Journal of Hydrology 520 (2015) 439–447

Contents lists available at ScienceDirect

Journal of Hydrology

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

Statistical modelling and power analysis for detecting trends in total suspended sediment loads

You-Gan Wang [a,*], Shen S.J. Wang [a], Jason Dunlop [b]

[a] Centre for Applications in Natural Resource Mathematics (CARM), School of Mathematics and Physics, The University of Queensland, Queensland 4072, Australia
[b] Water Assessment and Systems, Environmental Monitoring and Assessment Sciences, Science Division, Department of Science, Information Technology, Innovation and the Arts (DSITIA), 41 Boggo Road, Dutton Park, QLD 4102, Australia

ARTICLE INFO

Article history:
Received 24 July 2014
Received in revised form 27 October 2014
Accepted 30 October 2014
Available online 5 November 2014
This manuscript was handled by Geoff Syme, Editor-in-Chief

Keywords:
Environmental monitoring
Trend detection
Power
Pollutant loads
Suspended sediment
Uncertainty

SUMMARY

The export of sediments from coastal catchments can have detrimental impacts on estuaries and near shore reef ecosystems such as the Great Barrier Reef. Catchment management approaches aimed at reducing sediment loads require monitoring to evaluate their effectiveness in reducing loads over time. However, load estimation is not a trivial task due to the complex behaviour of constituents in natural streams, the variability of water flows and often a limited amount of data. Regression is commonly used for load estimation and provides a fundamental tool for trend estimation by standardising the other time specific covariates such as flow. This study investigates whether load estimates and resultant power to detect trends can be enhanced by (i) modelling the error structure so that temporal correlation can be better quantified, (ii) making use of predictive variables, and (iii) by identifying an efficient and feasible sampling strategy that may be used to reduce sampling error. To achieve this, we propose a new regression model that includes an innovative compounding errors model structure and uses two additional predictive variables (average discounted flow and turbidity). By combining this modelling approach with a new, regularly optimised, sampling strategy, which adds uniformity to the event sampling strategy, the predictive power was increased to 90%. Using the enhanced regression model proposed here, it was possible to detect a trend of 20% over 20 years. This result is in stark contrast to previous conclusions presented in the literature.

1. Introduction

Excessive contaminant runoff form agricultural catchments is a phenomenon that occurs around the world and is of international concern (Arnold and Fohrer, 2005; Loucks et al., 2005; Davies and Simonovic, 2011). In an attempt to reduce contaminant loads, restrictions on land use and changes in land management have been introduced by government and local councils (European Union, 2000; USEPA, 1997, 2003). Of these programs, the Reef Water Quality Protection Plan (hereafter referred to as Reef Plan) (Reef Water Quality Protection Plan Secretariat, 2013) provides an excellent example of how management practices have been introduced with the specific aim of reducing contaminant loads entering coastal waters adjacent to the Great Barrier Reef, Australia. In this part of Australia, recent estimates of sediment export from coastal catchments of the Great Barrier Reef suggest that sediment runoff has increased by 5.5 times post European development (Kroon et al., 2012). The Reef Plan (Reef Water Quality Protection Plan Secretariat, 2013) identifies targets to reduce contaminant loads. The target for Total Suspended Solids (TSS) is that there will be a minimum 20 per cent reduction in sediment load by 2018 at the end-of-catchments. Critical to the success of such a program is the ability to demonstrate that management actions have been successful in achieving its goals.

Despite a clear need to demonstrate the success of management practices, it is often difficult to estimate loads with sufficient accuracy and to detect trends in loads within limited timeframes. Water quality parameters are naturally variable and monitoring programs typically involve the collection of a limited number of discrete samples during ambient and flow event conditions. The degree of error associated with load estimates was highlighted in the results presented in a study by Kuhnert et al. (2012) that used advanced regression techniques to estimate loads in the Burdekin River. In that study, the error associated with load estimates (calculated as 80th percentiles of load estimates) were in some

* Corresponding author at: Centre for Applications in Natural Resource Mathematics (CARM), School of Mathematics and Physics, The University of Queensland, St Lucia 4071, Australia. Tel.: +61 7 33652236.
E-mail address: you-gan.wang@uq.edu.au (Y.-G. Wang).

http://dx.doi.org/10.1016/j.jhydrol.2014.10.062
0022-1694/

instances >80% of average annual TSS load estimates. A separate study by Kroon et al. (2012) of TSS loads from the Burdekin River reported a 275% difference between 80th percentiles of annual TSS loads. Given the high degree of error associated that could potentially be associated with load estimates, there is a need to evaluate the ability to detect trends as specified by the Reef Plan. A previous study by Darnell et al. (2012) investigated whether the targets specified in Reef Plan were achievable with current approaches to monitoring and modelling. In that study, Darnell et al. (2012) found low (<40%) predictive power to detect a change in TSS in the historic monitoring data collected at the Burdekin River. The study by Darnell et al. (2012) suggested that historic monitoring for the Burdekin was insufficient to demonstrate that a target of 10% reduction in sediment would be detected within 20 years.

However, the apparent lack of power in current monitoring approaches as reported by Darnell et al. (2012) may not simply be due to insufficient sampling. In that study, specific error structure was used to quantify temporal correlation and predict a concentration flow relationship. While this approach is logical, it is possible that the random effects estimation technique used by Darnell et al. (2012) may not have adequately represented the inter-annual physical processes affecting TSS such as hysteresis and exhaustion. Instead, we propose that the relationships between TSS and flow may be more accurately defined through the use of alternative modelling approaches that better incorporate flow and concentration relationships including fluvial processes such as first flush, exhaustion and hysteresis. The inclusion of fluvial processes into statistical models used to estimate loads is likely to improve predictive power substantially, although the effectiveness of such approaches has not previously been tested.

The study by Darnell et al. (2012) also did not utilise turbidity data available for the Burdekin River. According to Grayson et al. (1996), high resolution turbidity data provides a useful means of improving correlations between discharge and concentration to determine sediment loads. There have also been a number of studies including those by Adin and Elimelech (1989), Gippel (1995), Grayson et al. (1996), Meybeck (1993), Packman et al. (1999), and Minella et al. (2008) that have all demonstrated a strong correlation between turbidity and TSS. As turbidity can be collected at high temporal resolution, it can be used to improve the estimation of TSS. Despite the potential for its use in load estimation, turbidity data are not routinely used in a management context. This may be because turbidity data may not be collected at all monitoring stations, or that turbidity probes can be prone to failure when deployed in rivers. Practical issues with turbidity probes include excessive silting, bio-fouling, loss of calibration, vandalism and damage incurred during high flow events. Such issues can incur a significant maintenance cost. The use of turbidity measures to estimate sediment loads for the Burdekin River was evaluated by Mitchell and Furnas (2001). That study illustrated some of the challenges with collecting turbidity data in large, sub-tropical river systems such as the Burdekin River that experience highly variable flow. An issue associated with turbidity is that turbidity probes are typically placed in a fixed location low on the stream bank to ensure probes remain submerged during low and no flow periods. Because the Burdekin River experiences a large range of bank heights, probes in a fixed location may not adequately reflect cross sectional variability in TSS across the range of river heights. In addition, the effect of particle size distribution on measures of turbidity requires calibration to ensure turbidity measures can reflect TSS results (Mitchell and Furnas, 2001). Despite these uncertainties, Mitchell and Furnas (2001) indicated that the increase in sampling frequency achieved using turbidity probes compared with manual sampling techniques, was likely to allow much better precision in year-to-year load estimates.

In this study, a series of load estimation models were evaluated and compared using TSS and turbidity data from the Burdekin River in Queensland, Australia. The Burdekin catchment has a relatively long record of flow and sediment concentration data collection (1987—present) and has been the focus of previous studies aiding comparisons with the results of the present study. This paper attempts to identify appropriate statistical models by estimating and comparing the power associated with each method in detecting trends. The models evaluated here were compared with Reef Plan targets to allow a comparison with the previous study by Darnell et al. (2012).

Because the number and timing of event and base flow discrete samples can affect the accuracy of load estimates, the influence of sampling strategies on resultant load accuracy was also considered. Historically, the number of samples collected annually has not been fixed and the majority of samples have been collected during high flow events. Darnell et al. (2012) used negative binomial models to fit and simulate the random annual sample size. To evaluate the influence of sample number and uniformity on estimates of predictive power, we used four sampling scenarios with varying characteristics of sample numbers and sample timing. These included (i) a fixed sample size with uniform sampling, (ii) a random sample size (following negative binomial distribution) with uniform sampling, (iii) a random sample size with flow dependent sampling, and (iv) random sample size with regularly optimised sampling.

2. Methods

2.1. The statistical model

Regression models are perhaps the most commonly used methods in load estimation. The most widely recognised of these is a simple regression model known as the rating curve method (Brown et al., 1970; Colby, 1956; Ferguson, 1986; Miller, 1951; Thomas, 1985, 1988; Verhoff et al., 1980; Walling, 1977). A novel generalisation of this model was developed by Cohn et al. (1992) who introduced a 7-parameter model. More recently, new statistical methods have also been developed expressing the relationship between constituent concentration and flow. For example, Wang et al. (2011) developed a regression model that incorporated patterns of flow history in the regression model as well as temporal and spatial correlations in the uncertainty evaluation. The model described by Wang et al. (2011) provided a framework for quantifying or testing for trends in pollutant load over time that resulted in improved load estimates. These models belong to the framework of the well-known linear models, which can be expressed with explicit inclusion of time trend δ as,

$$y_t = X_t\beta + \delta t + \epsilon_t, \tag{1}$$

where y_t is the log-transformed concentration at time t, X_t is the covariate matrix representing the intercept and other predictive variables such as log (flow) and seasonal effects, and ϵ_t is the corresponding error term. In this paper, we consider the above linear model for hypothesis testing on δ assuming X_t values are observed and investigate how prediction can be improved when additional hydrological variables (and turbidity) are included in the concentration variable, 'X'.

For the rating curve method, the matrix X consists of a constant and log-transformed discharge. Akin to other applications, prediction of concentration can be improved by adding more useful explanatory variables so that the error variance becomes smaller (as more variance is explained). For example, in the 7-parameter model developed by Cohn et al. (1992), the matrix X consists of a constant, a quadratic polynomial to logarithm of flow, a quadratic