Journal of Hydrology 542 (2016) 581–588

Contents lists available at ScienceDirect

Journal of Hydrology

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

Technical Note

Constraining continuous rainfall simulations for derived design flood estimation

F.M. Woldemeskel [a], A. Sharma [a,*], R. Mehrotra [a], S. Westra [b]

[a] School of Civil and Environmental Engineering, The University of New South Wales, Sydney, NSW, Australia
[b] School of Civil, Environmental and Mining Engineering, The University of Adelaide, Adelaide, SA, Australia

A R T I C L E I N F O

Article history:
Received 17 March 2016
Received in revised form 27 July 2016
Accepted 9 September 2016
Available online 10 September 2016
This manuscript was handled by G. Syme, Editor-in-Chief, with the assistance of Ashok Mishra, Associate Editor

Keywords:
Rainfall simulation
Rescaling
Intensity frequency duration

A B S T R A C T

Stochastic rainfall generation is important for a range of hydrologic and water resources applications. Stochastic rainfall can be generated using a number of models; however, preserving relevant attributes of the observed rainfall—including rainfall occurrence, variability and the magnitude of extremes—continues to be difficult. This paper develops an approach to constrain stochastically generated rainfall with an aim of preserving the intensity-durationfrequency (IFD) relationships of the observed data. Two main steps are involved. First, the generated annual maximum rainfall is corrected recursively by matching the generated intensity-frequency relationships to the target (observed) relationships. Second, the remaining (non-annual maximum) rainfall is rescaled such that the mass balance of the generated rain before and after scaling is maintained. The recursive correction is performed at selected storm durations to minimise the dependence between annual maximum values of higher and lower durations for the same year. This ensures that the resulting sequences remain true to the observed rainfall as well as represent the design extremes that may have been developed separately and are needed for compliance reasons. The method is tested on simulated 6 min rainfall series across five Australian stations with different climatic characteristics. The results suggest that the annual maximum and the IFD relationships are well reproduced after constraining the simulated rainfall. While our presentation focusses on the representation of design rainfall attributes (IFDs), the proposed approach can also be easily extended to constrain other attributes of the generated rainfall, providing an effective platform for post-processing of stochastic rainfall generators.

1. Introduction

Continuous rainfall sequences at a subdaily resolution are important in the estimation of short-duration floods and pollutant load, and are commonly used for planning, design and management of urban water systems (Makhnin and McAllister, 2009; Sivakumar and Sharma, 2008; Westra et al., 2012). Continuous subdaily rainfall sequences are particularly important for flood estimation, providing one of the primary means to estimate the catchment's moisture conditions prior to the extreme (flood-producing) rainfall event (Berthet et al., 2009; Michele and Salvadori, 2002; Pathiraja et al., 2012). Despite its importance, subdaily rainfall is generally available at only a small number of locations and often contains a large percentage of missing data, mainly due to the cost and time required to collect data at such resolution. For this reason, stochastic generation models and rainfall disaggregation procedures are commonly used as an alternative way to obtain suitable subdaily rainfall data for a range of design and planning related applications.

A number of stochastic rainfall generation approaches have been investigated in the literature, with the most suitable approaches for a particular application depending on the required spatial and temporal scale of the rainfall data (Pui et al., 2012). Comprehensive reviews of annual, monthly and daily rainfall generation methods can be found in Srikanthan and McMahon (2001) and Sharma and Mehrotra (2010). A number of methods for generating subdaily rainfall through disaggregation procedures are also available, which include canonical and microcanoical models, Poisson cluster models as well as nonparametric based models (See Westra et al. (2012) for a brief review of these approaches). In a recent study, Mehrotra et al. (2012) and Westra et al. (2012) developed a regionalized stochastic model to generate daily and subdaily continuous rainfall sequences throughout Australia. This method involved generating daily rainfall sequences based on data from nearby stations, followed by daily to sub-daily disaggregation

* Corresponding author.
E-mail address: a.sharma@unsw.edu.au (A. Sharma).

http://dx.doi.org/10.1016/j.jhydrol.2016.09.028
0022-1694/© 2016 Elsevier B.V. All rights reserved.

to generate 6 min rainfall sequences—again borrowing information from the nearby stations.

Prior to using stochastically generated rainfall data for hydrological applications, it is important to test the data against important characteristics of the observed data at various scales of aggregation (Srikanthan and McMahon, 2001). In general, the following important characteristics are often sought to be preserved: sample mean, variance, coefficient of skewness, auto-correlation, extremes, and dry and wet spell length. Mehrotra et al. (2012) and Westra et al. (2012) tested the aforementioned daily and subdaily rainfall generation models at five locations across Australia. Although the models successfully reproduced a range of statistics, biases were found in the intensity-frequency relationships for short (subdaily) durations, as these are seldom designed to be reproduced by the generators.

The intended application of the stochastically generated rainfall sequences will inform the selection of the most important characteristics that should be preserved. For planning and designing of infrastructure, accurate representation of extreme rainfall statistics—commonly represented using intensity-frequency-duration (IFD) relationships—is critical. Given the importance of continuous simulation in a range of hydrological applications, the issue of mismatch between selected design attributes that are pre-specified, and the same attributes ascertained through the stochastically generated sequences, becomes important. For instance, design IFD relationships are often prepared through procedures that involve selective spatial and temporal smoothing, correction for geographic factors, and fitting of assumed probability distribution functions to maintain consistency across space. On the other hand, stochastically generated sequences often mimic the historical record while making different assumptions about the underlying probability distribution based on whether or not parametric or nonparametric alternatives are used (Mehrotra and Sharma, 2006). The issue then becomes how to maintain consistency between the designs resulting from these different sources of information, an issue that has significant implications in the context of design flood estimation. We address this here by formulating a method to constrain stochastically generated rainfall data to preserve pre-specified IFD relationships. The stochastically generated rainfall is obtained based on the method described by Westra et al. (2012) and Mehrotra et al. (2012). Two steps are involved: (i) the annual maximum rainfall is rescaled so that the difference between the generated and target IFDs is below a predefined tolerance; and (ii) the remaining rainfall data (i.e., all the rainfall data other than the annual maximum) are rescaled so that the average rainfall from the original stochastic sequences is maintained. The annual maximum rainfall is adjusted across multiple durations so as to provide a reasonable representation of the target IFDs that the approach seeks to mimic. In refining the algorithm used to constrain the IFD statistics, we also explore the following questions: Do we need to rescale the rainfall across all durations or selected specific target durations? To what extent does the rescaling at short duration (say 6 min) affect the rescaling at larger duration (say 30 min) or vice versa? Whether the number of realisations significantly affect the estimated rainfall rescaling factors or not? Note that annual maximum and annual extreme rainfall is synonymously used throughout the paper.

2. Method

Rainfall rescaling factors for annual maximum and the remaining rainfall data are estimated recursively at multiple durations. A flow chart describing the application of the rainfall rescaling is illustrated in Fig. 1 with more detailed explanation given below.

Step-1: Calculate the annual maximum and identify the remaining data. For a selected recursion (e.g. $r = 1$) and target duration (e.g. $D = 6$ min), calculate annual maximum and identify the remaining rainfall data of raw continuous rainfall sequences for all the realisations considered (N).

Step-2: Calculate ensemble mean. Estimate the ensemble mean of the annual maximum and the remaining rainfall data across all the realisations.

Step-3: Estimate rescaling factors. Factors to rescale rainfall are estimated in steps 3a and 3b. More details about these are provided in Sections 2.1 and 2.2, respectively.

Step-4: Rescale rainfall data. rescaling of the generated rainfall is carried out by multiplying with the factors estimated in step-3.

Step-5: Evaluate rescaled rainfall data. The rescaled rainfall sequences are evaluated by applying an objective function to the IFD relationships before and after rescaling. The analysis ends if the objective function is reduced below a pre-specified tolerance; otherwise, step-1 to 5 are repeated based on the next recursion and/or target duration. The objective function is described in more detail in Section 2.3.

2.1. Annual maximum rainfall rescaling factor (f_{ex})

To estimate rescaling factors for annual maximum rainfall, a ratio (r_{AEP}) between the target (IFD^T_{AEP}) and generated (IFD^G_{AEP}) IFD is estimated at each of the exceedance probabilities (Eq. (1)). Here AEP references to the annual exceedance probability of rainfall and overall twenty-one AEP values (1, 5, 10, 15, 20... 90, 95, and 99 years) are considered. The target IFD here is based on the observed data but can, in general, be independently specified, while the generated IFD is estimated empirically based on an ensemble mean of the generated rainfall sequences. Then, a polynomial regression function (g) is developed between the target IFD (IFD^T_{AEP}) and the ratio (r_{AEP}) (Eq. (2)). Finally, rescaling factors at each of the extreme rainfall ranks (f^n_{ex}) (here 'n' and 'ex' represent 'rank' and 'extreme', respectively) are estimated using the function g (Eq. (3)).

$$r_{AEP} = \frac{IFD^T_{AEP}}{IFD^G_{AEP}} \tag{1}$$

$$r_{AEP} = g(IFD^T_{AEP}) \tag{2}$$

$$f^n_{ex} = g(R^n_{ex}) \tag{3}$$

Finally, the rescaled annual extreme rainfall at each of the ranks is estimated by multiplying the raw annual maximum rainfall by the correction factors (f^n_{ex}) for all the realisations. Fig. 2 presents the sequence of processes involved in the estimation of the rescaling factors. Fig. 2 an illustrates an example of the ratio (r_{AEP}) and the function g fitted between the annual maximum rainfall and r_{AEP} at Alice Spring station. The annual maximum 6 min rainfall before and after rescaling is shown in Fig. 2b. As the rescaling factors are less than one, the overall mean annual maximum 6 min rainfall reduces from 7.1 mm to 5.5 mm after rescaling. The non-annual maximum rainfall thus needs to be rescaled to preserve the overall mean of the rainfall, as described in Section 2.2.

2.2. Non-extreme rainfall rescaling factor (f^n_{no-ex})

The ensemble mean of the non-extreme ('*no-ex*') rainfall across all realisations for each rank ('n', sorted from smallest to largest) is denoted by R^n_{no-ex}, and the corresponding non-extreme rainfall rescaling factor is denoted by f^n_{no-ex}. Eq. (4) shows the rainfall mass