Science of the Total Environment 449 (2013) 410–417

Contents lists available at SciVerse ScienceDirect

Science of the Total Environment

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

Uncertainty analysis of pollutant build-up modelling based on a Bayesian weighted least squares approach

Khaled Haddad [a], Prasanna Egodawatta [b], Ataur Rahman [a], Ashantha Goonetilleke [b,*]

[a] School of Computing, Engineering and Mathematics, University of Western Sydney, Building XB, Locked Bag 1797, Penrith, NSW 2751, Australia
[b] Science and Engineering Faculty, Queensland University of Technology, GPO Box 2434, Brisbane 4001, Australia

HIGHLIGHTS

► Water quality data spans short time scales leading to significant model uncertainty.
► Assessment of uncertainty essential for informed decision making in water quality.
► Bayesian approach and Monte Carlo simulation appropriate for assessing uncertainty.

ARTICLE INFO

Article history:
Received 6 November 2012
Received in revised form 23 January 2013
Accepted 27 January 2013
Available online 26 February 2013

Keywords:
Uncertainty analysis
Bayesian analysis
Monte Carlo simulation
Stormwater quality
Pollutant build-up
Stormwater pollutant processes

ABSTRACT

Reliable pollutant build-up prediction plays a critical role in the accuracy of urban stormwater quality modelling outcomes. However, water quality data collection is resource demanding compared to streamflow data monitoring, where a greater quantity of data is generally available. Consequently, available water quality datasets span only relatively short time scales unlike water quantity data. Therefore, the ability to take due consideration of the variability associated with pollutant processes and natural phenomena is constrained. This in turn gives rise to uncertainty in the modelling outcomes as research has shown that pollutant loadings on catchment surfaces and rainfall within an area can vary considerably over space and time scales. Therefore, the assessment of model uncertainty is an essential element of informed decision making in urban stormwater management. This paper presents the application of a range of regression approaches such as ordinary least squares regression, weighted least squares regression and Bayesian weighted least squares regression for the estimation of uncertainty associated with pollutant build-up prediction using limited datasets. The study outcomes confirmed that the use of ordinary least squares regression with fixed model inputs and limited observational data may not provide realistic estimates. The stochastic nature of the dependent and independent variables need to be taken into consideration in pollutant build-up prediction. It was found that the use of the Bayesian approach along with the Monte Carlo simulation technique provides a powerful tool, which attempts to make the best use of the available knowledge in prediction and thereby presents a practical solution to counteract the limitations which are otherwise imposed on water quality modelling.

1. Introduction

Modelling plays a critical role in the design of stormwater treatment systems and consequently in stormwater management decision making. However, as noted by Liu et al. (2012a), all modelling approaches are subject to uncertainty due to a diversity of reasons. As Bertrand-Krajewski et al. (2002) have pointed out, assessment of model uncertainty is an essential element of informed decision making in urban stormwater modelling. Uncertainty associated with stormwater quality modelling can be divided into two primary types, namely: (1) uncertainty due to input data and measurements; and (2) uncertainty due to simplification of reality in relation to the replication of pollutant processes (Bertrand-Krajewski, 2007; Freni et al., 2009; Gaume et al., 1998; Haydon and Deletic, 2009; Miguntanna et al., 2010). These types of uncertainties are referred to as epistemic (Merz and Thieken, 2005). This arises from the incomplete conceptual understanding of the systems under study. This can be attributed to the reliance on models that are simplified representations of the true complexities of natural processes, as well as the limited datasets available for testing, validation, hypotheses and/or simulating the systems.

A range of studies have focused on assessing uncertainty resulting from input data and measurements (for example, Bertrand-Krajewski, 2007; Liu et al., 2012a, 2012b; Sohrabi et al., 2003). Uncertainty associated

* Corresponding author. Tel.: +61 731381539; fax: +61 731381170.
E-mail address: a.goonetilleke@qut.edu.au (A. Goonetilleke).

0048-9697/$ – see front matter
http://dx.doi.org/10.1016/j.scitotenv.2013.01.086

with stormwater quality models has also been investigated by numerous researchers (for example, Dotto et al., 2012; Freni et al., 2008; Kanso et al., 2005). However, these studies have commonly focused on assessing the overall modelling uncertainty rather than the uncertainty associated with modelling specific pollutant processes, namely, pollutant build-up and wash-off. Consequently, this provides only limited understanding of the accuracy of the embedded equations for replicating pollutant processes. Therefore, even if the overall model uncertainty is known, the primary causes are difficult to pinpoint.

It is in this context that the lack of adequate data availability and the accuracy to support the development of stormwater quality models are points of concern. This stems from the fact that the available datasets span only relatively short time scales unlike water quantity data (van der Sterren et al., 2012). Therefore, the accuracy of water quality modelling outcomes is not necessarily very reliable as it constrains the ability to take due consideration of the variability associated with pollutant processes and natural phenomena (Stewart, 2000). Research has shown that pollutant loadings on catchment surfaces and rainfall within an area can vary over space and time scales and are typically not known with any reliable accuracy (UNESCO, 2005; Smith et al., 1997). More reliable models may be established if observed datasets over longer time periods are used. However, such data is scarce in the water quality arena due to the high cost associated with data collection, which constrains long term monitoring programmes.

This is compounded by the fact that the limited conceptual understanding as noted above leads to parameter uncertainty in water quality models. For example, considering pollutant build-up, the desirable approach is the development of a regression model that can be used for build-up prediction as a function of the number of antecedent dry days to encompass different scenarios. In this case, not only is there uncertainty associated with the build-up values (dependent variable), but there can also be uncertainty related to the number of antecedent dry days (predictor variable).

It is in this context, that this paper discusses a Bayesian statistical approach for assessing the mathematical replication of pollutant build-up on roof surfaces. This will enhance the understanding of model uncertainty associated with pollutant processes to enable the practical application of this knowledge to stormwater quality modelling.

A primary goal of this paper was to quantify the different aspects of uncertainty associated with water quality prediction models by: (i) increasing the availability of synthetic data via Monte Carlo simulation; and (ii) using efficient parameter estimation methods. To carry this out in a statistically meaningful manner, a number of different parameter estimation techniques including ordinary least squares regression (OLSR), weighted least squares regression (WLSR) and Bayesian WLSR (BWLSR) using a Monte Carlo simulation framework have been investigated. The errors arising in both the dependent and predictor variable data have also been considered by assuming that the data can be described by appropriate probability distributions.

Roof surfaces were selected for this study for a number of reasons. Despite the fact that road surfaces are commonly regarded as the primary pollutant source in an urban environment (Herngren et al., 2006; Sartor and Boyd, 1972), the pollutant contribution from roof surfaces is little understood and is also a significant pollutant source (Egodawatta et al., 2012). As noted by Egodawatta (2007), in the urban catchments he investigated the total roof areas were found to be 2–3 times larger than the total road area. Furthermore, understanding of pollutant processes on roof surfaces is important as rainwater harvesting is being increasingly considered as an alternative water supply source in water deficient regions (Imteaz et al., 2011; Khastagir and Jayasuriya, 2011; van der Sterren et al., 2009, 2012). As pointed out by Egodawatta et al. (2009), pollutant build-up on road and roof surfaces follows the same exponential relationship, but with different coefficients. Therefore, the application of uncertainty analysis to roof surfaces is easily extendable to road surfaces.

2. Materials and methods

2.1. Data collection

The data collection methodology has been described in detail in Egodawatta et al. (2009). Briefly, the pollutant build-up data was collected using model roofs as test plots (3 m^2). The model roofs were mounted on a scissor lift arrangement as shown in Fig. 1. As such, they could be raised to the typical roofing height to allow pollutant accumulation naturally and then lowered to ground level for sampling. This approach enabled overcoming the practical difficulties inherent in sample collection from actual roofs. Corrugated steel and concrete tiles were used for cladding as these are the most widely used roofing types in the study region. The model roofs were placed in a residential area with a few major roads in the vicinity.

Pollutant build-up samples were collected from the roof surfaces by washing the surface with approximately 7 L of deionised water using a soft brush and the runoff was collected. Build-up investigations were conducted for 1, 2, 3, 5, 7, 14 and 21 antecedent dry days. It was hypothesised that build-up primarily varies with the antecedent dry days (Ball et al., 1998; Sartor and Boyd, 1972).

Suspended solids were adopted as the indicator pollutant. Suspended solids are not only a significant stormwater pollutant in its own right, but also acts as a mobile substrate for the transport of other stormwater pollutants such as heavy metals and hydrocarbons to receiving waters (Herngren et al., 2005; Sartor and Boyd, 1972). Testing for total suspended solids (TSS) concentration was undertaken according to test method no. 2540D (APHA, 2005).

2.2. Data analysis

Ordinary least squares regression (OLSR) is generally adopted to develop empirical relationships such as between pollutant build-up and relevant explanatory variables (for example, in Rahman et al., 2002). In the past, water quality prediction models have commonly adopted the OLSR estimators. These estimators are appropriate and statistically efficient if all the observations have equal weights.

The appropriate application of the OLSR method requires a number of assumptions to be satisfied, such as: (i) independence,

Fig. 1. A model roof used for pollutant build-up data collection.