Journal of Hydrology 496 (2013) 71–78

Contents lists available at SciVerse ScienceDirect

Journal of Hydrology

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

Scale-dependent synthetic streamflow generation using a continuous wavelet transform

Jun Niu [a,*], Bellie Sivakumar [b,c]

[a] Department of Civil Engineering, The University of Hong Kong, Pokfulam, Hong Kong, China
[b] School of Civil and Environmental Engineering, The University of New South Wales, Sydney, NSW 2052, Australia
[c] Department of Land, Air and Water Resources, University of California, Davis, CA 95616, USA

ARTICLE INFO

Article history:
Received 28 March 2013
Received in revised form 9 May 2013
Accepted 10 May 2013
Available online 21 May 2013
This manuscript was handled by Geoff Syme, Editor-in-Chief

Keywords:
Synthetic streamflow generation
Continuous wavelet transform
Morlet wavelet
Random permutation
Pearl River basin

SUMMARY

This study develops a scale-dependent synthetic data generation method for streamflow by using a continuous wavelet transform. The detailed information of streamflow variability across different timescales embedded in the data is obtained from the continuous wavelet transform. To take into account the time-dependent flow magnitudes, the wavelet coefficients are simply separated into two basic categories, namely high-flow part and low-flow part. The data reconstruction is based on the random permutation of the separated wavelet coefficients for the two categories. The synthetic generation is performed at both the individual timescales and the multiple timescales. The Morlet wavelet transform is considered as a representative continuous wavelet transform, and generation of daily streamflow data is attempted. The method is applied to a streamflow series observed in the Pearl River basin in South China. The results indicate that the proposed method: (1) is suitable for scale-controlled generation of streamflow time series and (2) provides reliable information as to the extent of spectral properties present in the original data that need to be preserved.

1. Introduction

In hydrology, synthetic data generation is an important step for providing a basis for undertaking a variety of water-related design, operation, and diagnostic studies. The basic characteristics of hydrologic time series can be described in terms of, for example, (1) asymmetric and marginal probability distributions; (2) persistent large amplitude variations at irregular time intervals and frequency-dependent amplitude variations; (3) long memory, nonlinear dependence, and time irreversibility; and (4) nonlinear dynamic and chaotic properties (see, for example, Salas et al., 1980; Jayawardena and Lai, 1994; Lall and Sharma, 1996; Smith et al., 1998; Sivakumar et al., 2001; Whitcher et al., 2002). The synthetic generation of hydrologic time series (e.g. rainfall, streamflow) can be achieved by employing any of the following models: shot noise model (Weiss, 1977), fragments model (Srikanthan and McMahon, 1980), autoregressive moving average model (Box et al., 1994), artificial neural networks (ANNs) (Raman and Sunilkumar, 1995), stochastic disaggregation model (Valencia and Schaake, 1973; Tarboton et al., 1998; Acharya and Ryu, in press), Markov chain model (Aksoy, 2003), bootstrapping method (Lall and Sharma, 1996; Srinivas and Srinivasan, 2005), and wavelets (Bayazit and Aksoy, 2001), among others.

Hydrologic data, and data in earth sciences at large, are often nonstationary. Since traditional Fourier transform methods do not contain any information on the time dependence of the associated signal, they cannot provide any local information regarding the time evolution of its spectral characteristics (Lau and Weng, 1995). Wavelet transforms enable us to obtain expansions of a signal using the time–frequency atoms, called wavelets, that have good properties of localization in both time and frequency (time-scale) domains (Foufoula-Georgiou and Kumar, 1994; Kumar and Foufoula-Georgiou, 1997). The localized fluctuations that are inherently present in the nonstationary processes can be effectively reflected by wavelet analysis. The wavelets generally refer to either orthogonal or nonorthogonal wavelet functions (Torrence and Compo, 1998). The discrete wavelet transform is applied when an orthogonal basis (e.g. Haar, Coiflet, and Daubechies) is employed, while the use of a nonorthogonal wavelet function (e.g. Mexican hat, Beta, and Morlet) implies the possible use of the continuous wavelet transform. The various kinds of wavelet transform applications in hydrology include the feature characterization of precipitation (Kumar and Foufoula-Georgiou, 1993a,b; Kumar, 1996; Özger et al., 2010; Mishra et al., 2011a) and discharge (Coulibaly and Burn, 2004; Labat, 2006), rainfall–runoff relations (Labat

* Corresponding author. Tel.: +852 2859 1970.
E-mail addresses: niuj@hku.hk, niuj.hku@gmail.com (J. Niu).

0022-1694/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jhydrol.2013.05.025

et al., 2000), drought-related forecasting (Özger et al., 2011, 2012), and teleconnection analysis between regional hydrologic variability and climatic patterns (Marković and Koch, 2005; Zhang et al., 2007; Özger et al., 2009; Labat, 2010a; Mishra et al., 2011b; Niu, 2013), among others.

The use of wavelets for synthetic generation of hydrologic data was, to our knowledge, first attempted by Bayazit and Aksoy (2001), as a nonparametric data generation tool. The fundamental idea in their approach is the decomposition of time series into the details in time–frequency domain and then reconstruction by properly changing the details to generate new time series. Bayazit and Aksoy (2001) used the simplest wavelet function, the Haar wavelet, to synthetically generate annual and monthly streamflow series, and demonstrated its capability in generating non-skewed data. Comparing this method with an autoregressive model and five other models, Bayazit et al. (2001) demonstrated the advantages of the wavelet-based method in preserving the statistical characteristics (i.e. the mean value and the autocorrelation) of the observed time series. The method was then also successfully applied for generation of reservoir storage (Aksoy, 2001) and rainfall data (Ünal et al., 2004).

In the study by Bayazit and Aksoy (2001), the annual or monthly streamflow time series, with its length equal to some power of 2, was decomposed into the wavelet details (i.e. wavelet coefficients) at the corresponding different levels of resolution. In order to preserve the autocorrelation features of the series, a random selection of the decomposed wavelet coefficients at different resolutions was implemented for formation of the first element of the series. The subsequent elements were then generated one by one by the summation of the sequenced wavelet details, in which each detail value used was right next to those of the previous step at each resolution. Wang et al. (2011) adopted the Trous algorithm (Shensa, 1992; Aussem et al., 1998) to decompose and reconstruct the observed daily streamflow time series. The first two levels of the decomposed wavelet details (coefficients) and the remaining approximation series were used for the reconstruction. Compared to the reconstruction strategy in Bayazit and Aksoy (2001), the difference in Wang et al. (2011) was the random sampling. The details and approximation were divided into a number of sub-series based on a yearly period. Then, the random sampling was conducted from those sub-series at both the detail level and the approximation level. Wang et al. (2011) reported that their method avoids assumptions of probability distribution types (e.g. Normal) and of the dependence structure (linear or nonlinear).

These and other wavelet-based methods for synthetic generation of hydrologic data are essentially based on the discrete wavelet transform. The discrete wavelet transform produces wavelet spectrum that contains discrete 'blocks' of wavelet power, which is certainly useful to compactly represent the associated signals. At the same time, however, the features of 'compact representation' simplify the variations at longer timescales. To reveal (for the original signal) and reconstruct (for synthetic generation) the variability at different timescales, the continuous wavelet transform is preferable, since it offers smooth, continuous variations in wavelet amplitude. To this end, the present study makes the very first attempt to use the continuous wavelet transform for synthetic data generation in hydrology. To illustrate the utility of the continuous wavelet transform, the study employs the Morlet wavelet (Morlet et al., 1982a,b) for synthetic generation of daily streamflow series in the Pearl River basin in South China. To take into account the salient features of streamflow (including flow magnitude, the issue of data length, and simplicity in application), different random permutations of wavelet details are also presented in the data reconstruction process.

2. Wavelet transform

The wavelet transform has shown its promise in diverse scientific fields (e.g. for hydrology, see Labat, 2010b), with its capability in analyzing variability properties for both stationary and non-stationary time series at different timescales. Mathematically, a wavelet transform decomposes a time series x_t in terms of "daughter" wavelets $\psi(t, s)$ derived from a "mother" wavelet function $\psi_0(t)$ by the timescale (s) dilation and time position (t) translation:

$$\psi(t,s) = \frac{1}{s^{1/2}}\psi_0\left(\frac{t'-t}{s}\right) \tag{1}$$

where $s^{1/2}$ is an energy normalization factor to keep the energy of daughter wavelets the same as the energy of the mother wavelet. The wavelet transform of the time series x_t is defined as the convolution integral of x_t and a dilated and translated version of $\psi_0(t)$:

$$W(t,s) = \frac{1}{s^{1/2}}\int \psi^*\left(\frac{t'-t}{s}\right)x_t\,dt \tag{2}$$

where ψ^* is the complex conjugate of ψ defined on the time and scale.

To be admissible as a wavelet function, it must have zero mean and be localized in both time and timescale (frequency) domains (Farge, 1992). Furthermore, several factors should be considered in choosing the wavelet function, such as complex or real, width, and shape. The complex Morlet wavelet function, chosen for this study, is capable of (1) providing a good balance between time and frequency localizations and (2) capturing the oscillatory features because the complex function has more oscillation waves (Torrence and Compo, 1998). The complex Morlet wavelet function, shown in Fig. 1, consists of a plane wave modulated by a Gaussian:

$$\psi_0(\eta) = \pi^{-1/4}e^{i\omega_0\eta}e^{-\eta^2/2} \tag{3}$$

where η is the non-dimensional time parameter; ω_0 is the non-dimensional frequency, with value of 6 to satisfy the admissibility condition (Farge, 1992).

Consider a time series x_t (e.g. streamflow) observed at an equal time interval δt (e.g. daily) over a period of time $t = 1, \ldots, T$. For the purpose of convenience, the timescales of wavelet transform are written as fractional powers of two:

$$s_j = s_0 2^{j\delta j}, \quad j = 0, 1, \ldots, J \tag{4}$$

$$J = \delta j^{-1}\log_2\left(\frac{T\delta t}{s_0}\right) \tag{5}$$

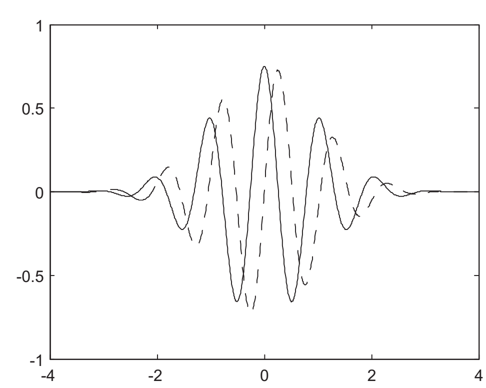

Fig. 1. The real part (solid) and imaginary part (dashed) for the Morlet wavelet ($\omega_0 = 6$) in the time domain.