



Research papers

A joint stochastic-deterministic approach for long-term and short-term modelling of monthly flow rates

Milan Stojković^{a,*}, Srđan Kostić^a, Jasna Plavšić^b, Stevan Prohaska^a^a Institute for Development of Water Resources “Jaroslav Černi”, Jaroslava Černog 80, Belgrade, Serbia^b University of Belgrade Faculty of Civil Engineering, Bulevar Kralja Aleksandra 73, Belgrade, Serbia

ARTICLE INFO

Article history:

Received 28 June 2016

Received in revised form 10 November 2016

Accepted 13 November 2016

Available online 17 November 2016

This manuscript was handled by A. Bardossy, Editor-in-Chief

Keywords:

Joint stochastic-deterministic modelling

Cross-correlation transfer function

Artificial neural network

Polynomial regression

Climatic input

The Great Morava River

ABSTRACT

The authors present a detailed procedure for modelling of mean monthly flow time-series using records of the Great Morava River (Serbia). The proposed procedure overcomes a major challenge of other available methods by disaggregating the time series in order to capture the main properties of the hydrologic process in both long-run and short-run. The main assumption of the conducted research is that a time series of monthly flow rates represents a stochastic process comprised of deterministic, stochastic and random components, the former of which can be further decomposed into a composite trend and two periodic components (short-term or seasonal periodicity and long-term or multi-annual periodicity). In the present paper, the deterministic component of a monthly flow time-series is assessed by spectral analysis, whereas its stochastic component is modelled using cross-correlation transfer functions, artificial neural networks and polynomial regression. The results suggest that the deterministic component can be expressed solely as a function of time, whereas the stochastic component changes as a nonlinear function of climatic factors (rainfall and temperature). For the calibration period, the results of the analysis infer a lower value of Kling-Gupta Efficiency in the case of transfer functions (0.736), whereas artificial neural networks and polynomial regression suggest a significantly better match between the observed and simulated values, 0.841 and 0.891, respectively. It seems that transfer functions fail to capture high monthly flow rates, whereas the model based on polynomial regression reproduces high monthly flows much better because it is able to successfully capture a highly nonlinear relationship between the inputs and the output. The proposed methodology that uses a combination of artificial neural networks, spectral analysis and polynomial regression for deterministic and stochastic components can be applied to forecast monthly or seasonal flow rates.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Prediction of flow rates tends to be the focus of hydrological research, given its great significance in terms of the operation of reservoirs that have to meet a number of requirements such as hydropower generation, water supply, flood protection and water quality. Additionally, reservoir operation rules must deal with multi-annual and intra-annual variability of the hydrologic process so that limited water resources can be shared (Yevjevich, 1984; Pekarova et al., 2003; Stojković et al., 2016). As such, there is a justified need for thorough investigation of the inner structure of flow rates, to enable “control” of their behaviour, i.e. predict temporal changes of flow rates considering the natural variability in hydrologic and meteorological time series (Loucks, 2005). One way of

doing this is to derive convenient mathematical relations among the input factors and the flow rate, which allow for relatively simple flow rate estimation rate based on observed inputs (Abudu, 2009; Abudu et al., 2011; Nourani et al., 2011; Kostić et al., 2016), or just as a function of time using long-term statistics (Pekarova et al., 2003; Stojković et al., 2016). The uncertainty connected with such approaches is strongly conditioned by the periodicity of the temporal changes in the flow rate or by the nature and selection of climatic inputs (Loucks, 2005; Taormina and Chau, 2015).

Hydrological models are commonly defined as deterministic or stochastic, or a combination of the two (Fleming, 1975). It is intuitively clear that deterministic models enjoy great popularity, primarily because they are much easier to derive in comparison to stochastic models, and they are simpler to use (same output for the same input). This is corroborated by a variety of deterministic approaches, including empirical, conceptual and physically-based

* Corresponding author.

E-mail address: milan.stojkovic@jcerni.co.rs (M. Stojković).

(Singh and Frevert, 2002). However, studies in the last decade show that deterministic modelling of flow rate variation leaves a certain part of the examined time series unexplained, and this error can be significantly amplified when deterministic models are used to predict the distant future. Fortunately, these uncertainties in hydrologic modelling of flow rate time-series can be overcome by stochastic models (Efstratiadis et al., 2014; Ilich, 2014) or joint stochastic-deterministic models (Koutsoyiannis, 2000), which have extensively been applied in the last decade.

So far, flow time-series have mostly been examined using a linear stochastic model such as ARMA (Autoregressive Moving Average), ARIMA (Autoregressive Integrated Moving Average) or SARIMA (Seasonal Auto Regressive Integrated Moving Average) (Salas et al., 1980; Valipour, 2015). Such models are capable of capturing a linear statistical dependence among several successive time lags. Their prediction can be improved by coupling with empirical model decomposition, an approach that is useful for analysing time series which are often non-stationary (Wang et al., 2015). Although widely used, the main drawback of such an approach lies in the fact that the autocorrelation function of the modelled hydrologic time series rapidly decreases and cannot be used for modelling of a long-memory hydrologic process (Koutsoyiannis, 2000).

Opposite to ARMA models, an alternative way of modelling river flows as a dependent process driven by influential climatic parameters is to use the cross-correlation Transfer Function (TF) (Box et al., 2008). The application of TFs reveals the physically meaningful mechanisms that underlie the rainfall-runoff relationship (Young, 2005). This approach has been widely applied for modelling monthly flow rates as a function of climatic inputs (Abudu, 2009; Abudu et al., 2011) or for flow prediction in ungauged basins (Palanisamy and Workman, 2014). It should be noted that even though the application of transfer functions allows one to infer the main physical mechanism behind the flow time-series, the main limitation of this method lies in its capability to capture only linear relations between the inputs (climatic factors) and the outputs (flow rate). Given the fact that these relations are, by definition, strongly nonlinear (which explains their stochastic-unpredictable nature), the application of transfer functions for this purpose can lead to ambiguous results. One way to deal with the “wild and uncontrollable” nature of hydrological processes is to engage some of the soft computing methods, including techniques such as artificial neural networks (ANN) (Nourani et al., 2011; Valipour et al., 2013; Gholami et al., 2015), the support vector machine (Wu et al., 2014), evolutionary polynomial regression (Guistolisi and Savić, 2006), singular spectrum analysis (Chau and Wu, 2010) and genetic programming (Akbari-Alashti et al., 2015). The results of previous studies have confirmed that ANN could capture a nonlinear relationship between climatological inputs and monthly flows (Abudu, 2009). A population-based optimisation algorithm can be deployed (Chen et al., 2015) to improve the reliability of an ANN model. Moreover, Abudu et al., 2011 developed a hybrid model based on TF and ANN for modelling of monthly flows during the spring-summer runoff season. The results of their analysis indicate that the recorded high quantiles of monthly flow rates are well reproduced by the modelled time series. For the same reason, Wu et al., 2009 proposed three data-preprocessing techniques (moving average, singular spectrum analysis, wavelet analysis), coupled with ANN to improve the estimate of daily flows. Considering model performance and complexity, they suggest that the optimal model is the coupled one that combines moving average and ANN.

As outlined above, there are many models for flow estimation that are based on long-term flow statistics (Pekarova et al., 2003; Pekarova and Pekar, 2006; Stojković et al., 2016) or which use the conventional mathematical approach based on dependence

among inputs and the output at a few successive lags (Abudu, 2009; Abudu et al., 2011; Nourani et al., 2011; Kostić et al., 2016). The former approach preserves the characteristics of the hydrologic process in the long-run, while the latter models can reproduce the short-term hydrologic pattern. The objective of the present paper is to introduce an approach that preserves both long-term and short-term characteristics of hydrologic time-series on two time scales (annual and monthly), using a disaggregated time series. The underlying assumption is that the hydrologic time series can be decomposed into deterministic, stochastic and random parts. In the present paper, the deterministic part consists of the composite trend, macro-periodic component and stationary seasonal component. These components describe low-frequency and seasonal variability, while high-frequency variability is represented by the stochastic component. In the present research, the stochastic component is modelled by TFs that describe linear relations among the flow time-series and climatic data. For the same purpose, artificial neural network (ANN) and polynomial regression (PLR) are used as techniques capable of modelling the complex relations between monthly flows and climatic inputs. In such cases, ANN and PLR serve as tools for examining the accuracy of the approach based on transfer functions, with the ultimate goal of deriving the most reliable model for the stochastic component of monthly flow time-series.

2. Methodology

The present research relies on the basic assumption that monthly flow time series can be decomposed into deterministic, stochastic and random parts (Yevjevich, 1972), i.e. monthly flow rates Q_t represent the sum of three components:

$$Q_t = Det_t + Stoch_t + error_t \rightarrow Q_t \\ = [Q_T + Q_P + Q_S]_t + [Y]_t + [e]_t \quad t = 1, 2, \dots, N \quad (1)$$

where Q_T is the composite trend, Q_P is the long-term periodic component, Q_S is the seasonal component, Y_t is the stochastic component and N denotes the sample size. Addend e_t is the error term (random time series) having a zero mean, constant variance σ_e^2 and covariance function $C_\tau = 0$ for lag $\tau > 0$.

One should emphasise that the composite trend (Q_T) and long-term periodic component (Q_P) describe a long-term property of the hydrological process expressed as large-scale variations, whereas intra-annual (seasonal) variations are represented by the seasonal component (Q_S). Alongside these fluctuations, the stochastic component represents high-frequency changes in monthly flows. It should be noted that the stochastic component is characterised by a substantial serial correlation and its auto-correlation function rapidly decreases with the time lag. The last term in Eq. (1) is the random component (e_t), which represents an independent, uncorrelated time series (white noise).

2.1. Modelling of the deterministic component

In the present paper, the authors model the deterministic component separately from the stochastic component, according to Eq. (1), as a combination of the composite trend Q_T , macro-periodic component Q_P and seasonal component Q_S . The composite trend and macro-periodic component constitute the annual deterministic component, which is modelled on an annual time-scale, whereas the seasonal component is modelled on a monthly time scale. After the annual deterministic component is modelled, it is downscaled to monthly intervals, so as to subtract it from the monthly flow time-series. The first and second order residuals, Q' and Q'' , are derived in this way. The latter contains a significant

Download English Version:

<https://daneshyari.com/en/article/5771367>

Download Persian Version:

<https://daneshyari.com/article/5771367>

[Daneshyari.com](https://daneshyari.com)