Applied Soft Computing 45 (2016) 187–196

Contents lists available at ScienceDirect

Applied Soft Computing

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

Suspended sediment concentration estimation by stacking the genetic programming and neuro-fuzzy predictions

Ehsan Shamaei, Marjan Kaedi*

Faculty of Computer Engineering, University of Isfahan, Hezar-Jerib St., Isfahan 81746-73441, Iran

ARTICLE INFO

Article history:
Received 14 February 2015
Received in revised form 14 January 2016
Accepted 24 March 2016
Available online 14 April 2016

Keywords:
Suspended sediment prediction
Stacking method
Linear genetic programming
Neuro-fuzzy
Neural networks

ABSTRACT

In the new decade due to rich and dense water resources, it is vital to have an accurate and reliable sediment prediction and incorrect estimation of sediment rate has a huge negative effect on supplying drinking and agricultural water. For this reason, many studies have been conducted in order to improve the accuracy of prediction. In a wide range of these studies, various soft computing techniques have been used to predict the sediment. It is expected that combining the predictions obtained by these soft computing techniques can improve the prediction accuracy. Stacking method is a powerful machine learning technique to combine the prediction results of other methods intelligently through a meta-model based on cross validation. However, to the best of our knowledge, the stacking method has not been used to predict sediment or other hydrological parameters, so far. This study introduces stacking method to predict the suspended sediment. For this purpose, linear genetic programming and neuro-fuzzy methods are applied as two successful soft computing methods to predict the suspended sediment. Then, the accuracy of prediction is increased by combining their results with the meta-model of neural network based on cross validation. To evaluate the proposed method, two stations including Rio Valenciano and Quebrada Blanca, in the USA were selected as case studies and streamflow and suspended sediment concentration were defined as inputs to predict the daily suspended sediment. The obtained results demonstrated that the stacking method greatly improved RMSE and R^2 statistics for both stations compared to use of linear genetic programming or neuro-fuzzy solitarily.

1. Introduction

In the new decade due to lack of water resources, it is vital to have an accurate sediment prediction for different purposes such as dam service life evaluation or water management. In order to predict the sediment, because of its paramount importance and its nonlinear nature, there have been great challenges for water and hydraulic structures engineers. Sediment has severe effects on hydraulic and geomorphologic features of the rivers and hydraulic structures. Furthermore, prediction of sediment has detrimental effects on the dams' reservoirs. So, incorrect estimation of sediment rate reduces the amount of waters stored behind the dams and consequently has a huge negative effect on supplying drinking water and agricultural water. This is why scientists and experts are constantly seeking for new ways to predict the sediment accurately and suggesting various empirical formulas for this purpose [1–4]. Accuracy of these empirical formulas was evaluated by some researchers [5,6]. Due to some innate errors in measuring the input parameters, applying empirical methods leads to uncertainties in the final results [7]. The nonlinear and seasonal nature of the related parameters in estimating sediment, inaccurate measurement and lack of sufficient data are some factors that cause uncertainties in the results obtained by empirical models [8].

Besides the empirical formulas, the sediment rating curve has been used for estimating the sediment in the both gauged and ungauged stations as well [9–13]. Heng and Suetsugi [13] showed that using sediment rating curve could increase the uncertainties. Thus, it is essential to use a method which is able to estimate the sediment more accurately and confidently. Soft computing methods, which are widely used in various fields, are able to help us in this way as well. In this study, among various methods of soft computing, our attention is focused on three methods of soft computing. These methods are artificial neural network, neuro-fuzzy, and genetic programming. Researches done by these methods for predicting sediment will be reviewed in Section 2.

In order to improve the prediction accuracy of soft computing methods, a combination of several predicting algorithms can be used. These methods, called ensemble learning [14], have such different varieties as Bayesian model averaging, bagging, boosting, model tree ensembles, and stacking. In some previous studies, to

* Corresponding author. Fax: +983137934038.
E-mail address: kaedi@eng.ui.ac.ir (M. Kaedi).

http://dx.doi.org/10.1016/j.asoc.2016.03.009
1568-4946/

Table 1
Summary of the reviewed studies sorted by the year of publication.

	Year	Author	Input variables	Predicted quantities	Soft computing methods
1	2002	Nagy et al. [17]	Streamflow Sediment	Sediment	Neural network
2	2003	Cigizoglu and Alp [15]	Rainfall River flow Sediment	Sediment	Neural network
3	2005	Kisi [33]	Streamflow Sediment	Sediment	Neuro-fuzzy
4	2006	Raghuwanshi et al. [16]	Runoff Rainfall Temperature	Sediment Runoff	Neural network
5	2007	Alp and Cigizoglu [21]	Rainfall Sediment	Sediment	Feed-forward back propagation and radial basis function neural networks
6	2008	Kisi [23]	Flow Sediment	Sediment	Neural network
7	2009	Rajaee et al. [30]	Discharge Sediment	Sediment	Feed-forward back propagation and neuro-fuzzy (Sugeno)
8	2009	Cobaner et al. [34]	Rainfall Streamflow Sediment	Sediment	Adaptive neuro-fuzzy inference system
9	2009	Kişi [35]	Streamflow Sediment	Sediment	Combination of fuzzy model and differential evolution
10	2009	Altunkaynak [38]	Streamflow Sediment	Sediment	Genetic algorithm
11	2010	Kişi [24]	Streamflow Sediment	Sediment	Neural differential evolution
12	2010	Zhang et al. [39]	Critical shear stress Resuspension	Optimizing the parameters of sediment transport	Genetic algorithm
13	2010	Kisi and Guven [42]	Streamflow Sediment	Sediment	Linear genetic programming
14	2011	Melesse et al. [22]	Precipitation Discharge	Sediment	Multi-layered Perceptron neural network
15	2011	Shiri and Kişi [44]	Streamflow Sediment	Sediment	Gene expression programming, wavelet-gene expression programming, wavelet-neural network, and wavelet neuro-fuzzy
16	2012	Singh et al. [20]	Rainfall Runoff Sediment	Sediment	Back propagation and radial basis function neural networks
17	2012	Kisi et al. [25]	Streamflow Sediment	Sediment	Combination of neural network and artificial bee colony algorithm
18	2012	Kisi et al. [41] Aytek and Kisi [7]	Streamflow Sediment	Sediment	Genetic programming
19	2012	Özger and Kabataş [37]	Rainfall Streamflow Sediment	Sediment	Gene expression programming
20	2013	Haddadchi et al. [19]	Sediment	Sediment load	Feed-forward back propagation neural network
21	2013	Rathinasamy et al. [51]	Streamflow	Streamflow	Bayesian model averaging (combining different wavelet models)
22	2013	Erdal and Karakurt [52]	Streamflow	Streamflow	Bagging and stochastic gradient boosting (combining tree regression models)
23	2014	Schnier and Cai [53]	Streamflow	Streamflow	Model tree ensembles and Bagging
24	2014	Kumar et al. [8]	Vegetative flow Incipient shear Total bed load	Threshold of sediment	Multi-gene genetic programming
25	2015	Schnier and Cai [53]	Sediment	Sediment load	combined wavelet and fuzzy logic techniques
26	2015	Pulido et al. [48]	Daily sediment Flow	Daily sediment load	Combination of feed-forward and radial basis function neural networks
27	2015	Mabu et al. [49]	Current Wave	Suspended sediment concentration	Combination of numerical models and neural network
28	2016	Soto et al. [47]	Discharge Rainfall Temperature	Sediment concentration	Neural networks
29	2016	Hosseini and Mahjouri [50]		Rainfall Runoff	Combination of support vector machine and geomorphologic-based neural network