Ain Shams Engineering Journal (2015) xxx, xxx–xxx

Ain Shams University
Ain Shams Engineering Journal
www.elsevier.com/locate/asej
www.sciencedirect.com

CIVIL ENGINEERING

Computational approaches for annual maximum river flow series

Harinarayan Tiwari *, Subash Pd. Rai, Nayan Sharma, Dheeraj Kumar

Department of Water Resources Development and Management, Indian Institute of Technology Roorkee, India

Received 7 February 2015; revised 7 June 2015; accepted 26 July 2015

KEYWORDS
Probability;
Wavelet;
ANN;
Kosi;
Peak discharge

Abstract Studies of annual peak discharge and its temporal variations are widely used in the planning and decision making process of water resources management. Very recently, soft computing techniques are gaining ground for time series analysis of hydrological events such as rainfall and runoff. In this study Artificial Neural Network (ANN) has been used in combination with wavelet to model the annual maximum flow discharge of rivers. The results of ANN-Wavelet (WANN) model indicate overall low coherence (R^2 = 0.39) better than ANN (R^2 = 0.31) in isolation. In the present analysis, the authors also conceded a probabilistic distributional analysis of river flow time series which has greater potential to better reflect peak flow dynamics. The results highlight that the overall performance of probability distribution models is superior to WANN model. Instead of that WANN is better than probabilistic models to find the global maxima of the series.

1. Introduction

River discharge measurements in general are carried out on discrete basis [1–3] and in particular on daily basis in India. Data and information on annual peak discharge supplements flood management, reservoir planning and irrigation scheduling [4,5]. Peak discharge is the consequence of summing up of the all contributing discharges from river tributaries. In hydrology, studies related to peak events (river discharge) are deemed necessary given its use in various statistical analyses [1–3].

Estimation of discharge is primarily carried out using two types of mathematical approaches: multivariate approach and univariate approach [6–8]. Physical hydrological modeling is a multivariate approach to estimate the peak discharge using hydro-meteorological data (rainfall, temperature, etc.) and geomorphological data (slope, soil type, etc.) [9,10]. Physical hydrological modeling requires enormous amounts of data and moreover it is a time-consuming process [11,12]. The recent trend is a marked shift from physical hydrological modeling to the use of soft computing techniques, which is gaining significance in a short spell of time [13,14]. Multivariate approach using soft computing techniques (ANN, SVM, etc.) is preferred over physical modeling owing to its limitations of time consumption and data volumes [15,16]. Some multivariate approaches may lead to underestimation of the events as well as the increased uncertainty associated with the given event [17–19]. In multivariate approach, underestimation may be attributed to cumulative effect of uncertainty

* Corresponding author.
E-mail address: haribit31@gmail.com (H. Tiwari).
Peer review under responsibility of Ain Shams University.

ELSEVIER Production and hosting by Elsevier

http://dx.doi.org/10.1016/j.asej.2015.07.016
2090-4479 © 2015 Faculty of Engineering, Ain Shams University. Production and hosting by Elsevier B.V.

Notations

Q	discharge, m^3/s	w	weights
p	time delay, Year	ε_t	fluctuations at time, t

existence in individual isolated factor. To minimize the cumulative uncertainty, univariate approach is useful given the point that all the external factors have previously influenced the production of the observed time series. It can also be said that the time series embodies all the information required to model the underlying generating process.

Soft computing has numerous applications in hydrological time series analyses, be it multivariate approach or univariate approach [20]. Stationary and non-stationary time series can be analyzed using soft computing techniques. The Artificial Neural Network (ANN) is a soft computing technique that comprises both linear and nonlinear concepts and can be operated with dynamic input–output system. Artificial Neural Network (ANN) is an influential processing tool which has been widely used in water resources research [21,22]. Processing time series components of water resources projects (WRP) with ANN requires a preprocessing stage for data reduction such as Wavelet Transform (WT) in order to gain advantages in training time and also to pass up the redundancy in input data. This helps to obtain a model with better generalization abilities. This is the prime reason for better performance of WANN in different aspects of water resources management [23,24].

Probability distribution is the process of statistical inference surveyed data. It is mainly of two types: parametric distribution and nonparametric distribution. Parametric distributions are probability distributions that can be described using an equation with a finite set of parameters. For a specified parametric distribution, the parameters are estimated by fitting to data. In the field of hydrology the concept of probabilistic distribution can be applied frequently [25,26].

There is a growing need to critically evaluate the annual maximum discharge (using univariate approach) in the river to assist the knowledge base for better planning and management of water resources projects. This promoted the use of ANN and Wavelet-ANN combination to study the annual maximum discharge. In the present study, analyses of extreme flood near the Kosi Mahasetu have been performed using annual maximum discharge time series over the 51 year period from 1964 to 2014. The goal of this study was to determine appropriate probability distributions for describing annual maximum stream flow series for the Kosi River. In this paper, three parameters based (shape, scale and location) study was carried out to characterize the maximum flow of Kosi River.

The major research findings of this study revolve around as follows:

(1) Trend analysis and autocorrelation analysis to detect time period of similar events.
(2) Performance evaluation of ANN and Wavelet-ANN time series analysis of annual maximum discharge series.
(3) Evaluation of probability distribution model for annual maximum discharge and comparison of the results with Wavelet-ANN model.

2. Description of techniques

2.1. Artificial neural network

Artificial Neural Networks (ANNs) may provide an alternative model to river discharge in areas which lack precise data and information about the internal hydrologic processes. ANN model has been developed with a correlation coefficient of 0.99 for the maximum daily river discharge [27].

ANN model employs nonlinear functional mapping on the past observations to predict the future values. ANN uses logistic hidden layer transfer function and two model parameters as connection weights [28].

$$Q_t = f\left(Q_{t-1}, Q_{t-2}, Q_{t-3} \ldots, Q_{t-p}, w\right) + \varepsilon_t \tag{1}$$

2.2. Wavelet analysis

The hypothesis of wavelet analysis was developed based on Fourier analysis. A signal is broken up into smooth sinusoids of unlimited duration in Fourier analysis [13]. A wavelet is a mathematical function which can be used to localize a given function in both space and scaling [29,30]. Wavelets can be utilized to extract information from diverse kinds of data; such as seismic, finance, heartbeat and hydrological [27,31–36]. Wavelet analysis is often used to learn evolutionary behavior to characterize fluctuated daily discharge time series [37,38]. The major improvement of wavelet transforms is their capability to concurrently acquire information on the time, location and frequency of a signal, while the Fourier transform provides only the frequency information of a signal.

The continuous Wavelet Transform (CWT) of a discharge time series $Q(t)$ is defined as follows:

$$W(\tau, s) = s^{-1/2} \int_{-\infty}^{+\infty} Q(t)\varphi^{*}\left(\frac{t-\tau}{s}\right)dt \tag{2}$$

$W(\tau, s)$ presents a two-dimensional representation of wavelet power under a different scale, where 's' is the wavelet scale, 't' is the time, 'τ' is the translation parameter and '*' denotes the conjugate complex function. The translation parameter 'τ' is the time step in which the window function is iterated.

2.3. WANN analysis

The Wavelet Artificial Neural Network (WANN) models are obtained combining two methods, Discrete Wavelet Transform (DWT) and ANN. The WANN model is an ANN model, which uses sub-time series components obtained using DWT on original data. The WANN model structure developed in the present study is shown in Fig. 1. For WANN model inputs, the original time series is decomposed into a certain number of sub-time series components (D_s). All component plays different role in the original time series and the behavior of each