Journal of Hydrology 490 (2013) 41–55

Contents lists available at SciVerse ScienceDirect

Journal of Hydrology

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

A geomorphology-based ANFIS model for multi-station modeling of rainfall–runoff process

Vahid Nourani [a,*], Mehdi Komasi [b]

[a] Dept. of Water Resource Eng., Faculty of Civil Eng., Univ. of Tabriz, Iran
[b] Dept. of Civil Eng., Faculty of Eng., Univ. of Ayatollah Boruojerdi, Boruojerd, Iran

ARTICLE INFO

Article history:
Received 16 November 2012
Received in revised form 20 February 2013
Accepted 20 March 2013
Available online 29 March 2013
This manuscript was handled by Andras Bardossy, Editor-in-Chief, with the assistance of Purna Chandra Nayak, Associate Editor

Keywords:
Artificial intelligence
Black box model
Spatiotemporal modeling
Fuzzy clustering
The Eel River watershed

SUMMARY

This paper demonstrates the potential use of Artificial Intelligence (AI) techniques for predicting daily runoff at multiple gauging stations. Uncertainty and complexity of the rainfall–runoff process due to its variability in space and time in one hand and lack of historical data on the other hand, cause difficulties in the spatiotemporal modeling of the process. In this paper, an Integrated Geomorphological Adaptive Neuro-Fuzzy Inference System (IGANFIS) model conjugated with C-means clustering algorithm was used for rainfall–runoff modeling at multiple stations of the Eel River watershed, California. The proposed model could be used for predicting runoff in the stations with lack of data or any sub-basin within the watershed because of employing the spatial and temporal variables of the sub-basins as the model inputs. This ability of the integrated model for spatiotemporal modeling of the process was examined through the cross validation technique for a station. In this way, different ANFIS structures were trained using Sugeno algorithm in order to estimate daily discharge values at different stations. In order to improve the model efficiency, the input data were then classified into some clusters by the means of fuzzy C-means (FCMs) method. The goodness-of-fit measures support the gainful use of the IGANFIS and FCM methods in spatiotemporal modeling of hydrological processes.

1. Introduction

The correct prediction of hydrological phenomena such as rainfall–runoff process can provide effective information for city planning, land use and water resource management of a watershed. It also plays an important role in mitigating the impacts of flood or drought on water resources systems. The rainfall–runoff modeling of a watershed is one of the most complex hydrological tasks to comprehend, as it is involved tremendous temporal and spatial variability. Therefore, many hydrological models have been developed in order to simulate such a complex process. In the cases with high rate of uncertainty and complexity where it is difficult to consider every effective physical parameter, it is not a surprising fact that black box models which convert inputs to output values in ways that have nothing to do with what happens in reality, may produce more accurate results than physical based models.

Artificial Intelligence (AI) approaches are such black box modeling tools that have lately found applications in a variety of areas including rainfall–runoff modeling. Nowadays, Artificial Neural Network (ANN), as a self-learning and self-adaptive function approximator, has shown great ability in modeling and predicting nonlinear hydrologic time series. Tremendous number of ANN models have been developed and described for modeling hydrological phenomena such as precipitation and rainfall–runoff processes (e.g., Hsu et al., 1995; Smith and Eli, 1995; Fernando and Jayawardena, 1998; Tokar and Johnson, 1999; Anmala et al., 2000; Nourani et al., 2009a,b). A comprehensive overview on concepts and applications of ANNs in hydrological simulations has been provided by ASCE (2000a,b). The success of ANN in any hydrological process modeling depends on the quantity and quality of the data used for training the model. The used model and data for the simulation of the rainfall–runoff process usually contain uncertainties. For example, an averaged value of the pointy measured rainfalls by the rain gauges over the watershed is usually assigned to whole basin. Subsequently, the utilization of such constant real number as the watershed rainfall in the ANN input layer can be a source of the uncertainty. In such situations, fuzzy theory may be employed to handle the uncertainties involved in the real world problems. From this point of view, hybrid ANN and fuzzy system is one of the researching focuses, which benefits the advantages of both ANN and fuzzy system namely ANFIS (Adaptive Neuro-Fuzzy Inference System). For instance, Bazartseren and Holz (2001) applied neuro-fuzzy approach to flood prediction and Gautam et al. (2001) developed a model for real-time forecasting of water levels using ANFIS approach. Inasmuch the application of ANFIS model has increased in the hydrological modeling, some researchers

* Corresponding author. Tel.: +98 914 403 0332; fax: +98 411 334 4287.
E-mail addresses: vnourani@yahoo.com, vnourani@umn.edu (V. Nourani), mehdi_komasi@yahoo.com (M. Komasi).

0022-1694/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jhydrol.2013.03.024

investigated the accuracy of ANFIS model with respect to the other models. For example, Aqil et al. (2007) used three different adaptive techniques (Levenberg–Marquardt feed forward neural network, FFNN, Bayesian regularization FFNN, and neuro-fuzzy) to examine the advantages of ANNs and neuro-fuzzy system in continuous modeling of daily and hourly discharges. They reported a detailed comparison about overall performance of the models and showed the superiority of the neuro-fuzzy model. Recently, Rajaee et al. (2009) applied ANN, ANFIS, multi-linear regression and conventional sediment rating curve models for time series modeling of suspended sediment concentration. The results demonstrated that ANFIS model presents a better performance in the prediction of suspended sediment concentration compared to other models. Hence, ANFIS model is an intelligent model which contains the learning capability of neural network and knowledge illustration of fuzzy logic system using linguistic expressions in a single framework and has showed accurate results for predicting the time series (Bazartseren et al., 2003; Jacquin and Shamseldin, 2006; Aqil et al., 2007). It is notable that there are a few studies about application of ANFIS to rainfall–runoff modeling. For examples, Nayak et al., 2004 reported that the ANFIS model not only fully preserves the potential of the ANN approach but also outperforms ANNs in terms of computational speed, forecast errors, and peak flow estimation. Tayfur and Singh (2006) developed fuzzy logic model, employing the triangular fuzzy membership functions, for event-based rainfall–runoff modeling. Jothiprakash et al. (2009) found that pure cause-effect ANFIS models are performing better than combined ANFIS models for the same number of time lagged input data sets for rainfall–runoff modeling of an intermittent river system. In the context of ANFIS-based hybrid models, an ANFIS model with autoregressive exogenous input (ARX) structure was proposed and examined by Gautam and Holz (2001) for modeling rainfall–runoff process of the Sieve basin in Italy. Furthermore, Nourani et al. (2011) applied the wavelet-ANFIS hybrid model to catch the seasonality patterns of rainfall and runoff time series in two watersheds at daily and monthly time scales.

In spite of suitable flexibility of ANN and ANFIS approaches for modeling hydrologic time series such as rainfall and runoff, sometimes, there is a shortage in predicting runoff at intermediate points or stations of a watershed. In such a situation, ANN and ANFIS models may not be able to predict runoff because the mentioned published ANN and ANFIS models consider a single runoff gauging station at the outlet of watershed as the model's output. Although the merit of ANN in temporal modeling of hydrological phenomena has been already confirmed by many researchers, only a few attempts have been done to assess the ANN ability in spatial and/or spatiotemporal modeling of hydrological processes. For instance, three alternative models were developed by Luk et al. (2001), based on the use of ANNs to implement the pattern recognition methodology which attempts to forecast short-duration rainfall at specific locations within a catchment. Mutlu et al. (2008) developed and compared ANN models to forecast daily flows at multiple gauging stations in an agricultural watershed. Feasibility of an ANN methodology for spatiotemporal groundwater level simulation in an aquifer was also evaluated by Nourani et al. (2008). The efficiency of spatiotemporal ANN was compared with two hybrid neural-geostatistics and multivariate time series-geostatistics models and it was found that the ANNs could provide accurate predictions (Nourani and Ejlali, 2012). Hence, few attempts have been made in hydrology to use ANN and ANFIS for spatiotemporal modeling (see e.g., Nasr and Bruen, 2008). Furthermore, because of the fluctuation, periodic pattern and paucity, involved in the rainfall and runoff time series, the data may not be employed without an appropriate data preprocessing. To contend with this problem, hydrologists usually use some data preprocessing techniques (e.g., classification, clustering and transformation) being coupled with the main models in order to partition the data into groups and to setup the separate models instead of one global model (Cannas et al., 2006; Wu et al., 2009; Nourani and Kalantari, 2010; Demirel et al., 2011; Sadri and Burn, 2011; Nourani et al., 2011; Tsai et al., 2012). It has been already concluded that the preprocessing of the hydrological data can significantly reduce the effort and computational time required for developing the ANN model and also it can increase the model's accuracy (Kisi, 2008).

In the field of clustering, as such a preprocessing approach, Ruspini (1970) introduced the idea of fuzzy clustering to construct a fuzzy partitioned clustering method. More recently, advances in fuzzy set literature and advent of powerful computational facilities have spurred the development and utility of fuzzy clustering methods for a variety of applications in different disciplines (Ramachandra Rao and Srinivas, 2006). A few attempts have been also made in hydrology to explore the potency of fuzzy clustering in hydrological modeling (e.g., Guler and Thyne, 2004; Wang et al., 2006; Sadri and Burn, 2011). In the current study, fuzzy C-means (FCM) clustering has been tested for rainfall–runoff modeling. FCM clustering was proposed by Bezdek (1981) as an improvement over the classic hard k-means clustering algorithm.

Considering the ability of ANFIS in modeling hydrological processes which usually involve some degrees of uncertainties, the objective of this study was to develop a new time-space integrated ANFIS model to simulate daily runoff at multiple stations within the Eel River watershed and it was aimed to improve the prediction capability by employing the pre-processed data obtained by means of FCM clustering method.

In conventional AI-based rainfall–runoff models, usually temporally variable parameters (as historical time series) of the process are considered as inputs of the model to predict runoff values at the outlet or several interior stations of a watershed. However in the proposed methodology, in addition to temporally variable parameters (rainfall and runoff time series), spatially variable parameters over the watershed (upstream drainage area, slope, curve number) were also imposed to the input matrix of the model to train a unique and integrated model (not several models, one for each station) to interpolate the process not only across the time but also over the space; so that the interaction of stations could be taken into consideration because of imposing all stations data in a unique framework. Such a spatiotemporal model, as a new generation of AI-based models, was firstly proposed by the authors, using ANN (not ANFIS) to predict river suspended sediment load (see, Nourani and Kalantari, 2010). In the current study, such a spatiotemporal model for modeling watershed rainfall–runoff process was adapted using ANFIS approach. Furthermore, the fuzzy clustering approach was also employed as a preprocessing method to improve the model ability to catch peak values of the discharge time series.

The remaining of the paper is organized as follows. First, brief descriptions of ANFIS model and FCM concept are presented. In the next sections, the proposed IGANFIS model, efficiency criteria and case study are introduced and then the performance of the model is presented and discussed in details. The final section presents the concluding remarks and identifies the area of future research.

2. Materials and methods

2.1. Adaptive Neuro-Fuzzy Inference System (ANFIS)

Each fuzzy system contains three main parts, fuzzifier, fuzzy data base and de-fuzzifier. Fuzzy data base contains two main parts, fuzzy rule base, and inference engine. In fuzzy rule base, rules related to fuzzy propositions are described (Jang et al.,