Using a data-driven approach to understand the interaction between catchment characteristics and water quality responses

A. Lintern[1], J.A. Webb[1], D. Ryu[1], S. Liu[1], U. Bende-Michl[4], P. Leahy[2], P. Wilson[3], A.W. Western[1]

[1]Department of Infrastructure Engineering, The University of Melbourne, Parkville, VIC, 3010. Email: a.lintern@unimelb.edu.au
[2] Applied Sciences Group, EPA Victoria, Macleod, VIC, 3085
[3]Department of Environment, Land, Water & Planning, East Melbourne, VIC, 3002
[4]Bureau of Meteorology, Melbourne, VIC, 3001

Key Points

- A Bayesian hierarchical modelling approach will be trialed to develop a predictive model of TN concentrations in rivers;
- spatial variability in TN concentrations throughout Victoria correlates positively with catchment fertilizer application rates, and negatively with catchment forest cover;
- the correlation between discharge and TN concentrations varies spatially; and
- this variability correlates with catchment climatic (temperature, rainfall) and topographic (catchment size, catchment elevation) characteristics of the water quality monitoring sites.

Abstract

To guide future management decisions, models are required to predict riverine water quality and how this is driven by catchment characteristics and management practices. However, there remains a lack of understanding of the importance of drivers of water quality variation between catchments. Here, we propose the use of a Bayesian hierarchical modelling approach to build predictive models of water quality. As an example, we present exploratory data analysis of Total Nitrogen (TN) concentrations in Victorian rivers to illustrate the catchment characteristics that appear to be influencing its spatial variability. These key characteristics were determined using frequentist statistical analyses using: (1) TN concentrations measured at 28 water quality monitoring sites and (2) the climatic, land use and topographic characteristics of these sites. Spatial variability in TN observations correlated positively with catchment fertilizer use and average temperature. TN concentrations correlated negatively to catchment forest cover, annual runoff, runoff perenniality, soil erosivity and catchment slope. The relationship between discharge and TN concentrations varied significantly between water quality sites, and we believe this spatial variability could be a result of climatic and topographic differences between the sites. The important spatial variables identified in this investigation will be used to develop the Bayesian hierarchical model.

Keywords

Catchment modelling, explanatory variables, nitrogen, spatio-temporal variability, statistical model, surface water quality.

Introduction

Across the world, a large number of streams, lakes and estuaries are experiencing degraded water quality, including increased salinity, sediment and nutrient levels (Nash, 1993). Degraded water quality in aquatic environments not only undermines ecosystem health, but can also have significant economic and social repercussions (e.g., Smith, VH et al., 1999). To develop successful mitigation strategies, we need models that can predict the water quality of aquatic systems based on catchment characteristics such as land use and management practices.

Although predictive catchment models have previously been developed, existing water quality modelling tools are: (1) statistical or empirical models (Hasani Sangani et al., 2014; Smith, RA et al., 1997) that can only be applied to the catchments for which they have been developed or calibrated; or (2) physically-based distributed models (Argent et al., 2009; Arnold and Fohrer, 2005) that require extensive datasets for their implementation. Furthermore most analysis has concentrated on temporal changes in water quality rather than identifying the causes of spatial differences in water quality. As such, existing models tend to perform poorly when applied to new catchments or catchments with limited monitoring data. We propose that these weaknesses in current models can be overcome using a Bayesian hierarchical modelling approach. The hierarchical or multi-level model structure will enable the inclusion of temporal and spatial influences at multiple scales on water quality (e.g., Webb and King, 2009). By using a Bayesian approach, the stochasticity inherent in water quality observations can be incorporated into the model (Clark, 2005). Spatial data of catchment characteristics and stream water quality data collected in river basins throughout Victoria over two decades, and Queensland (the Great Barrier Reef catchment) since 2006 will be used to develop the model.

This paper presents some of the preliminary data analysis conducted to inform development of the Bayesian hierarchical model. We use Total Nitrogen (TN) as a case study, as many Australian aquatic environments are nitrogen-limited (Harris, 2001). Specifically, the objective of this investigation is to identify the key spatial characteristics associated with changes in TN concentrations in Victorian rivers. The key factors are identified by frequentist statistical analysis of Victorian river water quality and spatial catchment characteristics data. We envisage that this investigation will illuminate the important spatial variables that should be incorporated in the Bayesian hierarchical predictive water quality model.

Methods

Water quality data from Victorian rivers between August 2010 and December 2014 were obtained from the Victorian Water Measurement Information System (Department of Environment, Land, Water and Planning Victoria, 2016). TN concentrations were calculated as the sum of Total Kjeldahl Nitrogen (TKN) and Nitrate and Nitrate (NO_x). 28 water quality monitoring sites, for which TN concentrations between August 2010 and December 2014 were available, were included in the statistical analysis (shown in Figure 1). These water quality monitoring sites are situated in independent catchments, which vary in size from 27 km^2 to 11027 km^2, and there is no transfer of surface flow from one site to the other. These sites are all located at or near the catchment outlet, and as such, reflect overall catchment characteristics. Due to the non-normality of the TN concentration data (assessed using the Shapiro-Wilk test, α=0.05), all water quality data were log-transformed to meet normality assumptions of the statistical analysis. Hereafter, when we mention analysis of TN values, we are referring to the transformed data. Mean daily discharge data were also extracted from the database and log-transformed.

All catchment characteristics data were obtained from the National Environmental Stream Attributes dataset (version 1.1.5) (Geoscience Australia, 2011). This database provides catchment characteristics (e.g., land use/land cover, climate, erosivity, topography, geology) for stream segments that are 1 to 20 km long, and for the catchments of these stream segments. The stream segments associated with each water quality monitoring site was determined and spatial data of these stream segments were extracted from the database.

First, we assessed whether there were significant differences in the mean TN concentrations observed at the 28 water quality sites using the Kruskal-Wallis test (α=0.05). We then attempted to explain these differences by identifying the spatial variables (i.e., land use/land cover, climate, topographic, discharge and geologic characteristics of the stream segments and their catchments) that correlate with spatial variability in mean TN concentrations observed at each of the 28 sites. The Spearman Rank correlation coefficient (α=0.05) was used for this analysis.

Second, we identified the relationship between discharge and TN at each of the 28 water quality monitoring sites using the Spearman Rank correlation coefficient (α=0.05). We explored how these relationships varied spatially by calculating the Spearman Rank correlation coefficient (α=0.05) between: (1) the correlation coefficient (ρ) between discharge and TN at each site and (2) the land use/land cover, climate, topographic, discharge and geologic characteristics of the stream segments and their catchments.

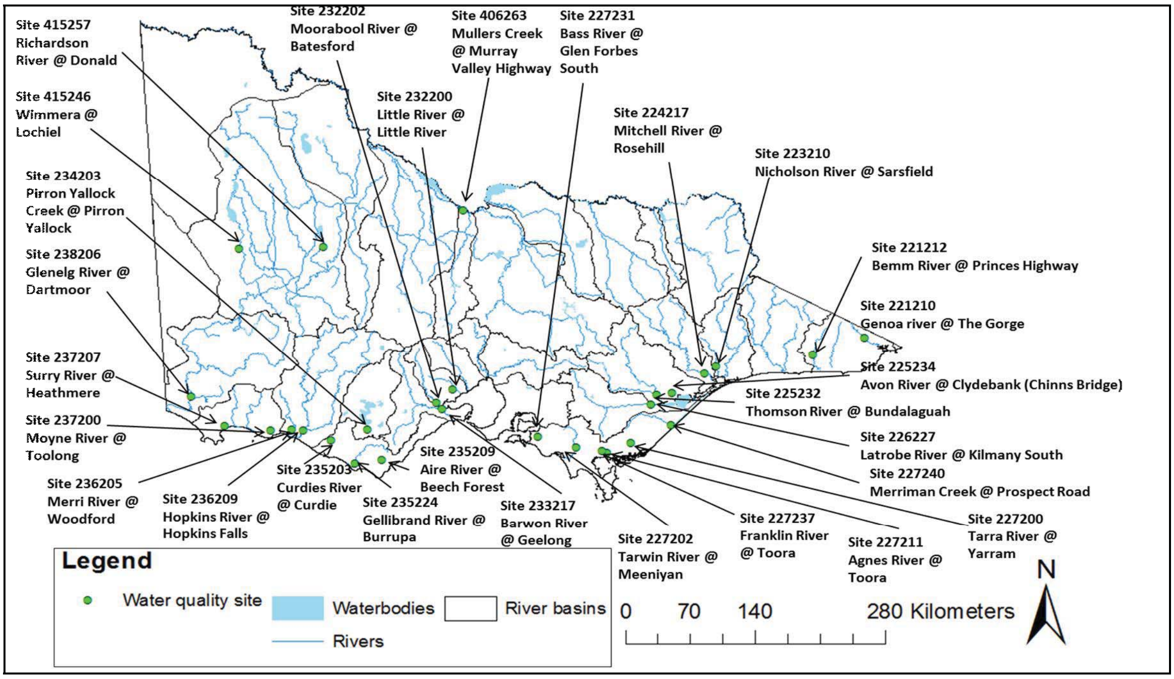

Figure 1. 28 Victorian water quality monitoring sites used for the statistical analysis.

Results and Discussion

Key factors influencing the spatial variability of TN concentrations

TN concentrations observed between August 2010 and December 2014 were found to vary spatially (Figure 2). Indeed, the Kruskal-Wallis test also indicates that the distributions of observed TN concentrations are statistically significantly different between water quality monitoring stations ($p<0.05$). Observed TN concentrations appear to increase from east to west across the state of Victoria (Figure 1 and Figure 2).

Mean TN concentrations at the 28 water quality sites correlated positively several land uses and temperature characteristics of the monitoring sites (Figure 3). The strongest positive correlation was with the percentage of catchment area with fertilisers applied (ρ=0.732). This is not surprising as the additional supply of nitrogen-containing fertilisers can be transported to receiving waters by runoff (Johnes et al., 1996). Mean TN concentrations also positively correlated with the catchment average annual temperature (ρ=0.571). This may be due to the fact that the nitrogen contained in organic matter is mineralised with increasing temperatures (Arheimer and Lidén, 2000). However, further investigation of soil organic matter levels is necessary to identify whether this is merely a statistical artefact. There is a cross-correlation between the catchment average annual temperature and the percentage of the catchment area applied with fertilisers (ρ=0.736, $p<0.05$). The strongest negative correlation was with the percentage of catchment area with forest cover (ρ=-0.812; Figure 3). There is a negative correlation between forest cover and areas applied with fertilisers (ρ=-0.887, $p<0.05$). Our results showed that if a large proportion of the catchment area is forested, there is less agricultural area within the catchment and less fertiliser usage.

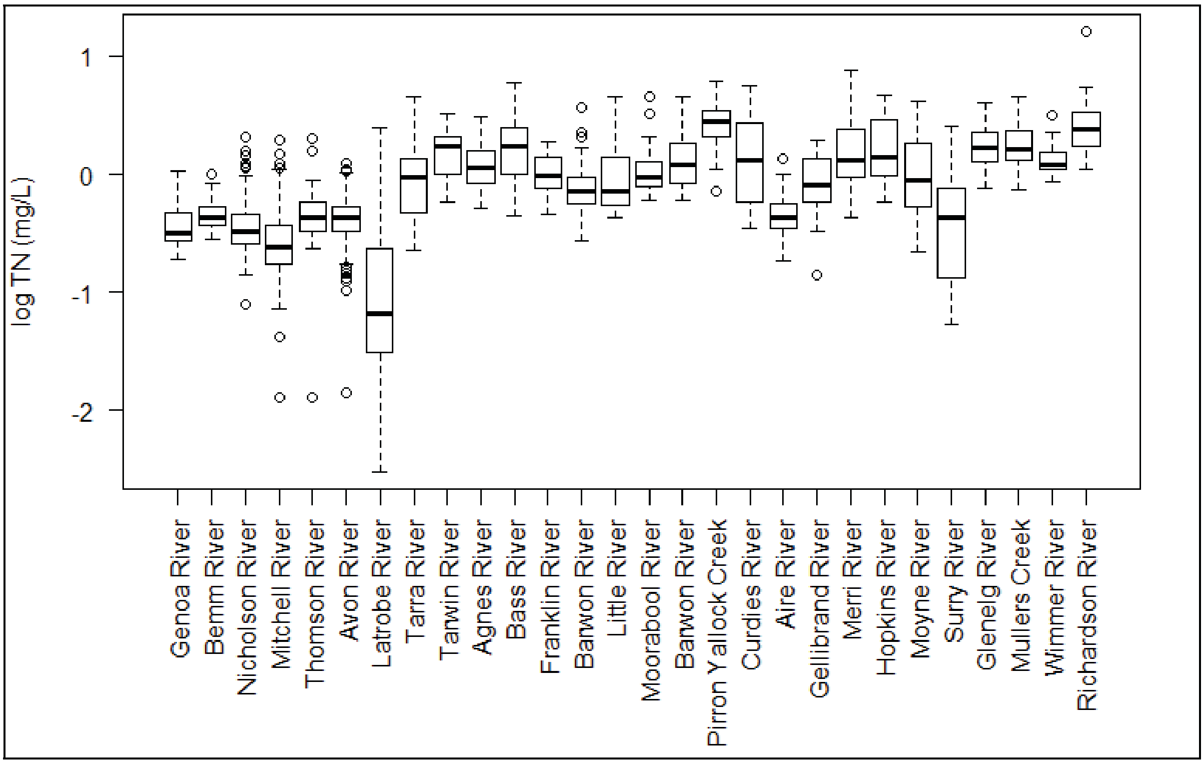

Figure 2. Boxplots showing distribution of TN concentrations at each water quality monitoring site.

There was also a strong negative correlation between TN and the maximum annual runoff depth (ρ=-0.627). This suggests that nitrogen is being diluted when the annual runoff volume increases (Whitehead et al., 2009). Finally, a negative correlation with the mean percentage contribution to total annual runoff during the six driest months of the year (ρ=-0.735) was also observed. It is likely that baseflow is a large component of summer runoff and it is possible that there are lower amounts of nitrogen in baseflow due to the denitrification that can occur in the soil. In addition, it is possible that due to the higher temperature, there are higher decay (or denitrification) rates of nitrogen (Hiscock et al., 1991).

Previous studies have argued that increased soil erosion potential and catchment slope increases the nutrient levels in rivers, because of adsorption of nitrogen to particulate matter (e.g., Bechmann et al., 2009). However our data analysis found negative correlations between TN and (1) the catchment average rainfall erosivity R factor ((MJ mm)/(ha hr yr)) (ρ=-0.649) and (2) the average catchment slope (ρ=-0.651). The negative correlation between the catchment slope and mean TN concentrations is likely to be cause by reduced rates of agricultural land uses on steeply sloping land (Arheimer and Lidén, 2000; Chang, 2008). Indeed, there is a negative cross-correlation between the average catchment slope and the percentage of the catchment area fertilised (ρ=-0.806, $p<0.05$).

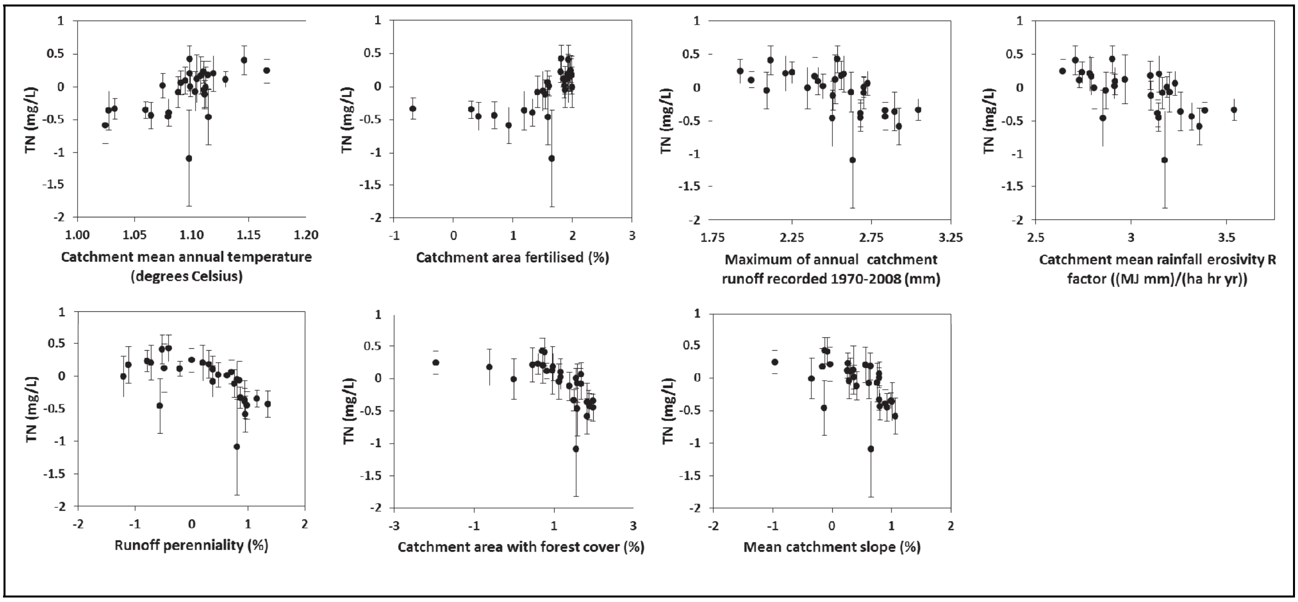

Figure 3. Scatter plots between mean log(TN) concentrations (August 2010 to December 2014) measured at 28 water quality monitoring sites and catchment characteristics. Error bars represent standard deviations in log(TN) concentrations.

Key factors influencing the relationship between discharge and TN concentrations in rivers

The previous section considered spatial variations in mean concentrations, we now turn our attention to the relationship between instantaneous discharge and concentration over time, and how the strength of this relationship varies between catchments. Previous studies have modelled TN concentrations in rivers as a function of the flow rate (e.g., Johnes et al., 1996). For Victorian rivers, we found that the temporal correlation between discharge and TN concentration varies spatially (Figure 4). The minimum Spearman rank correlation coefficient (ρ) is -0.423 (Richardson River, site 415257), the median is 0.689 (Genoa River, site 221210 and Agnes River, site 227211) and the maximum is 0.959 (Curdies River, site 235203).

Differences in the correlation coefficient (ρ) between discharge and TN at individual sites correlates with the local climate (Table 1). There was a positive correlation between the correlation coefficient and the average coldest month minimum temperature and the mean annual rainfall (mm) in the local stream segment and its valley bottoms. Further investigations are required to identify how specific nitrogen species behave in these catchments, to understand the mechanisms underlying these relationships.

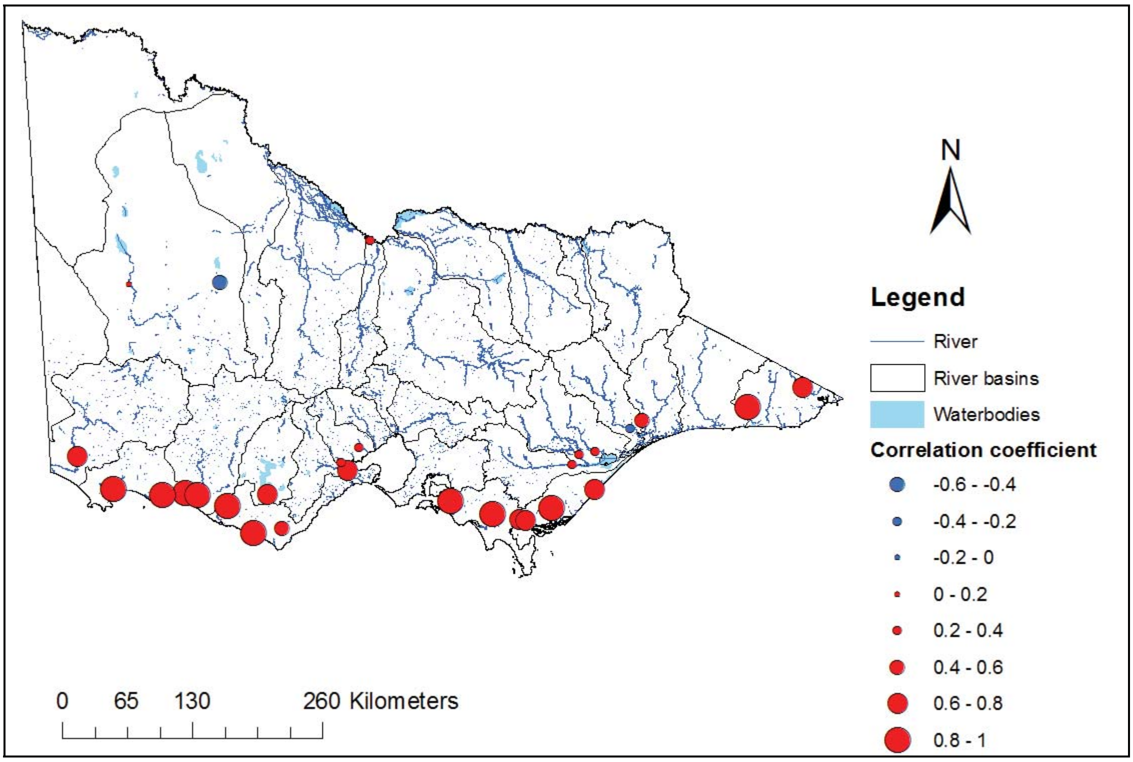

Figure 4. Map of water quality monitoring sites, showing the Spearman's Rank correlation coefficient (ρ) between log(TN) concentrations and discharge (ML/day) at each site.

Table 1: Spearman Rank correlation coefficients (ρ) between the temporal correlation between discharge and TN concentrations at each water quality monitoring site, and spatial catchment characteristics.

Explanatory Variable	Specific variable	ρ
Climate	Stream and environs annual average radiation	-0.741
	Stream and environs coldest month minimum temperature	0.708
	Stream and environs hottest month maximum temperature	-0.689
	Stream and environs mean annual rainfall (mm)	0.654
Topography	Maximum elevation in catchment (m)	-0.570
	Connected length (%)	0.679
	Total length (km)	-0.531

Catchment topography also appears to impact the spatial variability in correlations between TN concentrations and discharge (Table 1). TN concentrations and discharge are more strongly correlated when a greater proportion of the upstream flow path has no barrier (such as a dam or weir) blocking the flow. However, the correlation between TN and discharge becomes more negative with greater catchment length and higher maximum elevation in the catchment. When nitrogen is transported across greater distances to receiving waters, a larger proportion of the nitrogen can be lost by biogeochemical processes and deposition. Furthermore, when flow in the waterway is interrupted or held behind weirs and reservoirs, the energy of the flow decreases and nitrogen adsorbed to particles can be deposited upstream of the barrier. The decrease in flow energy might also enable cycling of nitrogen within the water storage. As a result, the TN concentrations may not necessarily vary as strongly with changing river discharge (Friedl and Wüest, 2002). The reduced correlation between TN and discharge alongside the increase in catchment elevation could be a result of lower nitrogen supply in catchments with higher elevations; there is generally less agriculture at higher elevation in the catchment (Chessman, 2003). When there is little fertiliser application, there is less nitrogen available in the catchment (from soil and atmospheric deposition) for transport to the receiving water. Consequently, increasing discharge and surface runoff may not necessarily increase TN concentrations in receiving waters, and may even dilute TN concentrations. Investigation of the relationship between specific nitrogen species concentrations and discharge at these 28 sites are required to better understand the mechanisms underlying spatial variability in the relationship between surface water nitrogen levels and discharge.

Conclusions

The objective of this investigation was to identify the key spatial variables associated with TN concentrations in rivers. Frequentist statistical analyses indicated that land use, climate, geology, land cover, and topographic catchment characteristics correlate strongly with the spatial variability in mean TN concentrations observed between August 2010 and December 2014 at 28 Victorian water quality monitoring stations. Mean TN concentrations were most strongly correlated to land use and land cover, specifically the proportion of the catchment area that is fertilised (positively), and the proportion of the catchment area covered by forest (negatively). However due to the negative cross-correlation between the proportion of the catchment area applied with fertilisers and the proportion of the catchment area covered by forest, it is likely that only one of these variables is required in any statistical predictive model of TN concentrations in lotic surface waters. Similarly it should be noted that other cross-correlations between spatial characteristics exist and make causal inferences more difficult. In addition the relationship between TN and discharge varied significantly across the 28 water quality monitoring sites. This variability is correlated to the topography and climate of water quality monitoring sites and their catchments. The spatial variability in the relationship between TN and discharge should be incorporated into future predictive models to reduce the amount of unexplained variation among sites.

Future work will incorporate the findings from this study into a statistical predictive model for TN concentrations. Similar exercises will also be undertaken for other water quality parameters (e.g., electrical conductivity, total suspended solids, total phosphorus).The development of these models will not only enable environmental managers to forecast future water quality transitions due to climate change or changing catchment management strategies, but they will also provide us with a robust understanding of the key explanatory variables affecting water quality in rivers.

Acknowledgements

This project has been supported by the Australian Research Council and the Victorian Environment Protection Agency, the Victorian Department of Environment, Land Water and Planning, the Bureau of Meteorology and the Queensland Department of Natural Resources and Mines through the linkage program (LP140100495). The authors would also like to offer the assistance provided by Matthew Johnson and Louise Sullivan in the compilation and analysis of the data.

References

Argent, R. M., Perraud, J.-M., Rahman, J. M., Grayson, R. B., & Podger, G. M. (2009). A new approach to water quality modelling and environmental decision support systems. *Environmental Modelling & Software*, *24*(7), 809–818.

Arheimer, B., & Lidén, R. (2000). Nitrogen and phosphorus concentrations from agricultural catchments—

influence of spatial and temporal variables. *Journal of Hydrology*, *227*(1–4), 140–159.

Arnold, J. G., & Fohrer, N. (2005). SWAT2000: current capabilities and research opportunities in applied watershed modelling. *Hydrological Processes*, *19*(3), 563–572.

Bechmann, M., Stålnacke, P., Kværnø, S., Eggestad, H. O., & Øygarden, L. (2009). Integrated tool for risk assessment in agricultural management of soil erosion and losses of phosphorus and nitrogen. *Science of The Total Environment*, *407*(2), 749–759.

Chang, H. (2008). Spatial analysis of water quality trends in the Han River basin, South Korea. *Water Research*, *42*(13), 3285–3304.

Chessman, B. C. (2003). New sensitivity grades for Australian river macroinvertebrates. *Marine and Freshwater Research*, *54*(2), 95–103.

Clark, J. S. (2005). Why environmental scientists are becoming Bayesians. *Ecology Letters*, *8*(1), 2–14.

Department of Environment, Land, Water and Planning Victoria. (2016). Victorian water measurement information system. Retrieved June 20, 2002, from http://data.water.vic.gov.au/monitoring.htm

Friedl, G., & Wüest, A. (2002). Disrupting biogeochemical cycles - Consequences of damming. *Aquatic Sciences*, *64*(1), 55–65.

Geoscience Australia. (2011). Environmental Attributes Database. Retrieved February 5, 2016, from http://www.ga.gov.au

Harris, G. P. (2001). Biogeochemistry of nitrogen and phosphorus in Australian catchments, rivers and estuaries: effects of land use and flow regulation and comparisons with global patterns. *Marine and Freshwater Research*, *52*(1), 139–149.

Hasani Sangani, M., Jabbarian Amiri, B., Alizadeh Shabani, A., Sakieh, Y., & Ashrafi, S. (2014). Modeling relationships between catchment attributes and river water quality in southern catchments of the Caspian Sea. *Environmental Science and Pollution Research*, *22*(7), 4985–5002.

Hiscock, K. M., Lloyd, J. W., & Lerner, D. N. (1991). Review of natural and artificial denitrification of groundwater. *Water Research*, *25*(9), 1099–1111.

Johnes, P., Moss, B., & Phillips, G. (1996). The determination of total nitrogen and total phosphorus concentrations in freshwaters from land use, stock headage and population data: testing of a model for use in conservation and water quality management. *Freshwater Biology*, *36*(2), 451–473.

Nash, L. (1993). Water quality and health. In Gleick, P. H. (Ed.), *Water in crisis : a guide to the world's fresh water resources*. New York: Oxford University Press.

Smith, R. A., Schwarz, G. E., & Alexander, R. B. (1997). Regional interpretation of water-quality monitoring data. *Water Resources Research*, *33*(12), 2781–2798.

Smith, V. H., Tilman, G. D., & Nekola, J. C. (1999). Eutrophication: impacts of excess nutrient inputs on freshwater, marine, and terrestrial ecosystems. *Environmental Pollution*, *100*(1–3), 179–196.

Webb, J. A., & King, E. L. (2009). A Bayesian hierarchical trend analysis finds strong evidence for large-scale temporal declines in stream ecological condition around Melbourne, Australia. *Ecography*, *32*, 215–225.

Whitehead, P. G., Wilby, R. L., Battarbee, R. W., Kernan, M., & Wade, A. J. (2009). A review of the potential impacts of climate change on surface water quality. *Hydrological Sciences Journal*, *54*(1), 101–123.