Journal of Hydrology: Regional Studies 4 (2015) 439–451

Contents lists available at ScienceDirect

Journal of Hydrology: Regional Studies

journal homepage: www.elsevier.com/locate/ejrh

A simulation based approach to quantify the difference between event-based and routine water quality monitoring schemes

J.S. Lessels[a,b,*], T.F.A. Bishop[b]

[a] Northern Rivers Institute, University of Aberdeen, Aberdeen AB24 3UF, UK
[b] Faculty of Agriculture and Environment, University of Sydney, Sydney, NSW, Australia

ARTICLE INFO

Article history:
Received 18 March 2015
Received in revised form 25 June 2015
Accepted 27 June 2015
Available online 24 August 2015

Keywords:
Australia
Water quality
Sampling
Linear mixed models
Simulation
Total phosphorous

ABSTRACT

Study region: South eastern Australia.
Study focus: This region is characterised with rainfall events that are associated with large exports of nutrients and sediments. Many water quality monitoring schemes use a form of event-based sampling to quantify these exports. Previous water quality studies that have evaluated different sampling schemes often rely on continuously monitored water quality data. However, many catchment authorities only have access to limited historical data which consists of event-based and monthly routine samples. Therefore there is a need to develop a method that assesses the importance of sampling events using information from limited historical data. This work presents a simulation based approach using unconditional simulation based on historical stream discharge. Such an approach offers site-specific information on optimal sampling schemes. A linear mixed model is used to model the relationship between total phosphorus and stream discharge and the auto-correlation of total phosphorus.
New hydrological insights for the region: The inclusion of event-based sampling improved annual load estimates of all sites with a maximum RMSE difference of 16.11 tonnes between event-based and routine sampling. Based on the accuracy of annual loads, event-based sampling was found to be more important in catchments with a large relief and high annual rainfall in this region. Using this approach, different sampling schemes can be compared based on limited historical data.

1. Introduction

The accuracy of water quality load estimates is directly related to the monitoring design used to collect the water samples. Ideally, the load of a water quality variable would be derived using continuously sampled data. However, it is not financially viable for most water quality monitoring programmes to collect continuous data (Bartley et al., 2012; Burt et al., 2011; Drewry et al., 2009). Therefore water quality sampling schemes must be designed to provide accurate load estimates with limited samples. Monthly sampling is commonly used throughout Australia (Bartley et al., 2012). However, monthly sampling often misses key rainfall events (Drewry et al., 2009), therefore many sampling schemes include a form of event sampling. Most

* Corresponding author.
E-mail addresses: jason.lessels@abdn.ac.uk (J.S. Lessels), thomas.bishop@sydney.edu.au (T.F.A. Bishop).

http://dx.doi.org/10.1016/j.ejrh.2015.06.020

studies examining the effect of different sampling schemes have been limited to catchments with access to continuous sampled data. Therefore there is a need for methods which can use site-specific historical water quality data to assess the effect of different sampling schemes on load estimation (e.g. event-based sampling).

Based on studies with access to sufficient data (Drewry et al., 2009; Johnes, 2007; Hopmans and Bren, 2007), it is generally accepted that a form of sampling is required during rainfall events to capture periods of high nutrient and sediment exports. This is especially important in catchment with large events separated by long periods periods of base-flow as large exports of nutrients and sediments occur during short rainfall events (Jones et al., 2011; Drewry et al., 2009; Gao, 2008; Hopmans and Bren, 2007). A study by Hopmans and Bren (2007) observed large sediment exports during events, with 70% of six years of sediment load being exported during a single event in south east Australia. One particular water quality property of interest is total phosphorus (TP), as in large concentrations TP can cause algal blooms (Davis and Koop, 2006; Kristiana et al., 2011). A study focusing on 17 streams in the UK, found 20% of annual TP was exported during a single day and the largest five events within a year contributed to 42% of the annual TP load (Johnes, 2007).

To improve monitoring schemes it is important to understand the relationship between catchment characteristics and water quality variable exports. This information is required to help the design of monitoring schemes in unmonitored catchments. The relationship between catchment characteristics (catchment size, slope, rainfall, stream discharge, land use and land cover) has been investigated by several studies (Ahearn et al., 2005; Banner et al., 2009; Mehdi et al., 2015; Saaltink et al., 2014; Sliva and Williams, 2001). Relationships between phosphorus and the proportion of land under agriculture (Ahearn et al., 2005) and catchment size (Ahearn et al., 2005; Johnes, 2007) have been found. Focussing on phosphorus, Banner et al. (2009) found a relationship between phosphorus, catchment topography and land use for 2 sites in Kansas, USA. In addition, urbanisation (Sliva and Williams, 2001) and population density (Johnes, 2007) have been found to have an effect on water quality. Johnes (2007) also found a relationship between the base-flow index and the uncertainty of TP load estimates, with catchments with a low base-flow index having larger uncertainty than streams with a high base-flow index.

Sampling schemes can be divided into two main categories; probability and non-probability based (de Gruijter et al., 2006). Probability based methods rely on known inclusion probabilities to provide unbiased estimates of the mean and its uncertainty. Non-probability sampling should use a model-based approach as the inclusion probabilities are unknown (de Gruijter et al., 2006; Lark and Cullis, 2004). Several probability based sampling schemes have been shown to provide accurate estimates of suspended sediments (Lewis, 1996; Thomas, 1985, 1988; Thomas and Lewis, 1993, 1995). However, non-probabilistic sampling schemes are more commonly used. One of the most commonly used schemes is to sample at equal intervals in time in combination with event sampling (for examples see Salles et al. (2008), Birkel et al. (2011)). This is one of the most common sampling schemes due to ease of implementation with available automatic sampling equipment.

Load estimation methods offer the ability to provide estimates over different time intervals (e.g. event-based or annually). The majority of these cannot be used with commonly applied sampling schemes (e.g. monthly or a combination of monthly and event-based sampling) as the sampling schemes are non-probabilistic, i.e. it is not possible to determine the probability of taking a sample at a specific time. Therefore average, ratio and regression based load estimation methods should not be under these situations as they require a form of probabilistic sampling, a requirement that has been noted by several water quality based studies (Cohn et al., 1992; Cohn, 2005; Cooper and Watts, 2002; Crawford, 1991; Thomas, 1985, 1988).

Linear mixed models (LMM) provide unbiased water quality load estimates without the assumption of probabilistic sampling schemes. Differing from simple linear models, LMM account for auto-correlation between samples within the error term of the model. Linear mixed models are regularly applied in soil science to account for the spatial auto-correlation between samples (Lark and Cullis, 2004). Temporal water quality data is similar to spatial soil data as water quality data has been shown to be auto-correlated through time (Kuhnert et al., 2012; Wang et al., 2011). Similarly to simple linear models, LMM also provide the ability to incorporate additional covariates (e.g. stream discharge and turbidity) to improve predictions (Lessels and Bishop, 2013)

A lack of data is a major limitation of water quality studies. Several studies have found strong linear trends between water quality variables and low cost continuously measured surrogates (e.g. stream discharge and turbidity) and have used these relationships to provide continuous predictions of water quality variables (Webb et al., 2000; Kim and Furumai, 2012; Lewis, 1996; Wang et al., 2011; Kuhnert et al., 2012). Webb et al. (2000) simulated numerous water quality variables using a linear relationship with continuously monitored stream discharge data. These simulations provided a method to explore the accuracy of various load estimation methods based on limited historical water quality data (Webb et al., 2000), However, the simulation method of Webb et al. (2000) did not examine the potential for temporal auto-correlation between the observed water quality samples and used a single realisation of the relationship with stream discharge. An alternative approach is to use unconditional Gaussian simulation to simulate data on a model fitted using a LMM. The advantage of this approach is that the model used to simulated the TP data is based on a valid statistical model describing the (co-)variation between water quality and discharge. The simulated data is based on the linear relationship with stream discharge while respecting the temporal auto-correlation of the observed water quality variables (Gebbers and Bruin, 2010).

In this work we simulate continuous TP using its relationship with stream discharge based on the LMM. Using these simulations we investigate the effect of using event-based sampling (in addition to routine sampling) in terms of the accuracy of annual load estimates. In addition, the effect of including event-based sampling is related to catchment characteristics as the information from this may help improve the design of monitoring schemes in unmonitored catchments. Therefore the aims of this work are to: