­A hydrological calibration procedure for stochastic rainfall model

# Abstract

# Introduction

Stochastic rainfall modelling involves the generation of sequences of rainfall at a scale of interest (e.g., sub-daily, daily, monthly, annual, multi-annual) that is statistically similar to observed rainfall timeseries, typically measured at rain gauges. Simulated rainfall is a primary input to hydrological model, for simulating streamflow. The simulated streamflow is then used to assess risks such as floods and droughts. Stochastic rainfall models (SRMs) are commonly fitted to observed rainfall data. However, when simulated rainfall data are used as input for hydrological models, it is possible that the simulated rainfall will not be translated to realisations of streamflow that is statistically similar to the observed one due to, for example, the incapability of SRMs to capture important rainfall attributes or limited understanding of the rainfall-runoff process. This paper introduces a new hydrological calibration procedure that allows SRMs to help capture streamflow statistics that cannot be achieved with common calibration approach (fitting with observed rainfall statistics) when SRMs are used to generate rainfall input for continuous hydrological simulation.

In catchment hydrology, the term continuous simulation is the simulation of the wet and dry condition of a catchment by estimating the loss in rainfall and generating streamflow at daily, hourly, or sub-hourly time scales. Generally, the process of continuous simulation requires a rainfall-runoff model with sequences of meteorological inputs such as rainfall, evapotranspiration or temperature depending on the specification of the model ([Beven, 2012](#_ENREF_8)). The rainfall time series, a primary input to rainfall-runoff model, can be (1) the observed rainfall data collected at a rain-gauge or a network of rain-gauges or (2) generated from a rainfall model. There has been a substantial amount of study on rainfall model development to capture both the temporal pattern including daily ([Richardson and Wright, 1984](#_ENREF_26), [Semenov and Barrow, 1997](#_ENREF_28), [Sharma and Lall, 1999](#_ENREF_29), [Srikanthan and McMahon, 2001](#_ENREF_30)), sub-daily ([Gupta and Waymire, 1993](#_ENREF_17), [Khaliq and Cunnane, 1996](#_ENREF_18), [Cowpertwait, 2006](#_ENREF_12)), year to year variation ([Thyer and Kuzera, 1999](#_ENREF_32), [Srikanthan and Pegram, 2009](#_ENREF_31)), as well as the spatial pattern ([Wilks, 1998](#_ENREF_34), [Qian et al., 2002](#_ENREF_24), [Leonard et al., 2008](#_ENREF_20), [Rasmussen, 2013](#_ENREF_25), [Baxevani and Lennartsson, 2015](#_ENREF_4), [Bennett et al., 2018](#_ENREF_5), [Evin et al., 2018](#_ENREF_14)) of the observed rainfall. While the rainfall-runoff model can be a lumped conceptual model, a semi-distributed model, or a distributed model ([Boughton and Droop, 2003](#_ENREF_10)). Continuous simulation allows the generation of long sequences of streamflow from which important flood or drought statistics can be extracted for the establishment of appropriate mitigation strategies, early warning, or long-term projection ([Linsley and Crawford, 1974](#_ENREF_21), [Adams and Howard, 1986](#_ENREF_1), [Boughton and Hill, 1997](#_ENREF_11), [Arnaud and Lavabre, 2002](#_ENREF_3), [Blazkova and Beven, 2002](#_ENREF_9), [Lamb, 2005](#_ENREF_19), [Viviroli et al., 2009](#_ENREF_33), [Berk et al., 2017](#_ENREF_7), [Rowe and Smithers, 2018](#_ENREF_27)).

However, one of many challenges to the continuous simulation approach is to assess the impact of uncertainties in rainfall data on rainfall-runoff models response. Poor understanding of the effect of rainfall variability or uncertainties in rainfall data on the output of rainfall-runoff models will hamper the capability of continuous simulation in providing reliable hydrological assessment ([Michaud and Sorooshian, 1994](#_ENREF_22), [Faurès et al., 1995](#_ENREF_15), [Andréassian et al., 2001](#_ENREF_2), [Cristiano et al., 2017](#_ENREF_13)). The sensitivity of rainfall-runoff models to the spatial variability of rainfall input has become an attention to researchers. [Wilson et al. (1979)](#_ENREF_35) found that poor spatial representation of rainfall input causes significant errors to the estimated hydrograph through an experiment on a small-sized catchment (68.6 km2). Similarly, [Michaud and Sorooshian (1994)](#_ENREF_22) examined the effect of rainfall sampling errors on peak-flow estimations on a 150 km2 semi-arid catchment and concluded that poor spatial representation of rainfall (inadequate rain-gauges network in point sampling) accounts for 58% underestimation in observed peak-flow. However, [Nicótina et al. (2008)](#_ENREF_23) pointed out mixed conclusions in the effect of rainfall spatial variability on rainfall-runoff models response from multiple mesoscale catchments, where some catchments tend to dampen the effect of spatial distributions while others tend to amplify it and concluded that the sensitivity of runoff response to the spatial distribution of rainfall is also affected by catchment characteristics such as hillslope and routing time distribution.

Sources of simulated streamflow deficiencies have also been reported to be caused by the deficiencies of simulated rainfall in different time periods. In particular, poor simulated streamflow within a month could be the result of rainfall deficiencies in the concurrent month, the preceding month, or a contiguous block of month ([Bennett et al., 2019](#_ENREF_6)). It was also found that in some cases, “good” simulated rainfall can create “poor” streamflow estimations while “poor” simulated rainfall can create “good” streamflow estimations ([Bennett et al., 2019](#_ENREF_6)). The case of “good” rainfall - “poor” streamflow has also been reported in ([Gao et al., 2020](#_ENREF_16)), in which a single site stochastic rainfall model was developed and was shown to capture well a range of rainfall statistics including the wet and dry spell distributions, the lower and the upper tails. However, when the model is used to generate rainfall input for a rainfall-runoff model, the streamflow estimation was shown to have relatively large underestimation in the high flow ranges (from the 90th to the 99th percentiles, with increasing magnitude of underestimation). It is evident that a rainfall model should not only preserve observed rainfall attributes but also be able to capture observed streamflow characteristics for practical hydrological application. Hitherto, studies on SRMs development generally considered observed rainfall statistics as benchmarks without any consideration on streamflow statistics. Hence, the aim of this paper is to develop a new approach to SRMs development that considers both rainfall and streamflow statistics.

The three key objectives of this paper are:

1. Present the hydrological calibration procedure for SRMs.
2. Demonstrate the calibration procedure with a single site rainfall model and a conceptual rainfall runoff model.
3. Evaluate the virtual hydrological calibration procedure with the traditional observed-rainfall calibration to identify rainfall attributes that could potentially affect the simulated streamflow.

This paper is structured as follows. Sect. 2 introduces the hydrological calibration procedure for SRMs. Sect. 3 presents the case studies that are used to demonstrate the hydrological calibration procedure. Results, discussion, and conclusion are in sect. 4, sect. 5, sect. 6 respectively.

# Hydrological calibration of SRMs

## Overview

A typical calibration procedure for SRMs involves computing some observed rainfall statistics which are then used as model parameter to generate simulated rainfall. Fig. 1 illustrates the general process of SRMs development.

Diagram

Description automatically generated

Figure 1. Schematic of SRMs development with observed rainfall data.

For example, the popular WGEN models ([Richardson and Wright, 1984](#_ENREF_26)) has a daily rainfall generator component that requires 4 rainfall statistics: the probability of dry-wet event (PDW), the probability of wet-wet event (PWW), the shape (α) and rate (β) parameters of the gamma distribution. The first two probabilistic parameters PDW and PWW control the wet/dry pattern of the rainfall using a 1st order Markov chain while parameters α and β controls the amount of rainfall occurs on wet days.

The occurrence parameters PDW and P­WW are computed using the following equations:

Where nDW is the number of wet days given a dry occurred previously, nDD is the number of dry days given a dry day occurred previously, nWW is the number of wet days given a wet day occurred previously and nWD is the number of dry days given a wet day previously.

The parameters α and β is computed by assuming the amount of rainfall on wet days has a gamma distribution. Then the parameters can be estimated using the maximum likelihood (MLE) approach or simply using the method of moment (MoM).

After obtaining the required parameter on a monthly basis (i.e., 4 parameters for each of the 12 month) sequences of daily rainfall can be simulated. The complete procedure are provided in ([Richardson and Wright, 1984](#_ENREF_26)).

Calibrating with observed rainfall data allows SRMs to preserve the observed rainfall statistics. The simulated rainfall time series can be used as input for hydrological models to produce streamflow time series for hydrological assessment. However, it is not necessarily given that simulated rainfall time series will translate to streamflow time series that preserve the properties of observed streamflow data. Therefore, another approach to SRMs development is to calibrate model parameters with streamflow statistics. Fig. 2 illustrates a schematic of the hydrological calibration procedure for stochastic rainfall models.

Diagram

Description automatically generated

Figure 2. Schematic of the hydrological calibration for SRMs with observed runoff

In comparison to the typical procedure of SRMs development, the hydrological calibration of SRMs requires an additional rainfall-runoff model component to simulate runoff estimates and observed streamflow data. The proposed steps for the hydrological calibration procedure are as follows.

1. Generating a sequence of simulated rainfall from the SRM with a pre-defined set of parameters.
2. Generating a sequence of simulated runoff using the rainfall runoff model with the simulated rainfall as input.
3. Computing a pair of statistics from the simulated runoff and the observed runoff (e.g., the flow duration curve).
4. Constructing an objective function with the simulated and observed runoff statistics (e.g., the sum of squared errors)
5. Finding the SRM parameters set that optimizes the objective function.

Diagram

Description automatically generated

Figure 3. Schematic of the hydrological calibration for SRMs with virtual observed runoff.

## Step 1 – rainfall simulation

**Estimating stochastic rainfall model and rainfall-runoff model parameters**: To initiate the experiment, the stochastic rainfall model will be calibrated with at-site observed rainfall data; while the rainfall-runoff model will be calibrated with at-site observed runoff data. This procedure will allow the stochastic rainfall model to simulate rainfall data that are similar to the condition at the site which could avoid potential divergence to the hydrological calibration procedure at later stages. While the set of parameters for the rainfall-runoff model will be fixed throughout the process after they are calibrated and evaluated with the observed runoff.

## Step 2 – streamflow simulation

**Simulating streamflow with simulated rainfall input:** the sequences of simulated rainfall will be used as input to the (already calibrated) rainfall-runoff model to generate sequences of simulated streamflow. Note that a separate aim will investigate the influence of the hydrological model on the overall method.

**Simulating virtual-observed streamflow with observed rainfall input**: A sequence of observed rainfall data will be used as input to the same rainfall-runoff model to generate a sequence of virtual-observed streamflow. This approach removes the possibility of errors from observed streamflow influencing the comparison

## Step 3 – objective function

Sum of square errors

Relative errors

## Step 4 – optimization

**Comparing simulated streamflow and virtual-observed streamflow**: The flow duration curve (FDC) will be the subject of the comparison. The FDC is computed from the streamflow sequences produced in the previous stage. The simulated FDC and the virtual observed FDC will be compared against each other forming an objective function using the sum of squares error (SSE) metric. The value of the objective function will be used to inform the calibration of the stochastic rainfall model parameters (i.e. minimizing the SSE by changing stochastic rainfall model parameters).

**Evaluating stochastic rainfall model performance**: To ensure the performance of the stochastic rainfall model in simulating rainfall input that preserves streamflow characteristics, the model will be verified with a virtual-observed FDC at a different time period (split-sample validation)

# Case study

# Results

# Discussion

* Rainfall attributes

# Limitations and opportunities

* Rainfall runoff models
* Objective function
* Feasibility – runtime
* Other Rainfall model (extend WGEN to capture seasonal pattern - harmonic function)

# Conclusions

* Limitation and future opportunities

*Data availability.* All the data used in this study can be requested by contacting the corresponding author Thien Nguyen at truonghuythien.nguyen@adelaide.edu.au.

*Author contributions.*

*Competing interest.* The authors declare that they have no conflict of interest.

*Acknowledgements.*

# Reference

ADAMS, B. J. & HOWARD, C. D. D. 1986. Design Storm Pathology. *Canadian Water Resources Journal,* 11**,** 49-55.

ANDRÉASSIAN, V., PERRIN, C., MICHEL, C., USART-SANCHEZ, I. & LAVABRE, J. 2001. Impact of imperfect rainfall knowledge on the efficiency and the parameters of watershed models. *Journal of Hydrology,* 250**,** 206-223.

ARNAUD, P. & LAVABRE, J. 2002. Coupled rainfall model and discharge model for flood frequency estimation. *Water Resources Research,* 38**,** 11-1-11-11.

BAXEVANI, A. & LENNARTSSON, J. 2015. A spatiotemporal precipitation generator based on a censored latent Gaussian field: SPATIOTEMPORAL STOCHASTIC GENERATOR. *Water resources research,* 51**,** 4338-4358.

BENNETT, B., THYER, M., LEONARD, M., LAMBERT, M. & BATES, B. 2018. A comprehensive and systematic evaluation framework for a parsimonious daily rainfall field model. *Journal of Hydrology,* 556**,** 1123-1138.

BENNETT, B., THYER, M., LEONARD, M., LAMBERT, M. & BATES, B. 2019. A virtual hydrological framework for evaluation of stochastic rainfall models. *Hydrology and Earth System Sciences,* 23**,** 4783-4801.

BERK, M., ŠPAČKOVÁ, O. & STRAUB, D. 2017. Probabilistic Design Storm Method for Improved Flood Estimation in Ungauged Catchments: PROBABILISTIC DESIGN STORM METHOD. *Water resources research,* 53**,** 10701-10722.

BEVEN, K. J. 2012. *Rainfall-runoff modelling : the primer,* Chichester, West Sussex ;, Wiley-Blackwell.

BLAZKOVA, S. & BEVEN, K. 2002. Flood frequency estimation by continuous simulation for a catchment treated as ungauged (with uncertainty). *Water resources research,* 38**,** 14-1-14-14.

BOUGHTON, W. & DROOP, O. 2003. Continuous simulation for design flood estimation—a review. *Environmental Modelling & Software,* 18**,** 309-318.

BOUGHTON, W. & HILL, P. 1997. A Design Flood Estimation Procedure Using Data Generation And A Daily Water Balance Model. Cooperative Research Centre For Catchment Hydrology.

COWPERTWAIT, P. S. P. 2006. A spatial–temporal point process model of rainfall for the Thames catchment, UK. *Journal of hydrology (Amsterdam),* 330**,** 586-595.

CRISTIANO, E., TEN VELDHUIS, M.-C. & VAN DE GIESEN, N. 2017. Spatial and temporal variability of rainfall and their effects on hydrological response in urban areas – a review. *Hydrology and Earth System Sciences,* 21**,** 3859-3878.

EVIN, G., FAVRE, A.-C. & HINGRAY, B. 2018. Stochastic generation of multi-site daily precipitation focusing on extreme events. *Hydrology and Earth System Sciences,* 22**,** 655-672.

FAURÈS, J.-M., GOODRICH, D. C., WOOLHISER, D. A. & SOROOSHIAN, S. 1995. Impact of small-scale spatial rainfall variability on runoff modeling. *Journal of Hydrology,* 173**,** 309-326.

GAO, C., BOOIJ, M. J. & XU, Y.-P. 2020. Development and hydrometeorological evaluation of a new stochastic daily rainfall model: Coupling Markov chain with rainfall event model. *Journal of hydrology (Amsterdam),* 589**,** 125337.

GUPTA, V. K. & WAYMIRE, E. C. 1993. A Statistical Analysis of Mesoscale Rainfall as a Random Cascade. *Journal of Applied Meteorology,* 32**,** 251-267.

KHALIQ, M. N. & CUNNANE, C. 1996. Modelling point rainfall occurrences with the modified Bartlett-Lewis rectangular pulses model. *Journal of Hydrology,* 180**,** 109-138.

LAMB, R. O. B. 2005. Rainfall-Runoff Modeling for Flood Frequency Estimation. Chichester, UK: John Wiley & Sons, Ltd.

LEONARD, M., LAMBERT, M. F., METCALFE, A. V. & COWPERTWAIT, P. S. P. 2008. A space-time Neyman–Scott rainfall model with defined storm extent. *Water resources research,* 44**,** W09402-n/a.

LINSLEY, R. & CRAWFORD, N. 1974. Continuous simulation models in urban hydrology. *Geophysical research letters,* 1**,** 59-62.

MICHAUD, J. D. & SOROOSHIAN, S. 1994. Effect of rainfall-sampling errors on simulations of desert flash floods. *Water Resources Research,* 30**,** 2765-2775.

NICÓTINA, L., CELEGON, E. A., RINALDO, A. & MARANI, M. 2008. On the impact of rainfall patterns on the hydrologic response. *Water Resources Research,* 44.

QIAN, B., CORTE‐REAL, J. & XU, H. 2002. Multisite stochastic weather models for impact studies. *International journal of climatology,* 22**,** 1377-1397.

RASMUSSEN, P. F. 2013. Multisite precipitation generation using a latent autoregressive model. *Water Resources Research,* 49**,** 1845-1857.

RICHARDSON, C. W. & WRIGHT, D. A. 1984. WGEN: a model for generating daily weather variables. U.S Department of Agriculture. *Agriculture Research Service,* ARS-8**,** 83.

ROWE, T. J. & SMITHERS, J. C. 2018. Continuous simulation modelling for design flood estimation - a South African perspective and recommendations. *Water S.A.,* 44**,** 691-705.

SEMENOV, M. A. & BARROW, E. M. 1997. Use of a Stochastic Weather Generator in the Development of Climate Change Scenarios. *Climatic Change,* 35**,** 397-414.

SHARMA, A. & LALL, U. 1999. A nonparametric approach for daily rainfall simulation. *Mathematics and Computers in Simulation,* 48**,** 361-371.

SRIKANTHAN, R. & MCMAHON, T. A. 2001. Stochastic generation of annual, monthly and daily climate data: A review. *Hydrology and earth system sciences,* 5**,** 653-670.

SRIKANTHAN, R. & PEGRAM, G. G. S. 2009. A nested multisite daily rainfall stochastic generation model. *Journal of hydrology (Amsterdam),* 371**,** 142-153.

THYER, M. & KUZERA, G. 1999. Modelling long-term persistence in rainfall time series: Sydney rainfall case study. *International Conference on Water Resources & Environment Research.* Brisbane, Qld.: Institution of Engineers, Australia.

VIVIROLI, D., MITTELBACH, H., GURTZ, J. & WEINGARTNER, R. 2009. Continuous simulation for flood estimation in ungauged mesoscale catchments of Switzerland – Part II: Parameter regionalisation and flood estimation results. *Journal of Hydrology,* 377**,** 208-225.

WILKS, D. S. 1998. Multisite generalization of a daily stochastic precipitation generation model. *Journal of hydrology (Amsterdam),* 210**,** 178-191.

WILSON, C. B., VALDES, J. B. & RODRIGUEZ-ITURBE, I. 1979. On the influence of the spatial distribution of rainfall on storm runoff. *Water Resources Research,* 15**,** 321-328.