Journal of Hydrology 552 (2017) 13–27

Contents lists available at ScienceDirect

Journal of Hydrology

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

Research papers

Event-based nonpoint source pollution prediction in a scarce data catchment

Lei Chen, Cheng Sun, Guobo Wang, Hui Xie, Zhenyao Shen *

State Key Laboratory of Water Environment Simulation, School of Environment, Beijing Normal University, Beijing 100875, PR China

A R T I C L E I N F O

Article history:
Received 13 April 2017
Received in revised form 14 June 2017
Accepted 18 June 2017
Available online 22 June 2017

Keywords:
Non-point source pollution
Scarce data
Event-based process
Unit hydrograph
LOADEST
Rainfall-runoff-pollutant relationship

A B S T R A C T

Quantifying the rainfall-runoff-pollutant (R-R-P) process is key to regulating non-point source (NPS) pollution; however, the impacts of scarce measured data on R-R-P simulations have not yet been reported. In this study, we conducted a comprehensive study of scarce data that addressed both rainfall-runoff and runoff-pollutant processes, whereby the impacts of data scarcity on two commonly used methods, including Unit Hydrograph (UH) and Loads Estimator (LOADEST), were quantified. A case study was performed in a typical small catchment of the Three Gorges Reservoir Region (TGRR) of China. Based on our results, the classification of rainfall patterns should be carried out first when analyzing modeling results. Compared to data based on a missing rate and a missing location, key information generates more impacts on the simulated flow and NPS loads. When the scarcity rate exceeds a certain threshold (20% in this study), measured data scarcity level has clear impacts on the model's accuracy. As the model of total nitrogen (TN) always performs better under different data scarcity conditions, researchers are encouraged to pay more attention to continuous the monitoring of total phosphorus (TP) for better NPS-TP predictions. The results of this study serve as baseline information for hydrologic forecasting and for the further control of NPS pollutants.

1. Introduction

Non-point source (NPS) pollution, especially in the Three Gorge Reservoir Region (TGRR) during specific rainfall events, is a major environmental threat (Lee et al., 2003). The accurate prediction of NPS pollution, which is not yet fully determined, is necessary to minimize the impacts of these events (Ongley et al., 2010; Yang and Jin, 2010). However, as many catchments around the world are poorly gauged or ungauged, necessary information in terms of measured rainfall, runoff and pollutant data is not always available for sites of interest. Thus, data scarcity has a negative impact on the accuracy of NPS predictions. Furthermore, poor model accuracy leads to limited knowledge regarding NPS simulation, and these models can not be used to support the optimal management of water quality. In recent years, NPS predictions of scarce data catchments have attracted considerable attentions and have also become the focus of watershed researchers (Sajikumar and Thandaveswara, 1999).

In a relative small catchment, rainfall is the most important forcing factor of runoff production process and NPS pollution process (Coulliette and Noble, 2008; Wu et al., 2010). Thus, rainfall-runoff-pollutant (R-R-P) relationship has become the key to the accurate prediction of NPS pollution (Kashani et al., 2016). Based on the numbers of physical processes involved, R-R-P models can be broadly classified as data, physically, and conceptually based. The structures of data-based models are derived from observed data without any prior assumptions on dominant hydrological or other physical-chemical processes. Physically based models build model structures from established principles, such as the conservation of energy, mass, and momentum. But their applications to scarce data catchments are often problematic due to the massive amounts of data that must be used and the numerous parameters that must be calibrated (Liu and Gupta, 2007; Quinn, 2004). As alternatives, conceptually based models use measured data and involve the researcher's understanding of dominant watershed processes. Due to their computational efficiencies, conceptually based models are widely used for various practical applications. For instance, Sherman (1932) made one notable contribution to conceptual hydrologic model by introducing the Unit hydrograph (UH), which is defined as the hydrologic response of a basin to a unit volume of effective rainfall. Other rainfall-runoff models include Linear regression, Hydrology backward, and Rainfall and runoff correlation diagrams. However, the application

* Corresponding author.
E-mail addresses: chenlei1982bnu@bnu.edu.cn, zyshen@bnu.edu.cn (Z. Shen).

http://dx.doi.org/10.1016/j.jhydrol.2017.06.034
0022-1694/

of Linear regression to assess runoff outputs is limited by large data requirements, so it is not appropriate for data-scarce condition. Hydrology backward has better application effect on yearly and monthly scale but the effect is not better in hourly scale. The application of the Rainfall and runoff correlation diagrams is dominated by stored-full runoff and it has limitation in precipitation runoff. In comparison, when flow data are available, the Load Estimator (LOADEST) can be used to establish a linear regression, and the logarithmic transform can be used to estimate constituent yields for hourly, daily, monthly and yearly intervals (Stallard, 2011). Different regression equations along with several modeling residual adjustment techniques have been established to provide insight into the water-quality trajectories of many river catchments. As well, other runoff-pollutant models include Pollution sediment method, Export coefficient model and Relevant relationship method. These traditional methods are appropriate for large scale analysis with requirement of constituent data in monthly and yearly time series. Generally, the construction of these conceptually based models involves the use of a proportion of measured data points from time series to identify their inherent structures and build regression equations between rainfall-runoff and runoff-pollutant data points. Thus, it is important to clearly define and quantify data requirements and to maximize information content for different methods, as such content is essential for model structures, particularly for catchment systems that exhibit complex, non-linear behaviors.

Nowadays, the development of data centers and satellite data observations has made acceptable rainfall and streamflow data sets for data-poor regions available. However, high-frequency time series of flow data, i.e., hourly or sub-hourly time scales, remain limited, especially with respect to event-based hydrological studies. Event-based NPS prediction can provide detailed information for the change of water quality, as well as the design of storm-based management practices. Water quality records are thus even scarcer because they have been obtained periodically. Comparatively, more studies have been conducted to quantify the impacts of scarce flow data on hydrological assessments and rainfall-runoff simulations (Beven and Freer, 2001; Di Piazza et al., 2011; Ebrahim and Villholth, 2016; Jakeman and Hornberger, 1993; Labat et al., 2000; Mango et al., 2011). Yadav et al. (2007) quantified the continuous hydrologic response observed in ungauged basins and showed that data scarcity impacted simulated flows and model calibration processes. Seo and Schmidt (2014) adopted the instantaneous UH in examining ungauged basins and showed that the simulated hydrograph was nearly identical to the measured flow. Studies have also shown that data scarcity has considerable impacts on the construction of runoff-pollutant load relationships (Brezonik and Stadelmann, 2002; Kim and Furumai, 2013; Maniquiz et al., 2010; Yi et al., 2015). Flores-Alsina et al. (2014) demonstrated that scenario analyses based on scarce data should consider the impacts of rainfall patterns, as such rainfall patterns were major forces that drove runoff production and mass transport. As rainfall patterns can have various impacts on NPS pollution, such as total nitrogen (TN) and phosphorus (TP) (Wu et al., 2012), it can be inferred that scarce data may have different effects on different LOADEST models. To date, a comprehensive account of the impacts of data scarcity on NPS simulations have not been reported, and few studies have considered its impacts on event-based R-R-P simulation.

This paper aims to improve knowledge on the data requirements of measured flows and water quality for NPS predictions by presenting an assessment of the impacts of scarce data on UH and LOADEST models. In order to achieve these goals, we propose three assumptions to be tested as followed. The classification of rainfall patterns and the establishment of data scarcity scenario are the basic for modeling with several indicators to evaluate the results. Specifically, this study assessed the following factors: (1) the characteristics of scarce data for specific rainfall patterns were discussed via field sampling and data investigations in the TGRR; (2) the impacts of scarce data on models accuracy were quantified using a case study of the TGRR in China; (3) a methodology was proposed for constructing a complete series of rainfall, runoff and pollutant loads.

2. Materials and methods

2.1. Description of the study area and monitoring program

In this study, Zhangjiachong catchment, which is a representative tributary in TGRR head of China, located in the southwestern area of Zigui County, is shown in the Fig. 1. Controlling the pollution of upstream of TGRR is benefit for environmental management, which can provide data information and methodology promotion. The catchment is a sub-watershed of the Moping River that drains a catchment area of 1.62 km^2 and flows directly into the Yangzi River. Its landscape is primarily mountainous, with an elevation ranging from 148 m to 530 m above the Yellow Sea level. Yellow brown soil accounts for 77% of the total area and is classified as analfisol according to the China Soil Scientific Database. Agricultural and forested land covers 27% and 43% of the total area, respectively. Major local crops include tea, corn, oilseed rape, and chestnut (Shen et al., 2009). Fertilizer use in this area is relatively high, increasing risks of nutrient loss into the stream. TN and TP were selected as target pollutants due to the high volumes of fertilizer applied on the agricultural land. The average annual temperature reaches approximately 18 °C. The average annual precipitation level reaches approximately 1439 m, which is affected by canyon topography significantly, and 80% of all precipitation in the area occurs between May and August, rendering soil erosion is common during wet seasons. We thus mainly focus on storm events occurring during this season.

Excessive cultivation and over-fertilization have caused the release of much N and P, resulting in serious eutrophication in the TGRR (Shen et al., 2014). Therefore, research regarding the control and management of NPS pollution in the TGRR has been of great significance for water resource protection. To assess the R-R-P simulation, a weather station (Skye Lynx Standard) was set within the catchment to provide a continuous record of climate data, including data on precipitation, wind speed, and temperature levels from 2011 to 2014. There is just one weather station in the Zhangjiachong catchment, because a well-located station might be sufficient for this small area. Water level measurements were collected from a float-operator sensor (WGZ-1) at the catchment outlet, and high-frequency sampling was used to record at approximately 15-min intervals. During the study period, the streamflow dynamics of 14 selected storm events were recorded, and stream discharge levels were estimated using a local rating curve. For the water quality analysis, water samples were collected at the catchment outlet from July of 2013 to September of 2014. Base flows were measured before runoff generation, and water samples were collected every 15 min in the first hour after runoff began and every 30 min over the following two hours. After water levels had stabilized, water samples were collected once every one hour until the end of the event. All water samples were placed into precleaned glass jars with aluminumfoil liners along the lids and stored at −20 °C during transportation to the laboratory for processing and analysis. TN and TP levels in the samples were measured via alkaline persulfate oxidation-UV spectrophoto metric and Potassium persulfate oxidation-molybdenum blue colorimetric methods. Finally, the recorded rainfall, flow and pollutant levels were used for the following analysis.