Journal of Hydrology 512 (2014) 240–253

Contents lists available at ScienceDirect

Journal of Hydrology

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

Comparison of regionalization approaches in parameterizing sediment rating curve in ungauged catchments for subsequent instantaneous sediment yield prediction

Sokchhay Heng *, Tadashi Suetsugi

Interdisciplinary Graduate School of Medicine and Engineering, University of Yamanashi, 4-3-11 Takeda, Kofu, Yamanashi 400-8511, Japan

ARTICLE INFO

Article history:
Received 18 June 2013
Received in revised form 15 February 2014
Accepted 1 March 2014
Available online 11 March 2014
This manuscript was handled by Konstantine P. Georgakakos, Editor-in-Chief, with the assistance of Attilio Castellarin, Associate Editor

Keywords:
Sediment yield
Ungauged catchment
Regionalization
Sediment rating curve
Multidimensional scaling
Lower Mekong Basin

SUMMARY

Observations of sediment loading are lacking for rivers in many parts of the world, particularly in developing countries. There is a need for better understanding of what models might be available to offer estimates of sediment yield in such poorly gauged or ungauged areas. Therefore, the aim of this research is to determine an ideal regionalization methodology for estimating sediment rating curves in ungauged catchments for instantaneous suspended sediment yield (*SSY*) prediction. A comparison of three regionalization approaches (catchment similarity, regression and spatial proximity) was carried out in 16 catchments, in the Lower Mekong Basin, after discarding one with the evidence that the data point is not homogeneous. The highest quality results are provided by the catchment similarity approach in which a single donor catchment was selected in accordance with an optimum catchment similarity index computed by the multidimensional scaling technique. The regression-based approach is intermediate. The sparse network of basin monitoring in the study area does not favor the spatial proximity approach, producing the worst regionalization solution. The overall predictive accuracy was further improved through a combined formulation of catchment similarity and regression. This reveals that different approaches have different advantages and therefore, using only one of them cannot fully encompass a wide range of catchment heterogeneity. Since the ideal regionalization methodology performs satisfactorily in all 16 sites, a regional model with a relatively simple geomorphic framework was established for estimating catchment-scale *SSY* in ungauged rivers of the basin.

1. Introduction

Hydrological modeling in ungauged basins has been regarded as one of the main challenges in hydrological sciences (Sivapalan et al., 2003; Saliha et al., 2011). In this context, a river catchment may be 'ungauged' with respect to catchment process variables such as rainfall, runoff and sediment transport. Many large river basins are commonly trans-boundary and the geographic distribution of gauging sites is largely non-uniform. For instance, the Lower Mekong Basin drains part of Lao PDR, Thailand, Cambodia and Vietnam. Thailand's monitoring infrastructure is more developed than the others and therefore, this country is very rich in historical records due to its extensive gauging network. In terms of data quality, observations of sediment transport in Thailand have proven to be the most reliable, complete and easy to access. The three neighboring countries all have problems with the extent of monitoring and data availability (Fuchs, 2004). Thus, model calibration in data-rich areas and transfer of information to ungauged regions, known as regionalization, is important.

Three popular regionalization approaches are catchment similarity, regression and spatial proximity. The regression-based approach consists of developing posteriori relationships between catchment descriptors and empirical model parameters calibrated on gauged areas. The catchment descriptors include not only the physical but also hydrological characteristics. The established relationships are afterward applied to parameterize the model in ungauged catchments of interest. A major drawback of this technique is the equifinality problem of model calibration having no unique set of parameters (Zhang and Chiew, 2009). The calibrated parameters depend on the specific conditions of the input datasets and parameterization techniques. Input errors and data inadequacies will lead to highly uncertain model parameters. Moreover, different sets of model parameters could also yield similar outputs.

The spatial proximity approach uses the entire parameter set transposed from the neighboring gauged catchment (the latter

* Corresponding author. Tel.: +81 080 3690 2013.
E-mail address: heng_sokchhay@yahoo.com (S. Heng).

http://dx.doi.org/10.1016/j.jhydrol.2014.03.003
0022-1694/

called donor catchment) with a rationale that catchments geographically close to each other should have similar characteristics. This technique is quite practical in areas with sparse monitoring because it does not require the use of catchment descriptors. However, the model does not work very well in cases of strong heterogeneity amongst the nearby catchments (Parajka et al., 2005). In mountainous areas, the complex topography could provoke different local rainfall and streamflow regimes. As more rainfall occurs at or near a mountain peak, much less rainfall could occur at the lower altitudes. This great spatial variation further affects the catchment response and streamflow conditions (Ly et al., 2013). Moreover, in developing countries, human activities including development of impervious areas and deforestation for agricultural purposes cause very diverse physical landscapes locally.

Similar to spatial proximity, the catchment similarity approach selects a donor catchment based on the similarity of catchment attributes. Here, the gauged catchment having the most similar characteristics to the target ungauged catchment is chosen as the donor catchment. The rationale of this technique is that catchments with similar attributes should behave similarly. The regionalization procedure based on both catchment similarity and regression is of interest generally because they associate catchment descriptors and are therefore very informative.

Parajka et al. (2005) compared the above-mentioned three regionalization approaches in the context of runoff modeling by applying the methods using a database extending across 320 Austrian catchments, and found that spatial proximity and catchment similarity perform similarly with median Nash–Sutcliffe model efficiency (*NSE*) equal to 0.67. These two methods are slightly superior to the regression approach (median $NSE = 0.65$). Oudin et al. (2008) tried to regionalize the GR4J and TOPMO rainfall–runoff model using a database from 913 French catchments. These results showed that spatial proximity is the best regionalization approach, with a median *NSE* equal to 0.73 and 0.71 for GR4J and TOPMO, respectively. Results obtained using the catchment similarity approach had scores of 0.71 and 0.69 for median *NSE*. Regression fared the worst, with median *NSE* equal to 0.68 for GR4J and 0.55 for TOPMO. They finally concluded that the inferiority of the two approaches associating catchment descriptors may be due to exclusion of soil type (a key descriptor) and that better results could be obtained by using the combination method. Based on 210 Australian catchments, Zhang and Chiew (2009) conducted a similar study and found that spatial proximity performs slightly better than catchment similarity. The predictive accuracy is marginally improved by the combination of these two approaches. A common conclusion drawn from these three earlier research efforts is that when working with a closely clustered set of catchments, the spatial proximity regionalization approach performs best.

Considering now the ungauged catchment problem with respect to sediment data, this issue is critical since there is a very limited number of monitoring stations around the world (Isik, 2013). Such ungauged catchments are usually located in headwater regions. Therefore, absence of the historical data records is generally due to inaccessibility, budget and technical constraints (particularly in developing countries), and historical lack of foresight in considering future developments in such areas. Due to the scarcity of monitoring data, very few sediment yield modeling studies in ungauged catchments have been carried out, and a majority of them work with an annual resolution (not continuous or time series prediction), e.g. Roman et al. (2012) and Isik (2013). Heng and Suetsugi (2013a) regionalized the parameters of an artificial neural network model that was used to predict monthly sediment yield in the Lower Mekong Basin. In their study, only one catchment attribute (topography) was accounted for in a regionalization process that was based on catchment similarity. As a result of validation in three catchments, the predictive accuracy in terms of determination coefficient (R^2) ranges from 0.59 to 0.64 with an average value of 0.61. To our knowledge, there is no comparative study to date that has considered a large set of catchments.

Although sediment is gauged in some areas, the sampling frequency is usually low (monthly or even larger time scale). Only suspended load, in most cases, is measured and measurement of bed load transport is lacking. Fortunately, the suspended portion is predominant and commonly account for more or less 90% of the total load transported by streams (Walling and Fang, 2003; Zhang et al., 2012). Accordingly, many modeling techniques have been developed to predict suspended sediment yield for a specific site of interest, where no data have been collected. The finest fraction of suspended sediment is often a non-capacity load and thus cannot be well predicted by sediment transport models that function using stream power or shear stress (Asselman, 2000; Warrick and Rubin, 2007). As an alternative, empirical models, e.g. sediment rating curve, are often taken into consideration. The sediment rating curve is the simplest and most popular tool for describing the relationship between water discharge and sediment flux (Isik, 2013). The rating curve is a 'black box' model, and it is advantageous because its parameters have physical meanings (Asselman, 2000; Horowitz, 2003; Yang et al., 2006; Sadeghi et al., 2007). Although many modeling tools are available, their predictive accuracy is still a major concern (Sivapalan et al., 2003). In the context of ungauged basin modeling, more reliable predictions could be achieved through uncertainty reduction in the regionalization processes (Almeida et al., 2012).

The main problem in ungauged basins is the problem of discontinuous and uneven sampling in space and time, leading to augmented uncertainty in model predictions (Sivapalan et al., 2003). In this regard, there are two groups of uncertainties, one related to local model calibration in gauged sites, and another related to the regionalization (spatial extrapolation) procedure (Wagener and Wheater, 2005). Uncertainties in the local modeling are associated with heterogeneity of input data; for instance, it may be influenced by multiple extreme events (e.g. flooding, land-slides and slope failures), measurement errors, and low sampling frequency. Another associated factor is process heterogeneity involving phenomena such as hysteresis (time delays between peak water discharge and peak sediment movement), dam-reservoir operation (irregular sediment supply conditions) and land-use changes (e.g. deforestation and changes of local agricultural practices). Uncertainties related to the regionalization procedure stem from arbitrary choices of catchment descriptors, similarity measures and selection of donor catchments.

The main objective of this research is to determine an ideal regionalization methodology for ungauged catchment modeling, in terms of monthly suspended sediment yield (*SSY*). The regionalization was conducted at two levels. The Level 1 regionalization is the use of three basic approaches: catchment similarity, regression and spatial proximity. In this level, the three approaches were compared on the basis of a single donor catchment, so as to highlight each individual approach's potential for prediction. The Level 2 regionalization is aimed to refine the overall predictive accuracy by introducing three different alternatives that associate the most outstanding approach from Level 1. They are based on the multiple donors, ensemble and combination method. A detailed description of these three alternatives/methods was presented in Section 2.6. Since the paper concentrates on the regionalization methodology, the sediment rating curve model was the sole technique employed for *SSY* simulation. A flow chart of the study is presented in Fig. 1. The case study was conducted in the Lower Mekong Basin. To our knowledge, this research is the first comparative study of regionalization approaches in the context of *SSY* modeling especially in this kind of data constraint region.