Journal of Hydrology 388 (2010) 480–490

Contents lists available at ScienceDirect

Journal of Hydrology

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

A daily stochastic weather generator for preserving low-frequency of climate variability

Jie Chen *, François P. Brissette, Robert Leconte

Department of Construction Engineering, École de technologie supérieure, Université du Québec, 1100 Notre-Dame Street West, Montreal, QC, Canada H3C 1K3

ARTICLE INFO

Article history:
Received 13 July 2009
Received in revised form 18 March 2010
Accepted 19 May 2010

This manuscript was handled by Andras Bardossy, Editor-in-Chief, with the assistance of K.P. Sudheer, Associate Editor

Keywords:
Weather generator
Precipitation
Low-frequency variability
Power spectra

SUMMARY

Weather generators are computer models that produce time series of meteorological data that have similar statistical properties as that of observed data. The past decade has seen a sharp and renewed increase in interest in weather generators, linked to their potential use in climate change studies. One appealing property of weather generators is their ability to rapidly produce time series of unlimited length, thus permitting impact studies of rare occurrences of meteorological variables. However, one problem with daily weather generators is that they underestimate monthly and inter-annual variances because they do not take into account the low-frequency component of climate variability. This research aims to present an approach for correcting the low-frequency variability of weather variables for weather generator and to assess its ability to reproduce key statistical parameters at the daily, monthly and yearly scales. The approach is applied to precipitation which is usually the variable displaying the largest inter-annual variability. The daily stochastic precipitation model is a Richardson-based weather generator that uses a first-order two-state Markov chain for precipitation occurrence and a gamma distribution for precipitation amounts. Low-frequency variability was modeled based on observed power spectra of monthly and annual time series. Generation of synthetic monthly and yearly precipitation data was achieved by assigning random phases for each spectral component. This preserved the power spectra, variances and the autocorrelation functions at the monthly and annual scales. The link to daily parameters was established through linear functions. The quality of these corrections was assessed through direct and indirect validation tests, with the direct validation focusing on comparing the means, standard deviations and autocorrelations of different weather series. The results showed that standard deviations of both monthly and annual precipitations were produced almost exactly. The proposed method also preserved the autocorrelation of annual precipitation. The indirect validation involved modelling the discharge of a river basin using a hydrological model driven by different precipitation series. The results showed that the corrected weather series significantly improved the variability of simulated flow discharges at the monthly and annual scales compared to those simulated using the data generated by the standard weather generator.

1. Introduction

A stochastic weather generator is a computer algorithm that uses existing meteorological records to produce a long series of synthetic daily weather data. The statistical properties of the generated data are expected to be similar to those of the actual data for a specified site. Unlike historical weather records, which may have missing data, the weather generator output provides a complete record for any desired period of time, thus enhancing the use of continuous hydrologic models (Kevin et al., 2005). Moreover, it can be used to generate daily weather data for ungauged areas through spatial interpolation of model parameters from adjacent gauged sites (Baffault et al., 1996). An important application of weather generators involves them serving as computationally inexpensive tools to produce multiple-year climate change scenarios at the daily time scale, which are used to assess the impact of future climate change (Semenov and Barrow, 1997; Wilks, 1992, 1999; Pruski and Nearing, 2002; Zhang et al., 2004; Zhang, 2005; Zhang and Liu, 2005; Minville et al., 2008). Model parameters of the weather generator can be readily manipulated to simulate arbitrary changes in mean and variance quantities for sensitivity analysis, or be deliberately modified to mimic changes in mean and variance as predicted by global climate models (GCMs) for impact assessment. Over the years, several weather generators have been developed, such as the Weather Generator (WGEN) (Richardson, 1981; Richardson and Wright, 1984), USCLIMATE (Hanson et al., 1994), Climate Generator (CLIGEN) (Nicks et al., 1995), Climate

* Corresponding author. Tel.: +1 514 396 8800.
E-mail address: chj092413@yahoo.com.cn (J. Chen).

0022-1694/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.jhydrol.2010.05.032

Generator (ClimGen) (Stockle et al., 1999), Long Ashton Research Station-Weather Generator (LARS-WG) (Semenov and Barrow, 2002), etc. While weather generators are good at preserving the precipitation quantity, they however underestimate low-frequency variations (e.g., Buishand, 1978; Johnson et al., 1996; Wilks, 1989,1999; Gregory et al., 1993; Katz and Parlange, 1993, 1998; Hansen and Mavromatis, 2001; Zhang and Garbrecht, 2003; Chen et al., 2009). This underprediction results from the simplifying assumption that climate, and more specifically, the daily precipitation process, is stationary. These models do not explicitly take into account aspects of low-frequency variability such as decadal oscillations, and thus underestimate monthly and yearly variances.

The low-frequency variability of precipitation depends on the daily precipitation occurrence and intensity processes, especially the variance of the daily precipitation amounts and number of wet days. Several studies have attempted to solve this drawback with weather generators. Wilks (1999) compared the variance of monthly precipitation generated by independent and identical (iid) Gamma distribution, Common-α Gamma distribution and Mixed Exponential distribution. The results showed that the iid Gamma distributions produced substantial overdispersion, and that the Common-a Gamma distribution brings only a slight improvement to this. By contrast, the overdispersion in wet-day variance produced by the Mixed Exponential distribution was substantially smaller, although not zero, meaning that using the Mixed Exponential distribution to represent wet-day precipitation amounts in stochastic weather models should bring about a substantial improvement in the simulation of inter-annual variability. Meanwhile, Wilks (1999) also compared the variance of the number of wet days in each month among different precipitation occurrence models, including first-, second-, third- and hybrid-order Markov models and Negative Binomial and Mixed Geometric distribution, as well as average percentage overdispersion of total monthly precipitation, for all combinations of precipitation occurrence models and precipitation intensity models. The results demonstrated that none of the combinations achieved complete recovery of the observed variance in monthly total precipitation, although increasingly complex component models did succeed in reducing the overdispersion or discrepancy between the synthetic and observed variability. This was unsatisfactory because although the complexity of the models was increased, it still did not take into account the low-frequency component of climate variability. These simple stationary models (whose statistics do not change from month to month and from year to year) cannot fully reproduce the variability of a nonstationary climate, which therefore makes the introduction of some degree of nonstationarity into these models appropriate.

Hansen and Mavromatis (2001) attempted to improve inter-annual variability characteristics by perturbing monthly parameters using a low-frequency stochastic model, and evaluated the effectiveness of the low-frequency component on low-frequency variability of the generated monthly climate at 25 locations in the continental USA. The results indicated that for monthly precipitation, the low-frequency correction reduced total error and eliminated negative bias of inter-annual variability, and reduced the number of station-months with significant differences between observed and generated inter-annual variability, but it over-represented the variability of precipitation frequency.

Dubrovsky et al. (2004) applied the monthly generator (based on a first-order autoregressive model) to fit the low-frequency variability based on the daily WGEN-like weather generator, Met-Roll. The results demonstrated that conditioning the daily generator on a monthly generator has the most positive effect, especially on the output of a hydrological model, and the variability of the monthly streamflow characteristics was better simulated. However, this method still could not reproduce the observed standard deviations and autocorrelations of monthly and annual precipitations exactly, because it did not specifically consider the inter-annual variability, thus indicating that schemes for correcting monthly variability have limited effect on the annual scale.

Wang and Nathan (2007) also provided a method for coupling daily and monthly time scales in the stochastic generation of rainfall series. The key feature of the method involves first generating two similar time series, one preserving key statistical properties at a finer time scale and the other at a coarser time scale. The finer time scale series is then adjusted to make it consistent with the coarser one. This method appears to perform well in that it satisfactorily preserved some key statistical properties at daily, monthly and even yearly scales. However, it was only tested for the coefficient of variation on Australian weather data. Other statistics, such as the autocorrelation of annual precipitation, are important for some applications.

Accordingly, this research aimed to present an approach for correcting the low-frequency variability of precipitation for the weather generator, assess its ability to reproduce key statistical parameters, and to compare it against Wang and Nathan's method.

2. Materials and methods

2.1. Introduction of a stochastic weather generator

Weather Generator École de Technologie Supérieure (WeaGETS), which is a WGEN-like three-variate (precipitation, maximum and minimum air temperature) single-site stochastic weather generator programmed in Matlab, was used as the basic stochastic weather generator in this study. This paper only focuses on precipitation generation.

The precipitation component of WeaGETS is a Markov chain for occurrence and a gamma distribution for quantity. A first-order two-state Markov chain is used to generate the occurrence of wet or dry days. The probability of precipitation on a given day is based on the wet or dry status of the previous day, which can be defined in terms of the two transition probabilities:

$$P01 = \Pr\{\text{precipitation on day } t \mid \text{no precipitation on day } t-1\} \quad (1a)$$

and

$$P11 = \Pr\{\text{precipitation on day } t \mid \text{precipitation on day } t-1\} \quad (1b)$$

Since precipitation either occurs or does not occur on a given day, the two complementary transition probabilities are $P00 = 1 - P01$ and $P10 = 1 - P11$.

For a predicted rain day, a two-parameter Gamma distribution is used to generate daily precipitation depth (Richardson, 1981). The probability density function for this distribution is:

$$f(x) = \frac{(x/\beta)^{\alpha-1}\exp[-x/\beta]}{\beta\Gamma(\alpha)} \quad (2)$$

where the variable x is the daily precipitation depth, α and β are the two distribution parameters, and $\Gamma(\alpha)$ represents the gamma function evaluated at α.

2.2. Correction of low-frequency variability and validations

The aim of the model is to specifically account for low-frequency variability by correcting daily precipitation at the monthly and yearly scales, using power spectra of observed time series at the same scales. The power spectra are computed using Fast Fourier Transforms (FFT). Wang and Nathan's (2007) method, which is arguably the best available for dealing with the low-frequency problem, was also programmed and used as a comparison method.