranges (MLR) were prescribed in 2005 and a draft Water Allocation Plan (WAP) was released for the western MLR in 2010 (AMLRNRMB, 2010a,b) and for the eastern MLR in 2011 (SAMDBNRMB, 2011). In addition to the WAP in the MLR the South Australian government has identified the need for improved water quality in the catchments through the Water Quality Improvement program and the Water for Good policy. In the current WAPs water quality is considered to be addressed when water volume is not limited i.e. water quality is not an issue if the flow regime is adequate. The current WAPs propose that during low flow conditions water may be diverted from upstream storage locations, however when and where such a strategy is implemented requires an understanding of the water quality in the MLR.

This risk assessment is expected to support policy decision making in the future by providing:

- a systematic, transparent, evidence-based process for determining the allocation of resources in the MLR that incorporates water quality in planning decisions which are fundamental to South Australian Government's Water for Good document and the South Australian EPA's Water Quality Improvement program; and
- a process for making informed decisions to target future monitoring in the MLR sub-catchments and for determining priority areas for investing in mitigation strategies based on a systematic risk ranking;
- a process for identifying which sub-catchments in the MLR would be suitable, and conversely which would not be suitable or present a water quality risk, for water diversion to provide water for environmental requirements during periods of low flow.

This report describes the outcomes of a tiered risk assessment (see Hayes et al. (2012)) of existing water quality data collected over various time scales and locations within the MLR. Initially five tiers, increasing in complexity and data requirement, were proposed, but due to limited data availability and current absence of ecological response functions for key endpoints, only three of the five tiers could be implemented. While all water quality data was collated, the tiered risk assessment focussed only on those parameters with an ANZECC/ARMCANZ guideline (ANZECC/ARMCANZ, 2000) and site/parameter combinations with more than 30 data points (in total across multiple years). Due to the importance of dissolved organic carbon (DOC) as a water quality issue for reservoir

management SA Water provided two guideline values for DOC for inclusion in the risk assessment.

Specific outputs from this study include:

- A systematic, evidence-based approach for ranking sub-catchments in the MLR on the basis of water quality;
- Maps of risk of exceedance “hotspots” across the MLR;
- An assessment of temporal trends in water quality parameters;
- A determination of the relationship between flow and concentrations for selected water quality parameters;
- An assessment of the suitability of using certain water quality parameters as surrogates for other parameters;

Two additional requests, and important outputs of the project, were identified during consultation with stakeholders: assessing the risk of exceedance before and after installation of a sedimentation pond in Cox Creek Uraidla in 2006; and the incorporation of flow and relationship with other water quality parameters.

The Tier 1 analysis identified sites at risk of exceeding the ANZECC water quality guidelines for 12 water quality parameters. The majority of sites exceeded (median calculated across all years) the ANZECC guidelines for most parameters. There were four water quality parameters in which all sites recorded a median below the ANZECC guideline. The median concentration for copper was exceeded for all sites, and for more than half the sites for Zinc. Surface water in the MLR however is known to have naturally elevated concentrations of metals, and these results are not therefore surprising. Other parameters were consistently exceeded across multiple sites and catchments. The Onkaparinga River catchment, which had 14 monitoring sites, and the monitoring sites along Cox Creek, were consistently ranked with high risk across all parameters.

The Tier 2 analysis characterised the distribution function of the water quality parameters considered and provided a more accurate assessment of exceedance risk. Due to strong auto-correlation, the EC data was not analysed using the standard Tier 2 assessment, but instead was analysed using time series analysis (this is included in Appendix B). Average Tier 2 Risk was calculated for all parameters and non-metal parameters (so excluding Copper, Lead and Zinc). For the 10 sites with metals data, the average risk for

non-metals increased for two sites, and decreased for eight sites.

The results of the Tier 2 risk assessment were used to create maps of exceedance risk “hotspots” for all parameters. The Onkaparinga catchment was consistently identified as a “hotspot” with monitoring sites within this catchment ranked high for risk of exceedance across all parameters. The highest risk of exceedance was recorded at Cox Creek at Uraidla for several parameters, namely Total Nitrogen (TN), NO_x and Suspended Solids (SS).

Table 7 shows the main results for Tier 2 risk for six key parameters ranked by average risk. The results indicate the Cox Creek sites as highest in overall risk, and ranked highly across the six key parameters.

An assessment was also made of temporal trends in data to determine if there were any times of the year when water quality increased and to identify optimal times for water quality intervention. However, while there were increased concentrations for some parameters in June and July at some sites, there were no consistent temporal patterns or trends evident in any of the water quality parameters.

The relationships between flow and water quality parameters was also explored for similar reasons. At several sites an increase in the concentrations of TN and NO_x was found with corresponding increase in flow. The positive relationship between increasing flow and increasing TN and NO_x concentrations potentially has implications for water management at these sites. The extent of this relationship throughout the rest of the MLR needs to be investigated.

Water quality monitoring can be expensive. Pairwise correlation between water quality parameter were therefore assessed to determine if there were any consistent relationships that might indicate identify any parameters that could act as surrogates for others. This analysis found no consistent correlations within the sites/parameters analysed.

A sedimentation pond was installed at the Cox Creek at Uraidla site in 2006. Tier 1 and 2 analyses were run separately for data before installation and after to determine the effectiveness of the sedimentation pond as a mitigation strategy to improve water quality. Results indicate substantial improvement in all water quality parameters after installation of the sedimentation pond. Two parameters, SS and Zinc showed medians which dropped below the ANZECC guideline after installation of the sedimentation pond. The Tier 1 assessment showed that most ratios of median to ANZECC guideline more than halved

after 2006. However, despite the decrease in the median the majority of measurements still exceeded the ANZECC guidelines post-2006 and consequently the Tier 2 risk assessment showed little decrease.

Tier 3 of the risk assessment required data that had been continuously observed at a frequent resolution (i.e. per second, per minute or per day). The MLR water quality was not collected at a sufficiently high resolution to allow this analysis so Tier 3 of the risk assessment was not implemented.

Tier 4 of the risk assessment was designed to assess the interactions between water quality parameters (specifically NO_x, TN, TP and SS) and other catchment and environmental parameters, namely land use, soil type, monthly rainfall and monthly flow, using more advanced statistical spatio-temporal models. The results from the Tier 4 assessment identified:

- a positive correlation between flow and TN;
- a positive correlation between the land use category irrigated perennial horticulture and TN; and
- multiple soil types that negatively influence water quality.

Tier 5 of the risk assessment required ecological response functions for key endpoints, but these were unavailable and this was therefore not implemented.

In this risk assessment the analysis and models were limited by the high variability in the frequency of data collection across the MLR and the poor spatial coverage. Future monitoring schemes would benefit from being explicitly tied to management objectives, and focus on consistent, frequent data collection on fewer water quality parameters and ideally at more sites. The ranking of the current monitoring locations based on the risk assessment in this study will assist state agencies in the development of a more carefully designed and implemented monitoring scheme. Furthermore, a more carefully designed monitoring scheme would provide data that could be better utilised in the Tier 4 assessment to identify any correlations, temporal trends, or associations between water quality catchment attributes and environmental information.

1 Introduction

Water allocation planning provides for the allocation and use of water, and for the transfer of water allocations. It is essential to protect the economic, social and environmental needs for future generations and to provide secure and equitable access to water for all users.

The Mt Lofty Ranges (MLR) are important socially, economically and ecologically to South Australia (SA). The MLR catchments provide significant water resources and there are a range of stakeholders using the resource, including the general community, agricultural landholders, secondary industries and potable water suppliers and consumers.

The water resources of the MLR were prescribed in 2005 and a draft Water Allocation Plan (WAP) was released for the western MLR in 2010 (AMLRNRMB, 2010a,b) and for the eastern MLR in 2011 (SAMDBNRMB, 2011). In addition to the WAP in the MLR the SA government has identified the need for improved water quality in the catchments through the Water Quality Improvement program and the Water for Good policy.

In the current WAPs water quality is considered to be addressed when water volume is not limited i.e. water quality is not an issue if the flow regime is adequate. However, this may not necessarily be the case, particularly during certain times of the year under low flow regimes, which usually occur during summer and autumn (December to May) in the Mediterranean climate of the MLR. During periods of low flow there is the potential for increased concentrations of contaminants in water moving off-site into streams due to decreased dilution.

The factors impacting upon surface water run-off and transport of contaminants are complex and include times of travel or run-off, the degree of mixing during transport and effects of deposition and re-entrainment (Schriever and Liess, 2007; Schulz, 2004; Wauchope, 1978). Nutrient levels in overland flow have been found to show seasonal dependency. For example, in a study on a grassland in the UK, Kurz et al. (2005) observed elevated levels of Phosphorus during summer, thought to be caused by the accumulation of nutrients at soils surface during dry periods, and lower concentrations in winter, due to increased dilution in large volumes of overland flow. In the Mt Lofty Ranges, edge-of-field studies have shown that while associations between the total load of Suspended Solids (SS), Total Nitrogen (TN), and Total Phosphorus (TP) have been observed for cherry, grape and apple production systems, there were also occasions when the concentrations

remained high when flow declined (Cox et al., 2012).

Furthermore, the current WAPs propose that during low flow conditions water may be diverted from upstream storage locations, however when and where such a strategy is implemented requires an understanding of the water quality in the MLR. Specific questions around this strategy of low flow diversions and water quality include:

- Can mapping water quality through the MLR, both temporally and spatially, improve the process for selecting which sub-catchments (and when) water will be diverted from during periods of low flow?
- Is the health of aquatic ecosystems only threatened at periods of low flow? Are there other times of the year (e.g. after first flush events) when water diversions could improve water quality in order to minimise impact on aquatic ecosystems (and impact on other stakeholders such as providers of potable water supply, agriculturalists etc.)?
- Could improved water quality through water diversions be used to identify locations from which to procure water that is more fit for purpose, and thus lead to cost savings?

This study collated the MLR water quality data held by numerous agencies and interrogated the data using a tiered risk assessment process. This study expanded the work done earlier using Source catchment modelling through the eWater CRC (Thomas et al., 2010; Fleming et al., 2010). In the Source Catchment modelling, event mean concentration and dry weather concentration values were used to parameterise and validate the Source Catchment Model for TN, TP and SS. This risk assessment utilised all available water quality data from the MLR and used the Australian water quality guideline values for freshwater aquatic ecosystems (ANZECC/ARMCANZ, 2000) as a threshold value to identify exceedances. The ANZECC/ARMCANZ (ANZECC/ARMCANZ, 2000) Water Quality Guideline, hereafter referred to as ANZECC guidelines, values were used as the threshold because these trigger values have been derived for a wide range of stressors (contaminants) and inherent within the ANZECC guidelines is the understanding that if the trigger value is not exceeded the risk of an impact is low. Conversely if the trigger value is exceeded there is some risk of an adverse biological impact (ANZECC/ARMCANZ, 2000).

The tiered assessment approach used in this study, detailed in the Methods section and in Cox et al. (2013), identified sub-catchments where the ANZECC guideline values for specific water quality parameters were exceeded and provided a ranking of the monitoring

locations. This study also developed maps of risk of exceedance for certain water quality parameters, investigated temporal trends in the water quality data and investigated the relationships between flow and selected water quality parameters.

Specific outputs from this study include:

- A systematic, evidence-based approach for ranking sub-catchments in the MLR on the basis of water quality;
- Maps of risk of exceedance “hotspots” across the MLR;
- An assessment of temporal trends in water quality parameters;
- A determination of the relationship between flow and concentrations for selected water quality parameters;
- An assessment of the suitability of using certain water quality parameters as surrogates for other parameters;

Using advanced statistical and modelling tools, this report seeks to identify and establish relationships, if any, between contaminant concentrations and landscape attributes such as land use, soil type, topography and flow. It is expected that the catchments dominated by intensive agricultural land use relying heavily on fertilisation would have higher concentrations of nutrients and suspended solids than those dominated by native vegetation. The overland flow conditions, as impacted by topography and soil type, together with seasonality is expected to influence water quality due to dilutions and concentrations of contaminants migrating off-site. While Electrical Conductivity (EC) is expected to reflect flow conditions, the SS are expected to be related to certain nutrient concentrations such as Phosphorus. Such correlations may allow identification of surrogate parameters. The following correlations are known to occur:

- $SS = f(\text{flow})$ (Cox et al., 2012)
- High Nitrogen, Phosphorus in catchments with large percentage intensive land use
- $SS \sim TP$ if SS/TP is low
- $SS \sim TN$ if NO_x/TN is low

Cox et al. (2012) show the factors impacting upon surface water run-off and transport of contaminants are complex and include times of travel of run-off, the degree of mixing during transport and effects of deposition and re-entrainment. Edge-of-field studies in the MLR show a general trend of high SS concentrations at the commencement of flow for the

season that decreased on the receding limb of the hydrograph. Results from three land uses (apples, cherries and grapes) however were variable between run-off events. However, it is important to note that exceptions to these trends have been observed (Cox et al., 2012).

This risk assessment is expected to support policy decision making in the future by providing:

- a systematic, transparent, evidence-based process for determining the allocation of resources in the MLR that incorporates water quality in planning decisions which are fundamental to South Australian Government's Water for Good document and the South Australian EPA's Water Quality Improvement program; and
- a process for making informed decisions to target future monitoring in the MLR sub-catchments and for determining priority areas for investing in mitigation strategies based on a systematic risk ranking;
- a process for identifying which sub-catchments in the MLR would be suitable, and conversely which would not be suitable or present a water quality risk, for water diversion to provide water for environmental requirements during periods of low flow.

2 Materials and Methods

2.1 Water quality data collation in the Mt Lofty Ranges

Over many years there has been a substantial amount of water quality data collected in the MLR, measuring various parameters including basic physico-chemical variables such as turbidity, dissolved oxygen, EC, pH and temperature as well as more investigation-specific parameters such as dissolved organic carbon (DOC) (Nelson et al., 1990; Varcoe et al., 2010) and nutrient and pesticide concentrations (Oliver et al., 2012; Cox et al., 2012). This data is held by different agencies and in various publications. This project collated this data and, contingent on data availability and quality, analysed the data to identify locations and times in those regions when water quality becomes an issue.

The custodians of the original datasets are listed in Appendix A. While all water quality data was collated, the tiered risk assessment focussed only on those parameters where a guideline value was available from the ANZECC guidelines for protection of freshwater aquatic ecosystems (ANZECC/ARMCANZ, 2000). A 90 % protection level was chosen in consultation with the stakeholders, whilst considering the state of the ecosystems in the study area. In the absence of an ANZECC value for SA the upper NSW value was used. There is no ANZECC guideline value for DOC but it is an important water quality parameter for SA Water for reservoir management. SA Water provided two guideline values for DOC, namely 5 and 10 mg/L, for the risk assessment. The water quality parameters considered in this risk assessment and the ANZECC guideline values used are given in Table 1. Only those site/parameter combinations with more than 30 data points (in total across multiple years) were used in the tiered risk assessment. Due to data issues (e.g. infrequent measurements across sites and missing data imputation) none of the data provided by SA EPA was included in the risk assessment. Note: a guideline value of 8 has been used for pH. Thus the risk is calculated as the risk of exceeding 8.

2.2 Tiered risk assessment process

A tiered approach for risk assessment was employed in this study, starting with a screening level assessment in Tier 1 and then progressively increasing level of accuracy, detail, and complexity, in Tiers 2 and Tier 4, where the data allows. A brief outline is provided for each of the tiers. A more detailed description of various methods is provided in Appendix C. Only a subset of results are presented in this report, all results are provided in a separate document (Ford et al., 2015).

Table 1: ANZECC guidelines for freshwater aquatic ecosystems (ANZECC/ARMCANZ, 2000). Guideline values for DOC were provided by SA Water.

Water Quality Parameter	Guideline Value
Dissolved Organic Carbon	5 or 10 mg/L
EC Corrected / Conductivity	2200 uS/cm
Reactive Phosphorus (Filterable)	0.04 mg/L
Total Phosphorus	0.10 mg/L
NOx	0.10 mg/L
Total Nitrogen	1.0 mg/L
Copper	0.0018 mg/L
Lead	0.0056 mg/L
pH	8.0
Suspended Solids	50 mg/L
Turbidity	50 NTU
Zinc	0.015 mg/L

It is worth noting that Tier 1 and 2 use the upper range (90 %) of the level of protection identified by the ANZECC guidelines for making the assessment. While the risk of exceedance will change depending upon the cut-off value used, the ranking of the locations will not change and so the results can still be used for a comparative assessment.

2.2.1 Tier 1

This is the simplest assessment and forms the first stage of the tiered risk assessment approach. Tier 1 compares the median to the relevant water quality guideline for each of the water quality parameters assessed. The ratio of the median to ANZECC guideline is calculated for each parameter and site combination. This provides a first rank for each of the sites and catchments. In addition, an indicator function is used which returns the value 1 if the median is above the ANZECC guideline and 0 otherwise. For a low tier risk assessment, both of these approaches provide an adequate initial screening tool.

2.2.2 Tier 2

Tier 2 determines the distribution function of water quality parameters. EC data was collected, in most cases, almost daily. Due to strong auto-correlation, EC was analysed using time series analysis (see Appendix B).

The first stage of Tier 2 involves fitting univariate distributions to each water quality parameter (except EC) individually. A likelihood ratio test was used to determine which of the following distributions best represented the data: Normal, Lognormal, Weibull, and a 3-parameter log-logistic distribution. Uncertainty in the relative risk predictions was quantified in this Tier using a simple bootstrapping procedure. The probability of loss (i.e. risk) in this Tier is given by the area of the univariate density function that lies above the water quality guideline.

The average risk was calculated by taking the average Tier 2 risk of exceedance of all water quality parameters. The overall average risk was compared with and without metals (i.e. excluding metals from the calculation of average risk). An increase in average risk when metals were excluded indicated that the remaining water quality parameters were the key drivers for overall risk. Conversely, a decrease in average risk when metals were excluded from calculation of average indicated that metals were key drivers for overall risk.

Following this, correlations for all pairwise parameter combinations were investigated. Any pairwise correlations with Kendall's tau greater than 0.5 or less than -0.5 were modelled using copulas to investigate the nature of the joint dependence between the two parameters. A copula links univariate marginal distributions to their full multivariate distribution. All four copulas were fitted to the pairwise combinations of data and the best fit was selected using log-likelihood: the Gaussian copula (to explore no tail dependence); the Frank copula (to explore both upper and lower tail dependence); the Clayton copula (to explore lower tail dependence); and the Gumbel (to explore upper tail dependence). The Gumbel copula, of most interest here, models upper tail dependence. That is an increase in the conditional probability of extreme (high) events in one parameter given extreme events in the other.

2.2.3 Tier 3

Tier 3 of the risk assessment requires data that has been continuously observed in order to develop a concentration-time curve by accumulating the actual time that the pollutant concentration occurred within a specific short time period (e.g. per second, per minute or per day). This approach is only possible for pollutants that have been continuously observed, at a very high temporal resolution because the method calculates the accumulated exposure time above and below water quality guidelines. To do this it must “bin” the data into appropriate time steps (usually hours or days) and the bin width can be no smaller than the resolution (per second, per minute, per day) of the observations. The utility of this approach diminishes quickly as the resolution of the observation deteriorates (per week, per month, per year) and may not therefore be applicable to all pollutants in all locations.

The water quality data from MLR was not collected at a sufficiently high resolution to allow this analysis, so this Tier was not implemented.

2.2.4 Tier 4

Tier 4 requires developing and implementing more advanced statistical spatio-temporal models. These kinds of models, in particular hierarchical Bayesian models, have become increasingly popular as their computational complexity has been made more tractable by the efficiency of computers and statistical software.

Hierarchical Bayesian models are generally defined through three stages: modelling (1) the data, (2) the process and (3) the parameters through a system of equations. Using this approach, the model can capture the variation in measurements coming from seasonality (temporal dependency), location (spatial dependency) and external features (covariates). Three models, with incremental amount of information modelled, were used to fit the data (see Appendix C): the first, a general (distance-wise) spatial dependency only model; the second included temporal modelling, allowing for both seasonality and lag dependence; and the final full model included the land use and soil type information for each of the catchments, together with the monthly flow levels. Results from the final model are use in this report.

The temporal behaviour of the water quality parameters was modelled using the

assumption that the temporal variation comes from two sources: in the first the temporal variation of the covariates, such as flow and rainfall, is induced by the observations; in the second the natural variation of the water quality parameters is due to unobserved (or unexplainable) covariates, or variable influence of observed covariates. The variation in this second source is estimated by extracting a mean behaviour across the stations after removing the temporal effect due to the (time-varying) covariates. These two sources are then combined and modelled as one overall temporal behaviour.

The spatial model follows the same structure, with two separate effects. The covariates with spatial correlation include the flows and the rainfall. The additional spatial effect is modelled according to the Kriging method, which allows the user to provide a structure to the spatial correlation while assuming that close sites are more correlated than distant sites.

Overall, we have one model with two components. The first component is dedicated to the influence of space-time varying covariates. The second component is dedicated to the fixed covariates with possible time-varying influences. These components are simple linear models. However, due to the difference between the number of sites and the number of covariates (in particular for the land use and the soil type covariates) as well as the format of the observed covariates, we cannot use the data as is. A data transformation, described below, was used in order to overcome this.

The land use and soil type information were collected in the aerial compositional format. This means that for every sub-catchment, we have the composition of the land use and soil type in terms of proportion of the total surface. In other words, the observations for each sub-catchment for the land use should sum to one, likewise for soil type.

A statistical procedure, principal components analysis (PCA), was used to group a set of the land use categories. This approach uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables. For example, in Tier 4 analysis (on TN, TP, SS and RP) there are 17 sites with data and 28 different land use categories. With so

few sites, and with many land use covariates, we cannot infer the influence of the 28 land use categories as we will undoubtedly not observe some of the associated variance. PCA is an established way of helping reduce the dimensionality of the explanatory covariates in these types of situations. Here it was used to aggregate the land use types. This reduced the number of covariates to a manageable, but still interpretable, number. The PCA aggregations are presented in the results, but use the naming conventions in Table 2. Extensive detail of the methods used in Tier 4 is supplied in Appendix C.

Table 2: Land use categories and abbreviations used in Tier 4.

	Land use categories	Abbreviations
1	Grazing modified pastures	GMP
2	Cropping	Cr
3	Plantation forestry	PF
4	Transport and communication	TC
5	Nature conservation	NC
6	Residential	Res
7	Other minimal uses	OMU
8	Managed resource protection	MRP
9	Reservoir/dam	RD
10	Grazing natural vegetation	GNV
11	Irrigated cropping	IC
12	Irrigated perennial horticulture	IPH
13	Irrigated modified pastures	IMP
14	Intensive animal production	IAP
15	Perennial horticulture	PH
16	Irrigated plantation forestry	IPF
17	Seasonal horticulture	SH
18	Services	Ser
19	Irrigated seasonal horticulture	ISH
20	Manufacturing and industrial	MI
21	Intensive horticulture	IH
22	Irrigated land in transition	ILT
23	Marsh/wetland	MW
24	Utilities	Util
25	Mining	Min
26	Waste treatment and disposal	WTD
27	River	Riv
28	Lake	Lak

3 Results

3.1 Screening level assessment of water quality in various catchments: Tier 1

Across three agencies (NRMB, SA Water and DEWNR), there were 88 different sites, in 19 catchments that have been or are being monitored to a varying extent. The analyses for the tiered risk assessment approach was restricted to sites and water quality parameter combinations with more than 30 measurements (in total across multiple years). Of the 88 sites, 61 sites only had data for one water quality parameter: 57 with data for EC; and 4 with data for DOC. Furthermore, more than 60% of the sites were located within five catchments and half of the catchments are represented by less than 3 sites (see Table 3). There were seven sites without specified catchments which are not included in Table 3. Onkaparinga River catchment features frequently across the results as this was the catchment with the most sites, and data. Overall, the spatial spread of the MLR water quality data does not allow for complete assessment.

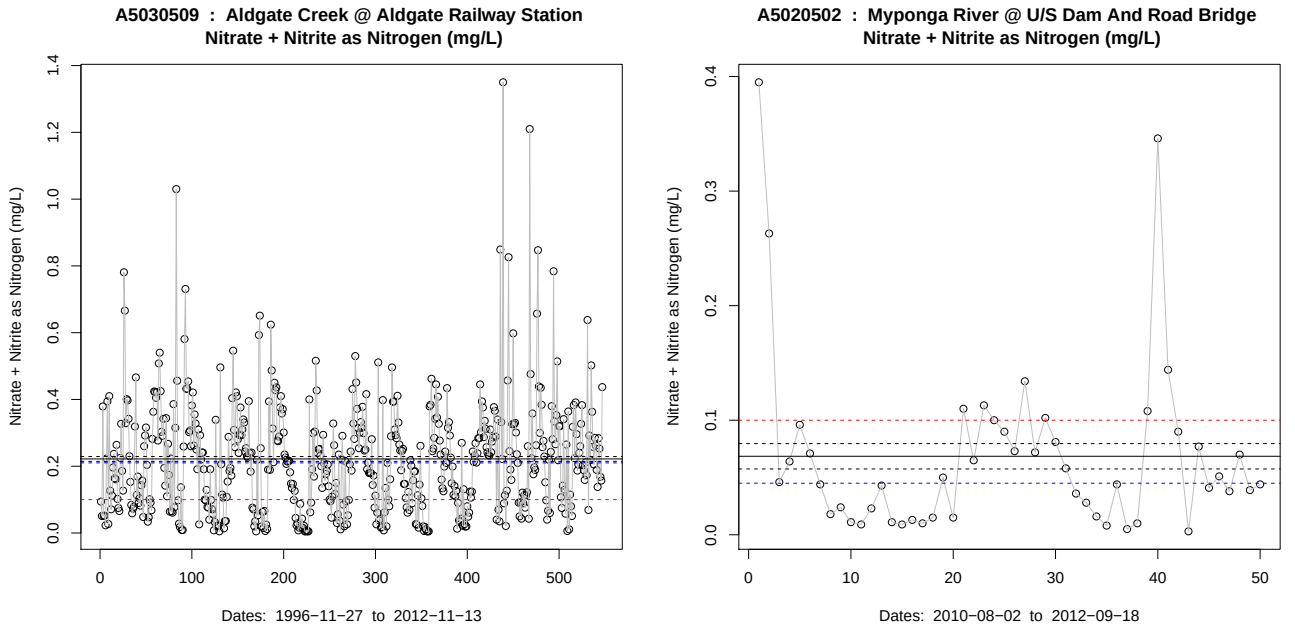
Table 3: Number of Sites per Catchment

Catchment	Number of sites
Onkaparinga River	14
Gawler River	13
Bremer River	11
Marne River	10
Torrens River	9
Angas River	4
Currency Creek	3
Finnis River	3
Saunders Creek	3
Deep Creek	2
Reedy Creek	2
Bungala River	1
Callawonga Creek	1
Hindmarsh River	1
Inman River	1
Myponga River	1
Tookayerta Creek	1
Yankalilla River	1

Tier 1 results provide an initial overview of the data, and the water quality of

the different sites and catchments. An example of the Tier 1 result is shown in Figure 1 for NO_x for Aldgate Creek (Figure 1a) and Myponga River (Figure 1b). The difference in overall medians and means and also quality of data is apparent with more than 15 years of data available for Aldgate Creek, versus only two years for Myponga River.

Table 4 shows the summary of the indicator variable (for median above or below the ANZECC guideline) used in Tier 1. Metal concentration measurements were recorded at 10 sites across the MLR. All of the sites recorded medians which exceeded the ANZECC guidelines for Copper, six out of 10 sites exceeded the ANZECC guideline for Zinc, and all were below the guideline for Lead. All four sites with data for DOC showed medians which exceeded the 5 mg/L guideline, but were below the 10 mg/L guideline. More than half the sites with data for TP and sites with data for NO_x recorded medians above the ANZECC guideline.



(a) Nitrate + Nitrite as Nitrogen at Aldgate Creek at Aldgate Railway Station (b) Nitrate + Nitrite as Nitrogen at Myponga River at U/S Dam and Road Bridge

Figure 1: Nitrate + Nitrite as Nitrogen (NO_x) for two sites with contrasting amount of data. Blue dashed line shows median, black solid and dashed lines show mean and standard error (respectively) and red line the ANZECC guideline.

Table 5 shows the number for all parameters, and all non-metal parameters (EC

Table 4: Number of Sites with median above/below the associated ANZECC water quality guideline

Parameter	Number above	Total Number
Copper (mg/L)	10	10
Dissolved Organic Carbon (mg/L) (5 mg/L)	4	4
Nitrate + Nitrite as Nitrogen (mg/L)	11	18
Zinc (mg/L)	6	10
EC Corrected/ Conductivity (uS/cm)	38	76
Total Nitrogen (mg/L)	4	9
Total Phosphorus (mg/L)	7	18
Reactive Phosphorus (mg/L)	2	12
pH	3	19
Dissolved Organic Carbon (mg/L) (10 mg/L)	0	4
Lead (mg/L)	0	10
Turbidity (NTU)	0	16
Suspended solids (mg/L)	0	18

excluded from all results in this table) exceeding the ANZECC guidelines, as a ratio of the total number of parameters (for all and non-metals respectively) measured for each site. The worst two sites had data only for pH. Cox Creek at Uraidla and Cox Creek u/s Brookes Road Bridge were the two worst sites in terms of number of parameters above the ANZECC guidelines. When metals were excluded both these sites had four of seven parameters exceeding the ANZECC guideline.

Table 5: Sites with number of water quality parameters (for all parameters, and for non-metals) above the associated ANZECC water quality guideline, as a ratio of parameters recorded at that site. EC is excluded from this table.

Catchment	Site ID	Site Name	#Above/ #Total (all)	#Above/ #Total (non- metals)
Bremer River	A4261173	Bremer R at Wanstead Rd	1/1	1/1
	A4261203	Lower Currency Ck	1/1	1/1
Onkaparinga River	A5030526	Cox Creek @ Uraidla	5/10	4/7
Onkaparinga River	A5031008	Cox Creek u/s Brookes Road Bridge	4/7	4/7
Onkaparinga River	1625	Clarendon Weir SP	1/2	1/2
Onkaparinga River	16250	Clarendon Weir Pump SP	1/2	1/2
Gawler River	7680	South Para Inlet Creek SP	1/2	1/2
Torrens River	79500	Gumeracha Forest Ford SP	1/2	1/2
Bungala River	A5011029	River Bungala u/s estuary	4/8	2/5
Torrens River	A5040508	Millbrook Res Intake Channel u/s Millbk Res	2/6	2/6
Onkaparinga River	A5030504	Onkaparinga R US Mt Bold	3/10	2/7
Onkaparinga River	A5031006	Cox Creek @ Woodhouse Wetland Inflow	2/7	2/7
Onkaparinga River	A5031007	Cox Creek @ Woodhouse Wetland Inflow	2/7	2/7
	A5051005	Smith Creek @ Womma Road	3/7	1/4
Hindmarsh River	A5011027	Hindmarsh River u/s estuary	3/8	1/5

Callawonga Creek	A5011030	Callawonga Ck U/S Mouth	1/5	1/5
Torrens River	A5040523	Sixth Creek Castambul	1/5	1/5
Torrens River	A5040525	Kersbrook Ck u/s Millbrook Reservoir	1/6	1/6
Onkaparinga River	A5030507	Lenswood Creek @ Lenswood	2/10	1/7
Onkaparinga River	A5030509	Aldgate Creek @ Aldgate Railway Station	3/10	1/7
Myponga River	A5020502	Myponga River @ U/S Dam And Road Bridge	0/6	0/6
Onkaparinga River	A5030502	Scott Creek @ Scott Bottom	2/10	0/7
Onkaparinga River	A5030506	Echunga Creek u/s Mt Bold Res.	2/8	0/5
Onkaparinga River	A5031005	Onkaparinga River U/S Estuary Old Noarlunga	1/8	0/5

3.2 Relative risk ranking of catchments: Tier 2

Tier 2 is methodologically more complex than the first tier and results in a more accurate quantification of the risk of exceeding the ANZECC guidelines for each of the water quality parameters at each site.

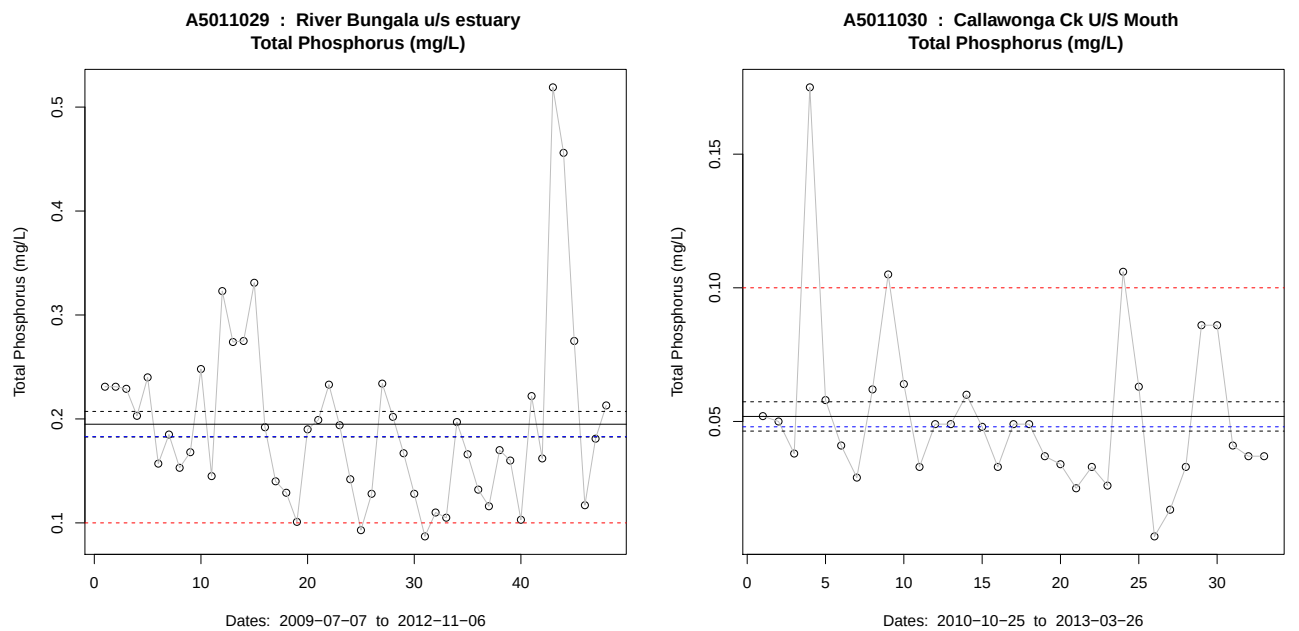
Table 6 shows data for all sites in Tier 2 (excluding sites with just EC data) including the total number of parameters analysed in Tier 2 for each site, and the average Tier 2 risk across these parameters, and the average risk excluding metals. This table ranks the sites in order of highest to lowest average risk of exceedance with metals excluded from average risk calculation. The relative change in average risk (calculated for the 10 sites with metals data) gives an indication of the percent increase or decrease in average risk when the metals are excluded. Of the ten sites with data for metals, two sites showed increased average risk when metals were excluded (see Table 6): Cox Creek at Uraidla showed a 15% increase in average risk when Copper, Lead and Zinc were excluded; and Onkaparinga R US Mt Bold a small 0.27% increase. In comparison, Onkaparinga River U/S Estuary Old Noarlunga showed more than a 50% reduction in average risk when metals were excluded. Bremer R at Wanstead Rd ranked highest for average risk but was based only on risk of exceedance in pH. Cox Creek u/s Brookes Road Bridge recorded the third highest average risk (0.6705), based on a Tier 2 risk assessment using seven water quality parameters.

Table 7 shows sites with data across the six key parameters: DOC, NO_x, Reactive Phosphorus (RP), SS, TN and TP. Sites are ranked by average Tier 2 risk. As shown in previous results, Cox Creek at Uraidla and Cox Creek u/s Brookes Road Bridge in the Onkaparinga catchment were ranked as the worst two sites when considering only these six key parameters.

Onkaparinga River catchment ranked high for risk of exceedance across all parameters; with Cox Creek at Uraidla the site with highest risk of exceedance in several of the water quality parameters (see Table 7). This site was in the top five sites, ranked by risk of exceedance, for all but Zinc, pH and Copper (where it had the lowest risk across all sites). It ranked third highest across the sites when ranked by average Tier 2 risk excluding metals (see Table 6). Bremer R at Wanstead Rd

and Lower Currency Ck ranked higher, however, both these sites only had data for pH. Cox Creek u/s Brooks Road Bridge ranked fourth. The four sites in Table 6 with two parameters in Tier 2 (site IDs 79500, 16250, 7680 and 1625) had data only for DOC with the average risk based on both the DOC 5 and 10 mg/L limits.

Figures 2a and b show the data for the two sites with highest and lowest Tier 2 risk for Total Phosphorus respectively. River Bungala, with the highest risk (0.86), had only two points below the ANZECC guideline across all years of recording; whereas Callawonga Ck U/S Mouth (0.052 Tier 2 risk) had only three measurements greater than the ANZECC guideline.



(a) Total Phosphorus measurements for River Bungala (b) Total Phosphorus measurements for Callawonga Creek

Figure 2: Total Phosphorus for two sites with contrasting water quality exceedance risk. The blue line shows median, the black lines show mean (solid) and standard error (dashed) and red line the ANZECC guideline.

Table 6: Sites with total number of water quality parameters in Tier 2 Risk assessment, average Tier 2 risk for all parameters with and without metals, and relative change in average Tier 2 risk when metals are excluded. The sites are listed from highest to lowest rank of average risk of exceedance excluding metals.

Catchment	Site ID	Site Name	No. param in Tier 2	Average Risk	Average Risk (excl metals)	Relative change in avg risk
Bremer River	A4261173	Bremer R at Wanstead Rd	1	0.9127	0.9127	
	A4261203	Lower Currency Ck	1	0.7430	0.7430	
Onkaparinga River	A5030526	Cox Creek @ Uraidla	10	0.6175	0.7075	0.1457
Onkaparinga River	A5031008	Cox Creek u/s Brookes Road Bridge	7	0.6705	0.6705	
Torrens River	79500	Gumeracha Forest Ford SP	2	0.6291	0.6291	
Onkaparinga River	16250	Clarendon Weir Pump SP	2	0.6126	0.6126	
Gawler River	7680	South Para Inlet Creek SP	2	0.5750	0.5750	
Bungala River	A5011029	River Bungala u/s estuary	8	0.5670	0.5458	-0.0374
Onkaparinga River	A5031007	Cox Creek @ Woodhouse Wetland Inflow	7	0.5091	0.5091	
	A5051005	Smith Creek @ Womma Road	7	0.5155	0.4720	-0.0844
Hindmarsh River	A5011027	Hindmarsh River u/s estuary	8	0.4464	0.4423	-0.0092
Onkaparinga River	A5031006	Cox Creek @ Woodhouse Wetland Inflow	7	0.4350	0.4350	
Onkaparinga River	A5030504	Onkaparinga R US Mt Bold	10	0.3860	0.3870	0.0026
Torrens River	A5040523	Sixth Creek Castambul	5	0.3785	0.3785	
Onkaparinga River	A5030507	Lenswood Creek @ Lenswood	9	0.3958	0.3653	-0.0771
Onkaparinga River	1625	Clarendon Weir SP	2	0.3619	0.3619	
Torrens River	A5040508	Millbrook Res Intake Channel u/s Millbk Res	6	0.3599	0.3599	
Onkaparinga River	A5030509	Aldgate Creek @ Aldgate Railway Station	10	0.4479	0.3566	-0.2038
Torrens River	A5040525	Kersbrook Ck u/s Millbrook Reservoir	6	0.2545	0.2545	
Callawonga Creek	A5011030	Callawonga Ck U/S Mouth	5	0.1842	0.1842	
Onkaparinga River	A5030502	Scott Creek @ Scott Bottom	8	0.3218	0.1755	-0.4546
Onkaparinga River	A5030506	Echunga Creek u/s Mt Bold Res.	7	0.3219	0.1529	-0.525
Myponga River	A5020502	Myponga River @ U/S Dam And Road Bridge	6	0.1426	0.1426	
Onkaparinga River	A5031005	Onkaparinga River U/S Estuary Old Noarlunga	8	0.2263	0.1031	-0.5444

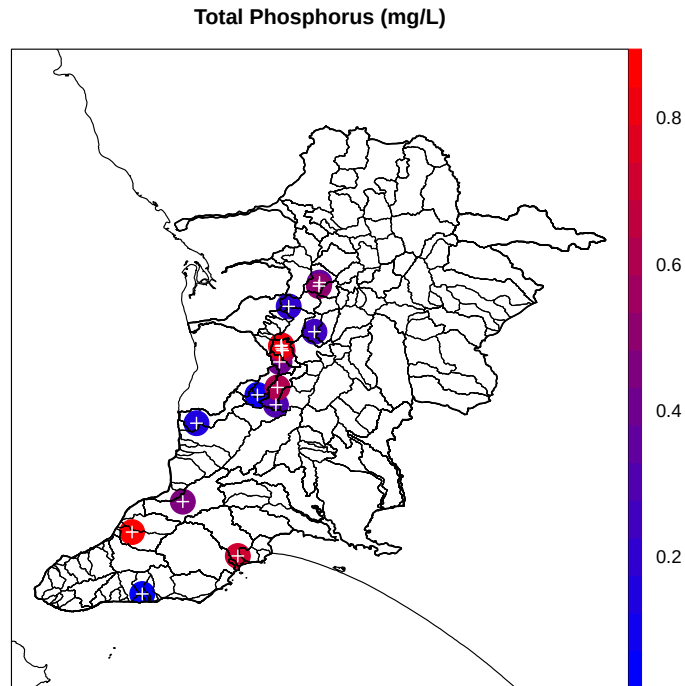
Table 7: Sites with Tier 2 Risk (and rank within parameter) for 6 key parameters, and average risk across parameters. Sites are ranked by average risk.

Catchment	Site ID	Site Name	DOC (5 mg/L)	DOC (10 mg/L)	NOx	Reactive P	SS	TN	TP	Avg Risk
Onkaparinga River	A5030526	Cox Creek @ Uraidla			1(1)	0.85(2)	0.6(1)	0.89(1)	0.84(2)	0.835
Onkaparinga River	A5031008	Cox Creek u/s Brookes Road Bridge			0.86(3)	0.88(1)	0.59(2)	0.87(2)	0.83(3)	0.805
Torrens River	79500	Gumeracha Forest Ford SP	0.82(2)	0.44(1)						0.629
Onkaparinga River	16250	Clarendon Weir Pump SP	0.98(1)	0.24(3)						0.613
Bungala River	A5011029	River Bungala u/s estuary			0.6(10)		0.36(7)		0.86(1)	0.606
Onkaparinga River	A5031007	Cox Creek @ Woodhouse Wetland Inflow			0.87(2)	0.47(3)	0.32(9)	0.7(3)	0.59(6)	0.591
Gawler River	7680	South Para Inlet Creek SP	0.81(3)	0.34(2)						0.575
	A5051005	Smith Creek @ Womma Road			0.57(12)		0.51(3)		0.59(5)	0.555
Onkaparinga River	A5031006	Cox Creek @ Woodhouse Wetland Inflow			0.73(4)	0.32(4)	0.22(12)	0.65(4)	0.51(8)	0.487
Hindmarsh River	A5011027	Hindmarsh River u/s estuary			0.26(14)		0.41(4)		0.74(4)	0.466
Torrens River	A5040508	Millbrook Res Intake Channel u/s Millbk Res			0.71(6)	0.25(5)	0.25(11)		0.51(9)	0.427
Onkaparinga River	A5030504	Onkaparinga R US Mt Bold			0.68(8)	0.21(6)	0.3(10)	0.36(7)	0.58(7)	0.425
Onkaparinga River	A5030509	Aldgate Creek @ Aldgate Railway Station			0.72(5)	0.05(9)	0.39(6)	0.5(5)	0.37(11)	0.406
Onkaparinga River	A5030507	Lenswood Creek @ Lenswood			0.59(11)		0.4(5)	0.41(6)	0.11(14)	0.378
Torrens River	A5040523	Sixth Creek Castambul			0.69(7)		0.34(8)		0.1(15)	0.376
Onkaparinga River	1625	Clarendon Weir SP	0.72(4)	0.01(4)						0.362
Torrens River	A5040525	Kersbrook Ck u/s Millbrook Reservoir			0.66(9)	0.07(8)	0.2(13)		0.19(12)	0.279
Myponga River	A5020502	Myponga River @ U/S Dam And Road Bridge			0.21(16)	0.1(7)	0(17)		0.39(10)	0.173
Onkaparinga River	A5030502	Scott Creek @ Scott Bottom			0.17(17)			0.24(9)	0.06(17)	0.158
Onkaparinga River	A5030506	Echunga Creek u/s Mt Bold Res.			0.1(18)		0.07(15)	0.32(8)	0.12(13)	0.153
Callawonga Creek	A5011030	Callawonga Ck U/S Mouth			0.23(15)		0.15(14)		0.05(18)	0.143
Onkaparinga River	A5031005	Onkaparinga River U/S Estuary Old Noarlunga			0.28(13)		0.01(16)		0.08(16)	0.120

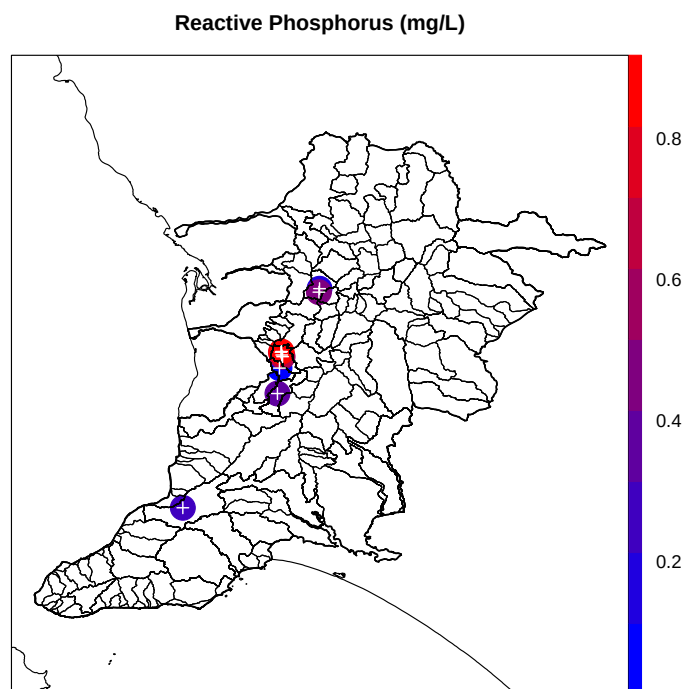
3.3 Water quality maps of risk of exceedance hotspots: Tier 2

Tier 2 risks were used to create maps of risk of exceedance hotspots across the MLR. The very poor spatial balance in the MLR data, however, prevents us from providing reliable interpolated risk surfaces. Previous attempts at simple interpolations provided very misleading results, and all further spatial analysis was restricted to the Tier 4 assessment. Risk maps for Tier 2 are therefore restricted to point-wise representations of risk of exceeding ANZECC guidelines. Examples of exceedance hotspots for TP, RP, TN, NO_x and SS have been presented below.

Figure 3a indicates the areas at high risk of exceeding the ANZECC guidelines for TP (high risk of exceedance is indicated by red and low risk by blue). Three major hotspots can be seen in Figure 3a. These hotspots are associated with the three highest rank catchments when ranked by Tier 2 risk of exceedance: one site in Bungala River catchment ranked highest (0.86 risk of exceedance); followed by two sites in the Onkaparinga catchment; and one site located in Hindmarsh River catchment. Figure 3b shows there was less data for RP, however high risk of exceedance can still be seen in the Onkaparinga catchment.



(a) Map for Total Phosphorus

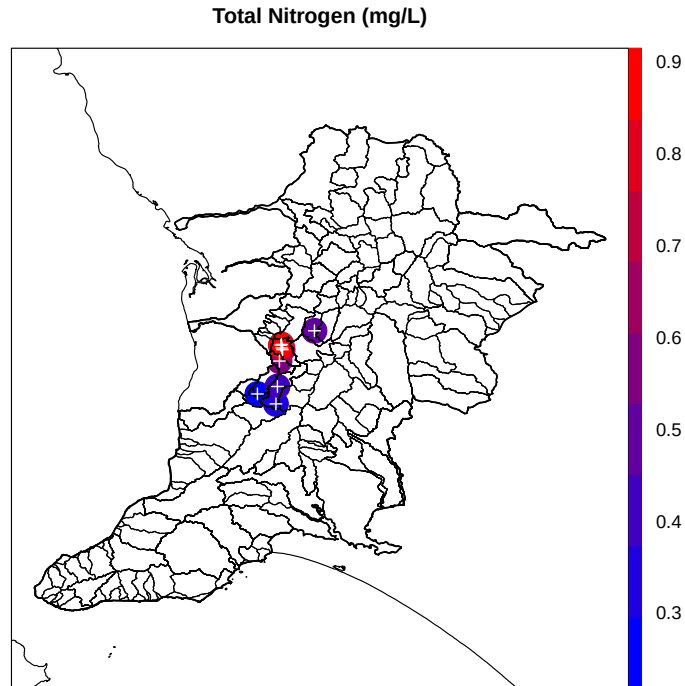


(b) Map for Reactive Phosphorus

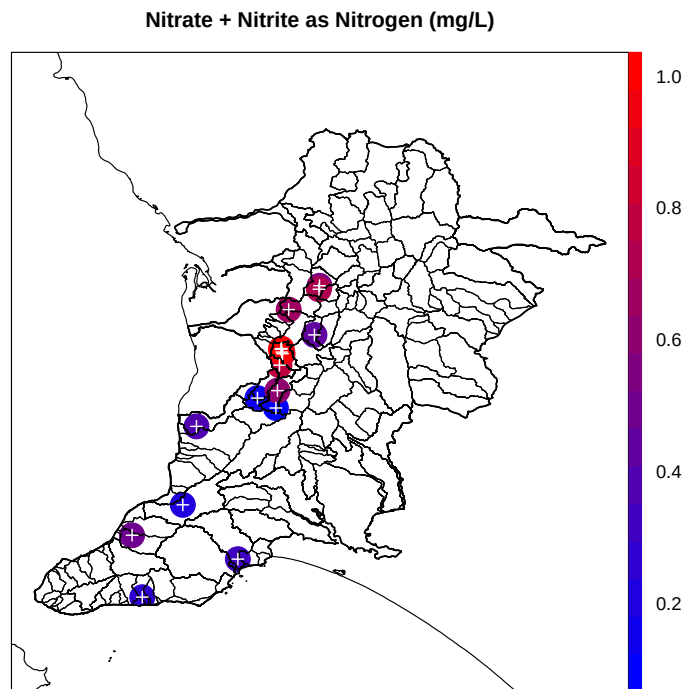
Figure 3: Visual representation of sites with risk of exceedance for Total Phosphorus and Reactive Phosphorus, increasing from low (blue) to highest (red) as determined in the Tier 2 assessment.

The map for TN (see Figure 4a) Tier 2 risk of exceedance indicates that the data was limited to sites within the Onkaparinga River catchment. However, even within this one catchment there was a substantial range of Tier 2 risk among the nine sites: Cox Creek at Uraidla recorded the highest Tier 2 risk (0.89) and Scott Creek at Scott Bottom the lowest (0.24).

There were 18 sites across six catchments with data for NO_x (Figure 4b). Sites from Onkaparinga catchment recorded the top five highest Tier 2 risk, followed by two sites from Torrens River catchment. Cox Creek at Uraidla again recorded the highest Tier 2 risk (1.00), and Echunga Creek u/s Mt Bold Res. the lowest (0.10).



(a) Map for Total Nitrogen



(b) Map Nitrate + Nitrite as Nitrogen

Figure 4: Visual representation of sites with risk of exceedance for Total Nitrogen and Nitrate + Nitrite as Nitrogen, increasing from low (blue) to highest (red) as determined in the Tier 2 assessment.

Figure 5 shows hotspots for high risk of exceedance for SS around Onkaparinga River and Hindmarsh River catchments. Cox Creek at Uraidla again ranked highest for risk of exceedance (0.60), followed by Cox Creek u/s Brookes Road Bridge (0.59), Smith Creek at Womma Road (0.51) and Hindmarsh River u/s estuary (0.41). Myponga River catchment recorded the lowest risk of exceeding ANZECC guidelines for SS ($1.35\text{E-}6$ at Myponga River at U/S Dam and Road Bridge).

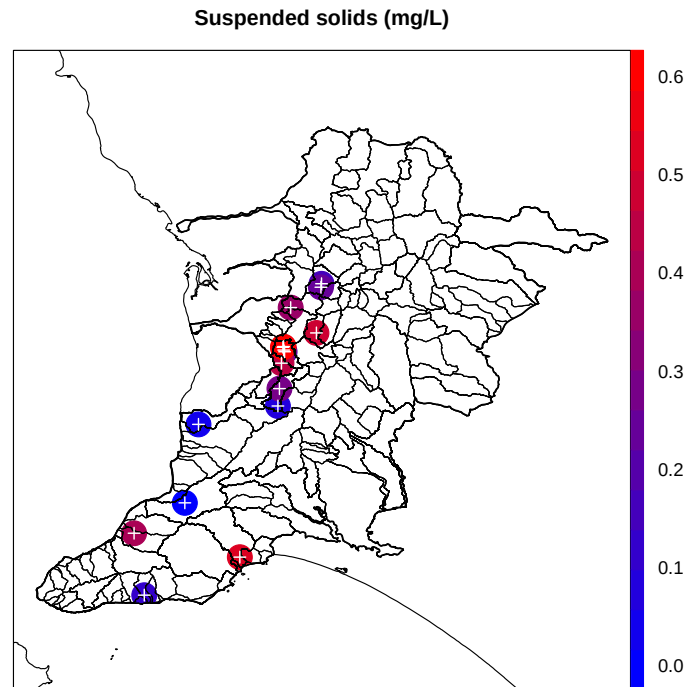
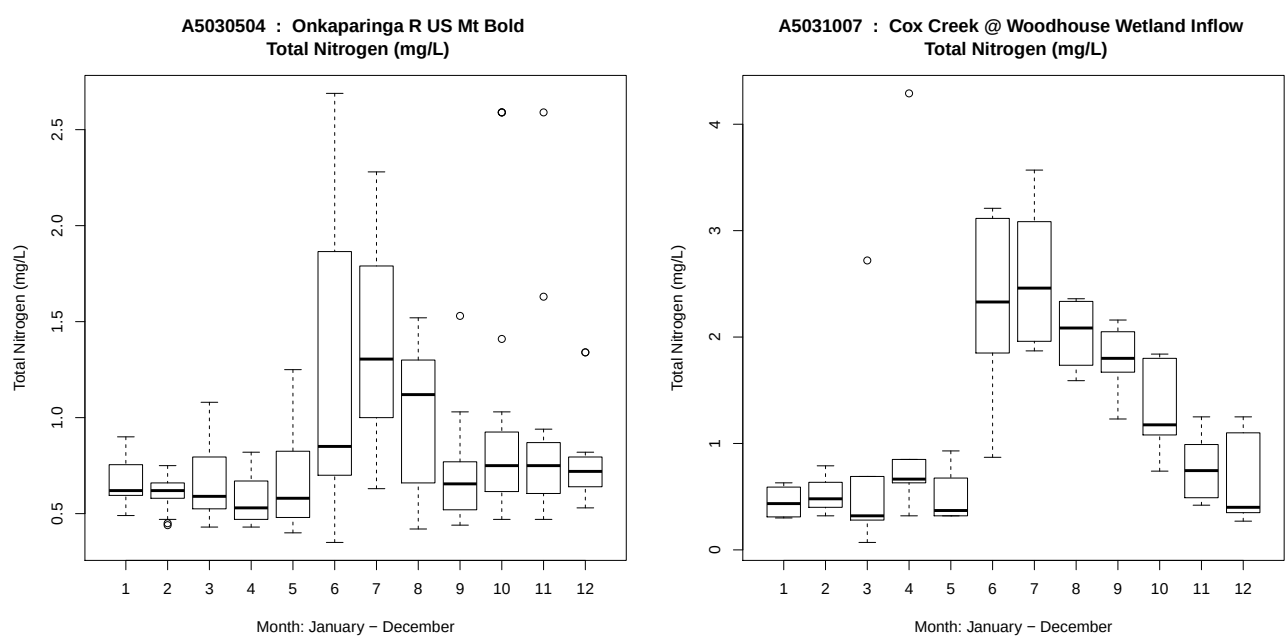


Figure 5: Visual representation of sites with risk of exceedance for Suspended Solids, increasing from low (blue) to highest (red) as determined in the Tier 2 assessment.

3.4 Temporal trends in water quality parameters: Tier 2

Investigation of temporal trends in the water quality parameters was additional to the original project plan and were investigated in order to identify any consistent annual patterns across sites and parameters. Boxplots by month for each of the pollutants (with the exception of EC) across all sites indicate no consistent temporal trends or peaks in data. EC data was collected consistently across all sites, and temporal trends were evident for most sites, as expected due to flow variations during the year.

A few sites showed some indication of increase in levels of TN and NO_x in June and July (Onkaparinga R US Mt Bold, Aldgate Creek @ Aldgate Railway Station Onkaparinga River, Cox Creek @ Woodhouse Wetland Inflow, Echunga Creek u/s Mt Bold Res., Cox Creek @ Woodhouse Wetland Inflow, Cox Creek u/s Brookes Road Bridge, Lenswood Creek @ Lenswood Onkaparinga River). However, this was not consistent across all sites or catchments. Figure 6 shows TN, by month, for two sites: the first five months show consistent stable levels of TN, with an increase in June and July, and dropping back down from August through to December. Two other sites showed similar peaks in July for three other parameters: Callawonga Creek U/S Mouth showed increase in Turbidity, TP and NO_x in July; likewise Sixth Creek Castambul showed increase Turbidity, TP, NO_x and also SS in July.



(a) Box plots of monthly Total Nitrogen for Onkaparinga Site (b) Box plots of monthly Total Nitrogen for Cox Creek Site

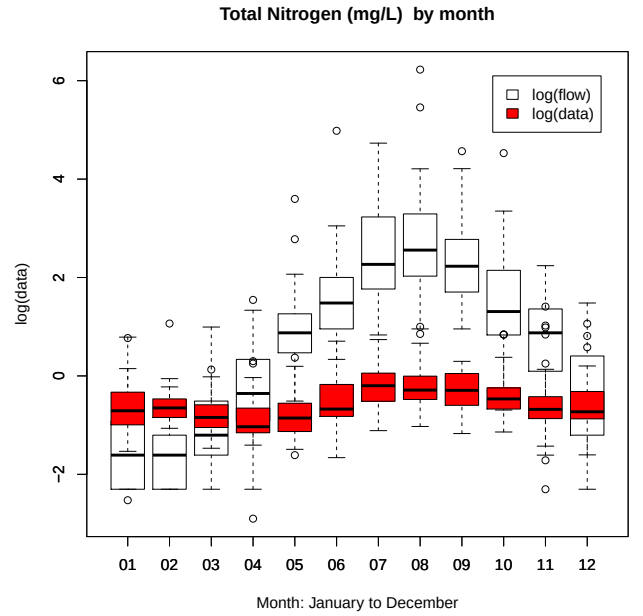
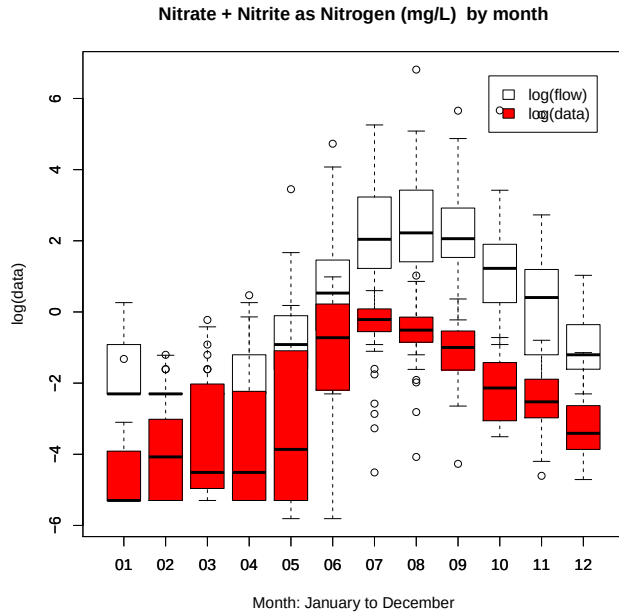
Figure 6: Box plots of monthly Total Nitrogen for two sites. Any values greater than 1.5 times the interquartile range are represented by a 'o'.

3.5 Effect of flow conditions on water quality

Analysis of flow data was additional to the original project plan and was considered for 12 sites across the MLR for which flow data was obtained. Box plots by month and year, and density plots were used to identify relationships between flow and other water quality parameters. Results were inconsistent across the sites: some sites showed no relationship between flow and any water quality parameters; other sites showed some indication of water quality parameters following increases and decreases in flow; others showed a decrease in water quality parameters with an increase in flow. Several sites showed inverse relationship between EC and flow, as expected due to dilution effect.

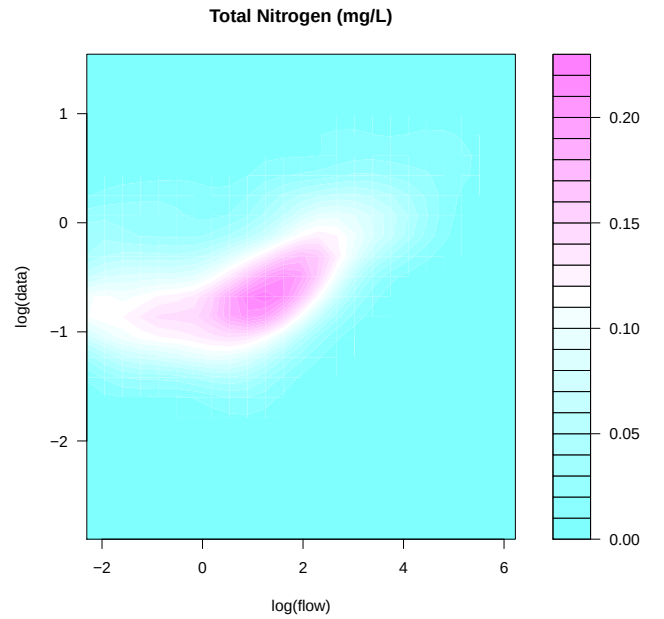
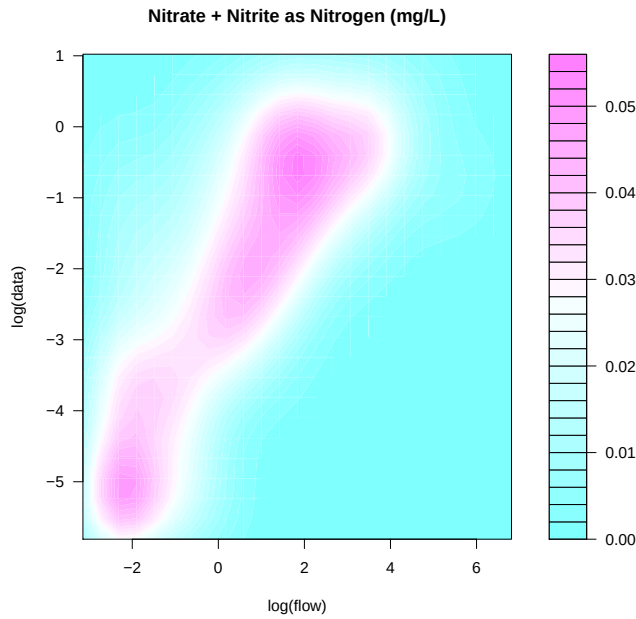
At several sites there was a corresponding change in some water quality parameters with increasing or decreasing flow, however these patterns were inconsistent due to limited amount of data for the water quality parameter. NOx showed a stronger pattern than most water quality parameters across several sites (see for example Figure 7a). The density plot indicates an increase in NOx with an increase in flow (Figure 7c). A similar pattern was observed in several other sites (Scott Creek at Scott Bottom, Echunga Creek u/s Mt Bold Res. and Aldgate Creek at Aldgate Railway Station). A similar pattern was observed for TN at site Scott Creek at Scott Bottom (Figure 7b). Figure 8a and b shows means (across all years) for all sites for flow and NOx, and flow and TN respectively. These figures suggest that across the monitored locations there is a positive relationship between flow and NOx and flow and TN. This conclusion is also borne out by the Tier 4 analysis (see Section 3.8). The positive relationship between increasing flow and increasing TN and NOx concentrations potentially has implications for water management at these sites. The extent of this relationship throughout the rest of the MLR needs to be investigated.

In some sites, some water quality parameters showed an inverse relationship with flow - that is a decrease in the water quality parameter with an increase in flow. This was seen for EC and metals across most sites, SS at Scott Creek at Scott Bottom and RP at site Cox Creek at Uraidla.



(a) Monthly log(Flow) and log(NOx) for Lenswood Creek at Lenswood

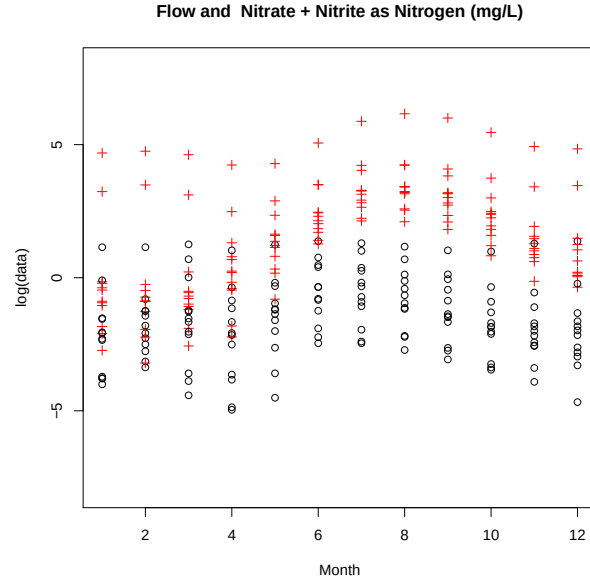
(b) Monthly log(Flow) and log(TN) for Scott Creek at Scott Bottom



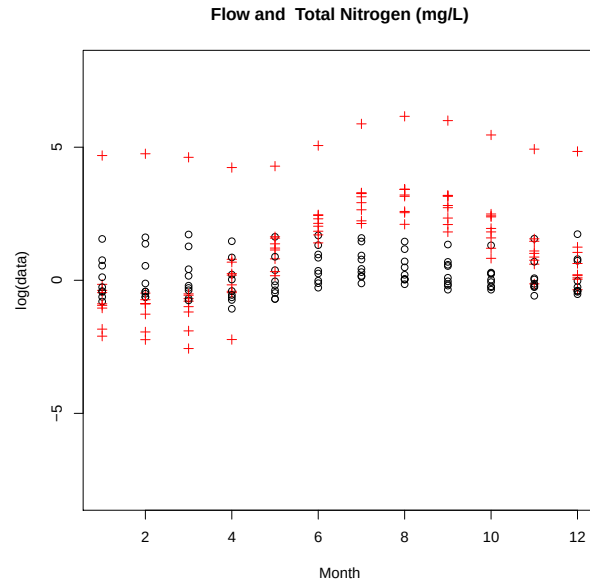
(c) Kernel density plot for log(Flow) and log(NOx) for Site Lenswood Creek at Lenswood

(d) Kernel density plot for log(Flow) and log(TN) for Scott Creek at Scott Bottom

Figure 7: Monthly box plots and kernel density plots for log(flow) with log(NOx) and log(Total Nitrogen) at two sites.



(a) Monthly $\log(\text{Flow})$ and $\log(\text{NOx})$ for all sites with data for both parameters



(b) Monthly $\log(\text{Flow})$ and $\log(\text{TN})$ for all sites with data for both parameters

Figure 8: Monthly means (across all years) for $\log(\text{flow})$ and $\log(\text{NOx})$ and $\log(\text{Total Nitrogen})$ for all sites. Red '+' indicates means for flow for a site, and black 'o' indicates mean for NOx and TN for a site.

3.6 Correlation between water quality parameters: Tier 2

Pairwise correlations between water quality parameters were also assessed as part of Tier 2. In addition the correlations were used to identify whether there were any consistent surrogate measurements, i.e. could one water quality parameter be used in place of another? This could be desirable from the perspective of cost and ease of measurement.

Pollutant combinations with pairwise correlations with Kendall's tau greater than 0.5 or less than -0.5 were modelled using copulas (see Section 2.2.2). Parameter combinations showing dependency with the Gumbel copula are of most interest here because this implies upper tail dependency i.e. stronger dependence between the parameters at high levels above their respective means. There were 27 strong pairwise correlations modelled with copulas; 15 of which were best modelled using the Gumbel copula (see Table 8). Figure 9 shows the Gumbel copula for two sites showing upper tail dependence. Sixth Creek Castambul (site ID A5040523) showed upper tail dependence for SS and TP; and Cox Creek @ Woodhouse Wetland Inflow (site ID A5031007) showed tail dependence for RP and SS. This means (for example) that the conditional probability of SS reaching levels in excess of the ANZECC guideline increases as the concentrations of TP increase.

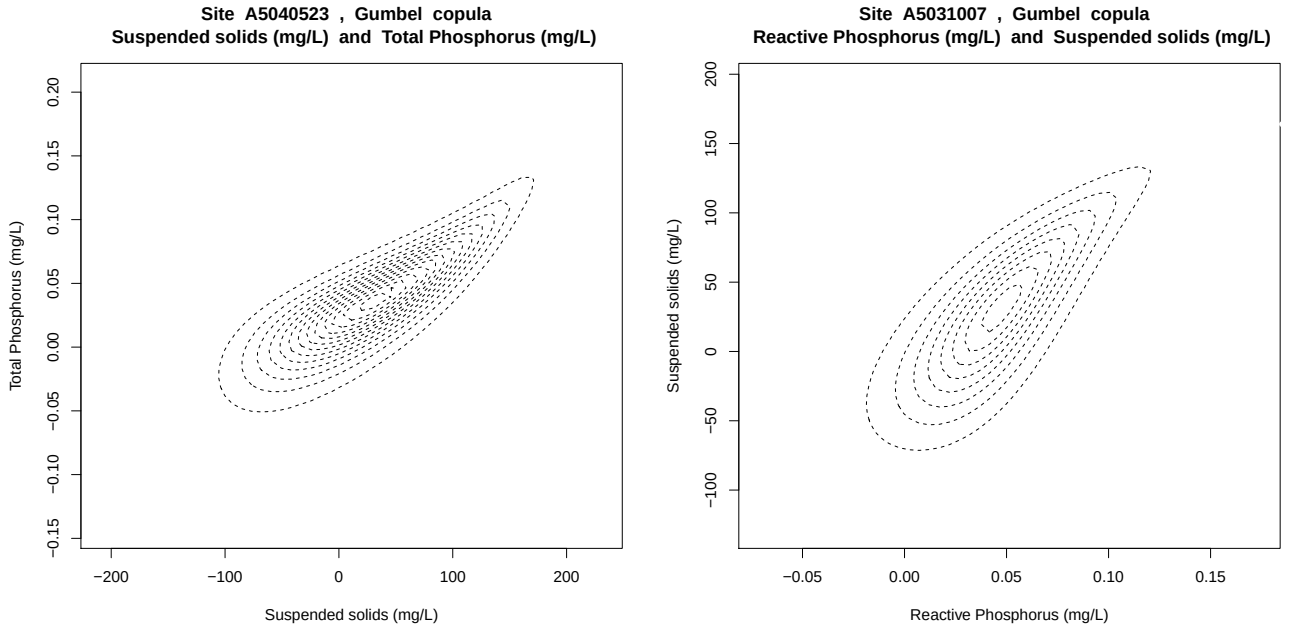
Across all sites there were several frequent pairwise parameter combinations that showed strong (Kendall's tau) correlations, namely:

- TP and RP showed strong correlations across five sites;
- TN and NO_x at four sites;
- Copper and Zinc was the most frequent, with strong correlation in eight sites;
- SS and TP in five sites;
- the other two combinations (SS and Turbidity; TP and Turbidity) appeared in only one site.

Overall, however, the results do not point to a consistent synergy across the MLR that would allow one parameter to be used as a surrogate for another.

Table 8: Summary of copulas for sites with pairwise parameter combinations (split by '/') where Kendall tau was greater than 0.5 or less than -0.5. Abbreviations for parameters are used for clarity: Cu = Copper; NO_x = Nitrate + Nitrite as Nitrogen; RP= Reactive Phosphorus; SS = Suspended Solids; Tb = Turbidity; TN = Total Nitrogen; TP = Total Phosphorus; Zn = Zinc.

Site	Gumbel	Clayton	Gaussian	Frank
A5011027	SS/Tb; SS/TP; Tb/TP			
A5040523	SS/TP			
A5020502	SS/TP			
A5030504	TP/RP		TP/NO _x	Cu/Zn
A5030526	RP/TP			NO _x /TN; Cu/Zn
A5031007	RP/TP; RP/SS; TP/SS		NO _x /TN	
A5031008	TP/SS		NO _x /TN	RP/TP
A5030507	TP/TN			
A5031006		NO _x /TN		
A5040525				TP/RP
A5031005	Cu/Zn			
A5051005	Cu/Zn			
A5030509	Cu/Zn			
A5030502			Cu/Zn	
A5030506			Cu/Zn	
A5011029			Cu/Zn	



(a) Gumbel Copula for Suspended Solids and Total Phosphorus for Sixth Creek Castambul (site ID A5040523) (b) Gumbel Copula for Reactive Phosphorus and Suspended Solids for Cox Creek @ Woodhouse Wetland Inflow (site ID A5031007)

Figure 9: Two Copulas modelling the joint correlation between two parameters for two different sites. The Gumbel copula is used to model strong upper tail dependence: the stronger the correlation contour, the stronger the tail dependence. The “narrowing” of the contours in the upper right quadrant indicates that the dependence between the two parameters becomes stronger as they become more extreme.

3.7 Effectiveness of sedimentation pond at Uraidla, Cox Creek

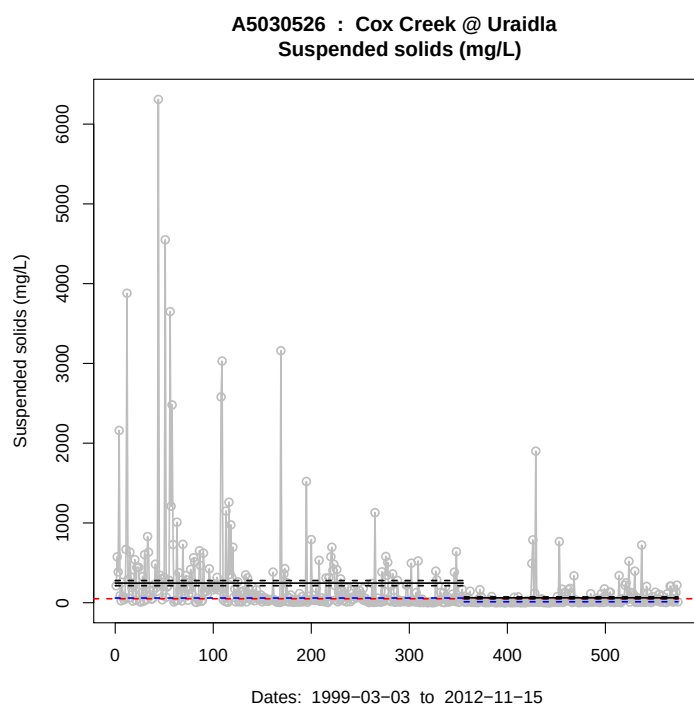
A sedimentation pond was installed in Cox Creek at Uraidla site in 2006 in order to decrease sediment load. In order to investigate any potential impacts of this sedimentation pond the Tier 1 and 2 assessments were repeated using the data split into that collected before 2006 and that collected since 2006.

Overall the results indicate a marked improvement across all water quality parameters after installation of the sedimentation pond. Despite the decrease in the median after installation, most water quality parameters that exceeded the ANZECC guideline pre-2006 remained above the guideline after the installation of the sedimentation pond (i.e. post 2006). However, both Zinc and SS showed medians which dropped below the ANZECC guideline after installation of the sedimentation pond (see Figure 10). The median SS before installation was 59 mg/L, and this dropped to 12 mg/L after installation of the sedimentation pond (mean from 245 mg/L to 61 mg/L).

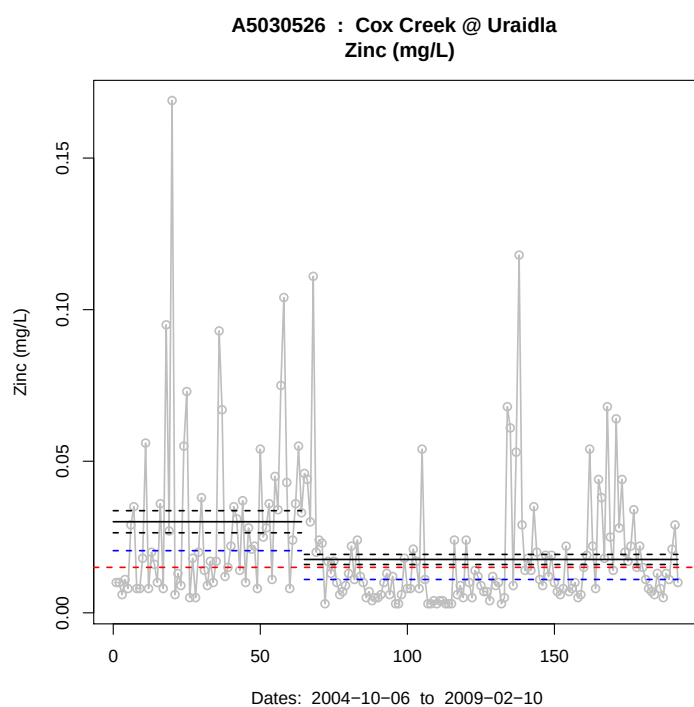
There was a consistent drop in the median for the water quality parameters, after installation of the sedimentation pond, however, the overall Tier 2 risk reduction is often small. The Tier 2 risks, presented in Table 9, show minimal decrease before and after 2006. Figure 10a indicates a marked drop in median TP before and after 2006, however the majority of recordings after 2006 are still above the ANZECC guideline, and as such there was minimal overall reduction in Tier 2 risk (from 0.73 to 0.72).

Table 9: Tier 2 Risk for Cox Creek, before and after installation (2006) of sedimentation pond. Note that pH and Turbidity were not recorded prior to 2006 so are not included in the table.

Parameter	Before 2006	After 2006
Nitrate + Nitrite as Nitrogen (mg/L)	1.00	0.90
Total Nitrogen (mg/L)	0.93	0.83
Total Phosphorus (mg/L)	0.73	0.72
Reactive Phosphorus (mg/L)	0.92	0.83
Copper (mg/L)	0.88	0.66
Lead (mg/L)	0.18	0.07
Zinc (mg/L)	0.62	0.37
Suspended solids (mg/L)	0.63	0.53



(a) Suspended Solids



(b) Zinc

Figure 10: Data for Suspended Solids and Zinc for Cox Creek at Uraidla. Split lines indicate medians (blue dashed line), means and standard errors (black solid and dashed lines respectively) for data recorded before and after the installation of the sedimentation pond in 2006. The red dashed line indicates ANZECC guideline.

3.8 Catchment characteristics and attributes associated with water quality parameters: Tier 4

The spatio-temporal model used in Tier 4 of the risk assessment examines the temporal patterns of the measurements, the spatial correlation between sites located nearby, and the effect of catchment related features on water quality parameters. The Tier 4 model was designed to investigate any relationships between the catchment features and the water quality data.

3.8.1 Key catchment characteristics, attributes and hotspots

The analysis was performed on data from 17 sites (except for TN, which was only recorded at nine sites), with a time-spread of the data from 1999 to 2010, for four water quality parameters: NO_x, TN, TP and SS. A subset of results for TN are presented here. All results are available in Appendix C. Because the number of land use categories (an important variable thought to influence water quality) far exceeds the number of sites, we used four principal components of the proportion of land use categories in each catchment in the model. For similar reasons, four soil type principal components, monthly rainfall, and monthly flow levels were also included in the model.

Figure 11 displays the predictions versus the observations, for each site for TN. The figure is a summary, as every measurement (and prediction) is time varying, but gives an indication of the quality of the model prediction depending on multiple factors. Figure 11 highlights that out of the nine sites being monitored and modelled, the model consistently under-estimates the concentration of water quality parameters for four sites: one from Lenswood Creek, one from Scott Creek, and two from Cox Creek (Woodhouse Wetland Outlet and Woodhouse Wetland Inflow). The error bars show reasonable overall magnitude, indicating that the model captures well the variation of the measurements.

The measurements and predictions for the Woodhouse Wetland Outlet site (Figure 12a) indicate that only a few measurements are available, from 2007 to 2010. The results indicate that the timing of spikes is captured, but the magnitude of the concentration peak is missed. The results in Figure 12a identify two important points. First, the lack of accuracy of the model for the Woodhouse Wetland Outlet site may be due to a lack of data, both in terms of water quality measurements and catchment features. However, given that the model is performing quite well on the other sites (see Appendix C), this failure indicates a different behaviour for this site. Secondly, one particular misfit of the data is of concern: in the case of the Woodhouse Wetland Inflow station (see Figure 12b),

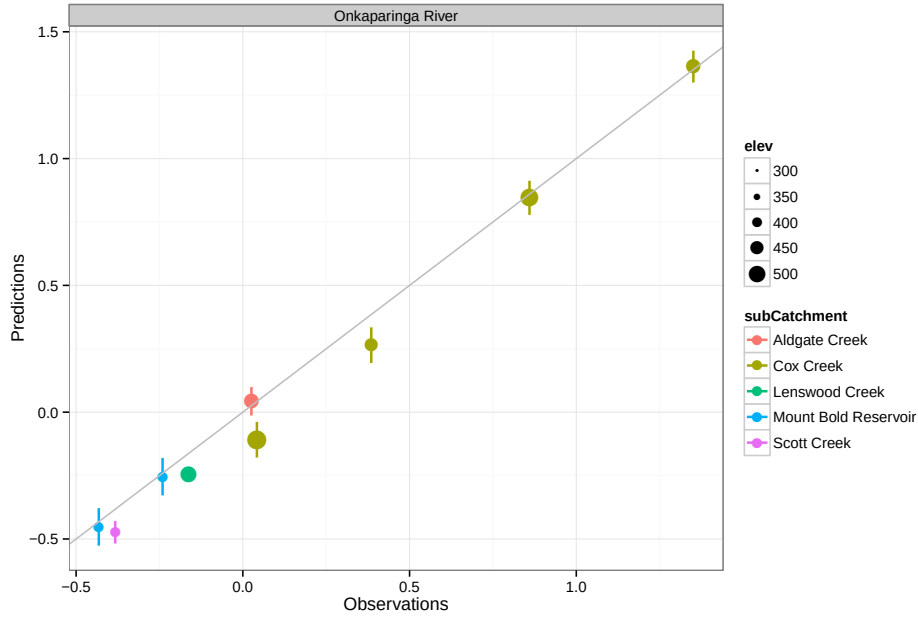
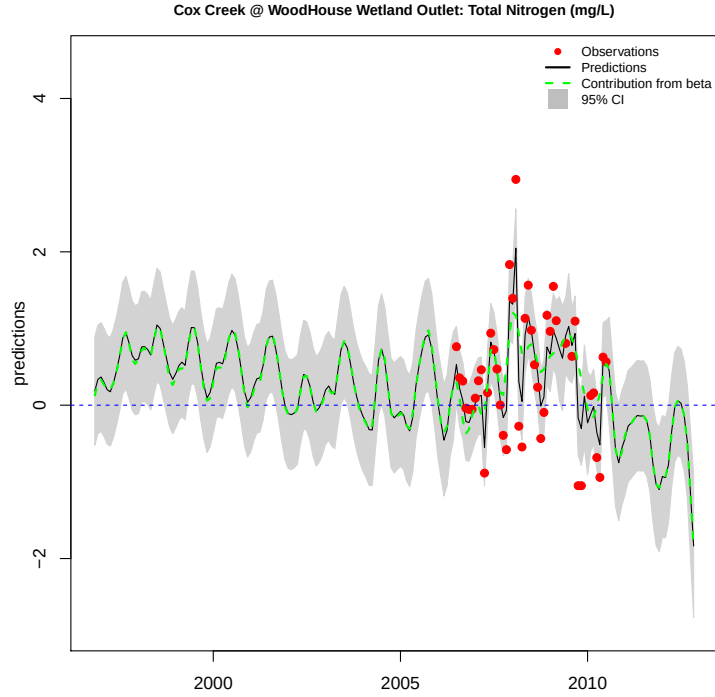


Figure 11: Prediction performance of the model for Total Nitrogen, at each site. Only one catchment (Onkaparinga River) has Total Nitrogen measurements. The vertical bar for each point is an error bar, stating the confidence interval for each prediction. Y-axis displays $\log(\text{predictions})$ and x-axis $\log(\text{observations})$. The legend 'elev' indicates elevation for each site.

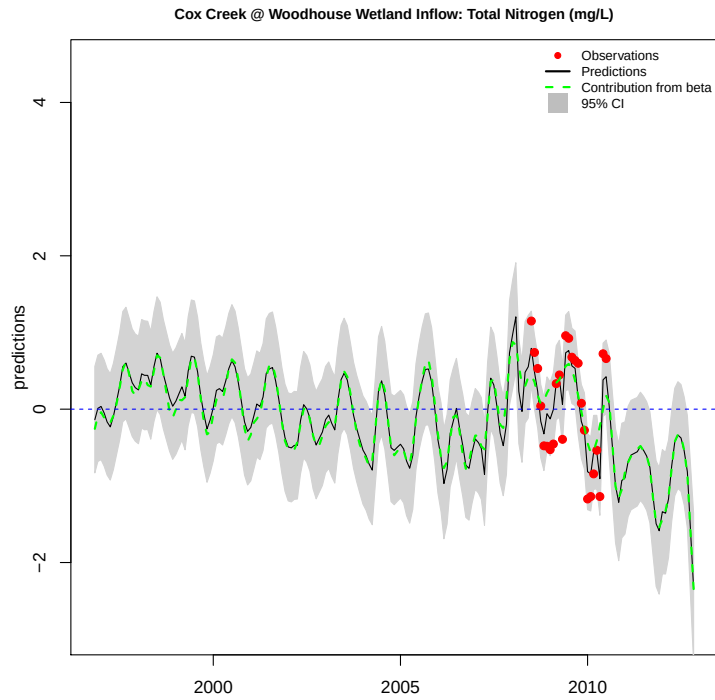
the model predicts a value below the recommended ANZECC guideline (which is 0 on the figure), when the actual measurements are slightly above this.

Figures 11 and 12 show that the model captures the variation of the water quality parameter fairly well, except for a few stations. The confidence in the model is important as it allows for the behaviour of the water quality parameters to be forecast for future events. In particular, it appears (from Figure 12) that after having been consistently over the guideline, that TN is predicted to drop under the ANZECC guideline value from 2010 onwards. Similar conclusions can be drawn from the other figures displayed in Appendix C.

The results indicate for NO_x that the land use category irrigated perennial horticulture is the only land use category with significant statistical influence. The results also indicated that the PIC and FOX soil types interact positively with this parameter. The six catchments analysed present the same spatial behaviour, with a small value for the spatial correlation parameter which indicates strong spatial consistency for this parameter amongst the sites. The conclusions are the opposite for soil types (for TN): the same soil types are significant, but their influence is negative. The land use influence is essentially negative, with the land use categories managed resources protection and grazing modified pastures displaying the main influence. The spatial consistency is also maintained, while



(a) Prediction performance of the model for Total Nitrogen at Cox Creek at WoodHouse Wetland Outlet.



(b) Prediction performance of the model for Total Nitrogen at Cox Creek at WoodHouse Wetland Inflow.

Figure 12: Prediction performance of the model for Total Nitrogen at two Cox Creek sites. The blue dotted line represents the log of the recommended ANZECC guideline. The y-axis shows $\log(\text{predictions})$ and the x-axis Year.

the flow levels are slightly positively significant. For TP, neither the soil type nor the land use proves significant, whilst flow levels do. Moreover, Cox Creek sub-catchment shows a different spatial behaviour: the nugget parameter is quite high, meaning a discontinuity in the spatial covariance between a site and its neighbours. Such behaviour is usually observed if one site follows a different protocol, or if some of its features are very different and were not adequately captured in the model (i.e. the information is most likely not available in the dataset). It is important to note here that a sedimentation pond was installed in Cox Creek at Uraidla in 2006. Data is available for this site prior to 2006, but only after 2006 for the other Cox Creek sites. The strength of the correlation between sites is time varying of the first order, indicating it follows the seasons. Finally, for SS, we observe the same kind of results as we did for NO_x: except for the land use, where the only significant component is essentially composed of Services, and this influence is negative.

3.8.2 Explanatory covariates: influence of landscape parameters

Figure 13 shows the results of the principal components analysis (PCA) used to reduce the number of land use and soil type categories. The results indicate the land use categories of influence defined by the PCA (PC1-PC4.LU); and the soil type categories by PC1-PC4.ST for the model on TN. These variables are defined based on a mathematical transformation used to avoid over-parameterisation. The seven main contributors to land use are: Grazing modified pastures (GMP); Irrigated perennial horticulture (IPH); Irrigated seasonal horticulture (ISH); Managed Res. protection (MRP); Nature Conservation (NC); Residential (Res); and Services (Ser). For soil type, the main contributors are: BRA, CAG, CLA, FOX, JUP, LEN and PIC.

The model used in the Tier 4 assessment emphasizes the influence of different landscape parameters (such as soil type, land use etc.) on the water quality measurements. The significant parameters are displayed in Figure 14.

Tier 4 results highlight the following outcomes for TN:

- flows are positively correlated with the TN measurements, which supports similar results from Tier 2;
- the land use category irrigated perennial horticulture is positively correlated with TN measurements. This suggests higher levels of TN are associated with catchments dominated by this land use category;

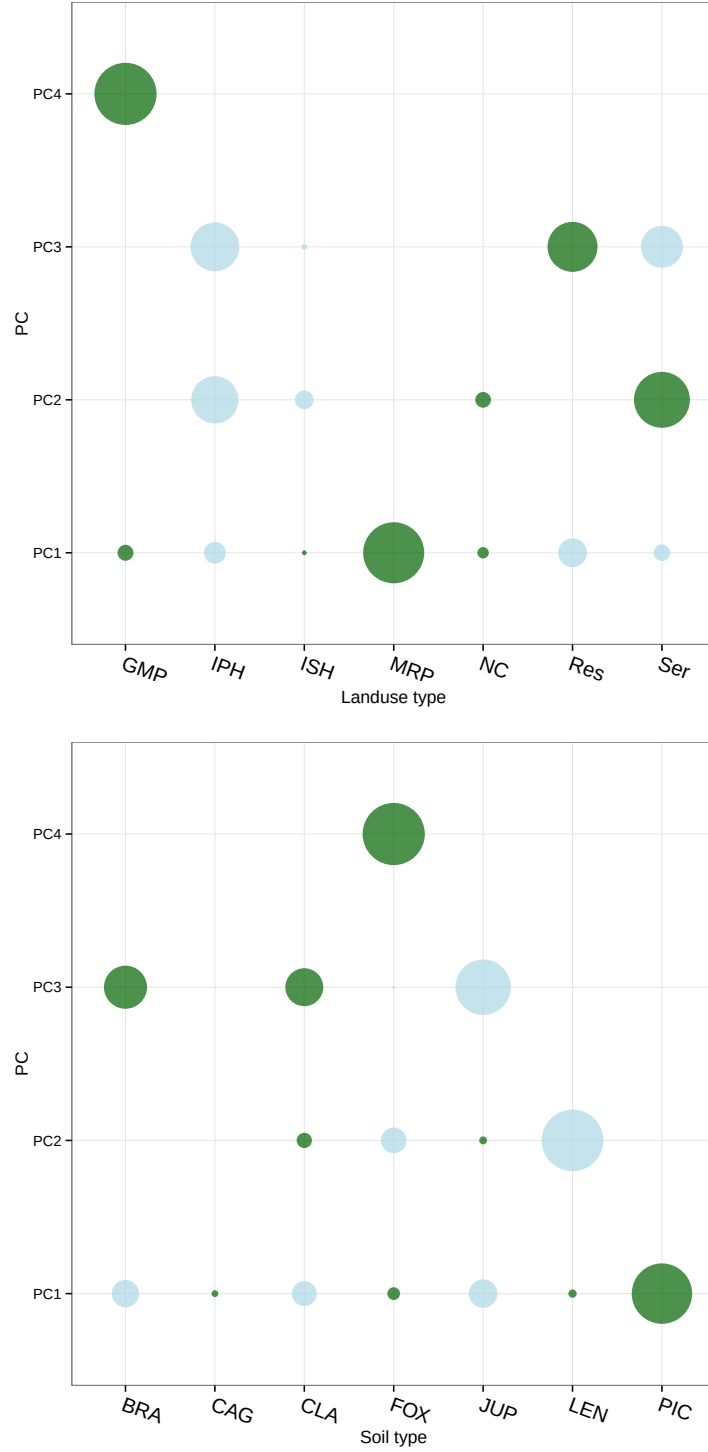


Figure 13: Land use and soil types PCA representation. The larger the circle, the bigger the weight. Green circles stand for a positive influence, and blue circles for a negative one.

- the land use categories grazing modified pastures, reservoir and dams, and managed resource protection are negatively correlated with the TN measurements, suggesting lower levels of TN in areas dominated by these land use categories;

- elevation does not have an influence on TN measurements, as partially observed in Figure 11.

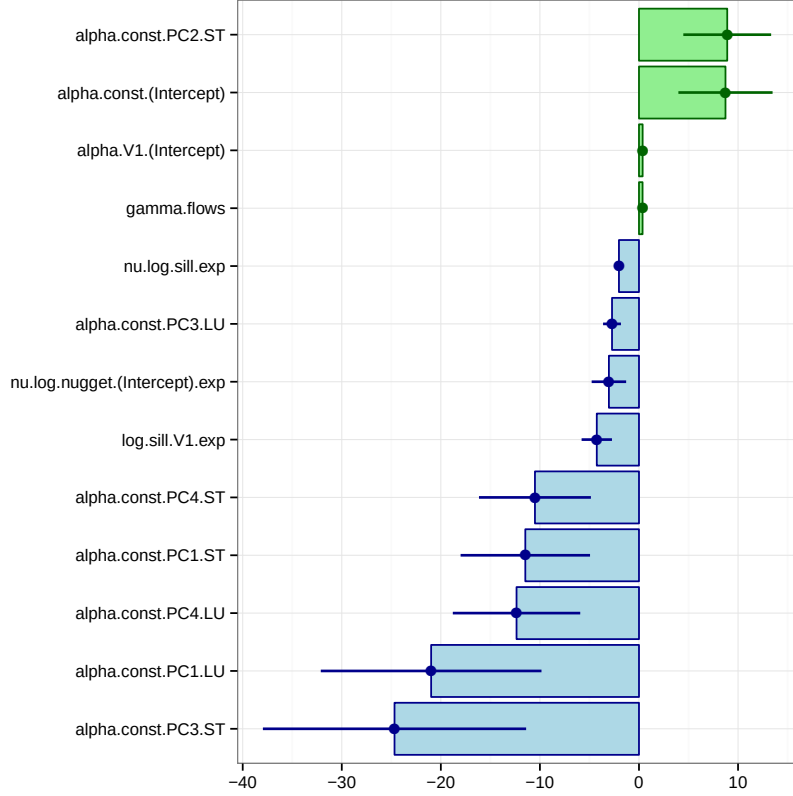


Figure 14: Estimated values (and error bars) for the different parameters of the water quality model. Only significant influences are displayed. The x-axis indicates the order of magnitude for parameters in the model. This is a proxy for increasing risk, green indicates positive effect on concentration, so larger values in green indicate high positive effect on concentration, and conversely for blue. The parameters preceded by “alpha” are the parameters of the fixed covariates component of the model. The terms “sill” and “nugget” are used to describe the spatial correlation. “V1” is the first order temporal influence (a constant influence over time).

Identifying which other sites (or catchments) may need to be monitored is a multifaceted problem. Looking at land use and soil type characteristics of the sites we modelled, PC1.ST was identified as a common variable to the poorly predicted sites (Cox Creek), as well as PC2.ST (Lenswood Creek) (see Figure 15).

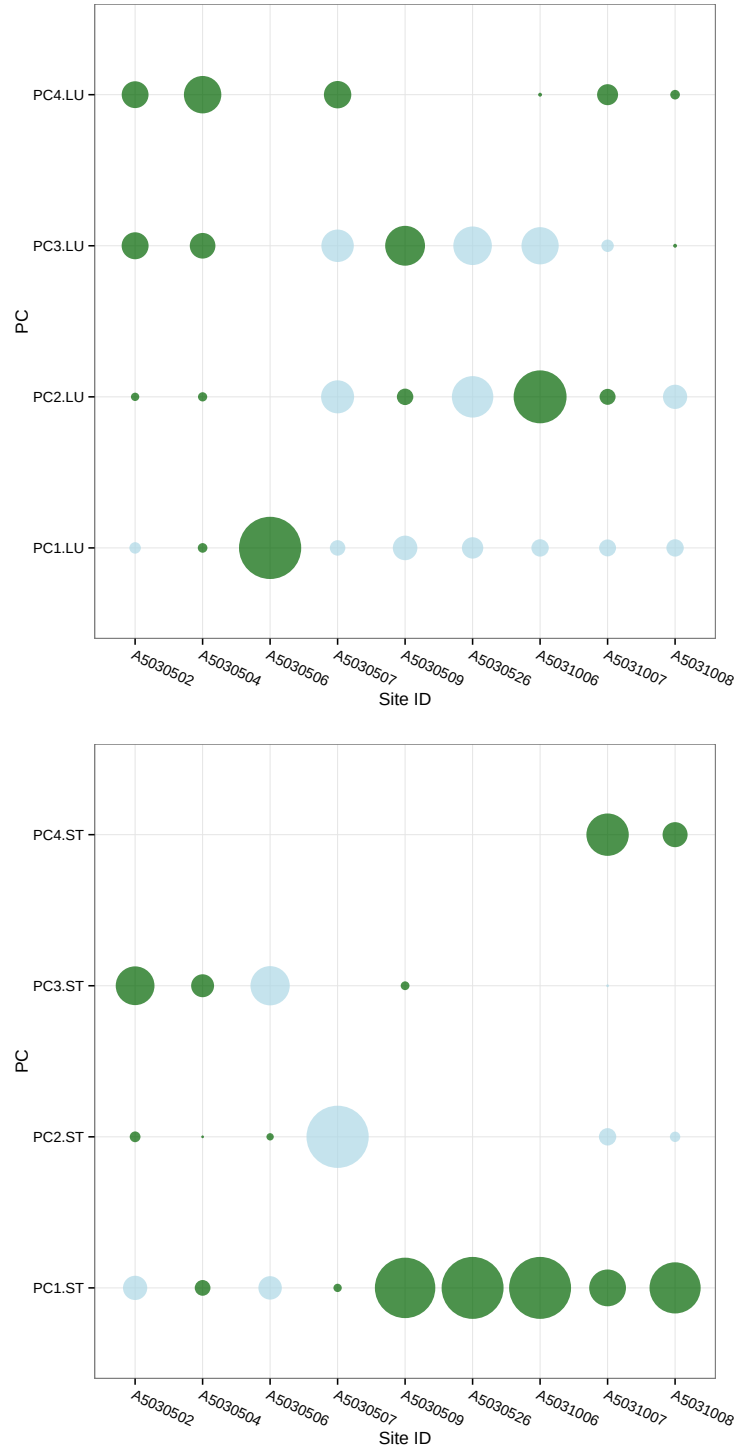


Figure 15: Land use and soil types composition of the monitored sites with sufficient data. The larger the circle, the bigger the weight. Green circles stand for a positive influence, and blue circles for a negative one. The y-axis for the two figures are four soil type and land use principal components respectively.

4 Discussion

This tiered risk assessment provides a transparent, data-driven probabilistic risk assessment. It uses a ratio-scale risk metric that provides managers with a risk estimate that indicates the extent to which a certain catchment is better than another in terms of water quality, providing an overall ranking for all sites and water quality parameters. This risk ranking can be used to guide the level of mitigation and investment required for each catchment and be used to guide the investment of resources to improve water quality across the MLR. The probabilistic foundation allows the application of uncertainty analysis techniques (see for example Morgan and Henrion (1990) and Frey and Burmaster (1999)) that can help guide future resource allocation and also clearly identify the impacts of data gaps on the risk outcomes.

4.1 Relative ranking of catchments (Tier 1 and Tier 2)

Tier 1 of the risk assessment provided an initial first overview of the data, and provided an assessment of the data and the water quality at each site. This initial stage highlighted the inconsistent monitoring across the sites. In some cases, there was regular monthly data collection for upwards of 10 years; in other cases, sporadic and inconsistent data collection for a year or two. Drawing robust conclusions across the MLR as a whole is difficult with such highly variable frequency of monitoring across different sites and the catchments within the region.

Nonetheless, the Tier 1 results indicated the majority of sites were exceeding the ANZECC guidelines in at least one, and often multiple, water quality parameters. All sites were below the ANZECC guideline for several water quality parameters: DOC (10 mg/L), Lead, Turbidity, and SS. In comparison all sites exceeded the ANZECC guideline for Copper and DOC (5 mg/L). The other parameters were consistently exceeded across several sites and catchments.

All 10 sites with metals data recorded medians which exceeded guidelines for Copper and more than half the sites for Zinc, but all were below the guideline for Lead. These results are not surprising as the MLR is known to have naturally elevated levels of metals such as Zinc and Copper, and in these circumstances exceedance of the national ANZECC guidelines may be an inappropriate endpoint (C. Jenkins, pers. comm.). Although the ANZECC guidelines do allow authorities to specify alternative limits that reflect region-specific circumstances, these were not available for the MLR region. For the other water quality parameters such as Nitrate and Phosphates, however, the results of Tier 1 con-

firms that the water quality in the MLR generally exceeds ANZECC guidelines for these parameters.

The Tier 1 results were sensitive to use of median or mean which suggests some periodic or occasionally large increases in concentrations. This was most noticeable for TP, Zinc, Turbidity and SS. These periodic increases, and the resulting distribution of the data, further supports the use of the more complex risk assessment in Tier 2 and Tier 4.

Tier 2 provided a more comprehensive, and more accurate quantification of the risk of exceeding ANZECC guidelines. Catchments ranked by Tier 2 risk assessment showed that the Onkaparinga River catchment was frequently ranked highest across the water quality parameters, representing poorest water quality of the sites assessed. This was in part due to the number of sites within the Onkaparinga River catchment (more than all other catchments), but also due to the water quality at the sites within this catchment: the Cox Creek sites were frequently ranked highest of all sites in multiple water quality parameters.

The average Tier 2 risk was calculated across all parameters. Removal of the metals data, and recalculation of the average risk, showed that of the 10 sites with measurements for metals, two showed an increase in average risk when metals were excluded, and eight a reduction. Cox Creek at Uraidla showed a 15% increase in average risk for non-metals versus all parameters, highlighting the overall poor water quality at this site. Risk of exceedance was high across all water quality parameters, and was evidently higher in non-metal parameters (such as NO_x, TP and TN). In comparison, Onkaparinga River U/S Estuary Old Noarlunga (also in the Onkaparinga catchment), had a lower overall average risk, but the main contributors to the risk were metals (3 of the 8 parameters measured there), with a notable 54% decrease in average risk when the metal data were excluded.

The maps of Tier 2 risk of exceedance hotspots, highlight that the Onkaparinga catchment was frequently identified as an area at high risk of exceeding the ANZECC guidelines, across all water quality parameters. This is partly due to the number of sites in the catchment and partly to the poorer water quality within the catchment.

4.2 Temporal patterns in water quality

Consistent temporal patterns were observed for EC in the majority of sites across the catchments. Time series analysis could be used for the EC data due to the large quantity of data collected at frequent intervals (see Appendix B). The results show strong seasonal

fluctuation across most sites. Monthly box plots, used for the remainder of the water quality parameters, did not show any consistent trend across sites or parameters. However, concentrations of some parameters (e.g. TN, NOx) increased in June and July, which may represent the first runoff events in the catchments.

4.3 Relationship between flow and water quality

The strongest positive relationship for flow with the water quality parameters assessed was found between NOx and flow. Since NOx is a measure of the soluble form of nitrogen this relationship is not unexpected. This highlights that contaminants are transported off-site in a soluble phase as well as attached to colloidal material and any mitigation strategies implemented in the MLR must deal with both transport processes and it is unlikely that one mitigation strategy alone will suit. In addition to strategies currently implemented in the region that trap sediment, such as buffer strips and sedimentation ponds, other strategies that minimise transport of soluble contaminants need to be considered. The positive relationship between increasing flow and increasing TN and NOx concentrations potentially has implications for water management at these sites. The extent of this relationship throughout the rest of the MLR needs to be investigated.

4.4 Correlation between water quality parameters

One of the objectives of the study was to determine whether any water quality parameters could act as surrogates for one another. If this was possible then savings could be made in current monitoring programs with a view to identifying easily measurable parameters which could serve as surrogates. Correlations were assessed in order to attempt to identify any strong, consistent, joint dependencies between parameters. Results indicate, as expected, strong joint dependency between metals, specifically Copper and Zinc across multiple sites; along with TN and NOx; and RP and TP. Although there were several strong correlations, no pairwise combination appeared consistently across all. Consequently, it is not possible from the correlations to recommend any water quality parameters that could act as a surrogate for another in this region.

4.5 Effectiveness of sedimentation pond at Cox Creek Uraidla

A sedimentation pond was installed at Cox Creek at Uraidla in 2006. In order to assess the impact of this mitigation measure, the data were analysed for before and after 2006. The results showed a consistent improvement across all water quality parameters. Tier 1 ratios

of medians to ANZECC guideline showed substantial change with most ratios dropping by more than half after installation of the sedimentation pond. However, although there was an overall consistent drop in median across the parameters following the installation of the sedimentation pond, the majority of measurements were still above the ANZECC guideline values. Two parameters, SS and Zinc showed medians which dropped below the ANZECC guideline after installation of the sedimentation pond.

4.6 Key catchments characteristics associated with water quality status

Tier 4 required developing and implementing more advanced statistical spatio-temporal models. Three models, with incremental amount of information modelled, were used to fit the data. The first one, a general (distance-wise) spatial dependency only model was considered, but this performed poorly. The second layer included temporal modelling, allowing for both seasonality and lag dependence. This proved considerably better than the first approach, allowing reasonable predictions for water quality. The full model included the land use and soil type information for each of the catchments, together with the monthly flow levels.

The land use information used in the Tier 4 assessment was collected in the aerial compositional format. The number of sites (together with the large number of possible covariates), however, far exceeded the number of locations with water quality data collected at sufficiently frequent intervals, resulting in a "sparse" covariance matrix. In order to improve the modelling, the number of land use classes was reduced using aggregation from a principal components analysis. The same method was used to reduce the number of soil type categories in the model. The results obtained from the Tier 4 model lead to the identification of features of influence (particular types of land use). Moreover, spatial and temporal consistency was observed in the results, despite poor quality observations. This suggests the model is well suited to this problem, and indicates that, with improved data quality, we could expect better results in terms of forecasting, but also modelling in terms of site features of influence.

4.7 Limitations and recommendations

One major outcome of the project, noticeable at each of the tiers was the inconsistency in the data. There were many sites which were not included in the tiered risk assessment due to the lack of data. The difficulty in interpreting results here, due to lack of data collected

at sufficiently frequent and consistent intervals, suggests that future monitoring schemes should focus on just a few key parameters (for example TP, TN, NO_x, SS) and ensure consistent, frequent, data is collected for these parameters. This would provide a better platform from which to investigate any underlying trends, correlations, joint dependencies and to further investigate effects of various land use.

We recommend that future monitoring studies would benefit from firstly establishing and clarifying management goals. Following this, determining clear objectives which relate to these management goals would help to design a monitoring program that meets the objectives. Designing a sample study before data collection, with a carefully chosen sample design, could help answer advanced questions such as the influence of water release, or the possibility to predict the water quality on certain sites without the need to monitor them.

5 References

- Aitchison, J. (2003). A concise guide to compositional data analysis. Technical report, In: 2nd Compositional Data Analysis workshop.
- AMLRNRMB (2010a). Draft Water Allocation Plan: Western Mount Lofty Ranges. Technical report, Adelaide and Mount Lofty Ranges Natural Resources Management Board.
- AMLRNRMB (2010b). Your guide to understanding the Western Mount Lofty Ranges Draft Water Allocation Plan. Technical report, Adelaide and Mount Lofty Ranges Natural Resources Management Board.
- ANZECC/ARMCANZ (2000). Australian and New Zealand Guidelines for Fresh and Marine Water Quality. Technical report, Australia and New Zealand Environmental Conservation Council / Agriculture Resource Management Council of Australia and New Zealand.
- Billheimer, D. and Guttorp, P. (1995). Spatial models for discrete compositional data. Technical report.
- Cox, J., Bald, M., Burch, M., Chittleborough, D., Cuddy, S., Deane, D., Fleming, N., Frizenschaf, J., Green, G., Halfyard, R., Holland, K., Kookana, R., Lomman, G., Oliver, D., Rassam, D., Saint, C., Savadamuthu, K., Skewes, M., van der Weilen, M., VanLaarhoven, J., and van Leeuwen, J. (2013). Water allocation plan and water quality improvement scoping study - Discussion paper. Technical report, Goyder Institute for Water Research Tehcnical Report Series No. 13/8, Adelaide, South Australia. ISSN: 1839-2725.
- Cox, J., Oliver, D., Fleming, N., and Anderson, J. S. (2012). Off-site transport of nutrient from three landuses in the Mt. Lofty Ranges South Australia. *Agricultural Water Management*, 106:50–59.
- Cressie, N. and Wikle, C. (2011). *Statistics for spatio-temporal data*.
- Fleming, N., Cox, J., He, Y., Thomas, S., and Frizenschaf, J. (2010). Analysis of total suspended sediment and total nutrient concentration data in the Mount Lofty ranges to derive event mean concentrations. Technical report, eWater Cooperative Research Centre Technical Report.
- Ford, J., Ickowicz, A., Oliver, D., Hayes, K., and Kookana, R. (2015). Results from Task 5 Tiered Water Quality Risk Assessment. Technical report, Goyder Institute for Water Research Technical Report Series No. 15/4, Adelaide, South Australia.

- Frey, H. and Burmaster, D. (1999). Methods for characterizing variability and uncertainty: Comparison of bootstrap simulation and likelihood-based approaches. *Risk Analysis*, 19(1):109–130.
- Gelfand, A. (2012). Hierarchical modeling for spatial data problems. *Spatial Statistics*, 1:30–39.
- Gelfand, A., Diggle, P., Guttorp, P., and Fuentes, M. (2010). *Handbook of spatial statistics*.
- Gelfand, A., Zhu, L., and Carlin, B. (2001). On the change of support problem for spatio-temporal data. *Biostatistics (Oxford, England)*, 2(1):31–45.
- Hayes, K., Oliver, D., Kookana, R., and Kroon, F. (2012). Development of ecological risk assessment methodology for terrestrial pollutants to the great barrier reef ecosystems. Technical report, CSIRO, Australia.
- Kurz, I., Coxon, C., Tunney, H., and Ryan, D. (2005). Effects of grassland management practices and environmental conditions on nutrient concentrations in overland flow. *Journal of Hydrology*, 304:35–50.
- Matern, B. (1986). *Spatial Variation.*, volume 36.
- Morgan, M. and Henrion, M. (1990). *Uncertainty: A guide to dealing with uncertainty in quantitative risk and policy analysis*. Cambridge University Press, Cambridge, England.
- Nelson, P., Cotsaris, E., Oades, J., and Bursil, D. (1990). Influence of soil clay content on dissolved organic matter in stream waters. *Australian Journal of Marine and Freshwater Research*, 41:761–774.
- Oliver, D., Anderson, J., Kookana, R., Cox, J., Fleming, N., Waller, N., and Smith, L. (2012). Off-site transport of pesticides from two horticultural landuses in the Mt. Lofty Ranges, South Australia. *Agricultural Water Management*, 106:60–69.
- Peterson, E., Theobald, D., and Ver Hoef, J. (2007). Geostatistical modelling on stream networks: Developing valid covariance matrices based on hydrologic distance and stream flow. *Freshwater Biology*, 52(2):267–279.
- Peterson, E. and Ver Hoef, J. (2010). A mixed-model moving-average approach to geostatistical modeling in stream networks. *Ecology*, 91(3):644–651.
- Sahu, S. and Mardia, K. (2005). Recent trends in modeling spatio-temporal data. *Proceedings of the special meeting on Statistics*, pages 69–83.

- SAMDBNRMB (2011). Draft Water Allocation Plan for the Eastern Mount Lofty Ranges Prescribed Water Resources Area. Technical report, South Australian Murray-Darling Basin Natural Resources Management Board.
- Schriever, C. and Liess, M. (2007). Mapping ecological risk of agricultural pesticide run-off. *Sci. Total Environ.*, 384:264–279.
- Schulz, R. (2004). Fieldsreview on exposure, effects, and risk mitigation of aquatic nonpoint-source insecticide pollution: a review. *J. Environ. Qual.*, 33:419–448.
- Skoien, J., Merz, R., and Blöchl, G. (2006). Top-kriging - geostatistics on stream networks. *Hydrology & Earth System Sciences*, pages 277–287.
- Thomas, S., He, Y., and Fleming, N. (2010). Progress report on the Mount Lofty Ranges source catchment application project. Technical report, eWater Cooperative Research Centre Technical Report.
- Tjelmeland, H. and Lund, K. (2003). Bayesian modelling of spatial compositional data. *Journal of Applied Statistics*.
- Varcoe, J., van Leeuwen, J., Chittleborough, D., Cox, J., Smernik, R., and Heitz, A. (2010). Changes in water quality following gypsum application to catchment soils of the Mount Lofty Ranges South Australia. *Organic Geochemistry*, 41:116–123.
- Ver Hoef, J., Peterson, E., Clifford, D., and Shah, R. (2014). SSN: an R package for spatial statistical modeling on stream networks. *Journal of Statistical Software*, 56(3).
- Wauchope, R. (1978). The pesticide content of surface water draining from agricultural fields - a review. *J. Environ. Qual.*, 7:459–472.
- Zhu, L. and Carlin, B. (2000). Comparing hierarchical models for spatio-temporally misaligned data using the deviance information criterion. *Statistics in Medicine*, 19(17-18):2265–78.
- Zhu, L., Carlin, B., and Gelfand, A. (2003). Hierarchical regression with misaligned spatial data: relating ambient ozone and pediatric asthma ER visits in Atlanta. *Environmetrics*, pages 537–557.

Appendices

Appendix A Data Sources

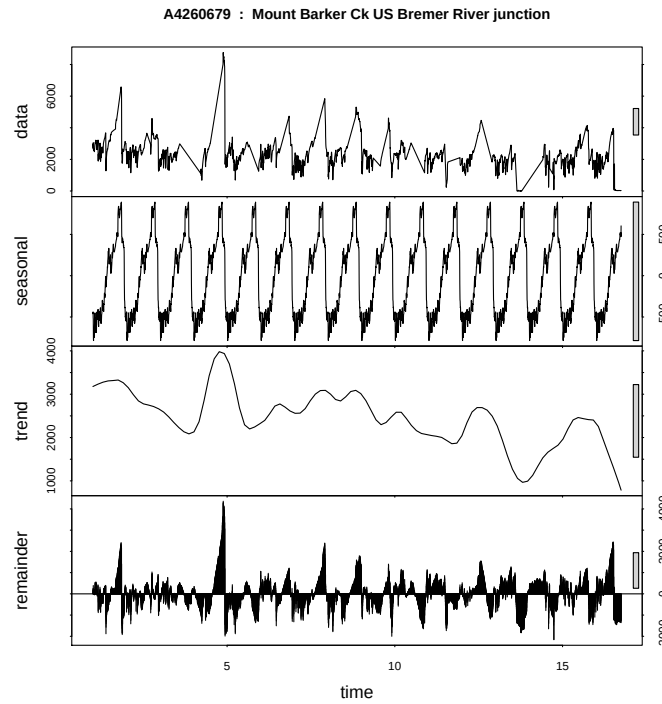
Custodian			Website	Contact name	Contact Phone	Contact email
SA Water			http://www.wdapp.com/Amlr.aspx	Jacqueline Frizenschaf	Phone: (08) 7424 1844 Fax: (08) 7003 1844 Mobile: 0427 797 196	Jacqueline.frizenschaf@sawater.com.au
SA Water – DOC data			None	Sean Lasslett	Phone: (08) 7424 2947 Fax: (08) 7003 2947 Mobile: 0467 807 792	Sean.lasslett@sawater.com.au
SA EPA			None	Shaun Thomas	Phone: (08) 8204 2023 Fax: (08) 8124 4673 Mobile: 0400 923 313	Shaun.thomas@epa.sa.gov.au
SA EPA			None	Clive Jenkins		Clive.jenkins@epa.sa.gov.au
SA EPA			None	Stephen Packer	Phone: (08) 84637809 Fax: (08) 81244673 Mobile: 0428 103 564	stephen.packer@epa.sa.gov.au
DEWNR data)	(NRM Board		http://www.waterconnect.sa.gov.au/SWA/Pages/default.aspx	Keith Smith		Keith.smith4@sa.gov.au
DEWNR data)	(NRM Board		http://www.waterconnect.sa.gov.au/SWA/Pages/default.aspx	Shane Johansen	Phone: (08) 8273 9123 Fax: (08) 8271 9585	Shane.johansen@sa.gov.au
DEWNR data)	(NRM Board		http://www.waterconnect.sa.gov.au/SWA/Pages/default.aspx	Steve Gatti	Phone:(08) 8273 9129 Fax: (08) 8271 9585 Mobile: 0409 126 175	Steven.gatti@sa.gov.au
Aquasave			None	Nick Whiterod	Phone: (08) 8555 0941 Mobile: 0409 023 771	nick.whiterod@aquasave.com.au
Leon van der Linden			None	Leon van der Linden		leonvanderlinden@sawater.com.au

Table 10: Summary of sources and contact details for Mt Lofty Ranges Water Quality Risk Assessment.

Appendix B Trends in Electrical Conductivity

Across the catchments, 76 sites had data recorded for EC. Although still irregular in some sites, this data was overall of higher quality (more consistent and frequent data collection) than the other water quality parameters. The time series analysis was used for EC to differentiate between seasonal patterns and overall trend.

An example of this analysis is presented in Figure 16 for Mount Barker Ck US Bremer River junction. The four panels show: all data; seasonal pattern; overall trend; and residuals. Some overall downward trend for EC at the Mount Barker Ck US Bremer River junction site can be seen in the third panel of Figure 16. The second panel in Figure 16 shows the strong seasonal fluctuations which was seen across most sites with data for EC.



(a) EC time series for site A4260679, Mount Barker Ck US Bremer River junction

Figure 16: EC time series analysis for Mount Barker Ck US Bremer River junction. The top panel shows all data; the second panel season; the third panel overall trend; and the bottom panel the remaining variance. The x-axis (time) is years of data.

Appendix C Tier 4 Spatio-temporal modelling: Methods and Results

C.1 Data

C.1.1 Water quality measurement frequency

	# stations	# Mode	Max #	Min #
Copper (mg/L)	10	193	193	32
Dissolved Organic Carbon (mg/L)	9	11	179	8
EC Corrected/ Conductivity (uS/cm)	79	546	5028	28
Iron (mg/L)	1	54	54	54
Lead (mg/L)	10	193	193	32
Nitrate + Nitrite as Nitrogen (mg/L)	18	86	743	32
pH	20	76	770	26
Reactive Phosphorus (mg/L)	12	594	594	50
Suspended solids (mg/L)	18	86	628	32
Total Dissolved Solids (by EC) (mg/L)	13	743	743	50
Total Kjeldahl Nitrogen (mg/L)	18	86	743	32
Total Nitrogen (mg/L)	9	666	666	74
Total Phosphorus (mg/L)	18	86	741	32
Turbidity (NTU)	17	49	101	26
Water Temperature (Deg Celcius)	15	49	97	25
Zinc (mg/L)	10	193	193	33

Table 11: Discrepancy of the parameter measurements. The first column states the number of stations measuring the parameter. The last three columns relate to the number of measurements (in time) per station.

C.1.2 Water quality measurement location

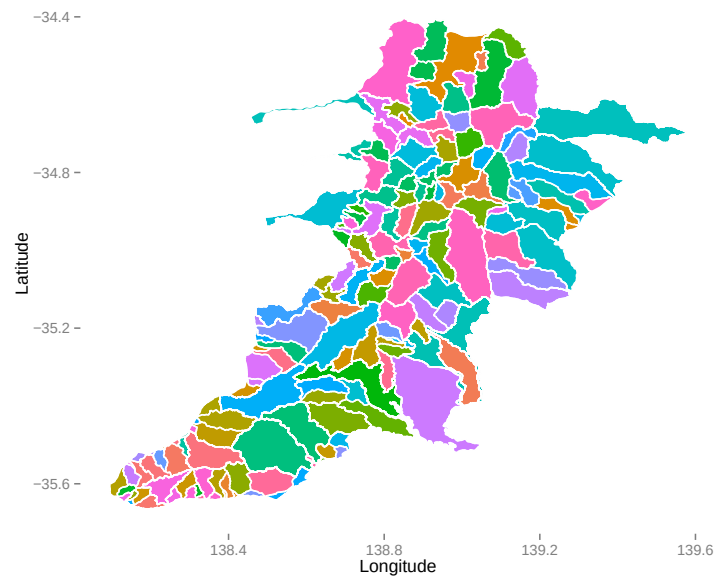
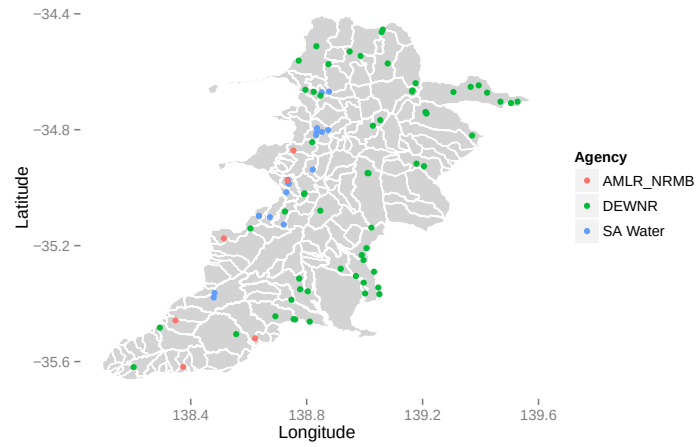


Figure 17: Mount Lofty Ranges map. Location of the WQ stations, and their agencies; and rainbow map of the different sub-catchments.

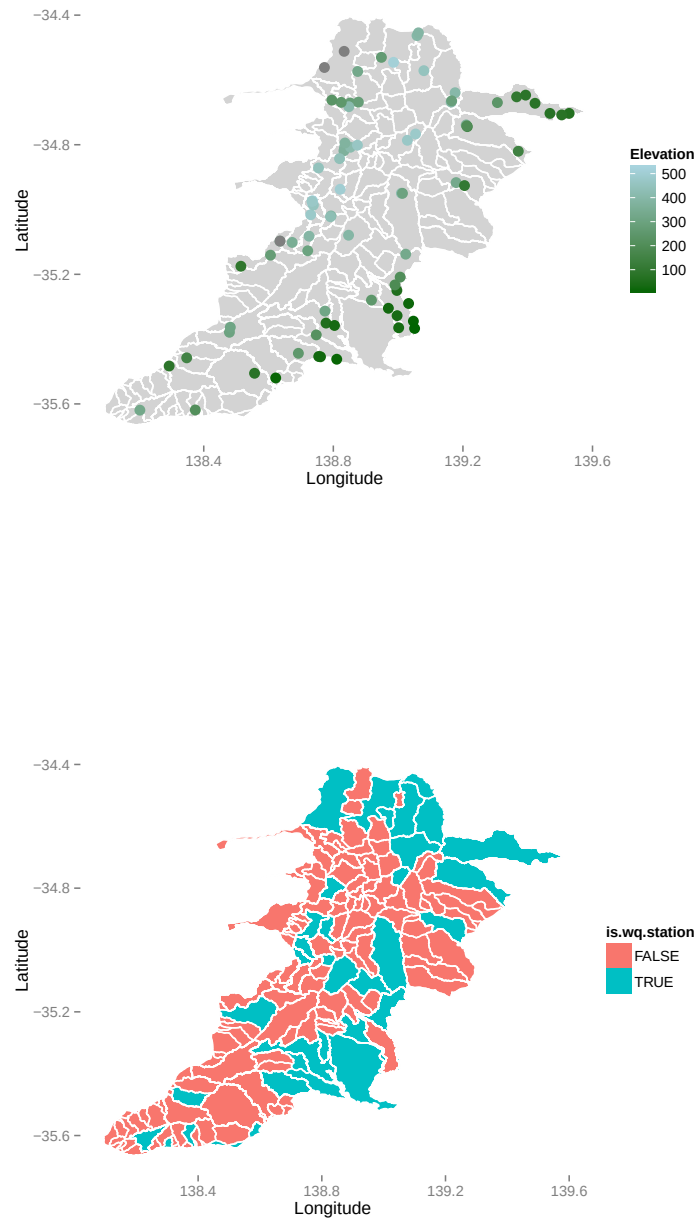


Figure 18: Mount Lofty Ranges map. Map of the sub-catchment with/without WQ stations; and elevation of the WQ stations.

C.2 The Spatio-temporal model

In this project, we are dealing with two different types of spatial data,

- Point-referenced data (water quality, precipitations)
- Areal data

Because of the different format, we are facing multiple issues,

1. Spatial change of support problem, between the water quality measurement and the catchments information, Gelfand et al. (2001); Zhu et al. (2003); Sahu and Mardia (2005);
2. Mathematical optimisation problem, as the land use information is given in proportion, and then has to be used as compositional data Billheimer and Guttorp (1995); Tjelmeland and Lund (2003); Aitchison (2003);
3. Temporal change of support problem, between the water measurements and the precipitations, Zhu and Carlin (2000); Gelfand et al. (2001);
4. Covariance modelling problem, as the correlation between the water quality measurements are not spatially uniform over the catchments Skoien et al. (2006); Peterson et al. (2007); Peterson and Ver Hoef (2010); Ver Hoef et al. (2014)

C.2.1 Overview of the spatio-temporal approach

The use of the Bayesian paradigm to model spatio-temporal data is increasingly popular, in particular the hierarchical Bayesian models Gelfand (2012), defined through three stages,

$$[\text{parameter} \mid \text{data}] = [\text{data} \mid \text{process, parameter}] \quad (1)$$

$$[\text{process} \mid \text{parameter}] \quad (2)$$

$$[\text{parameter}] \quad (3)$$

The second stage of the model is usually the place for the spatio-temporal random effect Gelfand et al. (2010); Cressie and Wikle (2011).

Let t denote the temporal unit, and $\mathbf{Z}_t = (Z(s_1, t), \dots, Z(s_n, t))'$ the observed point referenced data at sites $s_i, i = 1 \dots n$. We will consider a pure error term $\epsilon_t = (\epsilon(s_1, t), \dots, \epsilon(s_n, t))'$ (also called nugget effect in the dedicated spatial statistics literature). This error is assumed independently normally distributed $N(\mathbf{0}, \sigma_\epsilon^2 \mathbf{I}_n)$ where σ_ϵ^2 is the unknown pure error variance, and $\mathbf{I}$ the identity matrix. The spatio-temporal random effects will be denoted by $\eta_t = (\eta(s_1, t), \dots, \eta(s_n, t))'$, and these will be assumed to follow $N(\mathbf{0}, \Sigma_\eta)$ independently in time, where $\Sigma_\eta = \sigma_\eta^2 S_\eta$, σ_η^2 being the site invariant spatial variance, and S_η is the spatial correlation matrix obtained from the often used general Matérn correlation function Matern (1986),

$$\kappa(s_i, s_j; \phi, \nu) = \frac{1}{2^{\nu-1} \Gamma(\nu)} (2\sqrt{\nu} \|\phi\| s_i - s_j)^\nu K_\nu(2\sqrt{\nu} \|\phi\| s_i - s_j), \quad (4)$$

where $\Gamma(\cdot)$ is the standard gamma function and K_ν the modified Bessel function of second kind with order ν . The parameter ϕ controls the rate of the decay of the correlation between the sites, and ν controls the smoothness of the random field. Finally, let $\mathbf{X}_t$ be the matrix of covariates, where some covariates may vary in space and time.

Spatial Model The simplest spatial model follows the general equations,

$$\mathbf{Z}_t = \mathbf{O}_t + \epsilon_t \quad (5)$$

$$\mathbf{O}_t = \mathbf{X}_t \beta + \eta_t \quad (6)$$

If we denote $\theta = (\beta, \sigma_\epsilon^2, \sigma_\eta^2, \phi, \nu)$ all the parameters of the model, and let $\pi(\theta)$ be the prior distribution specified later, the logarithm of the posterior distribution is given by,

$$\begin{aligned} \log \pi(\theta, \mathbf{O}|\mathbf{z}) &\propto -\frac{N}{2} \log \sigma_\epsilon^2 - \frac{1}{2\sigma_\epsilon^2} \sum_t (\mathbf{Z}_t - \mathbf{O}_t)' (\mathbf{Z}_t - \mathbf{O}_t) - \frac{T}{2} \log |\sigma_\eta^2 S_\eta| \\ &\quad - \frac{1}{2\sigma_\eta^2} \sum_t (\mathbf{O}_t - \mathbf{X}_t \beta)' S_\eta^{-1} (\mathbf{O}_t - \mathbf{X}_t \beta) + \log \pi(\theta) \end{aligned} \quad (7)$$

Introducing the temporal dependency The spatio-temporal model loosely differs from the spatial model,

$$\mathbf{Z}_t = \mathbf{O}_t + \epsilon_t \quad (8)$$

$$\mathbf{O}_t = \rho \mathbf{O}_{t-1} + \mathbf{X}_t \beta + \eta_t \quad (9)$$

In particular, the auto-regressive model requires specification on the initial term $\mathbf{O}_0$. Let μ and σ^2 be these parameters. If we denote $\theta = (\beta, \sigma_\epsilon^2, \sigma_\eta^2, \phi, \nu, \rho, \mu, \sigma)$ all the parameters of the model, and let $\pi(\theta)$ be the prior distribution specified later, the logarithm of the posterior distribution is given by,

$$\begin{aligned} \log \pi(\theta, \mathbf{O}|\mathbf{z}) &\propto -\frac{N}{2} \log \sigma_\epsilon^2 - \frac{1}{2\sigma_\epsilon^2} \sum_t (\mathbf{Z}_t - \mathbf{O}_t)' (\mathbf{Z}_t - \mathbf{O}_t) - \frac{T}{2} \log |\sigma_\eta^2 S_\eta| \\ &\quad - \frac{1}{2\sigma_\eta^2} \sum_t (\mathbf{O}_t - \rho \mathbf{O}_{t-1} - \mathbf{X}_t \beta)' S_\eta^{-1} (\mathbf{O}_t - \rho \mathbf{O}_{t-1} - \mathbf{X}_t \beta) \\ &\quad - \frac{1}{2} \log |\sigma^2 S_0| - \frac{1}{2\sigma^2} (\mathbf{O}_0 - \mu)' S_0^{-1} (\mathbf{O}_0 - \mu) + \log \pi(\theta) \end{aligned} \quad (10)$$

REMARKS

This simple spatio-temporal model is quite widely used for application such as air monitoring, but requires an important number of observations taken at a high frequency. Moreover, no seasonality is modelled, nor spatial-temporal variation of the parameters. This lack of flexibility in the model makes it too limited for our purpose.

Full spatio-temporal model The spatio-temporal model we use for the water quality monitoring purpose is defined by the following equations,

$$\mathbf{Z}(s, t) = \mathbf{O}(s, t) + \epsilon(s, t) \quad (11)$$

$$\mathbf{O}(s, t) = \sum_l \gamma_l \mathbf{M}_l(s, t) + \sum_i \beta_i(s) f_i(t) \quad (12)$$

The $\mathbf{M}_l(s, t)$ are spatio-temporal covariates; γ_l are coefficients for the spatio-temporal covariates; $f_i(t)$ is a set of (smooth) temporal basis functions, with $f_1(t) = 1$; and the $\beta_i(s)$ are spatially varying coefficients for the temporal functions. The $\beta_i(s)$ -coefficients are treated as spatial fields with a universal kriging structure, allowing the temporal structure to vary between locations:

$$\beta_i \sim \mathcal{N}(\alpha_i \mathbf{X}_i, \Sigma_{\beta_i}) \quad (13)$$

where $\mathbf{X}_i$ are design matrices, α_i are matrices of regression coefficients, and Σ_{β_i} are covariance matrices. The $\mathbf{X}_i$ matrices often contain geographical covariates (land use, soil type). This structure allows for different covariates and covariance structures in the each of the $\beta_i(s)$ fields; the fields are assumed to be a priori independent of each other. The residual space-time field, $\epsilon(s, t)$, is assumed to be independent in time with stationary, parametric spatial covariance Σ_ϵ .

Because this model is essentially a linear combination of Gaussian vectors, the likelihood can be easily expressed:

$$\log \pi(\theta, \mathbf{O}|\mathbf{z}) \propto -\frac{1}{2} \log |\tilde{\Sigma}| - \left(\mathbf{Z} - \begin{bmatrix} \mathbf{M} & F\mathbf{X} \end{bmatrix} \begin{bmatrix} \gamma \\ \alpha \end{bmatrix} \right)^T \tilde{\Sigma}^{-1} \left(\mathbf{Z} - \begin{bmatrix} \mathbf{M} & F\mathbf{X} \end{bmatrix} \begin{bmatrix} \gamma \\ \alpha \end{bmatrix} \right) \quad (14)$$

where $\tilde{\Sigma} = \Sigma_\epsilon + F\Sigma_\beta F^T$.

C.3 Land use information treatment

We need to acknowledge that the number of land use categories is much more important than the number of different catchment. Even more important, the matrix of compositional data can be

	Land Use Category
1	Grazing modified pastures
2	Cropping
3	Plantation forestry
4	Transport and communication
5	Nature conservation
6	Residential
7	Other minimal uses
8	Managed resource protection
9	Reservoir/dam
10	Grazing natural vegetation
11	Irrigated cropping
12	Irrigated perennial horticulture
13	Irrigated modified pastures
14	Intensive animal production
15	Perennial horticulture
16	Irrigated plantation forestry
17	Seasonal horticulture
18	Services
19	Irrigated seasonal horticulture
20	Manufacturing and industrial
21	Intensive horticulture
22	Irrigated land in transition
23	Marsh/wetland
24	Utilities
25	Mining
26	Waste treatment and disposal
27	River
28	Lake

Table 12: Listing of the secondary land use categories.

considered sparse. This problem has only been recently tackled. The idea is to apply a derived PCA approach developed specifically to take into account the constraints of compositional data.

DEFINITION: SPARSE PRINCIPAL COMPONENT ANALYSIS

Let X be a matrix of covariates, with too many features to be successfully fed into a regression model. The full principal component decomposition of X can be given as

$$T = XW \quad (15)$$

where W is the matrix of eigenvectors and T the matrix of component scores. Geometrically, W corresponds to X on a new coordinate systems, where the coordinate vectors are sorted by

decreasing variance. This implies a better representation of the points cloud. This is considered as a dimension reduction technique as we may only keep the most important coordinates, for example the ones representing 75% of the total inertia.

The sparse version is obtained by the solving the following optimization problem,

$$\min_{X \in \Theta} \left[\|X\|_{l_1} + \mu \|\text{diag} X^T R X - \text{diag} D\|_2^2 \right] \quad (16)$$

where Θ is an adequate orthonormal matrix manifold, R is the covariance matrix, D the diagonal matrix containing the r largest eigenvalues of X , l_1 is the tr-based norm and μ controls the importance of the two terms. The smaller μ , the sparser component loadings X .

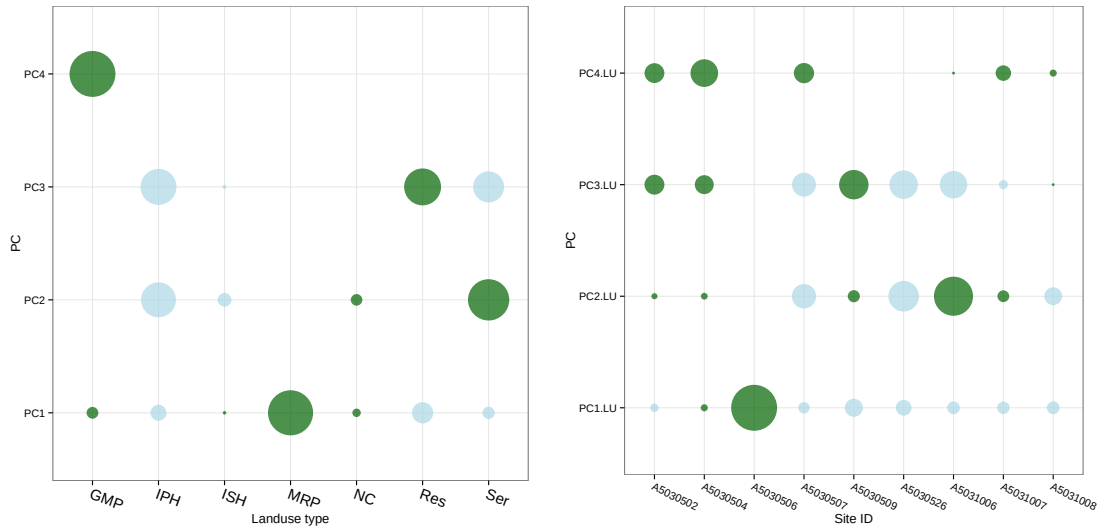


Figure 19: Land use principal components. The zero-value components have been removed from the figure.

Once the new matrix has been calculated, we may consider only the main components as covariates. Table 13 displays the inertia (percent of explained variance) explained by the different principal components. According to that table, considering the 4 first components covers for 70.4% of the variance.

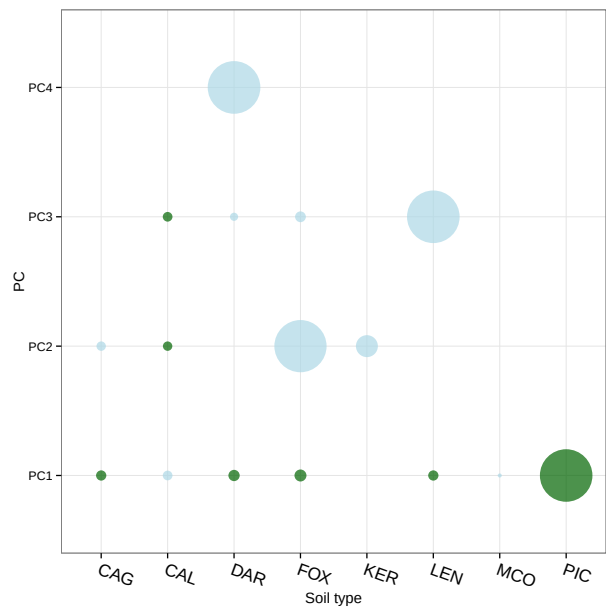
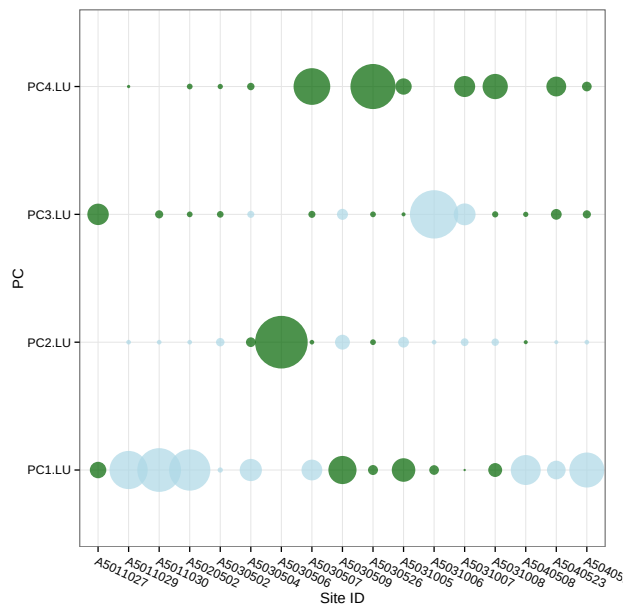
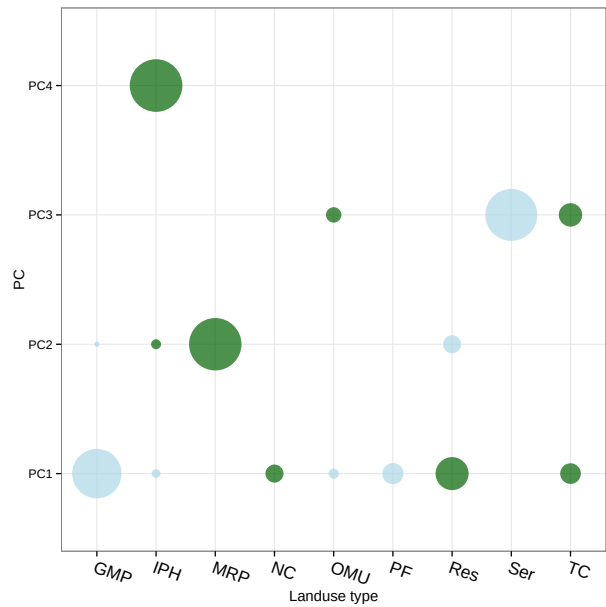
Vector	PC1	PC2	PC3	PC4	PC5	PC6
Perc. of exp. var.	39.1	12.9	11.4	7.0	5.6	5.4
Cum. perc. of exp. var.	39.1	52.0	63.4	70.4	76.0	81.4

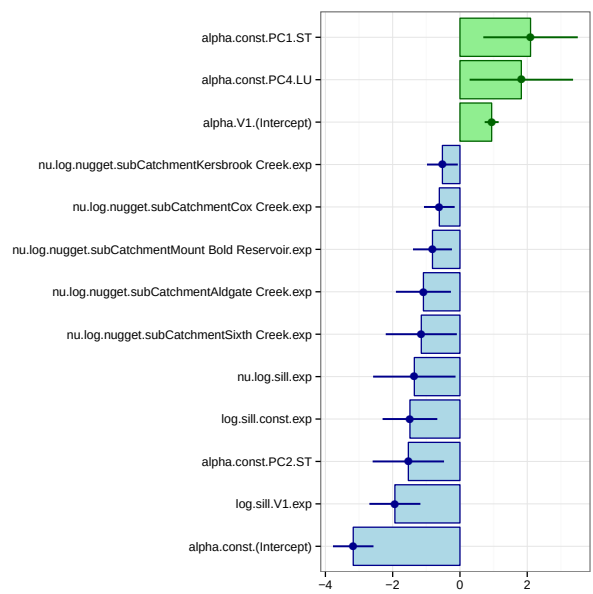
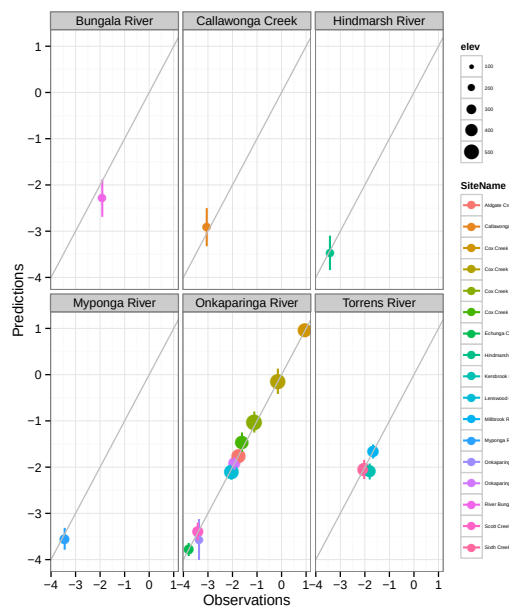
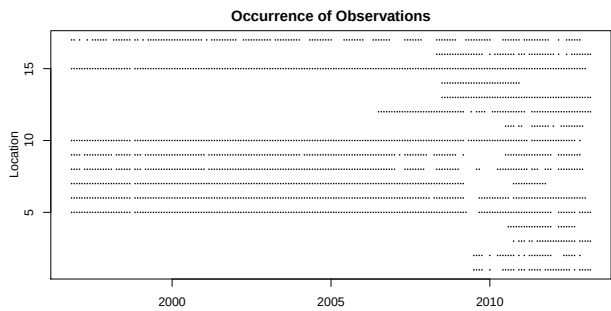
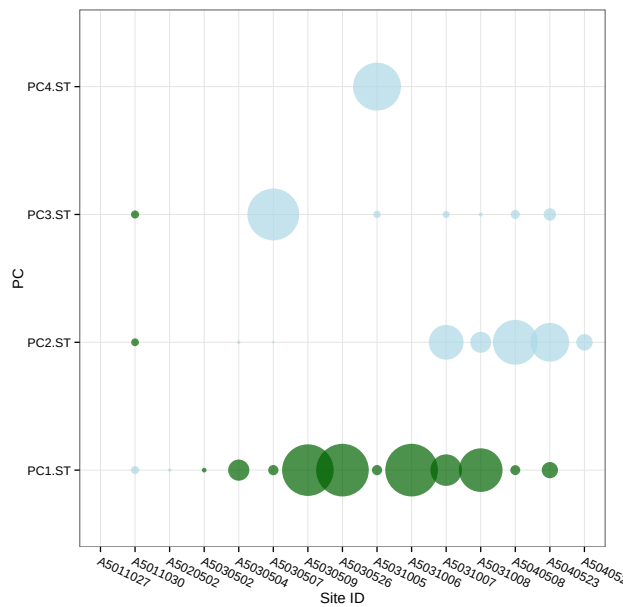
Table 13: Principal component decomposition. Percentage of explained variance.

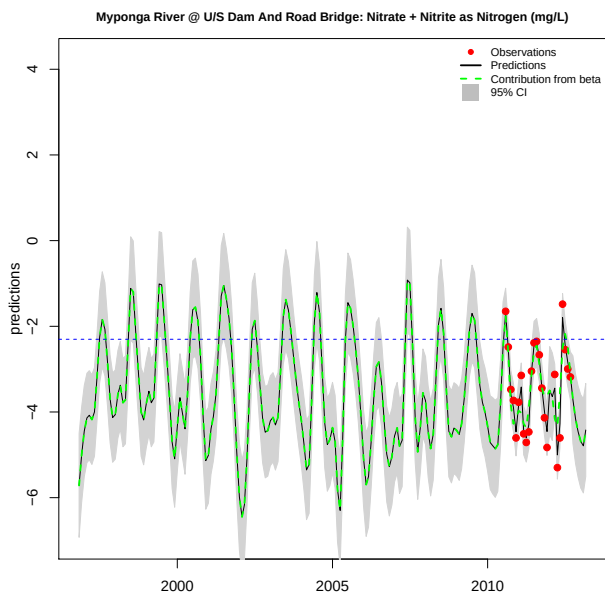
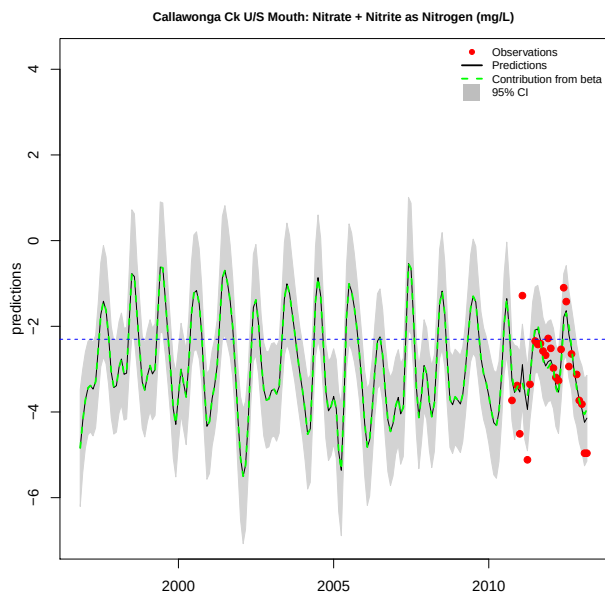
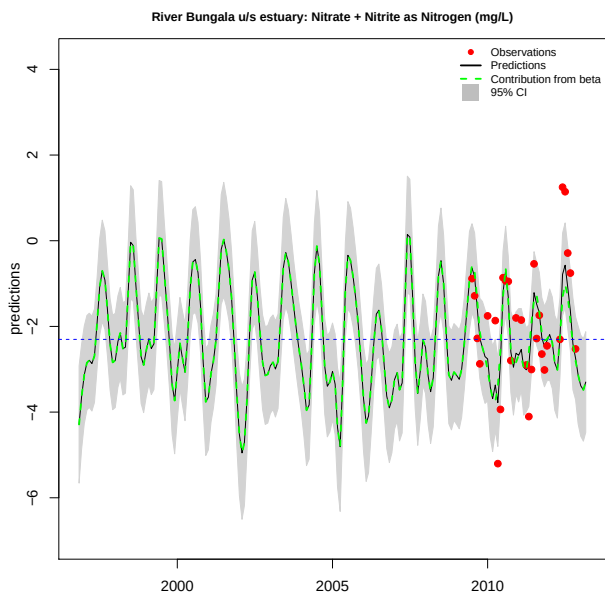
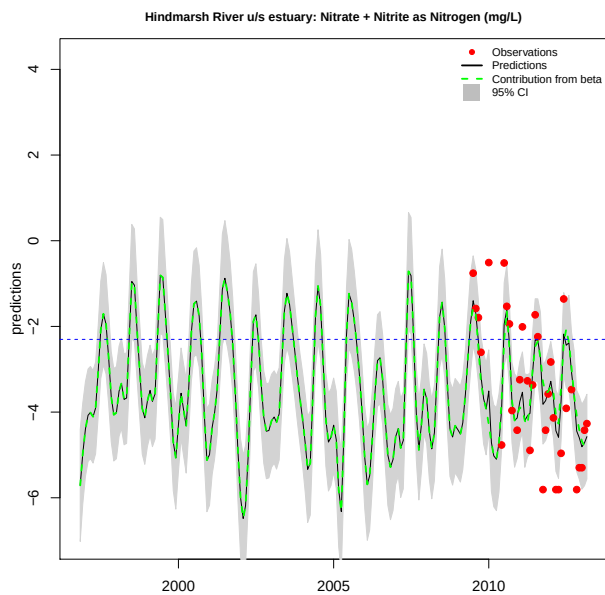
Figure 19 represents the composition of the 4 first principal components in the initial land use matrix coordinate system. It is worth noticing that the component 4 (and 1 almost) is taken directly from the initial coordinate system, which makes the interpretation even easier.

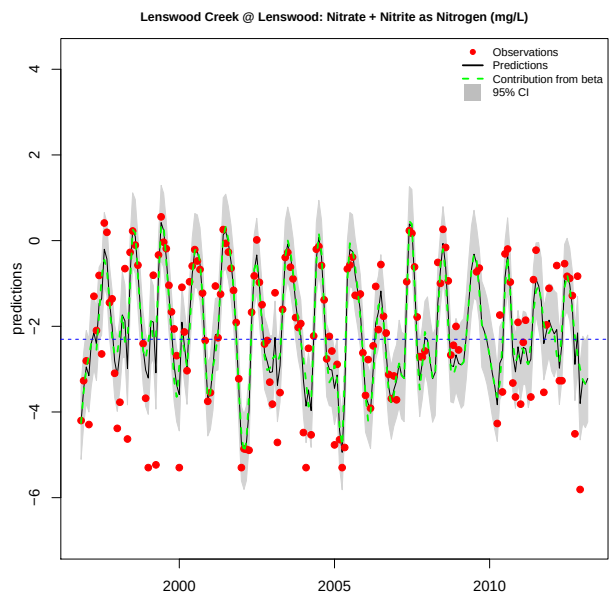
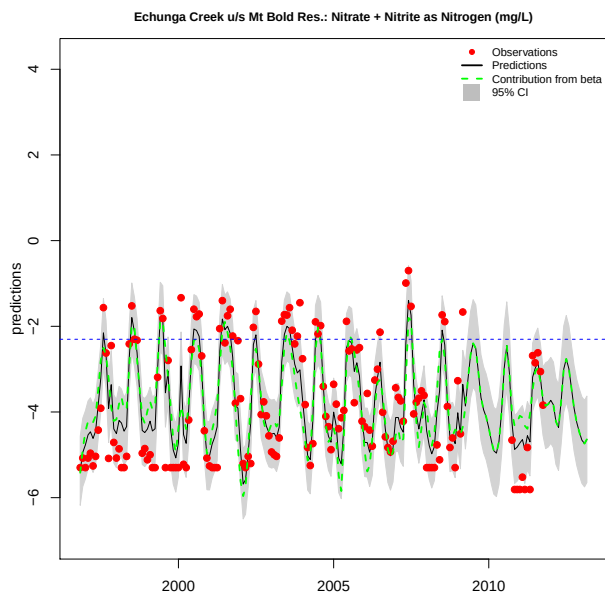
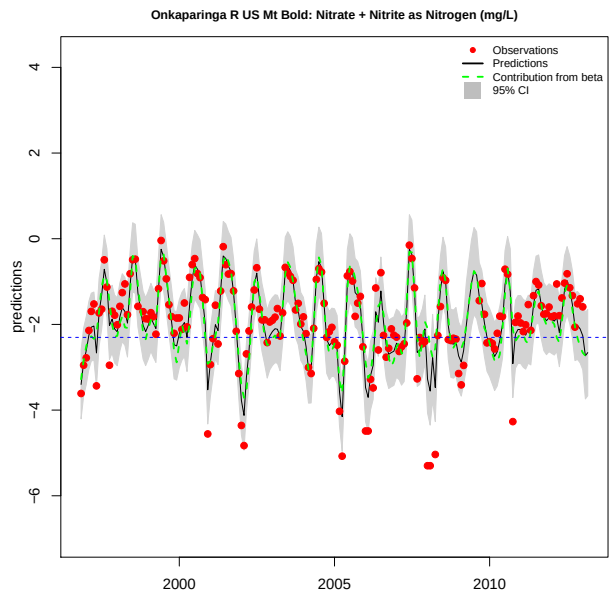
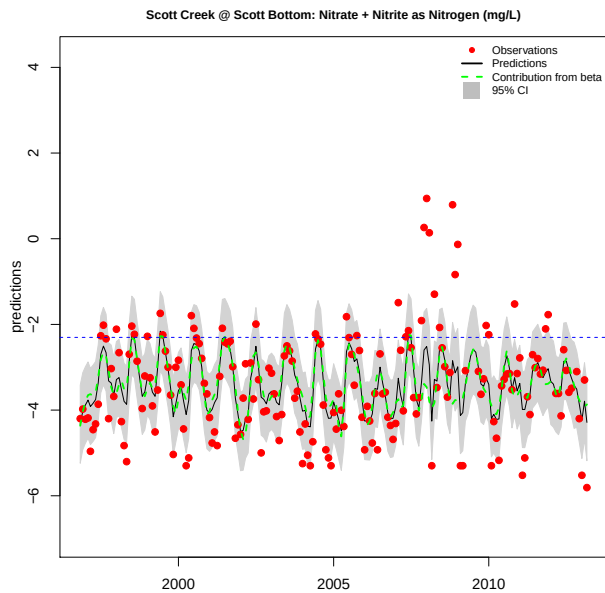
C.4 Results Nitrite + Nitrate as Nitrogen

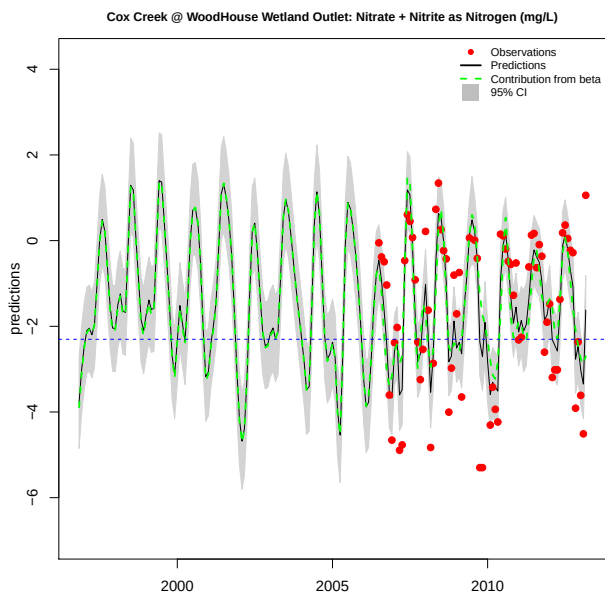
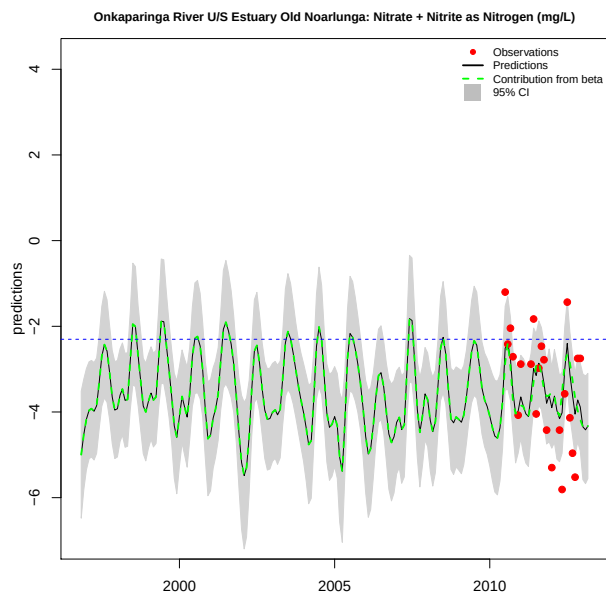
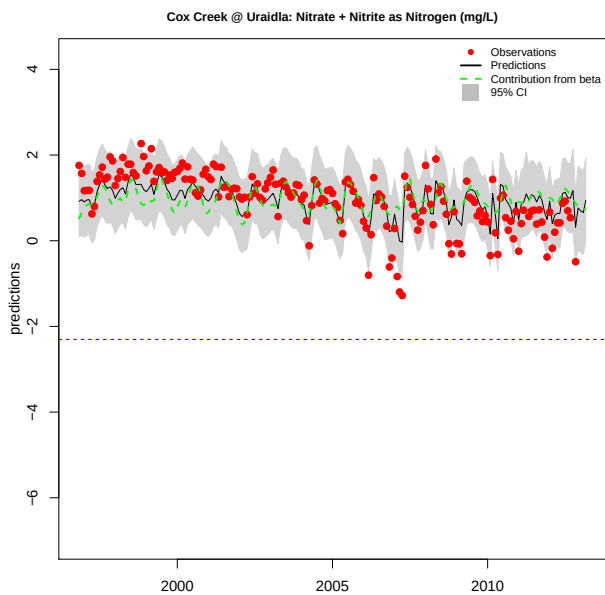
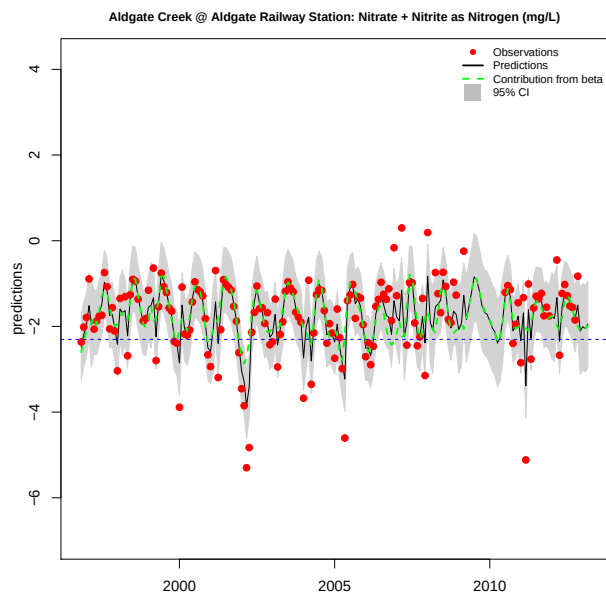
This sub-section lists all the results linked to the Nitrite + Nitrate as Nitrogen parameter. It includes 25 figures. The first eight figures (1 to 8, to be read clockwise from the top left figure of each page) describe the observations (Figure 1), the transformed land use and soil type covariates (Figures 2-5), the occurrences of observation per station (Figure 6), the model fitting summary (Figure 8) and the estimated parameters for the covariates (Figure 7). Then, for each station recording the Nitrite, a plot comparing the observation and the prediction of the model. The ANZECC guideline is shown as a dotted blue line.

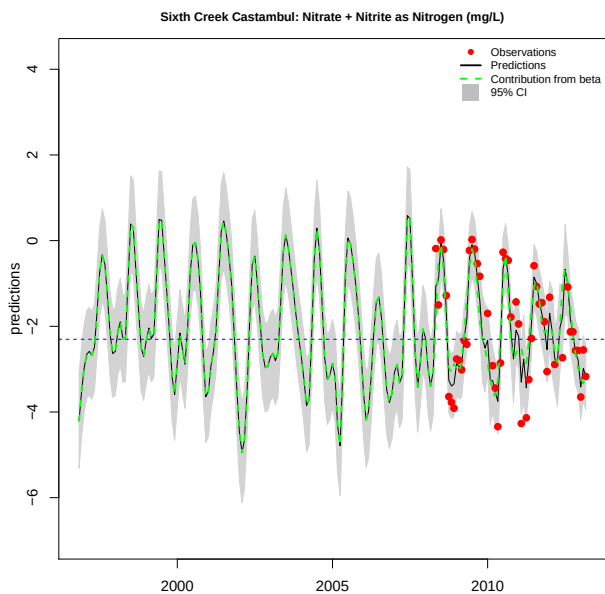
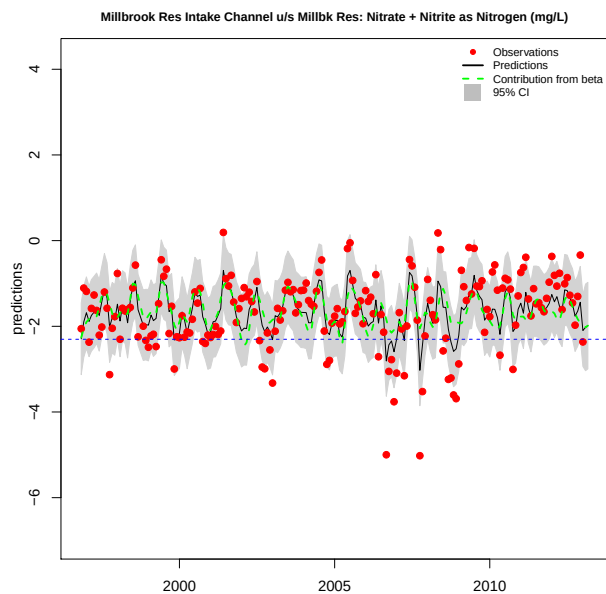
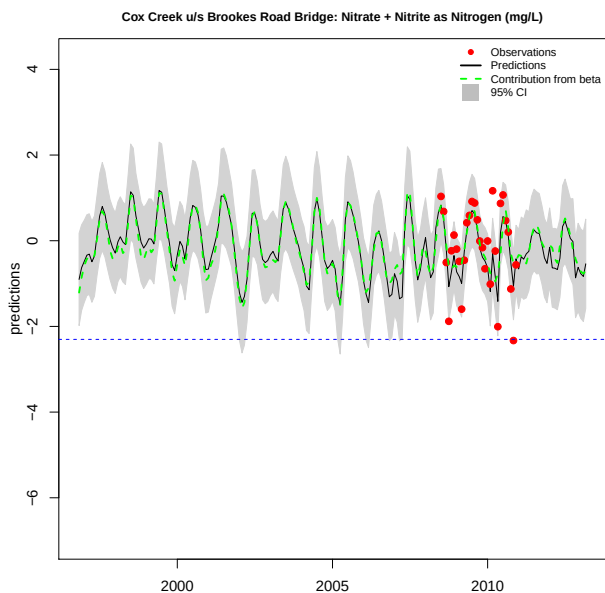
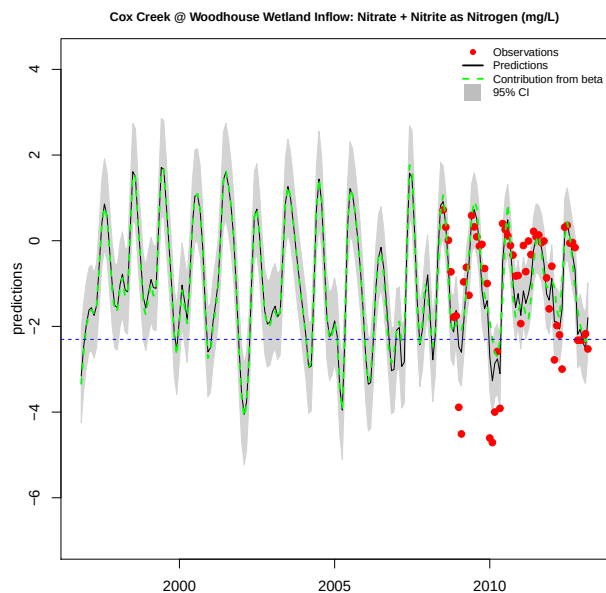


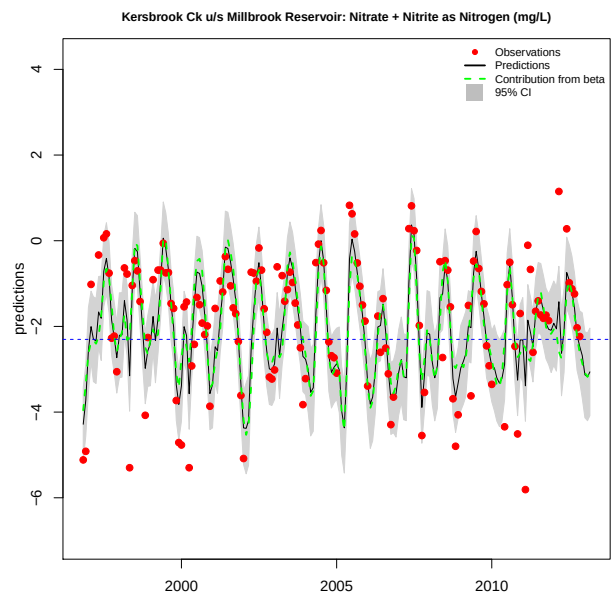






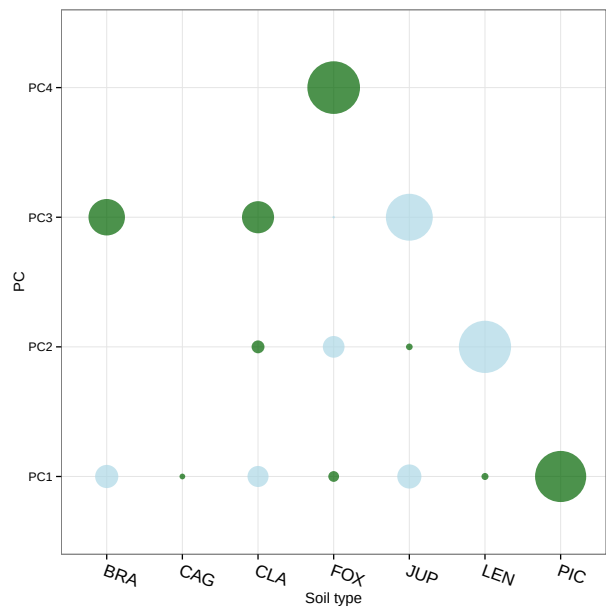
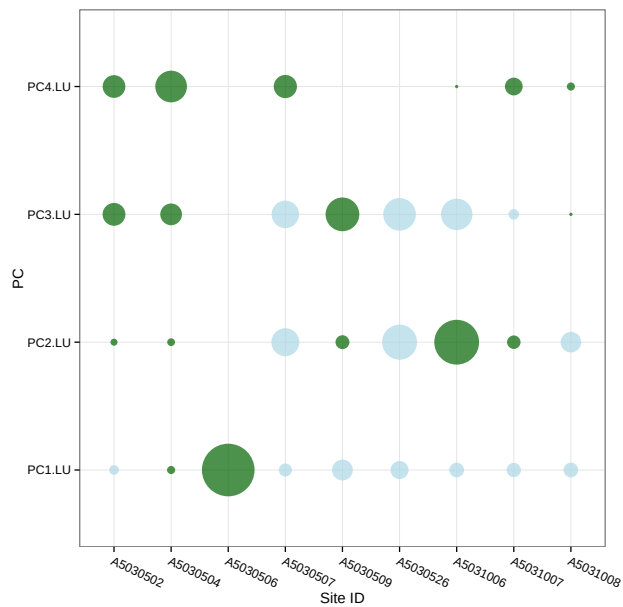
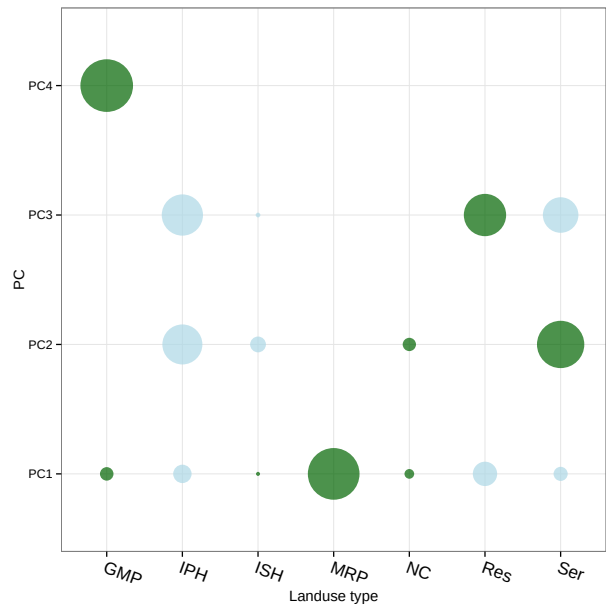
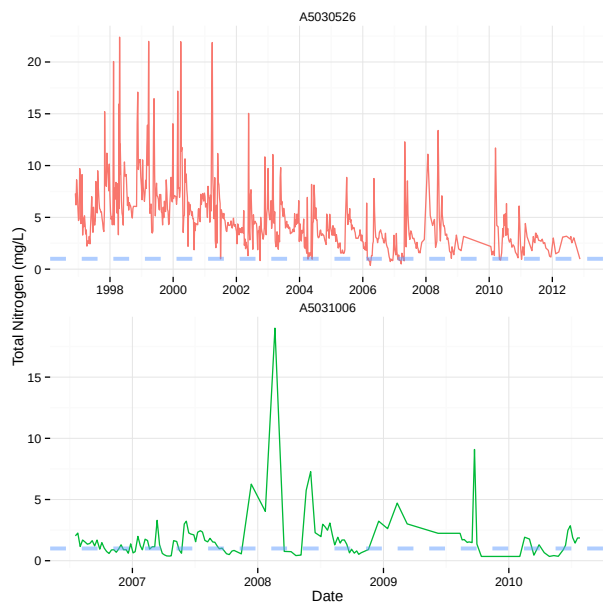


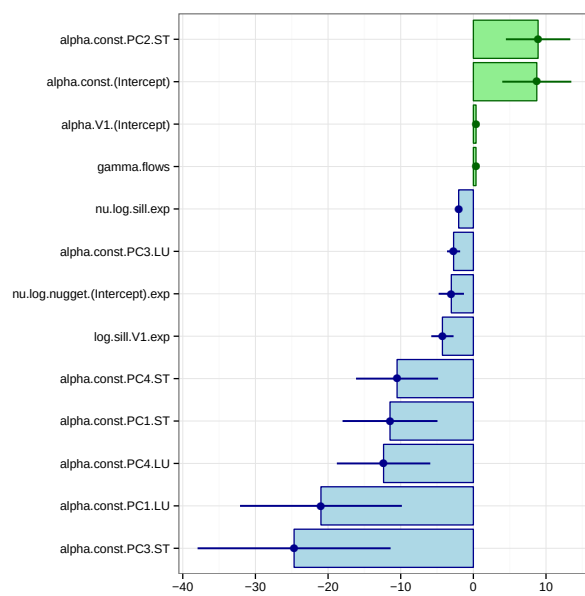
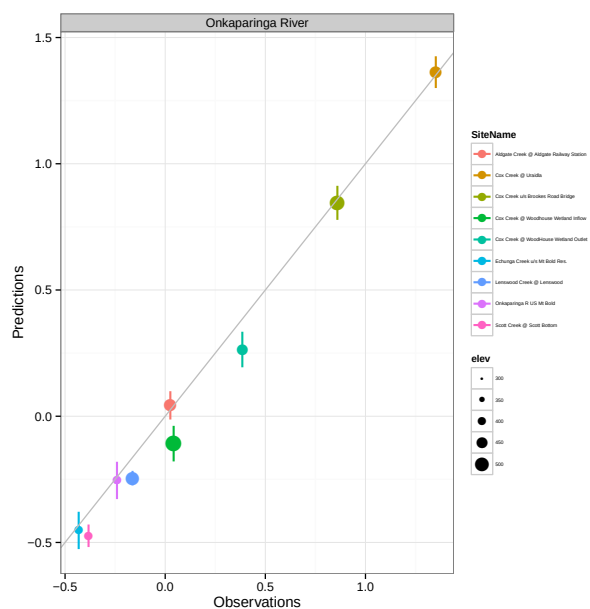
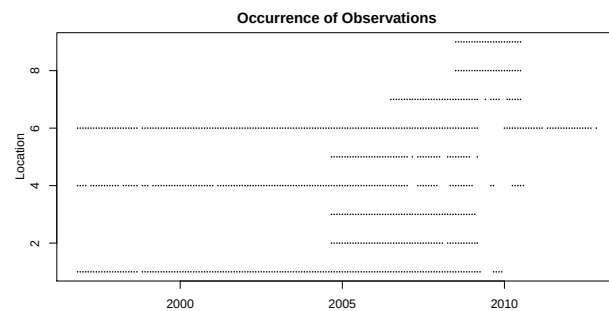
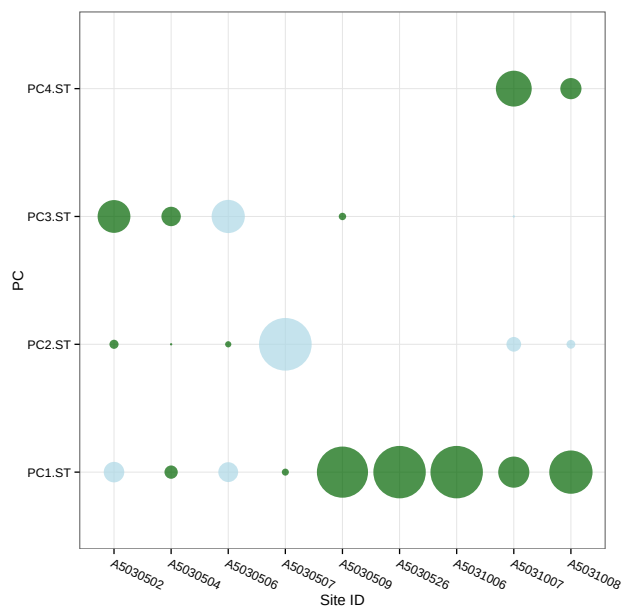


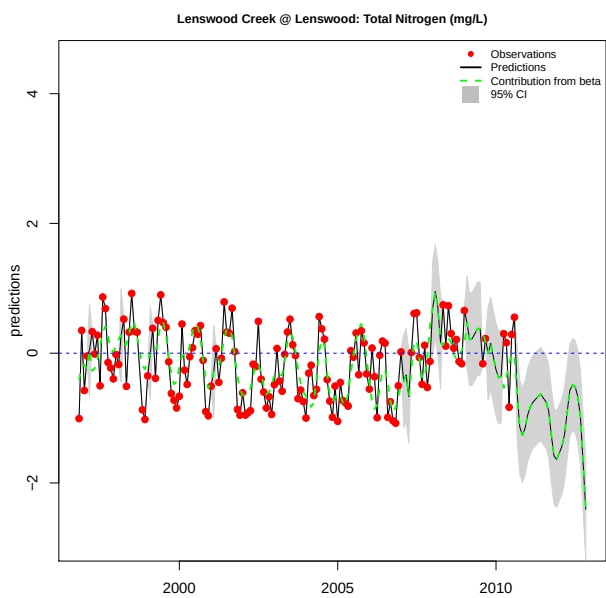
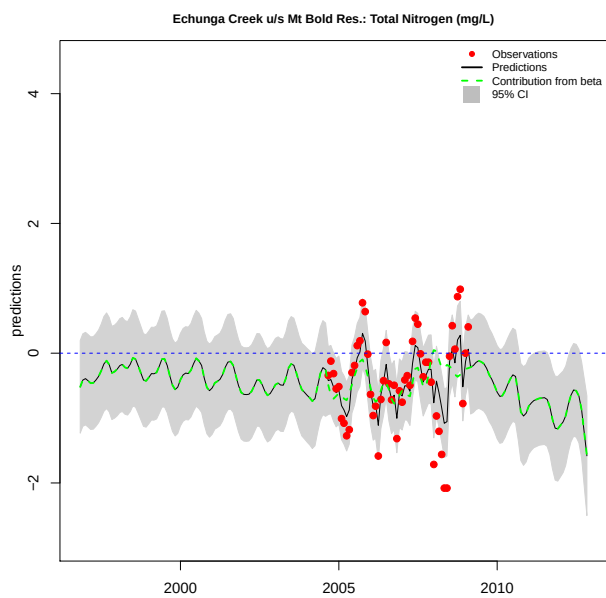
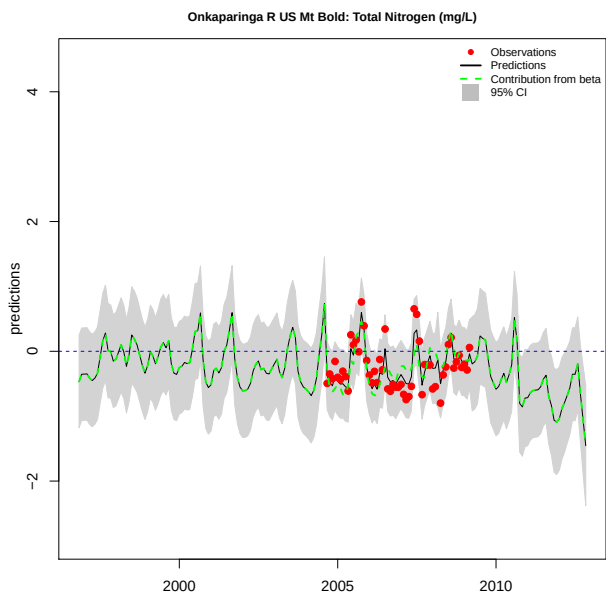
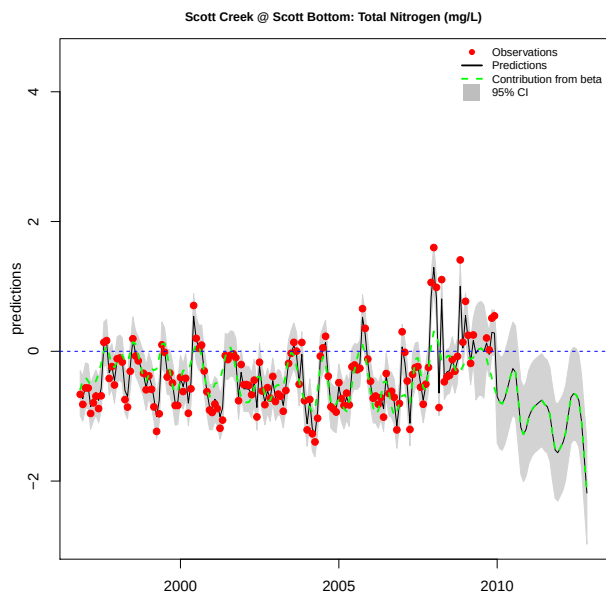


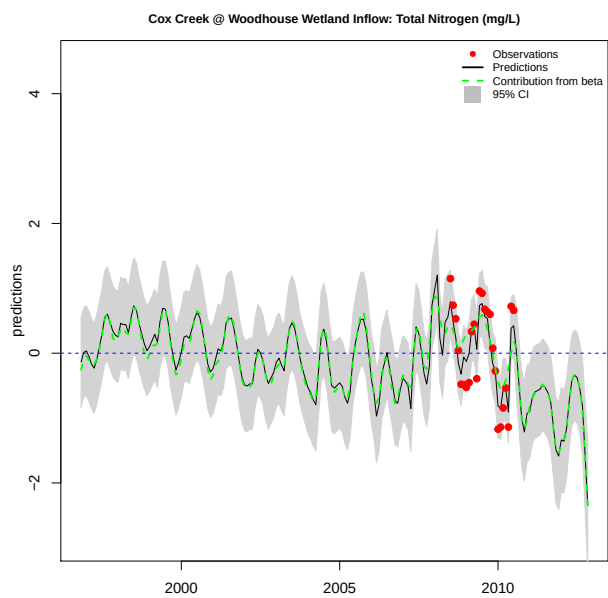
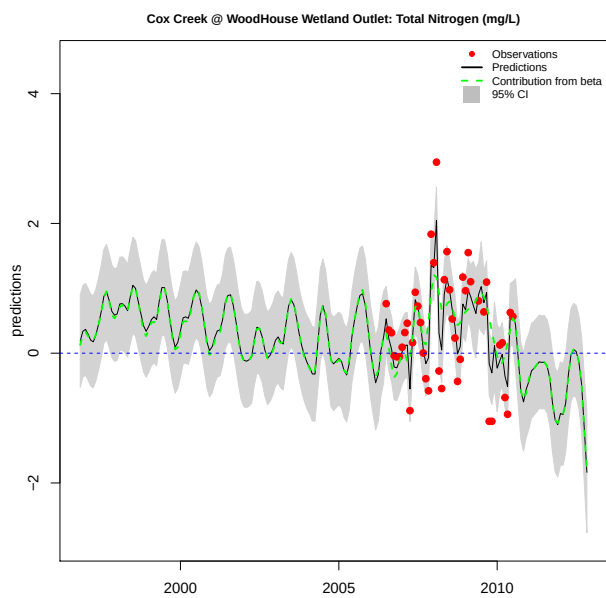
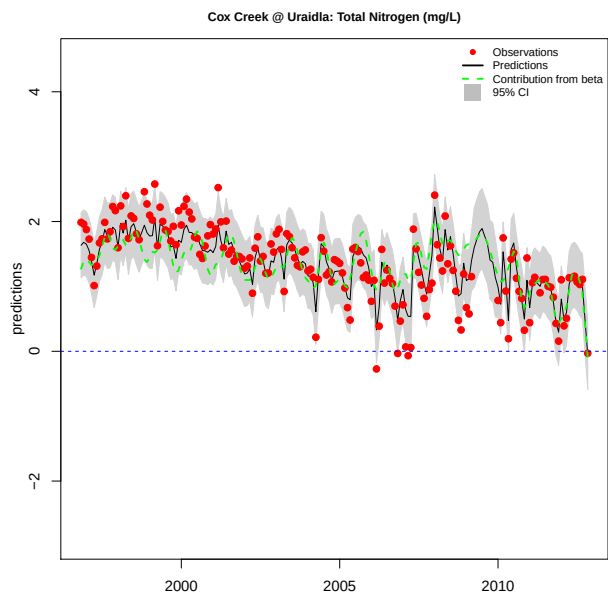
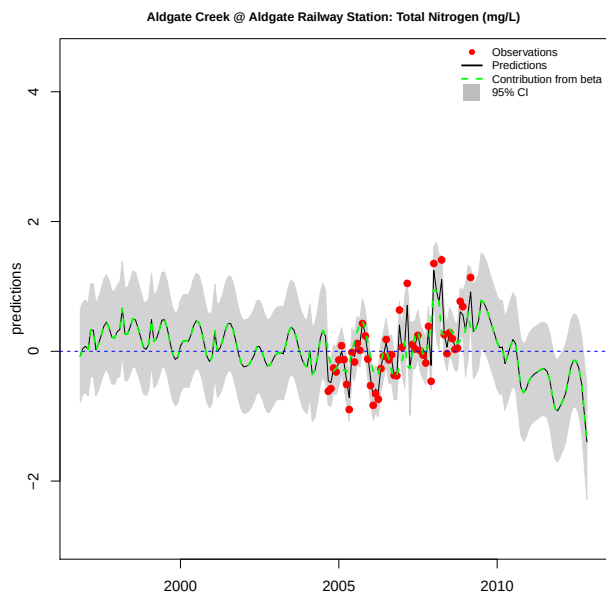
C.5 Results Total Nitrogen

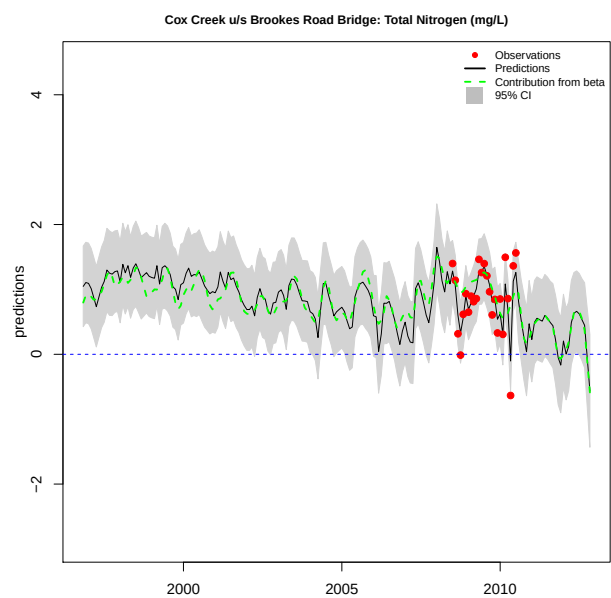
This sub-section lists all the results linked to the Total Nitrogen parameter. It includes 17 figures. The first eight figures (to be read clockwise from the top left figure of each page) describe the observations (Figure 1), the transformed land use and soil type covariates (Figures 2-5), the occurrences of observation per station (Figure 6), the model fitting summary (Figure 8) and the estimated parameters for the covariates (Figure 7). Then, for each station recording the Total Nitrogen, a plot comparing the observation and the prediction of the model. The ANZECC guideline is shown as a dotted blue line.





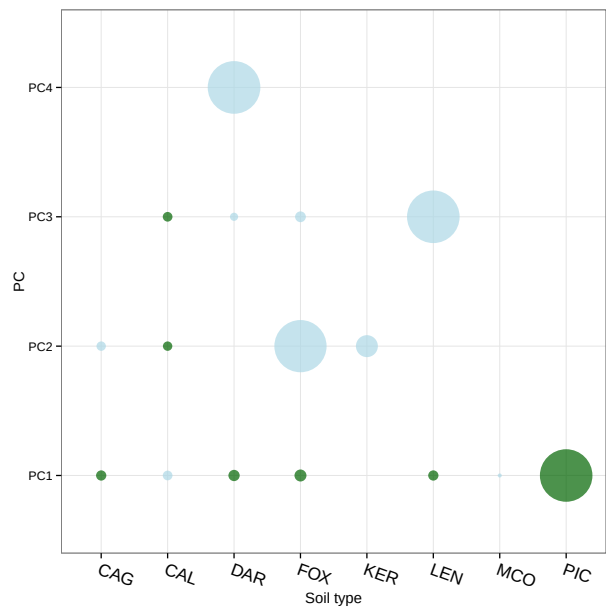
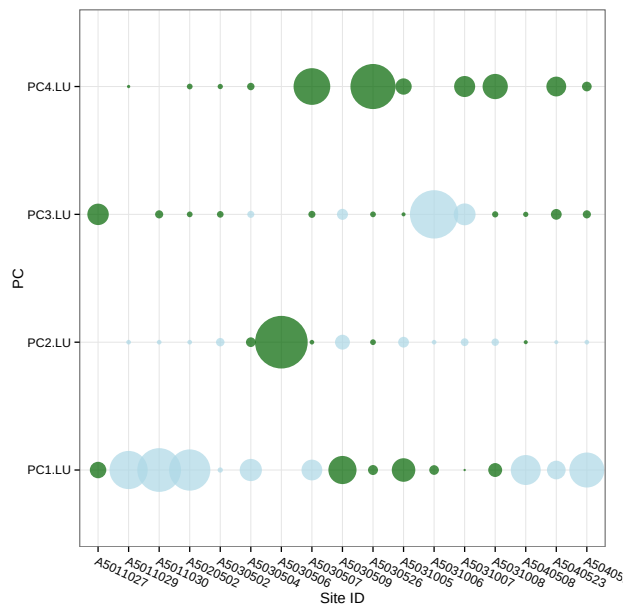
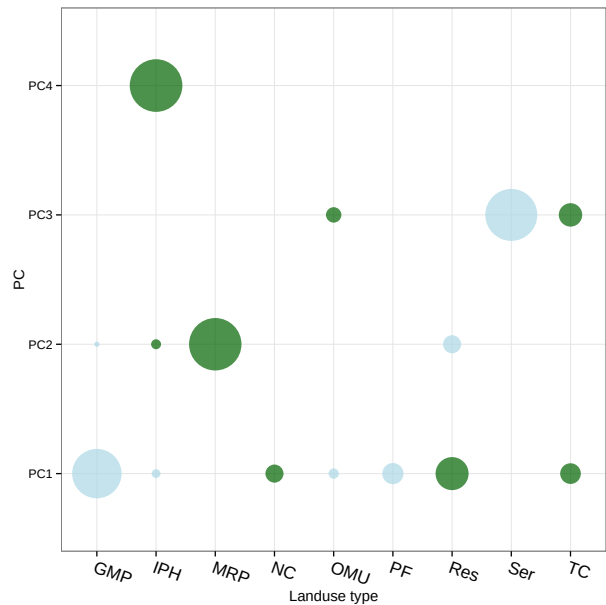
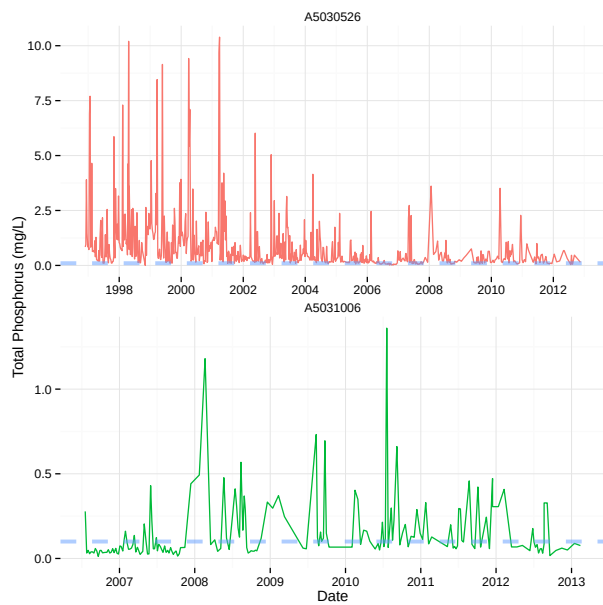


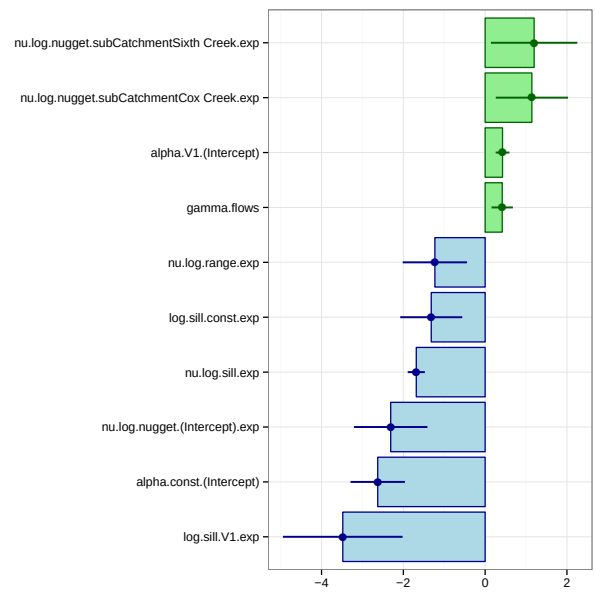
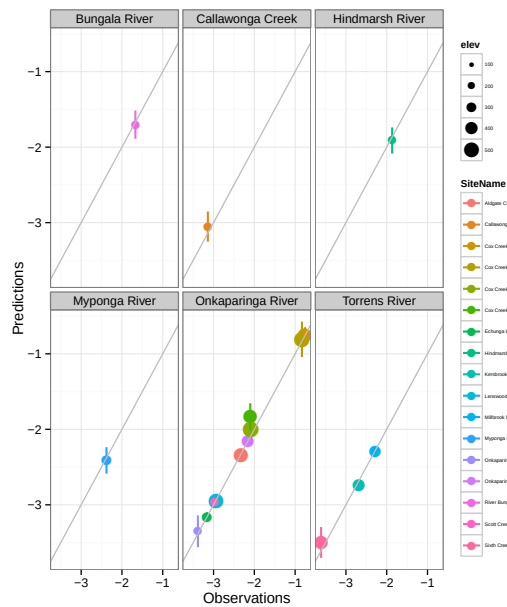
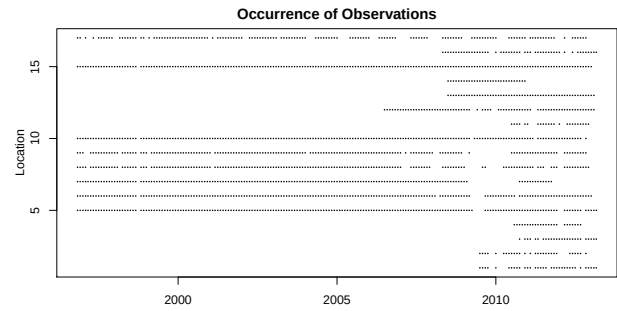
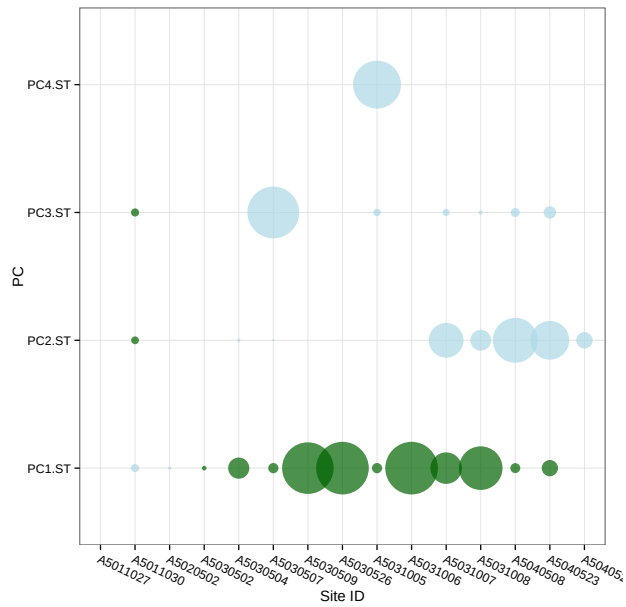


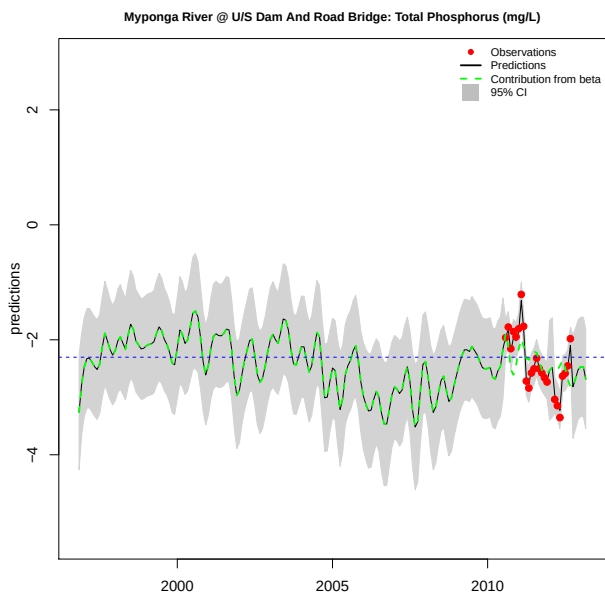
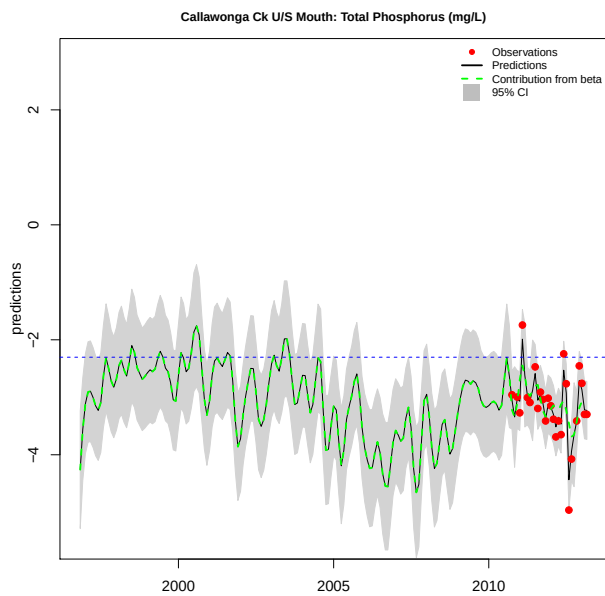
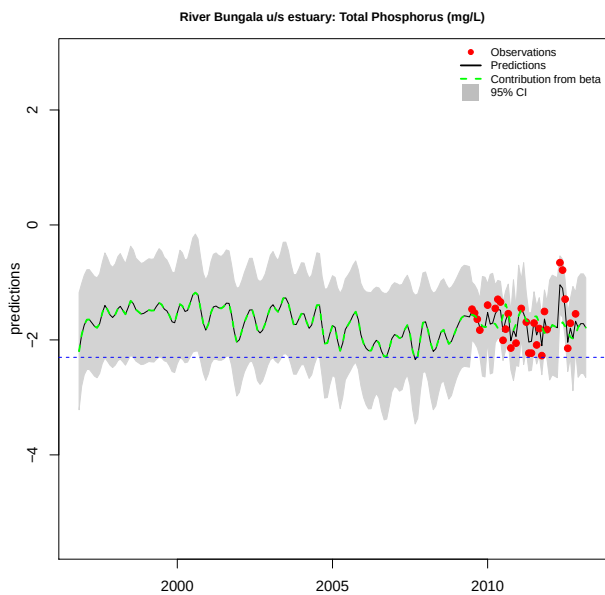
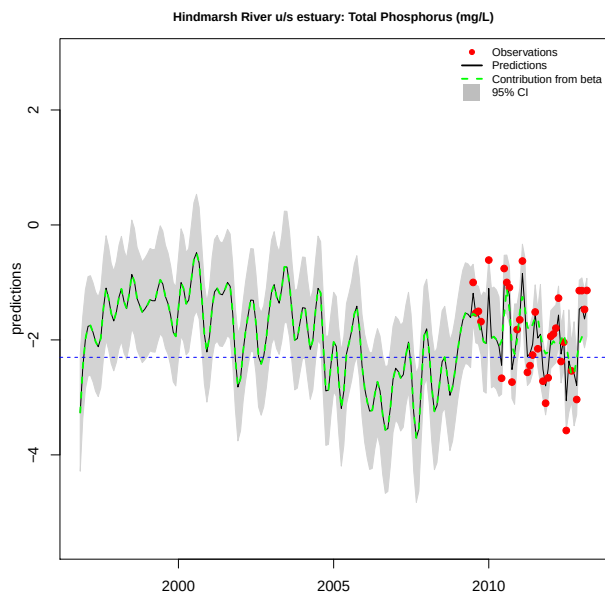


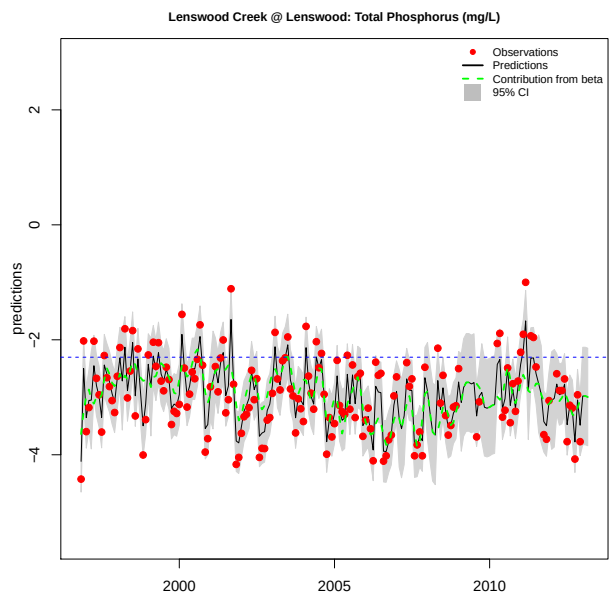
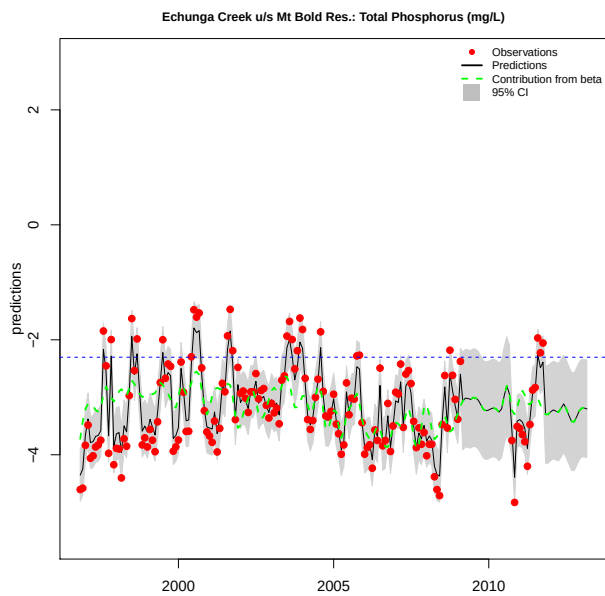
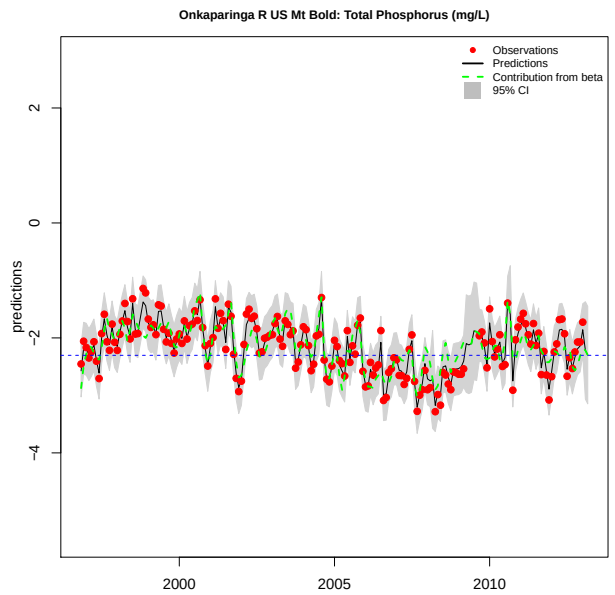
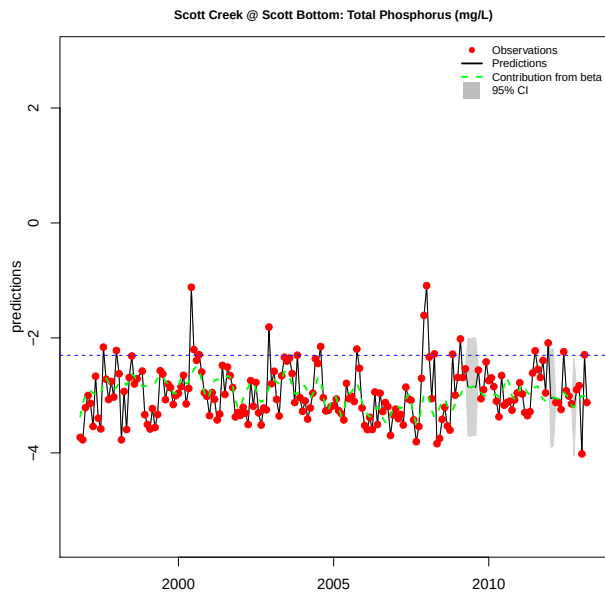
C.6 Results Total Phosphorus

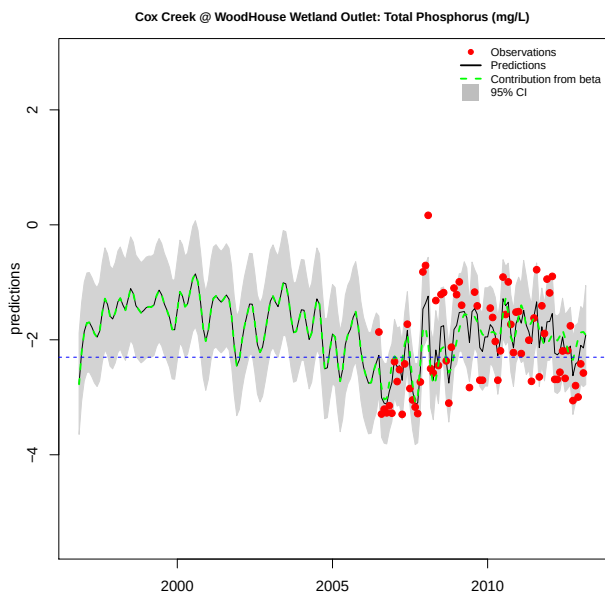
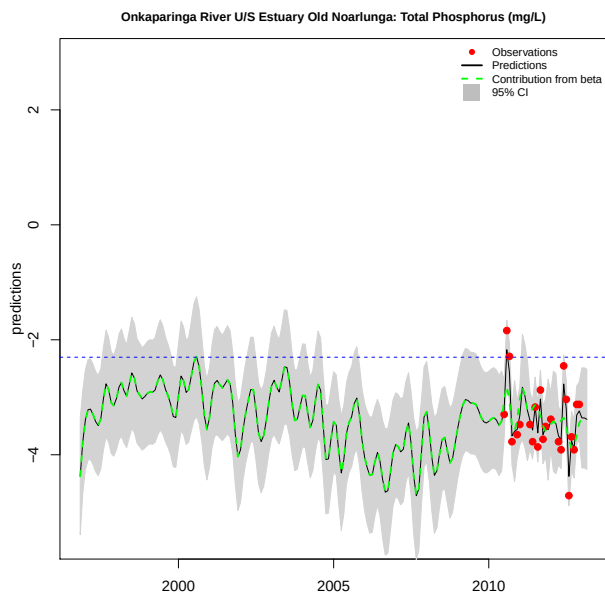
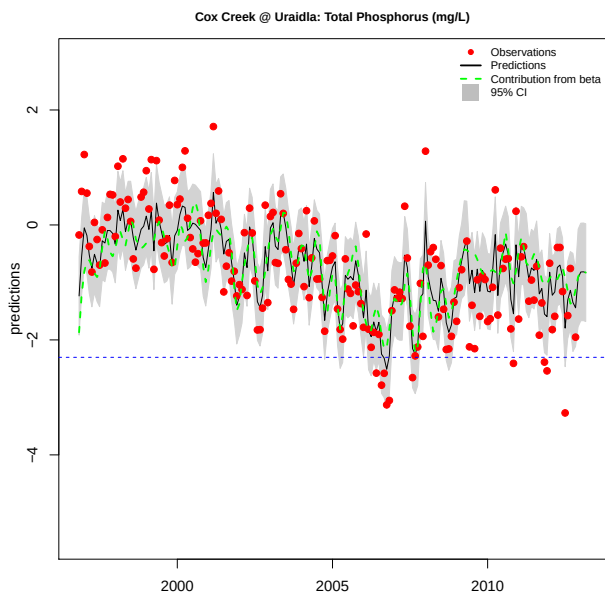
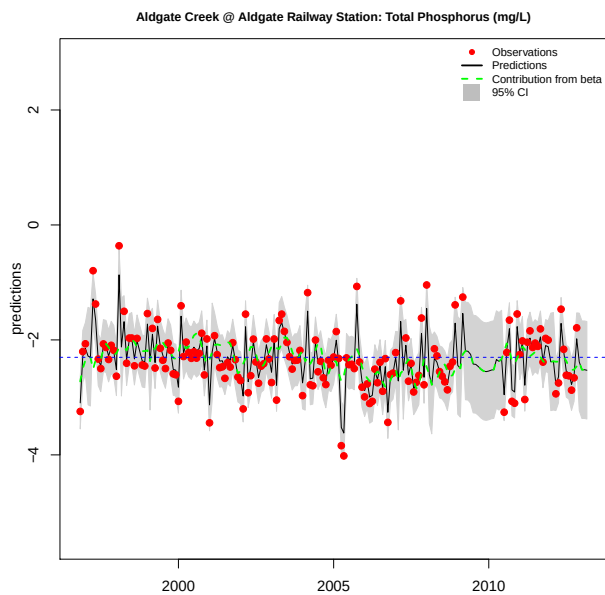
This sub-section lists all the results linked to the Total Phosphorus parameter. It includes 25 figures. The first eight figures (to be read clockwise from the top left figure of each page) describe the observations (Figure 1), the transformed land use and soil type covariates (Figures 2-5), the occurrences of observation per station (Figure 6), the model fitting summary (Figure 8) and the estimated parameters for the covariates (Figure 7). Then, for each station recording the Total Phosphorus, a plot comparing the observation and the prediction of the model. The ANZECC guideline is shown as a dotted blue line.

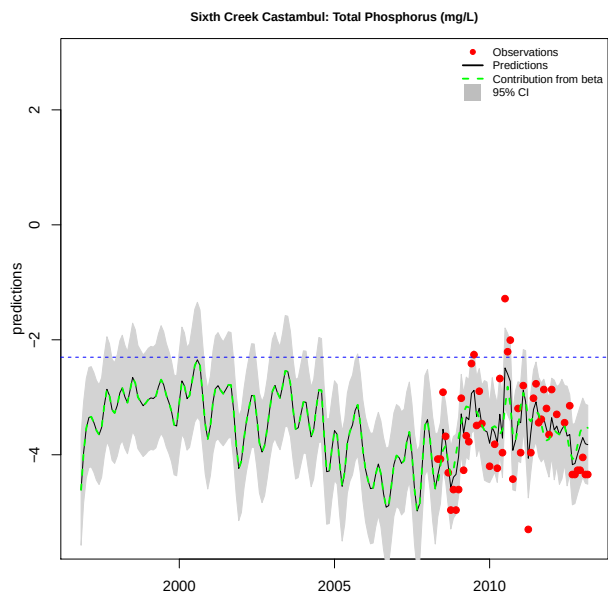
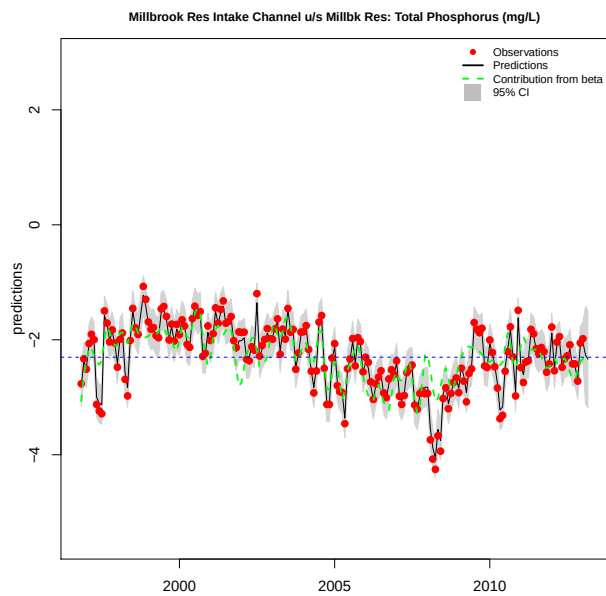
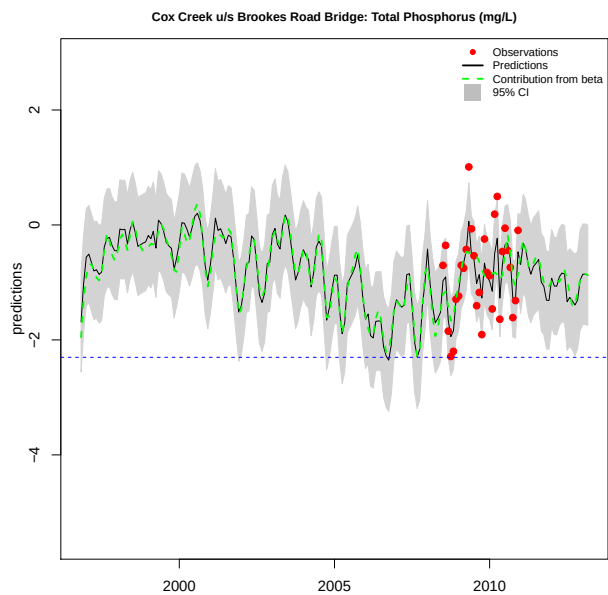
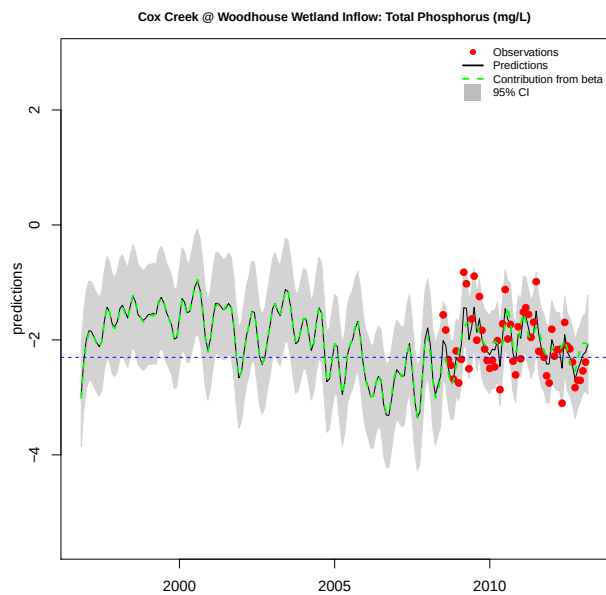


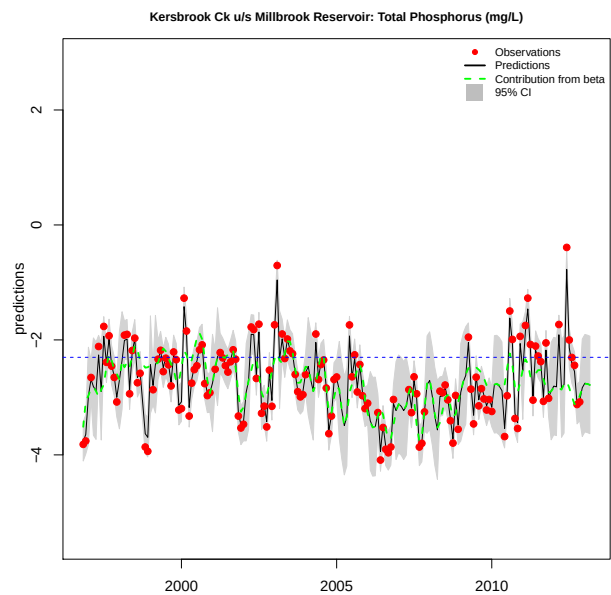






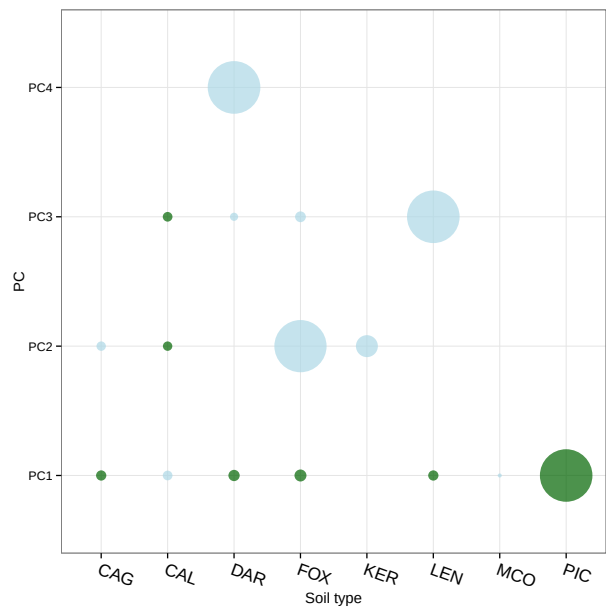
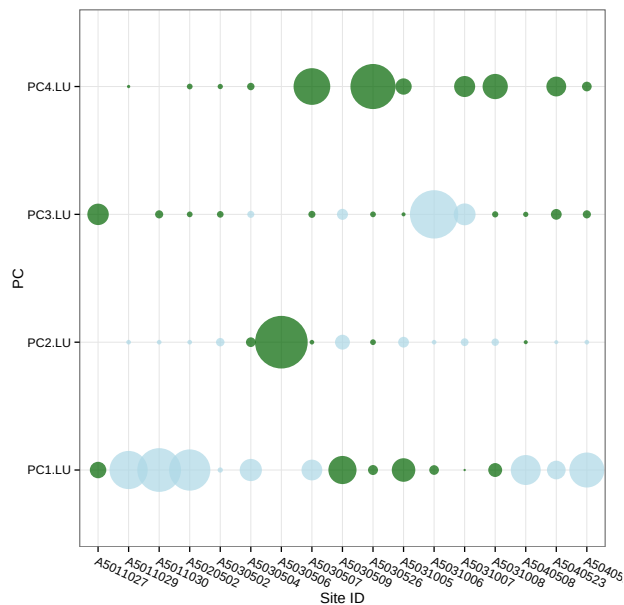
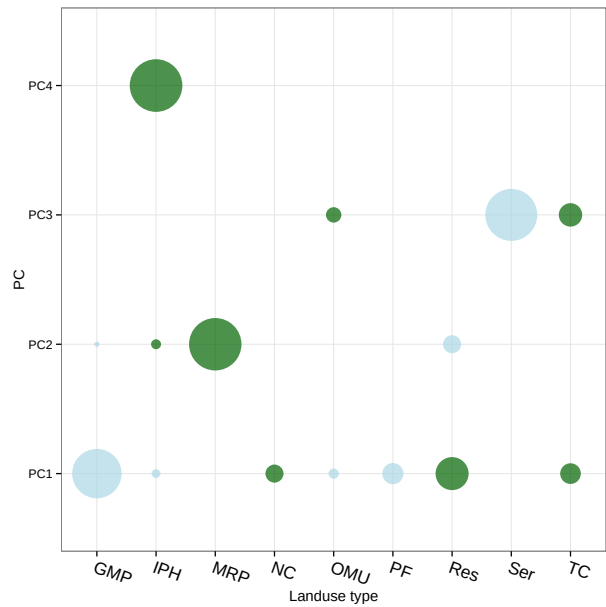


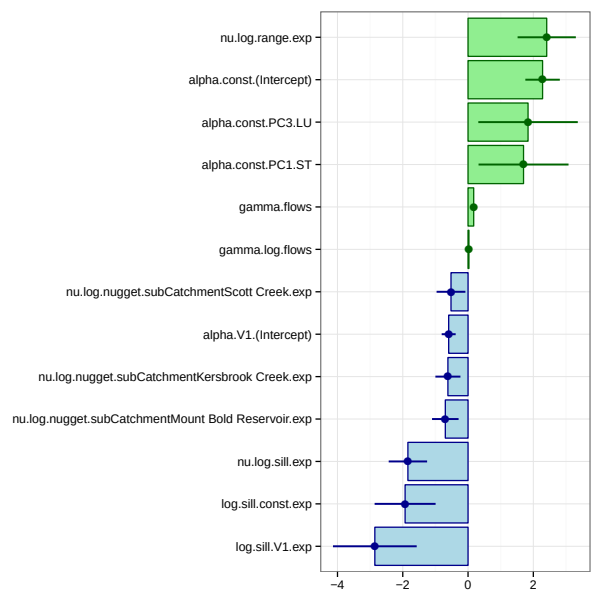
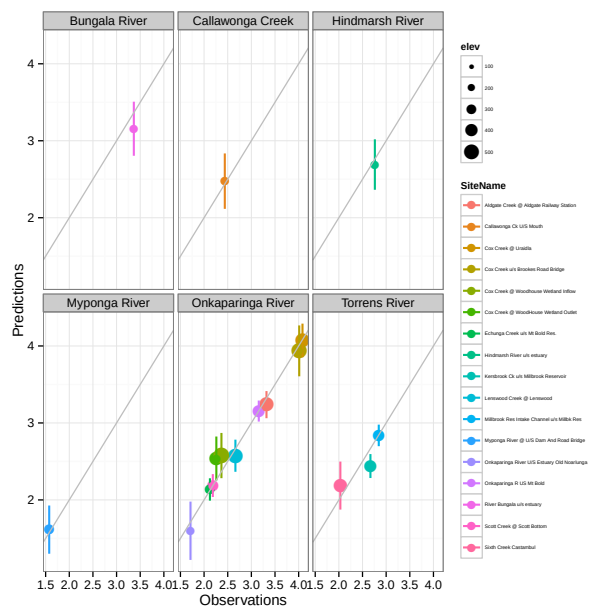
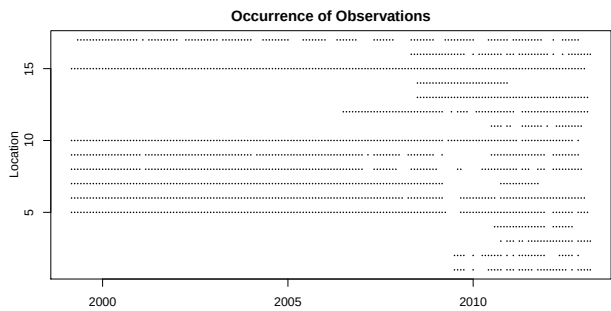
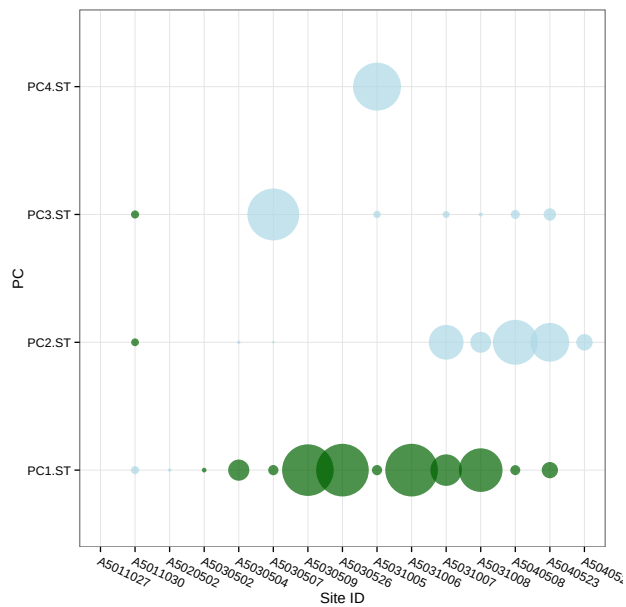


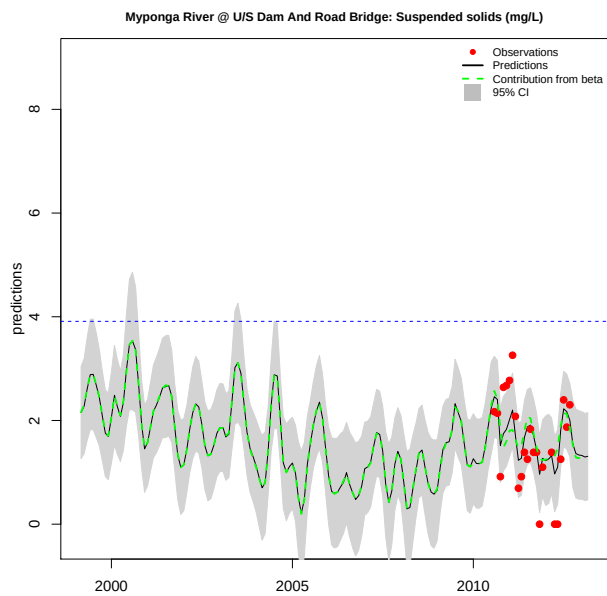
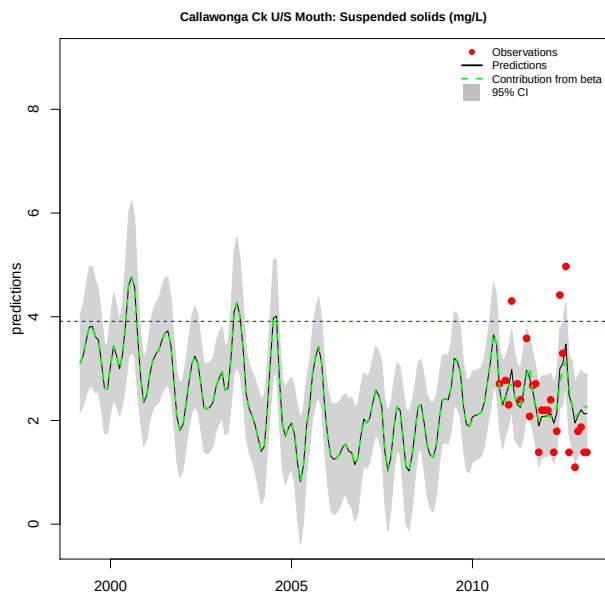
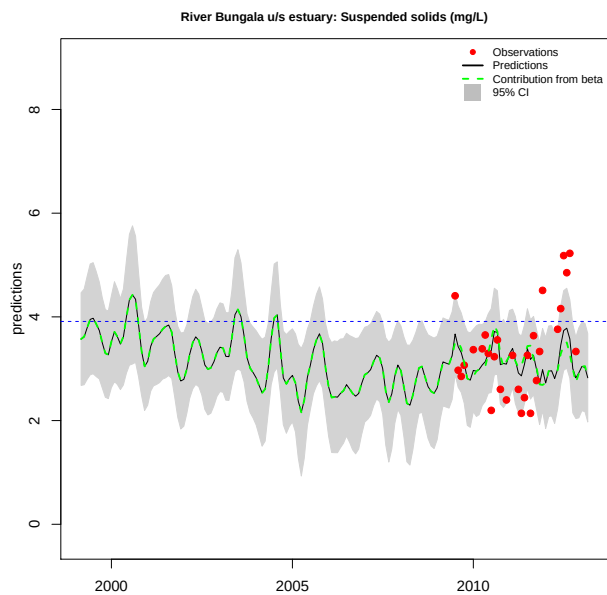
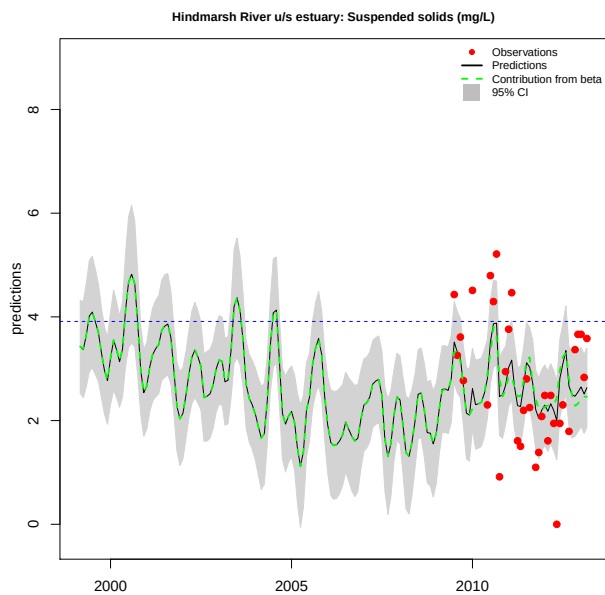


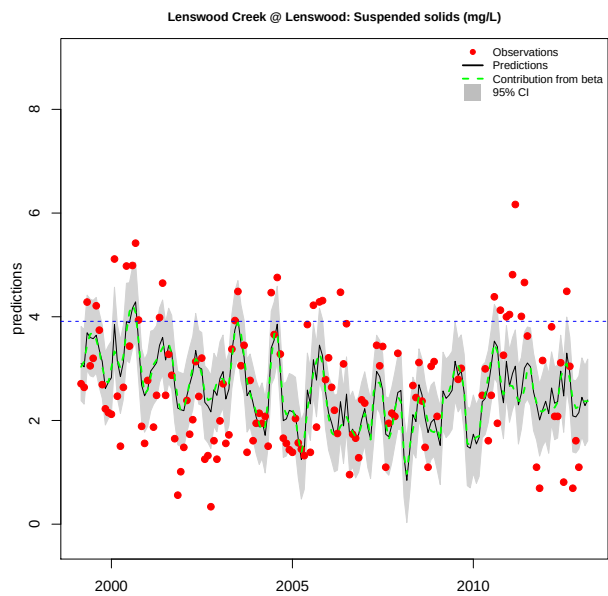
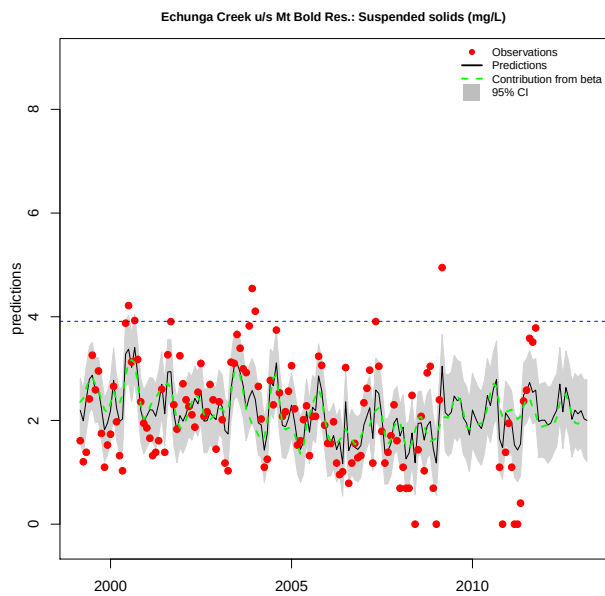
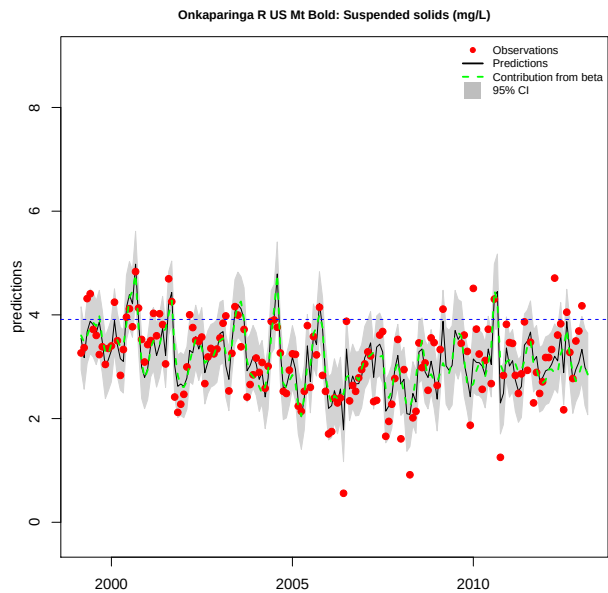
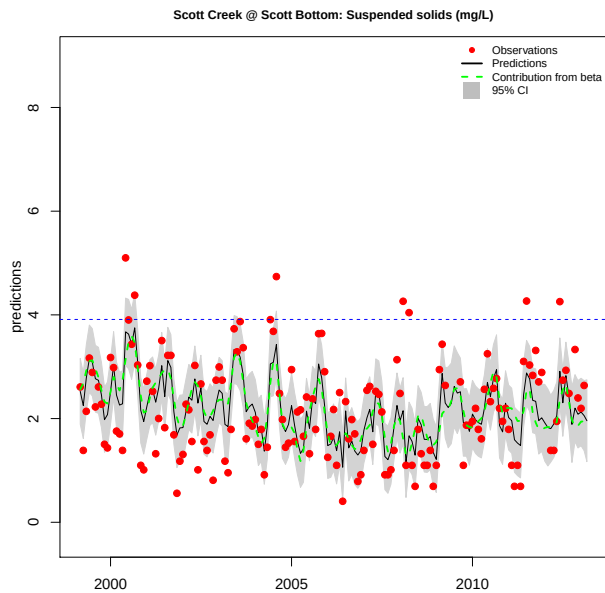
C.7 Results Suspended Solids

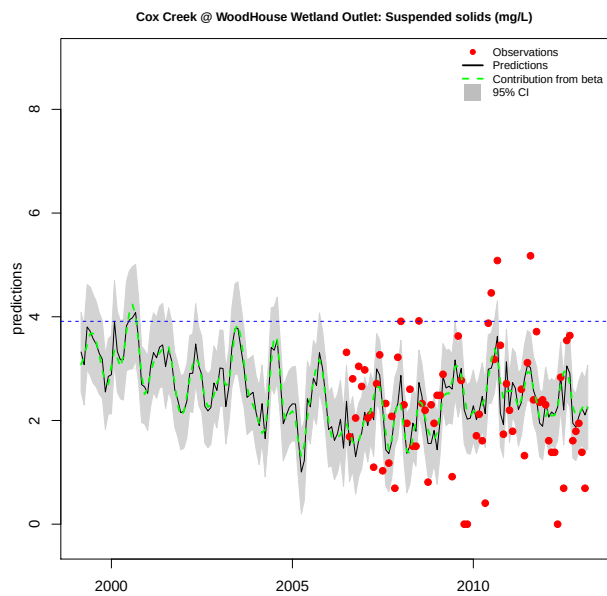
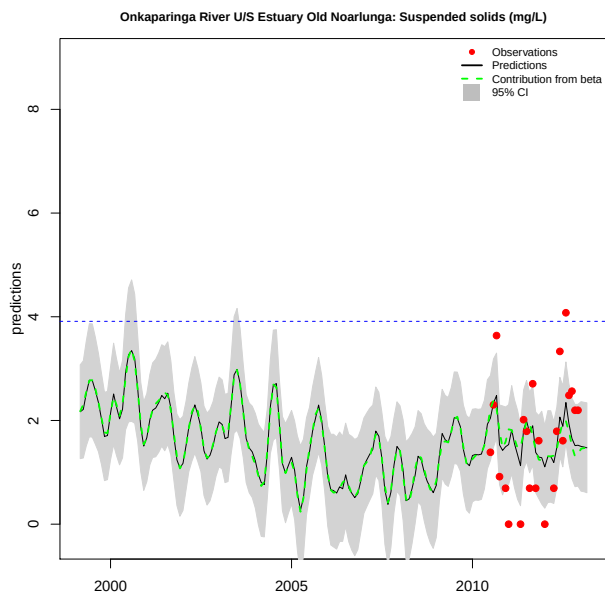
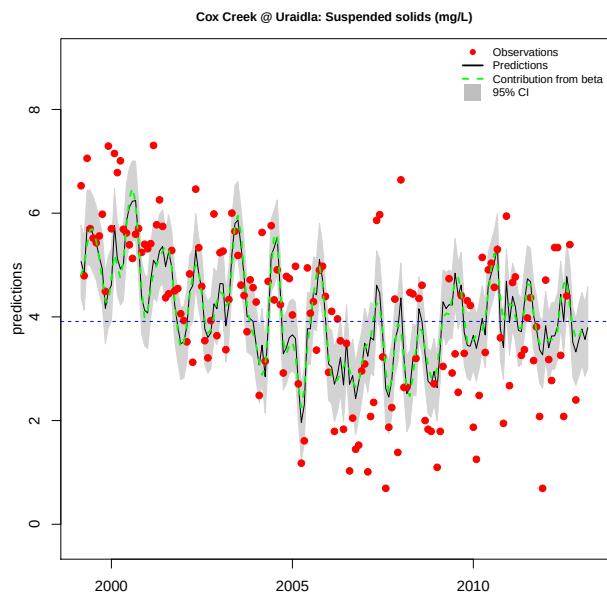
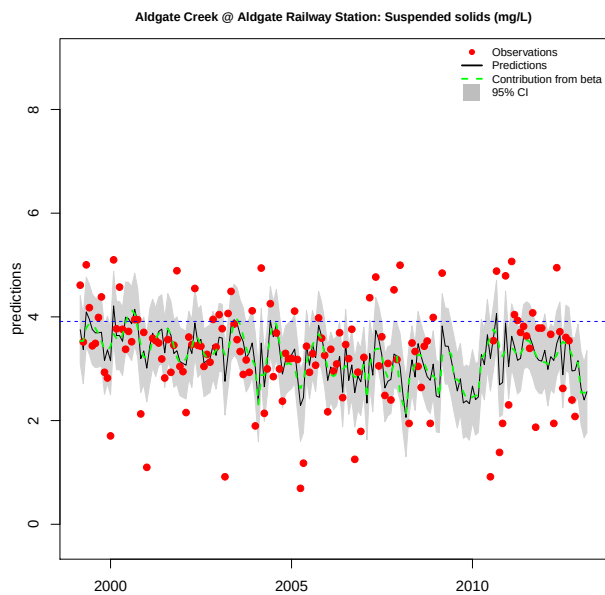
This sub-section lists all the results linked to the Suspended Solids parameter. It includes 25 figures. The first eight figures (to be read clockwise from the top left figure of each page) describe the observations (Figure 1), the transformed land use and soil type covariates (Figures 2-5), the occurrences of observation per station (Figure 6), the model fitting summary (Figure 8) and the estimated parameters for the covariates (Figure 7). Then, for each station recording the Suspended Solids, a plot comparing the observation and the prediction of the model. The ANZECC guideline is shown as a dotted blue line.

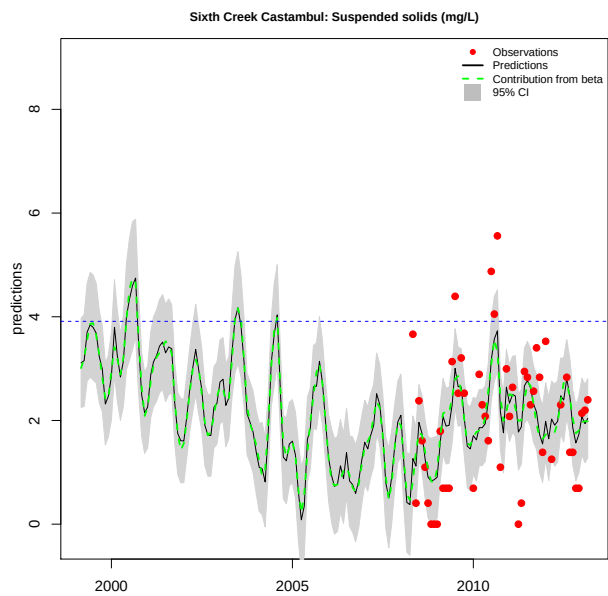
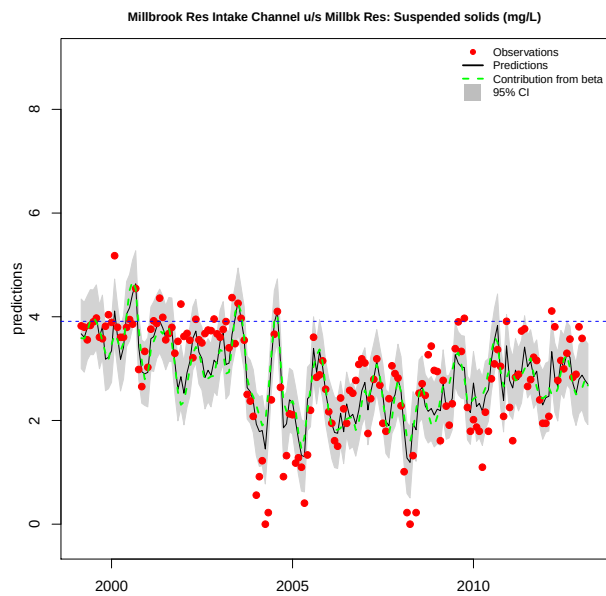
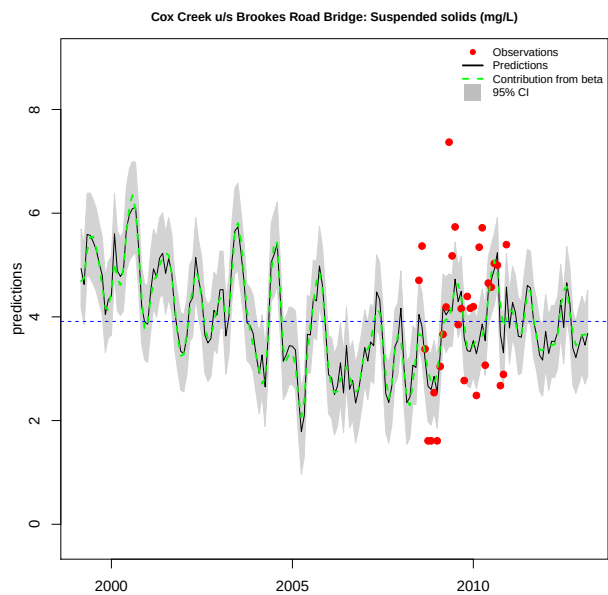
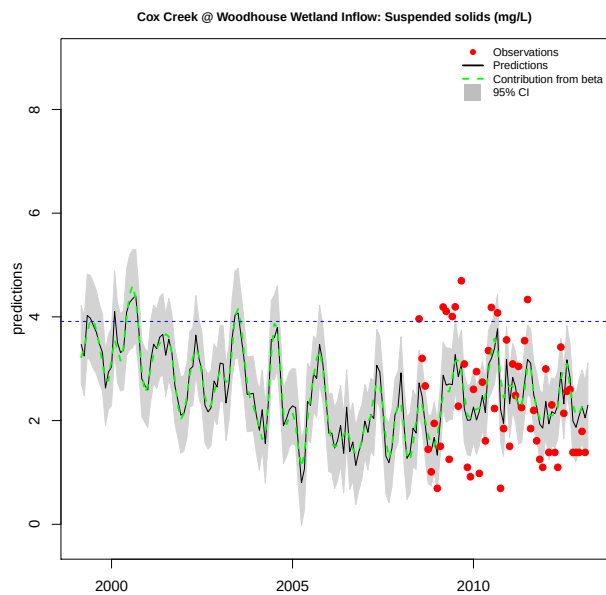


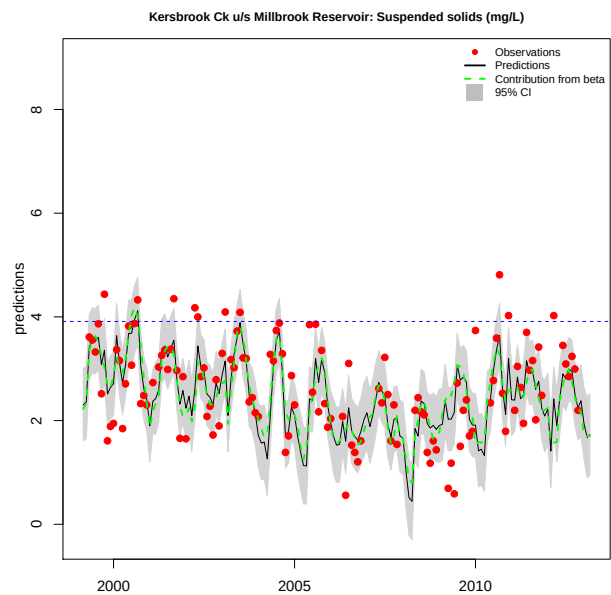


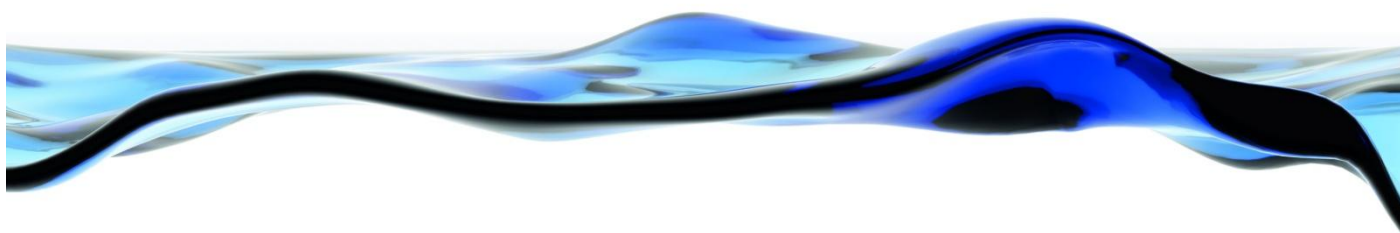












**Government
of South Australia**

Department of Environment,
Water and Natural Resources



THE UNIVERSITY
of ADELAIDE



University of
South Australia



Flinders
UNIVERSITY
ADELAIDE • AUSTRALIA

The Goyder Institute for Water Research is a partnership between the South Australian Government through the Department of Environment, Water and Natural Resources, CSIRO, Flinders University, the University of Adelaide and the University of South Australia.