Author/s:
Guo, D;Liu, S;Singh, D;Western, AW

Title:
Predicting quantiles of water quality from catchment characteristics

Date:
2021-01-01

Citation:
Guo, D., Liu, S., Singh, D. & Western, A. W. (2021). Predicting quantiles of water quality from catchment characteristics. Hydrological Processes, 35 (1), https://doi.org/10.1002/hyp.13996.

Guo Danlu (Orcid ID: 0000-0003-1083-1214)

# Predicting quantiles of water quality from catchment characteristics

**Authors:** Danlu Guo[1], Shuci Liu[1], Dhananjay Singh[1*], Andrew W Western[1]

**Affiliation:**

[1]Department of Infrastructure Engineering, University of Melbourne, Parkville, Australia

[*]Now at Senversa Pty Ltd.

**Corresponding Author's email:** Danlu.guo@unimelb.edu.au

**Graphical Table of Contents**

**Key words:** sediments, nutrients, salt, statistical model, stream water quality, spatial variation, quantile, extreme

**Running title:** Predicting quantiles of water quality

**Acknowledgements:** The Australian Research Council together with the Victorian Environment Protection Authority, the Victorian Department of Environment, Land Water and Planning, the Bureau of Meteorology and the Queensland Department of Natural Resources, Mines and Energy provided support for this study through the linkage program (LP140100495). The authors would like to thank Paul Wilson and Hannah Sleeth (Victorian Department of Environment, Land Water and Planning) for their help on data acquisition. We also appreciate the work of Shuangying Deng and Yuhan Zhou on preliminary data analyses.

**Data Availability:** All water quality data used for this study can be found on the Water Measurement Information System by Victorian Department of Environment, Land Water and Planning (http://data.water.vic.gov.au/monitoring.htm). Sources of other data are detailed in Table S1 of the Supporting Information.

# Abstract

Water quality is often highly variable both in space and time, which poses challenges for modelling the more extreme concentrations. This study developed an alternative approach to predicting water quality quantiles at individual locations. We focused on river water quality data that were collected over 25 years, at 102 catchments across the State of Victoria, Australia. We analysed and modelled spatial patterns of the 10th, 25th, 50th, 75th and 90th percentiles of the concentrations of sediments, nutrients and salt, with six common constituents: total suspended solids (TSS), total phosphorus (TP), filterable reactive phosphorus (FRP), total Kjeldahl nitrogen (TKN), nitrate-nitrite ($NO_x$), and electrical conductivity (EC). To predict the spatial variation of each quantile for each constituent, we developed statistical regression models and exhaustively searched through 50 catchment characteristics to identify the best set of predictors for that quantile. The models predict the spatial variation in individual quantiles of TSS, TKN and EC well (66%-96% spatial variation explained), while those for TP, FRP and $NO_x$ have lower performance (37%-73% spatial variation explained). The most common factors that influence the spatial variations of the different constituents and quantiles are: annual temperature, percentage of cropping land area in catchment and channel slope. The statistical models developed can be used to predict how low- and high-concentration quantiles change with landscape characteristics, and thus provide a useful tool for catchment managers to inform planning and policy making with changing climate and land use conditions.

**Key words:** sediments, nutrients, salt, statistical model, stream water quality, spatial variation, quantile, extreme

# 1 Introduction

Globally, there has been a large deterioration in water quality in streams and lakes. Total pollution loads generally increase due to non-point sources from agriculture and urban areas, as well as point sources from industry and municipal sectors (Best, 2019;Ongley et al., 2010). Pollution-related degradation in freshwater ecosystems is estimated to be responsible for a one-third decrease in biodiversity across the globe (UN Water, 2015). A recent assessment of rivers in Latin America, Africa and Asia suggests that one in six rivers have been severely affected by organic pollution (UNEP, 2016).

Knowledge of the low, normal, and high concentrations of various water quality constituents across space is valuable for interpreting monitoring data, informing water quality targets and prioritizing management focuses. For example, the Australian and New Zealand Environment and Conservation Council (ANZECC) Water Quality Guidelines defined the reference water quality conditions for any test site with the percentiles of concentrations from a reference site (80th or 20th percentile depending on the specific water quality parameters). Further management actions are recommended when the test site diverts from the corresponding reference condition (ANZECC and ARMCANZ., 2000). In the United States Environmental Protection Authority (US EPA), the trigger levels of river nutrient concentrations were specified with the 25th and 75th percentile concentrations of data collected at reference sites and regions (United States Environmental Protection Agency, 2001).

As illustrated, water quality management guidelines are often specified with quantiles of concentration; however, modelling of water quality concentration quantiles has received limited attention in the existing literature. Water quality models range from process oriented to statistical models. Process oriented models tend to be tuned to local catchment data, often with parameters that can purportedly be interpreted with physical meaning. In contrast, statistical models are more suited for data-based inference. Most statistical models tend to predict either spatial patterns of average conditions or time series at specific locations, although some spatiotemporal statistical approaches exist (e.g. Guo et al., 2020) that make it possible to predict timeseries and then calculate the quantiles at individual locations. Previous modelling studies have shown that differences in average ambient water quality conditions across locations can be explained by a set of catchment landscape characteristics (Chang, 2008;Lintern et al., 2018b;Onderka et al., 2012;Lintern et al., 2018a). This suggests that an alternative possibility is to predict concentration quantiles directly using catchment characteristics.

Only a few existing studies aimed to explicitly model the more extreme water quality conditions across multiple catchments with catchment characteristics (Tramblay et al., 2010a;Tramblay et al., 2010b). Tramblay et al. (2010a) analysed the annual maxima series of daily monitoring records for suspended sediments in 72 rivers in Canada and the United States. Statistical regression models were then developed to predict extreme concentrations with return periods of 2 and 20 years at individual sites, considering predictors such as catchment climate, topography, land cover and soil attributes. The same model structure was extended in Tramblay et al. (2010b) to predict suspended sediments at 19 river monitoring stations in California for a larger number of concentration statistics, including the annual return period of 10 years and the 99th percentile of all daily records at each site. Considering that quantile-based guidelines that are often used to identify alarming conditions and to trigger management actions (e.g. ANCEZZ and ARMCANZ, 2000; US EPA 2001), the utility of these models can be further extended through prediction of a wider range of quantiles across sites. Furthermore, with the limited water quality monitoring capacity in many parts of the world, daily monitoring records are scarce; this means that modelling the quantiles of all available records – instead of the corresponding annual maxima series – would be better suited to practical data availability.

Besides the ability to predict water quality quantiles, statistical models can also be used to identify the key catchment characteristics that influence extreme water quality conditions across space, which can help enhance the understanding of catchment processes. For example, the percentage of clay in the soils, precipitation intensity and forest cover were identified as the best predictors for extreme concentrations of suspended sediments, for 72 rivers in Canada and USA studied in Tramblay et al. (2010a). However, such understanding is highly limited due to the lack of models that explicitly focus on water quality quantiles. This provides another reason to improve the modelling of water quality quantiles with catchment characteristics.

This study aims to develop a spatial statistical model for water quality quantiles (10%, 25%, 50%, 75%, 90%) at a catchment scale, and to identify the key landscape characteristics that influence the spatial variation in the quantiles of each constituent. We use a set of 102 catchments with 25 years of monthly spot sampling from Victoria, Australia. Six constituents of common concern in river water quality are considered: total suspended solids (TSS), total phosphorus (TP), filterable reactive phosphorus (FRP), total Kjeldahl nitrogen (TKN), nitrate-nitrite ($NO_x$), and electrical conductivity (EC).

Specifically, we aim to address two research questions:

1. How well can we predict the spatial variation of individual water quality quantiles using catchment characteristics?
2. Which catchment characteristics most influence the spatial variation of individual concentration quantiles, and do these key drivers differ between constituents and quantiles?

# 2 Method

## 2.1 General Model Structure

For each of the six water quality constituents of interest (EC, TSS, TP, TKN, FRP, $NO_x$), we developed statistical models for the corresponding 10th, 25th, 50th, 75th and 90th percentiles of concentrations within all available records at individual sites. The spatial variation of each quantile was modelled by a multiple linear regression with catchment landscape characteristics as predictors. The general form of the model is:

$$C_i = \sum_{k=1}^{K} S_{k,i} \times eff.S_k \tag{1}$$

where $i = 1,2,3\ldots n$; $k = 1,2,\ldots K$

Here $C_i$ is a specific percentile (10th, 25th, 50th, 75th or 90th) of concentration for each constituent at catchment $i$. $S_{k,i}$ represents the value of the $k^{th}$ predictor (i.e., catchment characteristics) at catchment $i$, and $eff.S_k$ is the regression coefficient.

To develop these models, we first extracted data for water quality and catchment characteristics for our study region and pre-processed them as required by the linear regression models (Section 2.2). The model development is detailed in Section 2.3. For each quantile of each constituent, the best set of $K$ model predictors was selected, by comparing possible regression models consisting of all different combinations of catchment landscape characteristics. All possible model forms were evaluated with information criteria that considers both model performance and model complexity (Section 2.3.1). The performance and robustness of the final model were then evaluated, followed by assessment of the impacts of individual catchment characteristics on water quality quantiles (Section 2.3.2).

## 2.2 Data Collection and Pre-processing

All water quality data analyzed were extracted from the Water Measurement Information System surface water database, provided by the Victoria Department of Environment, Land, Water and Planning (Department of Environment Land Water and Planning Victoria, 2016a), publicly available at http://data.water.vic.gov.au/. The database consists of water quantity and quality data at over 400 monitoring sites in Victoria, which were collected since 1990. For each site, spot water quality samples were taken at a monthly time-step and analyzed by National Association of Testing Authority accredited laboratories. From these >400 sites, our analyses focused on 102 sites, because they balance having a long consistent period of continuous record (1994-2019) for the six focused constituents with having a large number of monitoring sites (Figure 1). These sites are within catchments that are predominantly natural and rural land uses, with areas ranging from 5 to 16,000 km$^2$. At 98% of the sites, the contribution from point discharges are less than 10% of the average annual flows (Lintern et al., 2018b).

**Figure 1: a) Maps of the 102 water quality monitoring sites analyzed in this study and their catchment boundaries. Insert shows location of the state of Victoria in Australia. Panels b), c) and d) illustrate the spatial distribution of the elevation, annual average rainfall and annual average temperature across Victoria, respectively.**

To develop predictive models, we used the full monthly spot-sampled time-series of concentration for each constituent to calculate concentration quantiles. A preliminary check on the representativeness of these spot-sampled water quality data was performed by comparing the flows at the sampling time with the corresponding flow duration curves for individual monitoring sites (Figure S3, Supporting Information). The results suggest that the water quality samples generally well represent the full flow regimes for each site, although the extreme flows are sometimes sparsely sampled (e.g., the high flows for 415207 and low flows for 406207). For all constituents other than EC, there are records below the detection limit (DL), defined as the '*minimum concentration detected for which there is 95% confidence of accuracy and therefore is accurate enough to report*' (Australian Water Technologies, 1999). Therefore, prior to the quantile calculations, any below-DL records were set equal to half of the DL concentration for the corresponding constituent, following the conventional approach for analyzing below-DL water quality concentration data (Helsel, 1990). The proportion of data below-DL for individual sites is summarized for each constituent in Figure S1 in the Supporting Information. After adjusting the below-DL data, the concentration time-series at each site was used to calculate the 10th, 25th, 50th, 75th and 90th percentiles of concentrations for each constituent based on empirical ranking.

We selected 50 catchment landscape characteristics as potential explanatory variables for the spatial variation in water quality quantiles. This selection was informed by a comprehensive literature review that summarized the key catchment landscape characteristics known to influence water quality (Lintern et al., 2018a). These 50 characteristics span multiple categories including catchment land use, land cover, topographic, climatic, geological, lithological and hydrological characteristics. These 50 variables were derived using datasets obtained from Geoscience Australia (2004, 2011), Bureau of Meteorology (2012), the Australian Bureau of Rural Sciences (ABARES) (2016), Department of Environment Land Water and Planning Victoria (2016b) and the Terrestrial Ecosystem Research Network (2016). Table S1 in the Supporting Information includes detailed explanation of all 50 variables and their corresponding data sources. It is worth noting that within the land use variables, agricultural areas include the irrigated and dryland cropping areas, pastures, horticulture and forest plantations. Each land use variable is defined as percentage coverage of the total catchment area by the specific land use type (e.g., agriculture, urban etc.), estimated using a static dataset from 2005-2006 (Bureau of Rural Sciences, 2010) to represent the entire study period. This was informed by a preliminary analysis, which suggested that the key land uses in these catchments (i.e., agriculture, grazing, conservation) had less than 1% changes in between 1996 and 2011. Figure S2 (Supporting Information) shows the distribution of percentage coverage of each key land use across all the catchments. Similar to the land use, the land cover variables are defined by percentage coverage of catchment area by each vegetation type. Forest, shrub and woodland are differentiated by the height and density of vegetation, where forests refer to the tallest vegetation with the greatest density, and shrubs have the shortest vegetation with the lowest density.

For each constituent, the values of each quantile across catchments generally exhibited high skewness (Figure S4, Supporting Information); highly skewed distributions were also observed in a number of variables within the 50 catchment characteristics. Therefore, both the water quality quantiles and the 50 potential explanatory variables were transformed to increase the symmetry of individual variables, making them more suitable for use in the linear model structure (Eq. 1). Specifically, all concentration quantiles calculated were log-transformed following the conventions commonly applied in catchment water quality statistical analysis. The values of each catchment characteristic across 102 catchments were log-sinh transformed (Wang et al., 2012) (Eq. 2), considering to its ability to resolve the presence of zero values that exist in several variables (e.g., percentage area of individual land uses). The best log-sinh transformation parameters (*a* and *b*) were identified with the *GA* package in *R* (Luca Scrucca, 2019) to minimize the skewness of each variable across 102 catchments (i.e., symmetry is maximized):

$$y_{log-sinh} = \frac{1}{b}\log(\sinh[a + by_{raw}]) \tag{2}$$

Following the log-sinh transformation of individual potential predictors, these transformed values were also standardized. This ensured that the model coefficients can be interpreted as indicating the relative impacts of individual predictors. To further enable comparison of these impacts across different constituents and quantiles, the log-transformed water quality quantiles (i.e., model predictands) were also standardized.

## 2.3 Model Development

### 2.3.1 Selecting the best explanatory variables for water quality quantiles

To develop the linear regression model for each quantile of each constituent, we need to determine the best predictors (i.e., $S_k$ in Eq. 1) from the 50 potential explanatory variables (see Section 2.2). This predictor selection was achieved by combining two predictor selection techniques: Bayesian variable selection (BVS) and exhaustive search. The decision to combine two predictor selection approaches was based on:

1) The high computational requirement. Specifically, 50 catchment characteristics were to be evaluated as the potential explanatory variables for developing water quality quantile models (i.e., $3.04\times10^{64}$ possible combinations to be evaluated for each of the 5 quantiles of the 6 constituents), making it both challenging and inefficient to evaluate all possible combinations of the 50 variables at once with either of the BVS or the exhaustive search.
2) The strengths of individual predictor selection techniques. Bayesian variable selection uses Bayes' theorem to estimate the probability of individual predictors being included in the best regression models, and is thus useful for establishing the relative importance of multiple potential predictors (Bayarri et al., 2012;García-Donato and Martínez-Beneito, 2013). The exhaustive search (Guyon and Elisseeff, 2003;May et al., 2011;Saft et al., 2015) is a more comprehensive way of comparing regression models composed of all possible combinations of potential predictors, from which the best model or the group of best models can be identified. For exhaustive search, the evaluation of alternative models is often based on the Bayesian Information Criterion (BIC) (Schwarz, 1978), which considers both model performance and complexity and thus effectively penalizing models with more predictors.

For each quantile of each constituent, we first used BVS (implemented with *R* package *BayesVarSel* by Gonzalo Garcia-Donato and Forte (2017) to rank the probability of selecting each potential predictor. The top 50% predictors (i.e., 25 predictors) suggested by BVS were then included in an exhaustive search which evaluated all possible models composed with these candidate predictors. This involved: 1) fitting linear regression models with all possible combinations of the 25 predictors; 2) calculating the BIC for each possible model developed with a unique combination of predictors; and 3) identifying the best model that has the lowest BIC.

### 2.3.2 Model evaluation and interpretation

Once the final model was developed via predictor selection, performance of the model (research question 1) was evaluated based on the Nash-Sutcliffe Efficiency (NSE) (Nash and Sutcliffe, 1970), which describes percentage of observed variability explained by the model. We calculated the NSE for both:

1) log-transformed concentration quantiles; and
2) untransformed concentration quantiles, in which case model predictions were back transformed using inverse logarithm (i.e., exponential).

The first evaluation informs the structural validity of the model that operates in a transformed space (i.e., the modelling space). The second evaluation is at the scale of absolute concentrations of pollutants, which thus informs the predictive capacity of the model for practical uses.

The models developed with all available data (referred to as the 'full model' hereafter) were then cross validated in two ways to evaluate the robustness of 1) the predictor selection; and 2) the model performance. Specifically, the two cross validations were performed as follows:

1) *Evaluating the robustness of predictor selection.* To test this, we conducted a 10-fold cross-validation of the full predictor selection process for each constituent-quantile combination. Specifically, we randomly partitioned the 102 catchments to ten groups of similar size (10 or 11 members). We held out each group and performed the two-step predictor selection process (Section 2.3.1) with data from the remaining nine groups (92 or 91 sites) and then made predictions for the held-out group. This was repeated until all 10 groups are held out. The predictors selected for the cross-validation models were compared to those of the full model.
2) *Evaluating the robustness of model performance.* To test the performance of the models based on the final set of predictors (i.e. the final model), we ran another 10-fold cross-validation that was separate to the process mentioned in 1). Note that since this evaluation focused on the robustness of calibrated model performance instead of predictor selection, only the predictors selected in the full models were used. For each constituent-quantile combination, all held-out data were used to evaluate the performance of the cross-validation models, which was summarized by the NSE and compared to that of the corresponding full model.

To identify the key catchment characteristics that influence the spatial variation of water quality quantiles (research question 2), the model coefficients were extracted to summarize the impact of catchment characteristics on the low, medium and high quantiles of each constituent.

# 3 Results

We first present the spatial patterns of the quantiles of constituent concentrations that informed the model development (Section 3.1). The model performance in predicting these concentration quantiles are then presented, with the corresponding cross-validation results to assess the robustness of model calibration (Section 3.2). We then summarize the key catchment characteristics based on the predictor selection results for individual concentration-quantile models, for which the robustness was also evaluated with the cross-validation of the predictor selection process (Section 3.3).

## 3.1 Spatial Patterns of Water Quality Quantiles

Figure 2 shows the spatial variation of the 75th percentiles across the 102 catchments for each of the six constituents. As a common pattern across all constituents, low concentrations are often seen in mountainous areas in the north-east part of the state. The ‘hot-spots’ for higher concentrations show some constituent-specific patterns. Specifically, high concentrations for TSS, TP and TKN are mainly seen in streams in the north-west and south-west of the state. For FRP and $NO_x$, there is a ‘band’ of high concentration sites stretching from the north to south-west of the state. For EC, streams in the western half of the state have systematically higher concentrations than the eastern half.

**Figure 2: Spatial variation of the 75th percentile of concentrations for TSS, TP, FRP, TKN, $NO_x$, and EC across the 102 study sites, with elevation in the background. Colors show the ranges between inter-quartiles.**

Figures S5 to S8 in the Supporting Information show the spatial distributions of the other four percentiles (10th, 25th, 50th and 90th) for all constituents. The 90th percentiles generally show similar spatial patterns to those of the 75th percentile. The lower quantiles display much less distinct spatial variability, which is possibly due to the lower concentrations measured, meaning that more observations are close to the lower limit of measurable concentrations (i.e., the detection limit). Consequently, for a lower quantile we should expect a smaller number of distinct values (and thus spatial variation) across sites. A good example is seen for FRP, for which half of the sites share the same value for both the 10th percentile and the 25th percentile.

## 3.2 Model Performance in Predicting Water Quality Quantiles

The top row of Figure 3 shows the modelled versus observed values for the 75th percentiles for all six constituents in the log-transformed scale, which is the scale that the model was developed in. To explore the impacts of data transformation, model performance was also evaluated in a back-transformed scale, which is the scale in which the water quality measurements are taken (bottom row of Figure 3). Across constituents, spatial patterns are most well predicted for TKN, EC and then TSS. When predictions are back-transformed, these models are also the best-performing ones, with only small degradation in performance compared to the transformed results. For TP, FRP and $NO_x$ there is substantial degradation in model performance when predictions are back-transformed, which can be due to a small number of high concentration outliers – a more detailed discussion of potential reasons for these weaker models is provided in Section 4.1. These patterns of performance are consistent across all other constituent-quantile models, as summarized in Figures S9 and S10 for the transformed and back-transformed scales, respectively.

**Figure 3: Modelled versus observed 75th percentile of concentration for each constituents across all 102 sites, plotted in the log-transformed scale (top row) and the back-transformed scale (bottom row). The NSE (Nash-Sutcliffe Efficiency) values for each constituent-quantile model and the 1:1 lines (red dashed line) are also shown.**

The performance across all constituent-quantile models is summarized with the corresponding NSE in Figure 4, for both the log-transformed scale (x-axis) and the untransformed scale (y-axis). In the transformed scale, the models are able to explain more than 50% spatial variability in all quantiles for all constituents except FRP, for which over half spatial variability remain unexplained for the 10th, 25th, 50th and 90th percentiles. The performance of all models declines after back-transforming the simulated quantiles to the scales of measured concentrations. This is indicated by the points plotting below the 1:1 line in Figure 4. In both the transformed and untransformed scales, the predictive power for TKN, EC and then TSS are systematically higher than that for other constituents.

**Figure 4: Calibration performance (as the Nash-Sutcliffe Efficiency, NSE) across all 102 sites, in the scale of actual measurement i.e., untransformed scale (y-axis), versus the corresponding NSE values assessed in the scale of modelling i.e., log-transformed scale (x-axis).**

Although a systematic drop in model performance is seen for all constituents and quantiles after back-transformation, the extent of performance decline varies with constituents. For TKN, EC and TSS, where the log-transformed models show good performance across quantiles, the performance of the back-transformed models remain relatively high (most NSE > 0.5). In contrast, substantial performance loss after back-transformation is clearly seen for TP, FRP and $NO_x$. The loss of performance is generally greater for higher quantiles. This is illustrated by the greater vertical distances from the darker points to the 1:1 line compared to the lightly points, and is further discussed in Section 4.1.

The performance robustness for each model is summarized in Figure 5, by comparing the performance of the full model with those obtained from the 10-fold cross-validation (Section 2.3.2). For TSS, TKN and EC – where the full models performs substantially better – the cross-validation model performance is generally consistent with the corresponding full models. The good cross-validation model performance indicates that these models are likely robust across different locations within the study region. In contrast, lower robustness is seen for TP, FRP and $NO_x$, for which limited performance is also found for the full models. It is worth noting that the contrast between TSS and TP is surprising: while the full-model performance for the two constituents are generally comparable, TP has large performance declines for cross-validation, which is not seen for TSS. This is likely related to the more extreme outliers in the TP data, as discussed previously.

**Figure 5: Nash-Sutcliffe Efficiency (NSE) of the full model and all 10-fold cross-validation models, for each constituent-quantile combination.**

## 3.3 Key Catchment Characteristics for the Spatial Variation of Water Quality Quantiles and Their Impacts

Figure 6 shows the key catchment characteristics selected as the predictors of the full models for each constituent and quantile. For each constituent, only the predictors (i.e., catchment characteristics) that were selected in at least one quantile models are shown. The x-axes show the coefficients of individual predictors, which indicate the strengths and directions of their influences. The coefficient values are comparable across all predictors, constituents and quantiles, since they have been standardized before model fitting (Section 2.2). As a general pattern, catchment climate has the greatest influence on concentration quantiles, as shown by the larger magnitudes of coefficient values compared to other categories of predictors. Regarding specific predictors, a few key predictors are common to multiple constituents and quantiles:

1) *Percentage area of cropping land* (*PerCropping* in *land use*) shows positive relationships with concentrations for most quantiles of TSS, TP, TKN and EC;
2) *Higher annual temperature* (*AnnTemp* in *climate*) is generally associated with higher concentrations of quantiles for the phosphorus (TP, FRP) and nitrogen (TKN, $NO_x$);
3) *Mean channel slope* (*ChannelSlope* in *topography*) has negative relationships with most quantiles of phosphorus and nitrogen species.

**Figure 6: Catchment characteristics that have been identified as key predictors for the spatial variations of each quantile concentrations for each constituent (y-axis) and their corresponding impacts, summarized by their coefficient values (x-axis). Categories of catchment characteristics are differentiated by colors. All predictors and predictands were standardized, so that the coefficient values are comparable across predictors, constituents and quantiles.**

Apart from the abovementioned common predictors, a few other catchment characteristics have varying impacts across quantiles and constituents. The key characteristics that influence individual constituents are summarized below, with detailed interpretations in Section 4.2:

1) For TSS, all quantiles are positively correlated with the *percentage clay in soil* (*PercClay*) and *percentage area of shrub land* (*Shrubs*), and negatively correlated with the *percentage area of grass land* (*Grasses*). The *hottest month temperature* (*HotMonthTemp*) shows positive effects on all quantiles except for the 10th percentile. Spatial variation of the lower and median quantiles (10%, 25% and 50%) are dominated by similar catchment characteristics including the *maximum catchment elevation* and *storage*, which both have positive effects. *Valley bottoms* is a predictor characterizing the percentage of valley bottom area in a catchment. The higher quantiles (75% and 90%) are positively correlated with *catchment area*.
2) The two phosphorus species (TP and FRP) both show less association with catchment characteristics for the low and high quantiles (10th and 90th). For the medium quantiles (25%, 50% and 75%), both constituent concentrations generally increase for catchments with lower *erosivity*. In addition, these quantiles of TP are also commonly negatively correlated with the *cold season temperature* (*ColdQTemp*) and *runoff perenniality* (*RunPerenniality*). For FRP, while the models are very poor, the other common drivers for these medium quantiles are *percentage of catchment area used for pasture* (*PerPasture*) and *forest*, which both have negative effects.
3) The two nitrogen species (TKN and $NO_x$) are influenced by distinct sets of landscape characteristics. Most quantiles of TKN are positively correlated with the *percentage area used for pasture* (*PerPasture*), the *percentage of clay in soil* (*PerClay*), and the *mean cease to flow* (annual number of no flow days). For $NO_x$, while most quantiles increase with the *percentage of urbanized area*, the other key drivers display more ‘stratification’ between lower and higher quantiles. Specifically, the 10th and 25th percentiles increase with *average daily streamflow* (*AverageDayFlow)* while they decrease with the *standard deviation of daily streamflow* (*sdDayFlow)* and the *percentage of riparian zone fragmented* (*FraRipaZone*)– a measure of

vegetation cover gaps in the riparian zone (Department of Environment Land Water and Planning Victoria, 2014); the 50th and 75th quantiles increase with *warm season rainfall* and *temperature* (*WarmQRain* and *HotMonthTemp*), and decrease with *annual radiation* (*AnnRad*).

4) For EC, the key drivers are relatively consistent across quantiles except for the 90th. Concentrations generally increase with the *wet season rainfall* (*WetQRain*) and *warm season temperature* (*HotMonthTemp*), and decrease with the *cold season rainfall* (*ColdQRain*) the *runoff perenniality*.

The robustness of the key predictors selected is summarized in Figure 7, by comparing the predictors included in the full model (black dots) with those obtained from the 10-fold cross-validation of the predictor selection process (Section 2.3.2). A darker shade indicates a more frequent selection during the cross-validation. As the black dots largely coincide with where the darker cells are, the results suggest that the predictors selected for the full model are generally robust. Amongst constituents, TSS shows highest consistency of key predictors selected when different subsets of data (i.e., monitoring stations) are used. Therefore, the key catchment processes affecting TSS concentration quantiles are likely more similar across space compared to other constituents in this region. It is also worth noting the low robustness for the predictors selected for the 90th percentile of FRP and the 10th percentile of EC, suggesting that the spatial patterns of these quantities are likely highly heterogenous across space.

**Figure 7: Robustness of key predictors selected in the models of quantile for individual constituents. Black dots highlight catchment characteristics selected in the main models, for each constituent (panel) and each quantile (x-axes). Red shades show the number of times each predictor being selected in the 10 cross-validation models.**

# 4 Discussion

## 4.1 Predictive Power for Water Quality Quantiles and Practical Implications

As highlighted in Figures 3 and 4, with the exception of FRP, the models developed are generally capable of predicting the spatial variation of quantiles of sediments, nutrients and salinity within the study region. This provides a parsimonious alternative approach to predicting the expected values of concentration quantiles. The modelled outputs are directly comparable with water quality standards and guidelines (e.g. Australian and New Zealand Environment and Conservation Council and Zealand., 2000;United States Environmental Protection Agency, 2008) and thus can provide a quick assessment of the expected non-compliance of water quality conditions based on catchment characteristics.

Model performance is best for TSS, TKN and EC. This might relate to how readily different constituents participate in biogeochemical cycles, with more labile constituents being harder to predict. This is consistent with the previous findings from modelling the spatial variation of long-term average concentrations of constituents within the same study region (Lintern et al., 2018b). Conservative constituents are expected to be predominantly driven by the three fundamental catchment water quality processes: source, mobilization and delivery (Granger et al., 2010), for which clear spatial gradients can be identified across landscapes (Lintern et al., 2018a). For example, the source of sediments and nutrients in waterways can differ with soil type, land cover and land use practice (Collins et al., 2012;Hamilton and Miller, 2002;Holloway et al., 1998;van der Perk et al., 2007), whereas the transport of pollutants can vary with streamflow regime (Coynel et al., 2005;Ali and De Boer, 2007). In contrast, non-conservative constituents like FRP and $NO_x$ are influenced by additional biogeochemical processes. Various such processes have been highlighted in previous studies, such as the impacts of microbial activity in soils, organic carbon levels and atmospheric deposition on the storage and transport of FRP and $NO_x$ in streams (Camarero and Catalan, 2012;Rothwell et al., 2010;Fuss et al., 2016;Ahearn et al., 2005). These biogeochemical processes are likely more heterogeneous across space and time, and thus more difficult to capture (Srivastava et al., 2007).

Results in Figure 4 also suggest that, the deterioration of model performance after back-transformation is most evident for TP, FRP and $NO_x$. In part this just relates to the poorer performance in the transformed space as discussed above. In addition, limited model performance in the untransformed space may also relate to the small number of outliers with high concentrations evident in Figure 3. More generally, such performance deterioration can also occur where there is higher skewness in the raw data, as previously illustrated when modelling the spatiotemporal variability of water quality for same study region (e.g. Guo et al., 2020). The impacts of data skewness on performance is less clearly seen in this study when the focus is modelling quantiles, as the constituents having lower model performance (i.e., TP, FRP and $NO_x$) do not have systematically higher skewness compared to other constituents (Figure S4, Supporting Information). Specifically, as seen in Figure S4, apart from TKN, the raw data for all constituents and all quantiles depart clearly from the red lines that represent normal distribution, indicating that all constituents except TKN have comparably high data skewness. This suggests that the limited performance of quantiles models for TP, FRP and $NO_x$ is likely better explained by the low performance of the log-transformed models and occasional outliers, rather than high data skewness.

Model outliers indicate that concentration quantiles at the corresponding sites show 'unexpected' responses to catchment characteristics, relative to other sites. To explore this, we assessed the top 10 outlier sites that have the highest residuals from each of the 90th percentile models of TP, FRP and $NO_x$. For each constituent, we found no notable commonality across these outlier sites on their spatial characteristics. However, we note some commonality on the outlier sites across constituents – specifically, sites 233215 and 407255 are common outliers for all TP, FRP and $NO_x$, and both sites have concentration quantiles that are higher than expected. This likely indicates some unusual features about these catchments that have not been captured by the model, potentially the influence of a point source (see discussion in Section 4.3).

## 4.2 Key Catchment Characteristics that Influence Water Quality Quantiles

The models developed can help to infer the key landscape properties that are most closely linked to the spatial variation in constituent concentration quantiles (Section 3.3). These models represent the impact of key spatial drivers on the more extreme concentration conditions, thereby expanding previous modelling work on ambient water quality conditions (Lintern et al., 2018b). Interpreting these key spatial drivers of water quality quantiles along with their model coefficients can help enhance current understanding of hydrological processes within the study region.

As illustrated in Figure 6, across most quantiles and constituents (i.e., TSS, TP, TKN and EC), we found that a higher percentage of cropping area is the most common factor leading to higher concentrations across all quantiles. For the nutrient species (i.e., TP, FRP, TKN and $NO_x$), milder channel slope, along with higher annual temperature, were also associated with consistently higher concentrations across quantiles. The impacts of other landscape characteristics are more specific to individual constituents. This section discusses the impacts of different landscape drivers on stream water quality quantiles, starting with the more general influences.

### 4.2.1 General influences

*Cropping land:* Figure 6 shows consistent positive relationships between the area of cropping land within a catchment and the concentrations of pollutant constituents. This result indicates a substantial effect of cropping land, since approximately 80% of the catchments modelled have less than 5% area occupied by cropping land with none having more than 30% (Figure S2). It is worth noting that we also included the *percentage of agricultural land* as a potential predictor; however, it has a weaker impact on constituent concentrations than *cropping area*. The key distinction in defining these two land use types is that the agricultural land includes '*all primary production activities including plantation forests, grazing pastures, cropping and horticulture*', whereas the cropping land focuses only on production of cultivated crops (Australian Bureau of Rural Sciences (ABARES), 2016). The contrasting impacts of these two predictors suggests a potentially greater influence of cropping activities on the spatial patterns of stream water quality within our study region, compared with the impacts of general agricultural activities.

The existing literature has highlighted various dominant pathways in hydrological processes through which cropping can impact individual water quality constituents. For TSS and sediment attached nutrients, cropping activities can lead to increasing concentrations in waterways via: 1) increasing erosion associated with land preparation, such as tillage (Munodawafa, 2007;Zhao et al., 2001) and periods of low vegetation cover (Steegen et al., 2000); and 2) the creation of gullies due to agricultural practices, which also enhances erosion and thus transport of sediments (Newham et al., 2004;Nagle and Ritchie, 2004;Steegen et al., 2000).

Cropping and agricultural land uses are also frequently acknowledged sources of nitrogen and phosphorus in rivers (Wan et al., 2014;Dauer et al., 2000;Sileika et al., 2006). This is often linked to additional sources of fertilizers and manure applied to these land uses, which add to sources of nitrogen and phosphorus that can be mobilized and transported to nearby water bodies (Drewry et al., 2006;Zhu et al., 2012;Arheimer and Lidén, 2000).

Finally, elevated stream salinity is found to be closely associated with cropping and agricultural land uses in catchments. This is likely due to a correspondence between cropping lands and the more salinity prone regions of Victoria, which have higher source levels of salinity to be mobilized into waterways due to low runoff ratios, high levels of evapoconcentration and limited leaching. Further, these cropping lands are also more likely to experience removal of native vegetation and increases in groundwater recharge, which can further enhance salt mobilization (Ghassemi et al., 1991;Yihdego and Webb, 2012).

*Annual temperature:* We observed positive impacts of annual temperature on most quantiles of the phosphorus and nitrogen species, which is consistent with previous studies (Beklioğlu et al., 2017;Bucak et al., 2018). Higher temperature may enhance soil erosion due to drying of soil particles (i.e., desiccation), which in turn leads to higher potential for mobilizing soil-attached nutrients into receiving water bodies (Prosser et al., 2000). Higher temperature can also lead to a higher rate of mineralization of organic materials releasing both phosphorus and nitrogen (Butler et al., 2012;Grierson et al., 1999;Nadelhoffer et al., 1991), making nutrients easier to be mobilized into waterways. There is also a general pattern of more extensive agricultural land use at lower elevations that have higher temperature. Compared with other land uses, agricultural lands are likely associated with higher inputs of nutrients and sediments, which can be flushed into water bodies by rainfall and/or runoff and cause elevated nutrient concentrations in streams.

*Channel slope*: Our model suggests higher nutrient concentrations in streams with lower channel slopes. This is surprising since a steeper channel slope alone is expected to increase the risk of erosion both in-channel and at banks (Mosselman, 1998;Bracmort et al., 2006), which can lead to higher nutrient loads into waterways. One possible explanation of our result is that milder stream channel slopes are typically found in lowland catchments, which are more likely used for agriculture, grazing, and urban purposes. This correlation between channel slope and cropping/agricultural activities seen in our study region (Figure S11) is also commonly in other parts of the world (Chang, 2008;Ye et al., 2009;You et al., 2019).

These lowland catchments with intensive land use often have substantial sources of nutrients, potentially leading to higher nutrient concentrations in the receiving water bodies.

### 4.2.2 Other influences on specific constituents

*TSS:* We saw impacts on all concentration quantiles from soil clay content (positive), and percentages of shrub land (positive) and grass land (negative). Lower quantiles are influenced by storage (percentage of valley bottom area) and maximum catchment elevation (both positive), while higher quantiles are associated with catchment area (positive). This finding is consistent with Tramblay et al. (2010a), in which  soil clay is found to be one of the best predictors for extreme concentrations of sediments, for return periods of 2 and 20 years.

Higher clay contents may lead to higher TSS through enhanced surface runoff (Fu et al., 2015), enhancing the transport of suspended sediment (Oeurng et al., 2010), and due to a smaller soil particle size (less deposition) (Wu et al., 2017). The effects of percentages of grassland and shrub land on stream TSS are in the opposite directions. In our dataset, the former is defined as '*grasslands with tussock, hummock, reeds/rushes*' and the latter is defined as '*open and dry woodlands and shrublands with hummock or tussock grass, Melaleuca shrublands, lignum shrublands, saltbush and chenopods*' (Geoscience Australia, 2011). Therefore, catchments with a higher proportion of shrubland probably have more sparsely distributed vegetation and thus higher potential to mobilize both sediments and nutrients into waterways (Eldridge et al., 2011;Schlesinger et al., 1990). In contrast, grasslands generally have more even vegetation cover, which effectively prevents erosion of sediments and nutrients through surface runoff (Breshears et al., 2003).

The positive effects of catchment area on only the higher quantiles of TSS mean that the larger the catchment, the greater the extreme TSS concentrations. One possible explanation for this is that high TSS concentrations typically occur during higher flow conditions when there is significantly greater sediment transport capacity and stream power available. There are strong geomorphic links between area, slope and stream power, which can lead to a more distinct relationship between catchment area and TSS during higher flows.

*TP:* We saw impacts on most quantiles from erosivity (negative) and runoff perenniality (negative). For erosivity, we would expect the opposite correlation to our result (i.e., positive), as higher rainfall and thus higher erosivity would enhance the mobilization of sediments and associated nutrients (Granger et al., 2010). In explaining this result, we assessed the cross-correlations between the model predictors (Figure S12 in the Supporting Information) and noted strong negative correlations between erosivity and both the annual average temperature and the percentage of cropping land within catchment. These negative correlations, together with the previously discussed effects of temperature and cropping land on most quantiles of nutrients, may explain the negative impact of erosivity on quantile concentrations in our result.

On the other hand, the negative impacts of runoff perenniality on phosphorus mean that higher phosphorus concentrations are generally seen in the more intermittent (i.e. less perennial) rivers. This is expected, as rivers of low perenniality often experience more ‘pulses’ of high nutrient concentrations immediately following dry periods, which is acknowledged as the ‘first flush’ effects (Datry et al., 2014;Stutter et al., 2008). Another plausible explanation is that rivers in agricultural catchments are often more intermittent due to artificial water extraction (Döll and Schmied, 2012;Larned et al., 2010). This is illustrated with a negative correlation between the percentage of agricultural land in catchment and runoff perenniality in our study region (Figure S12).

*FRP:* Some caution is needed in interpreting the FRP results as the model predictive performance is generally low for all quantiles. Nevertheless, the key drivers and their impacts are generally similar to those for TP i.e., erosivity and runoff perenniality both show negative impacts. In addition, the percentages of pasture and forest lands in catchment both have negative effects on the intermediate concentration quantiles (25% to 75%).

*TKN:* Most quantiles of TKN increase with the percentage of pastureland in a catchment, the soil clay content and the number of no-flow days. For catchments with greater proportion of pasture land, higher nutrient concentrations are expected due to the application of fertilizer and particularly the deposition of organic nitrogen in faecal matter from livestock (Arheimer and Lidén, 2000;Drewry et al., 2006). High soil clay content can enhance sediment mobilization and transport (see previous discussions for

TSS), which are also expected to export more nutrients into waterways in particulate forms. Lastly, the positive effects of the number of no-flow days is similar to that of runoff perenniality as discussed for TP, which may indicate a ‘first flush’ mechanism that introduces concentration pulses after a dry period.

*$NO_x$:* Urbanization has positive impacts on most concentration quantiles. This can be explained as urban areas produce stormwater with more nitrate and, in some cases, might also be associated with discharge of treated waste water (Duncan et al., 2017;Zendehbad et al., 2019). Further, higher rate of fossil-fuel emission in urban areas may also contribute more oxidized nitrogen to streams via atmospheric deposition (Howarth et al., 2002); however, this is an unlikely explanation for our result, since the urban areas within our study region are predominantly only small towns.

*EC:* The positive effects of warm season temperature on EC in our results is likely due to higher aridity (and thus higher evapotranspiration) in the warmer regions, which leads to more accumulation of salts in soil through evapoconcentration (Poulsen et al., 2006). The negative effects of runoff perenniality on EC concentration can be due to saline groundwater development in drier parts of the study area (which often have lower runoff perenniality), thus becoming a substantial source of stream salinity (Peck and Hatton, 2003). Lastly, rainfall has mixed effects on EC concentrations: cold season rainfall generally having negative impacts, while positive effects are seen for the rainfall during wet season (which is highly positively correlated with cold season rainfall, see Figure S12). These rainfall impacts seem rather confusing when considering physical processes and are likely results of statistical artefacts as discussed subsequently.

### 4.2.3 Implications on catchment management

The key catchment characteristics identified are useful to inform planning and policymaking for catchment managers. If a key catchment characteristic identified is controllable (e.g., land use and land cover), management interventions can be designed to modify this catchment condition to improve water quality status. For example, higher proportions of cropping land in catchment is shown to lead to increases in the quantile concentrations of most pollutants, suggesting that catchment management in this region may seek to improve water quality by improving land management or restricting the area of cropping land.

Some natural catchment characteristics identified are difficult to modify, such as natural characteristics including hydrologic and climatic conditions. One example of this in the study results is annual temperature, which has positive impacts of on most quantiles of the phosphorus and nitrogen species. Management actions cannot be taken to directly influence these catchment conditions to improve water quality. Instead, this information is useful to highlight potential hotspots and help target management of water quality issues. The links between water quality and catchment conditions established in this study can also provide a rapid tool to estimate water quality status for unmonitored rivers.

## 4.3 Limitations

It is worth noting that all the key landscape characteristics that affect water quality were identified with: a) a process-based, comprehensive literature review to identify a broader range of potential characteristics that are expected to influence stream water quality (as detailed in Section 2.2); and b) a data-driven modelling approach to finalize the key driving variables for each constituent, based on statistical relationships between constituent concentrations and their predictors, which can be cross-correlated. As a result, the key factors that we identified might be surrogates for actual physical or chemical processes.

Cross-correlation between predictors can lead to issues in the statistical inferences (see Figure S12 for the correlations amongst all 50 spatial characteristics considered). An example here is the opposite effects of cold season and wet season rainfall on EC. These two predictors are extremely highly correlated (Spearman $\rho = 0.99$) yet their influences suggested in the models are in opposite directions, which are hence likely to be statistical artefacts. In general, most strong correlations (absolute Spearman $\rho > 0.7$) occur within individual categories of predictors, for example among the climate variables. Across different categories, correlations are generally weaker, although climate is often moderately correlated with other variables. Therefore, it is likely that the issue of one predictor acting as a surrogate for another primarily occurs within the same category of predictors.

We acknowledge that the landscape characteristics of catchments used to explain the spatial variability of water quality quantiles are likely not comprehensive. Due to limited data availability, some critical information is missing. Examples include the presence of wastewater treatment plants discharging to stream (Jin et al., 2017) and land use intensity including fertilizer application rate and stocking rate (Smith et al., 2013). The current models are limited with these data, which could be a reason that a few catchments become common outliers across different constituents (as discussed in Section 4.1). These data are generally less available across landscape and at a matching resolution to water quality modelling (e.g., catchment scale), which highlights a key monitoring need for future model developments.

The water quality quantiles models developed in this study do not explicitly address any temporal trends in water quality. To understand the extent of temporal trends, we estimated the direction and significance of trends in the six constituents at each site using the non-parametric Theil-Sen estimator (Theil, 1992;Sen, 1968). Within the 102 study catchments, significant trends ($p < 0.05$) are seen in about half of the catchments. Specifically, at a maximum of 60 sites for TSS and a minimum of 49 sites for EC. Figure S13 (Supporting Information) shows the time-series of three sites that have the greatest magnitudes of significant trends for each constituent. However, in relative to the large number of sites showing significant trends, the deterioration of model performance on sites with significant trends is only marginal (Figure S14, Supporting Information), indicating the likely limited impacts of temporal trends on the quantile models. The causes of these trends remain an open question. A previous study for the same region illustrated that TSS concentrations may have experienced systematic shifts due to the large-scale prolong drought between 1997-2009 in south-east Australia (the ‘Millennium drought’) (Guo et al., 2020). Further, changes in land use and management occurring at the finer spatial scales have been found to relate to trends in stream nutrient levels (Smith et al., 2013). Future studies can seek to explore temporal trends specifically by estimating water quality quantiles for different temporal periods at each site that experiences significant trends, and thus improve the representativeness of quantiles on the typical water quality conditions.

# 5 Conclusion

Using long-term large-scale water quality monitoring data from Victoria, Australia, statistical models were developed to predict the concentrations of the 10%, 25%, 50%, 75% and 90% quantiles for six key water quality constituents (TSS, TP, FRP, TKN, $NO_x$ and EC) using catchment landscape characteristics. Amongst these constituents, the spatial variation in all quantiles for the three more conservative constituents, TSS, TKN and EC, are well explained (66%-96%) by the models. In contrast, predictive capacity remains limited for the quantiles of the more reactive constituents, TP, FRP and $NO_x$, with 37%-73% spatial variation explained.

The models also identified key catchment characteristics that are closely related to spatial variation in the quantiles of water quality constituents. Across constituents, annual temperature, percentage of cropping land area in catchment and channel slope, are identified to be the most common factors that influence spatial variations of concentration quantiles. The selection of key catchment characteristics is generally robust as illustrated with a 10-fold cross-validation.

The models provide a parsimonious alternative approach to predicting concentration quantiles across catchments, and thus to identify poor water quality hot spots. The ability to estimate quantiles of constituent concentration is useful since quantiles of concentrations are often important water quality indicators used in management guidelines for water quality evaluation.

## Supporting Information legends

The Supporting Information contains supplementary figures and tables to the main manuscript.

## References

Ahearn, D. S., Sheibley, R. W., Dahlgren, R. A., Anderson, M., Johnson, J., and Tate, K. W.: Land use and land cover influence on water quality in the last free-flowing river draining the western Sierra Nevada, California, Journal of Hydrology, 313, 234-247, https://doi.org/10.1016/j.jhydrol.2005.02.038, 2005.

Ali, K. F., and De Boer, D. H.: Spatial patterns and variation of suspended sediment yield in the upper Indus River basin, northern Pakistan, Journal of Hydrology, 334, 368-387, https://doi.org/10.1016/j.jhydrol.2006.10.013, 2007.

ANZECC, and ARMCANZ.: Australian and New Zealand guidelines for fresh and marine water quality. Volume 1, The guidlines., 2000.

Arheimer, B., and Lidén, R.: Nitrogen and phosphorus concentrations from agricultural catchments—influence of spatial and temporal variables, Journal of Hydrology, 227, 140-159, https://doi.org/10.1016/S0022-1694(99)00177-8, 2000.

Australian and New Zealand Environment and Conservation Council, and Zealand., A. a. R. M. C. o. A. a. N.: Australian and New Zealand Guidelines for Fresh and Marine Water Quality. In: Volume 1: The Guidelines, 4, Canberra, 2000.

Australian Bureau of Rural Sciences (ABARES): Australian Land Use and Management (ALUM) Classification version 8. 2016 (Ed.), 2016.

Australian Water Technologies: Victorian water quality monitoring network and state biological monitoring programme manual of procedures, Australian Water Technologies, 68 Ricketts Rd, Mt Waverley VIC 3149, 1999.

Bayarri, M. J., Berger, J. O., Forte, A., and García-Donato, G.: Criteria for Bayesian model choice with application to variable selection, The Annals of statistics, 40, 1550-1577, 2012.

Beklioğlu, M., Bucak, T., Coppens, J., Bezirci, G., Tavşanoğlu, Ü. N., Çakıroğlu, A. İ., Levi, E. E., Erdoğan, Ş., Filiz, N., and Özkan, K.: Restoration of eutrophic lakes with fluctuating water levels: A 20-year monitoring study of two inter-connected lakes, Water, 9, 127, 2017.

Best, J.: Anthropogenic stresses on the world's big rivers, Nature Geoscience, 12, 7-21, 10.1038/s41561-018-0262-x, 2019.

Bracmort, K. S., Arabi, M., Frankenberger, J., Engel, B. A., and Arnold, J. G.: Modeling long-term water quality impact of structural BMPs, Transactions of the ASABE, 49, 367-374, 2006.

Breshears, D. D., Whicker, J. J., Johansen, M. P., and Pinder, J. E.: Wind and water erosion and transport in semi-arid shrubland, grassland and forest ecosystems: quantifying dominance of horizontal wind-driven transport, Earth Surface Processes and Landforms, 28, 1189-1209, 10.1002/esp.1034, 2003.

Bucak, T., Trolle, D., Tavşanoğlu, Ü. N., Çakıroğlu, A. İ., Özen, A., Jeppesen, E., and Beklioğlu, M.: Modeling the effects of climatic and land use changes on phytoplankton and water quality of the largest Turkish freshwater lake: Lake Beyşehir, Science of The Total Environment, 621, 802-816, https://doi.org/10.1016/j.scitotenv.2017.11.258, 2018.

Bureau of Meteorology: Geofabric V2 2012.

Butler, S. M., Melillo, J. M., Johnson, J., Mohan, J., Steudler, P. A., Lux, H., Burrows, E., Smith, R., Vario, C., and Scott, L.: Soil warming alters nitrogen cycling in a New England forest: implications for ecosystem function and structure, Oecologia, 168, 819-828, 2012.

Camarero, L., and Catalan, J.: Atmospheric phosphorus deposition may cause lakes to revert from phosphorus limitation back to nitrogen limitation, Nature Communications, 3, 1118, 10.1038/ncomms2125, 2012.

Chang, H.: Spatial analysis of water quality trends in the Han River basin, South Korea, Water Res., 42, 3285-3304, 10.1016/j.watres.2008.04.006, 2008.

Collins, A. L., Zhang, Y., McChesney, D., Walling, D. E., Haley, S. M., and Smith, P.: Sediment source tracing in a lowland agricultural catchment in southern England using a modified procedure combining statistical analysis and numerical modelling, Science of The Total Environment, 414, 301-317, https://doi.org/10.1016/j.scitotenv.2011.10.062, 2012.

Coynel, A., Seyler, P., Etcheber, H., Meybeck, M., and Orange, D.: Spatial and seasonal dynamics of total suspended sediment and organic carbon species in the Congo River, Global Biogeochemical Cycles, 19, 10.1029/2004gb002335, 2005.

Datry, T., Larned, S. T., and Tockner, K.: Intermittent Rivers: A Challenge for Freshwater Ecology, BioScience, 64, 229-235, 10.1093/biosci/bit027, 2014.

Dauer, D. M., Ranasinghe, J. A., and Weisberg, S. B.: Relationships between benthic community condition, water quality, sediment quality, nutrient loads, and land use patterns in Chesapeake Bay, Estuaries, 23, 80-96, 10.2307/1353227, 2000.

Department of Environment Land Water and Planning Victoria: Index of condition system, 2014.

Victorian water measurement information system: http://data.water.vic.gov.au/monitoring.htm, 2016a.

Department of Environment Land Water and Planning Victoria: Victorian water measurement information system. . 2016b.

Döll, P., and Schmied, H. M.: How is the impact of climate change on river flow regimes related to the impact on mean annual runoff? A global-scale analysis, Environmental Research Letters, 7, 014037, 10.1088/1748-9326/7/1/014037, 2012.

Drewry, J. J., Newham, L. T. H., Greene, R. S. B., Jakeman, A. J., and Croke, B. F. W.: A review of nitrogen and phosphorus export to waterways: context for catchment modelling, Marine and Freshwater Research, 57, 757-774, https://doi.org/10.1071/MF05166, 2006.

Duncan, J. M., Welty, C., Kemper, J. T., Groffman, P. M., and Band, L. E.: Dynamics of nitrate concentration-discharge patterns in an urban watershed, Water Resources Research, 53, 7349-7365, 10.1002/2017wr020500, 2017.

Eldridge, D. J., Bowker, M. A., Maestre, F. T., Roger, E., Reynolds, J. F., and Whitford, W. G.: Impacts of shrub encroachment on ecosystem structure and functioning: towards a global synthesis, Ecology Letters, 14, 709-722, 10.1111/j.1461-0248.2011.01630.x, 2011.

Fu, T., Chen, H., Zhang, W., Nie, Y., Gao, P., and Wang, K.: Spatial variability of surface soil saturated hydraulic conductivity in a small karst catchment of southwest China, Environmental Earth Sciences, 74, 2381-2391, 10.1007/s12665-015-4238-5, 2015.

Fuss, C. B., Driscoll, C. T., Groffman, P. M., Campbell, J. L., Christenson, L. M., Fahey, T. J., Fisk, M. C., Mitchell, M. J., Templer, P. H., Durán, J., and Morse, J. L.: Nitrate and dissolved organic carbon mobilization in response to soil freezing variability, Biogeochemistry, 131, 35-47, 10.1007/s10533-016-0262-0, 2016.

García-Donato, G., and Martínez-Beneito, M. A.: On Sampling Strategies in Bayesian Variable Selection Problems With Large Model Spaces, Journal of the American Statistical Association, 108, 340-352, 10.1080/01621459.2012.742443, 2013.

Geoscience Australia: Dams and water storages. . 2004.

Geoscience Australia: Environmental Attributes Dataset. 2011.

Ghassemi, F., Jakeman, A. J., and Nix, H. A.: Human induced salinisation and the use of quantitative methods, Environment International, 17, 581-594, https://doi.org/10.1016/0160-4120(91)90170-U, 1991.

BayesVarSel: Bayes Factors, Model Choice and Variable Selection in Linear Models. R package version 1.8.0.: https://CRAN.R-project.org/package=BayesVarSel, 2017.

Granger, S. J., Bol, R., Anthony, S., Owens, P. N., White, S. M., and Haygarth, P. M.: Chapter 3 - Towards a Holistic Classification of Diffuse Agricultural Water Pollution from Intensively Managed Grasslands on Heavy Soils, in: Advances in Agronomy, Academic Press, 83-115, 2010.

Grierson, P., Comerford, N., and Jokela, E.: Phosphorus mineralization and microbial biomass in a Florida Spodosol: effects of water potential, temperature and fertilizer application, Biology and fertility of soils, 28, 244-252, 1999.

Guo, D., Lintern, A., Webb, J. A., Ryu, D., Bende-Michl, U., Liu, S., and Western, A. W.: A data-based predictive model for spatiotemporal variability in stream water quality, Hydrol. Earth Syst. Sci., 24, 827-847, 10.5194/hess-24-827-2020, 2020.

Guyon, I., and Elisseeff, A.: An introduction to variable and feature selection, Journal of machine learning research, 3, 1157-1182, 2003.

Hamilton, P. A., and Miller, T. L.: Lessons from the national water-quality assessment, Journal of Soil and Water Conservation, 57, 16A-21A, 2002.

Helsel, D. R.: Less than obvious - statistical treatment of data below the detection limit, Environmental Science & Technology, 24, 1766-1774, 10.1021/es00082a001, 1990.

Holloway, J. M., Dahlgren, R. A., Hansen, B., and Casey, W. H.: Contribution of bedrock nitrogen to high nitrate concentrations in stream water, Nature, 395, 785-788, 10.1038/27410, 1998.

Howarth, R. W., Sharpley, A., and Walker, D.: Sources of nutrient pollution to coastal waters in the United States: Implications for achieving coastal water quality goals, Estuaries, 25, 656-676, 10.1007/BF02804898, 2002.

Jin, Z., Zhang, X., Li, J., Yang, F., Kong, D., Wei, R., Huang, K., and Zhou, B.: Impact of wastewater treatment plant effluent on an urban river, Journal of Freshwater Ecology, 32, 697-710, 2017.

Larned, S. T., Datry, T., Arscott, D. B., and Tockner, K.: Emerging concepts in temporary ‐ river ecology, Freshwater Biology, 55, 717-738, 2010.

Lintern, A., Webb, J. A., Ryu, D., Liu, S., Bende-Michl, U., Waters, D., Leahy, P., Wilson, P., and Western, A. W.: Key factors influencing differences in stream water quality across space, Wiley Interdisciplinary Reviews: Water, 5, e1260, doi:10.1002/wat2.1260, 2018a.

Lintern, A., Webb, J. A., Ryu, D., Liu, S., Waters, D., Leahy, P., Bende-Michl, U., and Western, A. W.: What Are the Key Catchment Characteristics Affecting Spatial Differences in Riverine Water Quality?, Water Resources Research, 54, 7252-7272, doi:10.1029/2017WR022172, 2018b.

Luca Scrucca: Package ‘GA’. In: The Comprehensive R Archive Network,, 2019.

May, R., Dandy, G., and Maier, H.: Review of input variable selection methods for artificial neural networks, in: Artificial neural networks-methodological advances and biomedical applications, InTech, 2011.

Mosselman, E.: Morphological modelling of rivers with erodible banks, Hydrological Processes, 12, 1357-1370, 10.1002/(sici)1099-1085(19980630)12:8<1357::Aid-hyp619>3.0.Co;2-7, 1998.

Munodawafa, A.: Assessing nutrient losses with soil erosion under different tillage systems and their implications on water quality, Physics and Chemistry of the Earth, Parts A/B/C, 32, 1135-1140, https://doi.org/10.1016/j.pce.2007.07.033, 2007.

Nadelhoffer, K., Giblin, A., Shaver, G., and Laundre, J.: Effects of temperature and substrate quality on element mineralization in six arctic soils, Ecology, 72, 242-253, 1991.

Nagle, G. N., and Ritchie, J. C.: Wheat field erosion rates and channel bottom sediment sources in an intensively cropped northeastern Oregon drainage basin, Land Degradation & Development, 15, 15-26, 10.1002/ldr.587, 2004.

Nash, J. E., and Sutcliffe, J. V.: River flow forecasting through conceptual models part I — A discussion of principles, Journal of Hydrology, 10, 282-290, https://doi.org/10.1016/0022-1694(70)90255-6, 1970.

Newham, L. T. H., Letcher, R. A., Jakeman, A. J., and Kobayashi, T.: A framework for integrated hydrologic, sediment and nutrient export modelling for catchment-scale management, Environmental Modelling & Software, 19, 1029-1038, https://doi.org/10.1016/j.envsoft.2003.11.006, 2004.

Oeurng, C., Sauvage, S., and Sánchez-Pérez, J.-M.: Dynamics of suspended sediment transport and yield in a large agricultural catchment, southwest France, Earth Surface Processes and Landforms, 35, 1289-1301, 10.1002/esp.1971, 2010.

Onderka, M., Wrede, S., Rodný, M., Pfister, L., Hoffmann, L., and Krein, A.: Hydrogeologic and landscape controls of dissolved inorganic nitrogen (DIN) and dissolved silica (DSi) fluxes in heterogeneous catchments, Journal of Hydrology, 450-451, 36-47, https://doi.org/10.1016/j.jhydrol.2012.05.035, 2012.

Ongley, E. D., Xiaolan, Z., and Tao, Y.: Current status of agricultural and rural non-point source Pollution assessment in China, Environmental Pollution, 158, 1159-1168, https://doi.org/10.1016/j.envpol.2009.10.047, 2010.

Peck, A. J., and Hatton, T.: Salinity and the discharge of salts from catchments in Australia, Journal of Hydrology, 272, 191-202, https://doi.org/10.1016/S0022-1694(02)00264-0, 2003.

Poulsen, D. L., Simmons, C. T., Le Galle La Salle, C., and Cox, J. W.: Assessing catchment-scale spatial and temporal patterns of groundwater and stream salinity, Hydrogeology Journal, 14, 1339-1359, 10.1007/s10040-006-0065-9, 2006.

Prosser, I. P., Hughes, A. O., and Rutherfurd, I. D.: Bank erosion of an incised upland channel by subaerial processes: Tasmania, Australia, Earth Surface Processes and Landforms, 25, 1085-1101, 10.1002/1096-9837(200009)25:10<1085::Aid-esp118>3.0.Co;2-k, 2000.

Rothwell, J. J., Dise, N. B., Taylor, K. G., Allott, T. E. H., Scholefield, P., Davies, H., and Neal, C.: A spatial and seasonal assessment of river water chemistry across North West England, Science of The Total Environment, 408, 841-855, https://doi.org/10.1016/j.scitotenv.2009.10.041, 2010.

Saft, M., Western, A. W., Zhang, L., Peel, M. C., and Potter, N. J.: The influence of multiyear drought on the annual rainfall ‐ runoff relationship: An Australian perspective, Water Resources Research, 51, 2444-2463, doi:10.1002/2014WR015348, 2015.

Schlesinger, W. H., Reynolds, J. F., Cunningham, G. L., Huenneke, L. F., Jarrell, W. M., Virginia, R. A., and Whitford, W. G.: Biological feedbacks in global desertification, Science, 247, 1043-1048, 1990.

Schwarz, G.: Estimating the dimension of a model, The annals of statistics, 6, 461-464, 1978.

Sen, P. K.: Estimates of the regression coefficient based on Kendall's tau, Journal of the American statistical association, 63, 1379-1389, 1968.

Sileika, A. S., Stålnacke, P., Kutra, S., Gaigalis, K., and Berankiene, L.: Temporal and Spatial Variation of Nutrient Levels in the Nemunas River (Lithuania and Belarus), Environmental Monitoring and Assessment, 122, 335-354, 10.1007/s10661-006-9186-9, 2006.

Smith, A. P., Western, A. W., and Hannah, M. C.: Linking water quality trends with land use intensification in dairy farming catchments, J. Hydrol., 476, 1-12, 2013.

Srivastava, P., Migliaccio, K., and Simunek, J.: Landscape models for simulating water quality at point, field, and watershed scales, Transactions of the ASABE, 50, 1683-1693, 2007.

Steegen, A., Govers, G., Nachtergaele, J., Takken, I., Beuselinck, L., and Poesen, J.: Sediment export by water from an agricultural catchment in the Loam Belt of central Belgium, Geomorphology, 33, 25-36, https://doi.org/10.1016/S0169-555X(99)00108-7, 2000.

Stutter, M. I., Langan, S. J., and Cooper, R. J.: Spatial contributions of diffuse inputs and within-channel processes to the form of stream water phosphorus over storm events, Journal of Hydrology, 350, 203-214, https://doi.org/10.1016/j.jhydrol.2007.10.045, 2008.

Terrestrial Ecosystem Research Network: Soil and landscape grid of Australia. 2016.

Theil, H.: A Rank-Invariant Method of Linear and Polynomial Regression Analysis, in: Henri Theil's Contributions to Economics and Econometrics: Econometric Theory and Methodology, edited by: Raj, B., and Koerts, J., Springer Netherlands, Dordrecht, 345-381, 1992.

Tramblay, Y., Ouarda, T. B. M. J., St-Hilaire, A., and Poulin, J.: Regional estimation of extreme suspended sediment concentrations using watershed characteristics, Journal of Hydrology, 380, 305-317, https://doi.org/10.1016/j.jhydrol.2009.11.006, 2010a.

Tramblay, Y., Saint-Hilaire, A., Ouarda, T. B. M. J., Moatar, F., and Hecht, B.: Estimation of local extreme suspended sediment concentrations in California Rivers, Science of The Total Environment, 408, 4221-4229, https://doi.org/10.1016/j.scitotenv.2010.05.001, 2010b.

UN Water: Wastewater Management: A UN-Water Analytical Brief, 2015.

UNEP: A Snapshot of the World's Water Quality: Towards a global assessment. United Nations Environment Programme, Nairobi, Kenya, 162pp, 2016.

United States Environmental Protection Agency: Rivers and Streams in Nutrient Ecoregion I, OFFICE OF WATER,

OFFICE OF SCIENCE AND TECHNOLOGY,

HEALTH AND ECOLOGICAL CRITERIA DIVISION., 2001.

United States Environmental Protection Agency: State adoption of numeric nutrient standards (1998-2008). 2008.

van der Perk, M., Owens, P. N., Deeks, L. K., Rawlins, B. G., Haygarth, P. M., and Beven, K. J.: Controls on catchment-scale patterns of phosphorus in soil, streambed sediment, and stream water, Journal of environmental quality, 36, 694-708, 10.2134/jeq2006.0175, 2007.

Wan, R., Cai, S., Li, H., Yang, G., Li, Z., and Nie, X.: Inferring land use and land cover impact on stream water quality using a Bayesian hierarchical modeling approach in the Xitiaoxi River Watershed, China, J Environ Manage, 133, 1-11, 10.1016/j.jenvman.2013.11.035, 2014.

Wang, Q. J., Shrestha, D. L., Robertson, D. E., and Pokhrel, P.: A log-sinh transformation for data normalization and variance stabilization, Water Resources Research, 48, W05514, doi:10.1029/2011WR010973, 2012.

Wu, X., Wei, Y., Wang, J., Xia, J., Cai, C., Wu, L., Fu, Z., and Wei, Z.: Effects of erosion degree and rainfall intensity on erosion processes for Ultisols derived from quaternary red clay, Agriculture, Ecosystems & Environment, 249, 226-236, https://doi.org/10.1016/j.agee.2017.08.023, 2017.

Ye, L., Cai, Q.-h., Liu, R.-q., and Cao, M.: The influence of topography and land use on water quality of Xiangxi River in Three Gorges Reservoir region, Environmental Geology, 58, 937-942, 10.1007/s00254-008-1573-9, 2009.

Yihdego, Y., and Webb, J.: Modelling of seasonal and long-term trends in lake salinity in southwestern Victoria, Australia, Journal of Environmental Management, 112, 149-159, https://doi.org/10.1016/j.jenvman.2012.07.002, 2012.

You, Q., Fang, N., Liu, L., Yang, W., Zhang, L., and Wang, Y.: Effects of land use, topography, climate and socio-economic factors on geographical variation pattern of inland surface water quality in China, PLOS ONE, 14, e0217840, 10.1371/journal.pone.0217840, 2019.

Zendehbad, M., Cepuder, P., Loiskandl, W., and Stumpp, C.: Source identification of nitrate contamination in the urban aquifer of Mashhad, Iran, Journal of Hydrology: Regional Studies, 25, 100618, https://doi.org/10.1016/j.ejrh.2019.100618, 2019.

Zhao, S. L., Gupta, S. C., Huggins, D. R., and Moncrief, J. F.: Tillage and nutrient source effects on surface and subsurface water quality at corn planting, Journal of environmental quality, 30, 998-1008, 10.2134/jeq2001.303998x, 2001.

Zhu, Q., Schmidt, J. P., and Bryant, R. B.: Hot moments and hot spots of nutrient losses from a mixed land use watershed, Journal of Hydrology, 414-415, 393-404, https://doi.org/10.1016/j.jhydrol.2011.11.011, 2012.

## Figures

**Figure 1: a) Maps of the 102 water quality monitoring sites analyzed in this study and their catchment boundaries. Insert shows location of the state of Victoria in Australia. Panels b), c) and d) illustrate the spatial distribution of the elevation, annual average rainfall and annual average temperature across Victoria, respectively.**

**Figure 2: Spatial variation of the 75th percentile of concentrations for TSS, TP, FRP, TKN, NOx, and EC across the 102 study sites, with elevation in the background. Colors show the ranges between inter-quartiles. See Figures S4 to S7 in the Supporting Information for summary plots of other quantiles.**

**Figure 3: Modelled versus observed 75th percentile of concentration for each constituents across all 102 sites, plotted in the log-transformed scale (top row) and the back-transformed scale (bottom row). The NSE (Nash-Sutcliffe Efficiency) values for each constituent-quantile model and the 1:1 lines (red dashed line) are also shown.**

**Figure 4: Calibration performance (as the Nash-Sutcliffe Efficiency, NSE) across all 102 sites, in the scale of actual measurement i.e., untransformed scale (y-axis), versus the corresponding NSE values assessed in the scale of modelling i.e., log-transformed scale (x-axis).**

**Figure 5: Nash-Sutcliffe Efficiency (NSE) of the full model and all 10-fold cross-validation models, for each constituent-quantile combination.**

**Figure 6: Catchment characteristics that have been identified as key predictors for the spatial variations of each quantile concentrations for each constituent (y-axis) and their corresponding impacts, summarized by their coefficient values (x-axis). Categories of catchment characteristics are differentiated by colors. All predictors and predictands were standardized, so that the coefficient values are comparable across predictors, constituents and quantiles.**

**Figure 7: Robustness of key predictors selected in the models of quantile for individual constituents. Black dots highlight catchment characteristics selected in the main models, for each**

**constituent (panel) and each quantile (x-axes). Red shades show the number of times each predictor being selected in the 10 cross-validation models.**