Journal of Hydrology 589 (2020) 125150

Contents lists available at ScienceDirect

Journal of Hydrology

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

Research papers

A stochastic rainfall model that can reproduce important rainfall properties across the timescales from several minutes to a decade

Dongkyun Kim[a,*], Christian Onof[b]

[a] Department of Civil Engineering, Hongik University, Seoul, South Korea
[b] Department of Civil and Environmental Engineering, Imperial College London, UK

ARTICLE INFO

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

Keywords:
Poisson cluster rainfall model
Rainfall variability
Timescale
Holistic approach

ABSTRACT

A stochastic rainfall model that can reproduce various rainfall characteristics at timescales between 5 min and one decade is introduced. The model generates the fine-scale rainfall time series using a randomized Bartlett-Lewis rectangular pulse model. Then the rainstorms are shuffled such that the correlation structure between the consecutive storms are preserved. Finally, the time series is rearranged again at the monthly timescale based on the result of the separate coarse-scale monthly rainfall model. The method was tested using the 69 years of 5-minute rainfall data recorded at Bochum, Germany. The mean, variance, covariance, skewness, and rainfall intermittency were well reproduced at the timescales from 5 min to a decade without any systematic bias. The extreme values were also well reproduced at timescales from 5 min to 3 days. The past-7-day rainfall before an extreme rainfall event, which is highly associated with the extreme flow discharge was reproduced well too. The rainstorm shuffling approaches introduced here may be adopted as a standard procedure in combination with any Poisson cluster rainfall model. The methods are simple and parsimonious, yet significantly reduce the systematic underestimation of rainfall variance at coarse scales, and improve the reproduction of skewness, and extreme rainfall depths values at a range of time-scales, thereby addressing well-known shortcomings of Poisson cluster rainfall models.

1. Introduction

Most natural and anthropogenic systems react sensitively to a distinct range of rainfall temporal variability. Fine-scale rainfall temporal variability (e.g. several minutes to a day) influences flash floods (Singh, 1997; Oh et al., 2016; Anh et al., 2020) and subsequent transport of contaminants (Marshall et al., 2000) and sediments (Tucker and Bras, 2000). Coarse-scale rainfall variability (e.g. several days to years) influences water shortage (Gommes and Petrassi, 1996), human health (Patz et al, 2005; Kovats et al, 2003), food insecurity (Ayoub, 1999) and the corresponding human adaptation (Barbier et al., 2009) and migration (Afifi et al., 2016, Milan and Ruano, 2014), as well as human adaptation strategies to recurring floods (David et al., 2017).

For this reason, most of the current system management strategies are established based on rainfall models designed to reproduce the rainfall variability at a limited range of time scales. However, the *real* natural and anthropogenic phenomena are the consequences of complex interactions of various components that are influenced by rainfall variability at a wide range of timescales. Therefore, a thorough understanding of the systems and comprehensive system management plans may only be achieved by employing a modelling framework that encompasses all relevant components based on one single rainfall model that captures the variability at all relevant timescales.

However, most rainfall models have an intrinsic limitation derived from their fundamental assumptions that do not precisely reflect the complex physical rainfall generation process, so they can reproduce the variability only within a limited range of timescales. For example, models based on the autoregressive process (Mishra and Desai, 2005; Modarres and Ouarda, 2014; Yoo et al., 2016) are good at reproducing the observed rainfall variability at timescales greater than 1 month, and the Markov chain models (Haan et al., 1976; Kwon et al., 2009), alternating renewal processes (Bernardara et al., 2007), and generalized linear models (Coe and Stern, 1982; Beecham et al., 2014; Chandler and Wheater, 2002) cannot reproduce the observed variability at timescales finer than 1 day. Poisson cluster rainfall models can reproduce the rainfall variability at timescales ranging from several minutes to several days (Marani et al., 2000; Park et al., 2019).

Several studies tried to overcome this issue by coupling multiple rainfall models. Koutsoyiannis and Onof (2001) suggested a novel coupling algorithm combining two seasonal autoregressive models with

* Corresponding author.
E-mail address: kim.dongkyun@hongik.ac.kr (D. Kim).

https://doi.org/10.1016/j.jhydrol.2020.125150
Received 20 December 2019; Received in revised form 29 May 2020; Accepted 1 June 2020
Available online 04 June 2020

different temporal resolutions. He argued that the recursive application of the algorithm can produce a rainfall time series preserving the first- to the third-order moments of the observed rainfall at hourly to daily timescales. Menabde and Sivapalan (2000) combined the coarse-scale alternating renewal process model with a fine-scale multiplicative cascade model. Their model reproduced the scaling behaviour of extreme events up to a temporal resolution of 5 min. Fatichi et al. (2011) combined an autoregressive model with a Poisson cluster rainfall model (Rodriguez-Iturbe et al., 1987, 1988). Their composite model showed improved performance in reproducing the interannual rainfall variability that the latter often fails to capture. Kim et al. (2013) disaggregated the monthly rainfall that is drawn from a Gamma distribution using the Poisson cluster rainfall model. They found that their composite approach helps reproduce not only the rainfall variability at hourly through yearly timescales, but also the statistical behaviour of rainfall annual maxima and extreme values at timescales ranging from 1 to 24 h. Paschalis et al. (2014) combined a Markov chain model or Poisson cluster rainfall model for large timescales (e.g. daily) and a multiplicative random cascade model for fine timescales (e.g. minute), which outperformed the individual models across a wide range of scales. Park et al. (2019) suggested a method to combine the Seasonal Auto-Regressive Integrated Moving Average (SARIMA) model for monthly rainfall generation and the Poisson cluster rainfall model for hourly rainfall generation. Their model successfully reproduced the mean, variance, covariance, and proportion of dry periods of the observed rainfall at 1 hourly to yearly timescales at 15 locations across the United States.

Another research avenue addressing this topic stems from the recognition that the statistical distribution characterizing observations at a given timescale are distinct from one another. Papalexiou et al. (2018) suggested an algorithm of disaggregating a coarse time series into any finer temporal aggregation level while keeping the statistical properties of both fine and coarse timescales. Their algorithm replaces the observations at coarse timescales with a set of randomly placed blocks. Here, the blocks are randomly drawn from a normal distribution. They employed the unique parametric algorithm of Papalexiou (2018) for the transformation between the parent normal distribution and the distribution of the target variable. Their model successfully disaggregated the 30 years of monthly precipitation observed at a ground gauge in Kentucky, USA, to an hourly one while preserving moments of order one to three of the depth distribution, as well as the proportion of dry periods at all intermediate timescales.

The aim of this study is to show how one can preserve the main advantage of Poisson-cluster models (Rodriguez-Iturbe et al., 1987; 1988; Cowpertwait, 1991; Onof and Wheater, 1993; Cowpertwait, 1995; Kaczmarska et al., 2014; Onof and Wang, 2019), i.e. their storm-cell structure emulating the organisation of observed rainfall, while reproducing statistics over a similar range of scales. Poisson cluster rainfall models generate the rainfall time series with the assumption that the rainstorms arriving according to a Poisson process contain a series of rainfall cells with random depths and durations (Fig. 1). “While the assumption of Poisson arrival of rainstorms shows its limitations when examining timescales that are much larger than the typical storm duration (of several days or more), this unique approach of conceptualizing rainfall based on the physical storm structure ensures that the model reproduces many statistical properties of the observed rainfall at timescales ranging from several minutes to several days (Olsson and Burlando, 2002). The performance of the model has been validated using rainfall data across the world (Onof et al., 2000), with a variety of climatological characteristics of Australia (Gyasi-Agyei and Willgoose, 1997), Belgium (Verhoest et al., 1997), Germany (Kossieris et al., 2018), Greece (Kossieris et al., 2015), Iran (Dodangeh et al., 2017), Ireland (Khaliq and Cunnane,1996), Korea (Kim et al., 2016), New Zealand (Cowpertwait et al., 2007); the UK (Onof and Wheater, 1993; Koutsoyiannis and Onof, 2001; Burton et al., 2008), the US (Koutsoyiannis and Onof, 2001; Kim et al., 2013). The model was also modified to consider rainfall diurnal cycle (Gyasi-Agyei and Willgoose, 1997).

However, Poisson cluster rainfall models have an intrinsic limitation in reproducing the rainfall variability at time scales coarser than several days, which leads to the underestimation of extreme values at large time scales. Before further investigating this matter, note that the variance of a time series at a coarse time scale consists of the two distinct components coming from the independence and the correlation of the fine-scale records according to the following equations:

$$Var(Y_k^{(nh)}) = nVar(Y_1^{(h)}) + 2 \sum_{i=(k-1)n+1}^{kn} \sum_{j=(k-1)n+1, j>i}^{kn} Cov(Y_i^{(h)}, Y_j^{(h)}) \quad (1)$$

where $Y_i^{(h)}$ represents ith value in a stationary rainfall time series at the aggregation interval h and n represents the degree of time series aggregation.

Note that the second term of the right-hand side of the equation represents the correlation between *all* fine-scale records for time lags smaller than or equal to the relevant coarse scale. If this correlation is underestimated by a model, the variance of the coarse-scale time series, the left-hand side of Eq. (1), will be also underestimated. To see why this might be the case, consider Fig. 1b, which shows the aggregated time series of the storm and the cell structure modelled by the Poisson cluster rainfall models (Fig. 1a). The figure shows that the values in the aggregated time series will be correlated with each other if they share the same rain cell (shaded in yellow colour in Fig. 1a). On the contrary, they will be independent if the values do not share the same rain cell. This happens in particular when these cells belong to different non-overlapping storms (the probability of storms overlapping is tiny). This means that Poisson cluster rainfall models have the inherent limitation of not being able to reproduce the fine-scale correlation[1] between rainfall values observed at distant times that is observed, for instance with monsoon rainfall (Singh et al., 1981), soil moisture recycling (Eltahir, 1998; Entekhabi et al., 1996; Kim and Wang, 2007) and as a result of large-scale global atmospheric circulation (Mooley and Parthasarathy, 1984; Carvalho et al., 2004; Berkelhammer et al., 2010). Equation (1) shows that this leads to the underestimation of the variance at coarse timescales.

This investigation also leads to the conjecture that, if Poisson cluster rainfall models are adjusted so that they can account for the correlations between rainfall values observed at distant times, the issue of underestimating large timescale rainfall variability will be resolved. The rainfall model being proposed here was developed based on this principle. The model is composed of three sub-modules each of which is designed to reproduce the rainfall correlation over a range of timescales (i.e. 5 min to a couple of days, a couple of days to one month, and one month to a decade) reflecting the realistic storm features associated with internal storm structure, summer monsoon, soil moisture recycling, and the large scale global atmospheric circulations. The separation of these ranges of timescales is loosely connected to observed breaks in the scaling behaviour of rainfall detected by multiscaling analyses of rainfall depths (e.g. de Lima and Grasman, 1999; Marani, 2005). The proposed model was tested with 69 years of 5-minute rainfall records observed in Bochum, Germany.

2. Methodology

2.1. Data description

This study used the 69 years of 5-minute rainfall data observed at Bochum, Germany for the period between January 1st, 1931 and

[1] This is also true of coarse-scale correlations; here we initially focus upon fine-scale correlations. Below, we will see how one can also increase correlation at coarser scales.

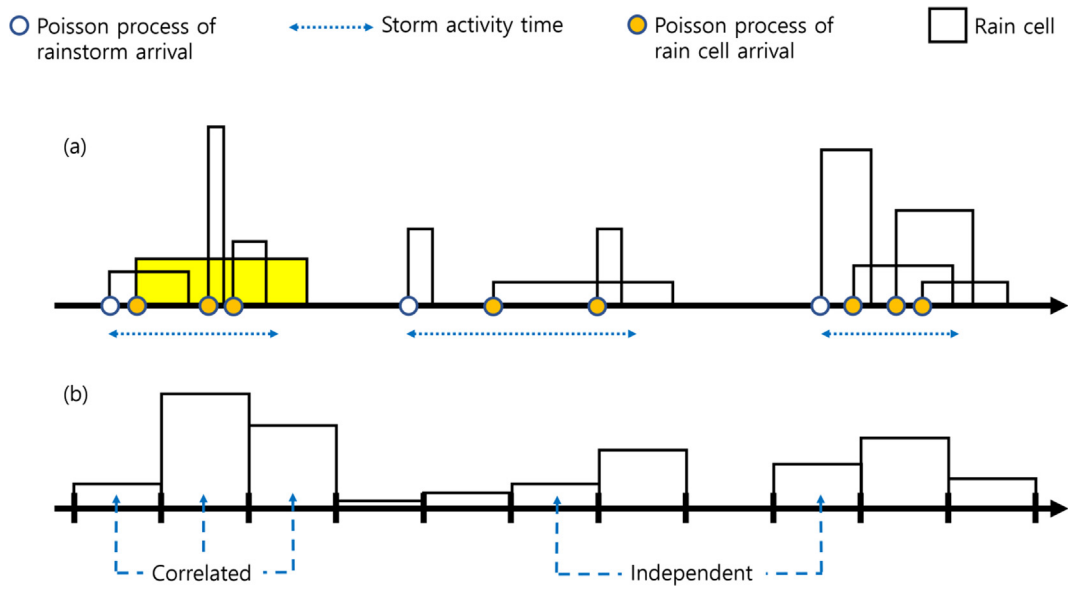

Fig. 1. (a) Schematic of the Poisson cluster rainfall model. (b) Aggregated time series over a given temporal interval. The aggregated values sharing the same rain cell (e.g. shaded in yellow) are correlated with each other while those not sharing the same rain cell are independent with each other.

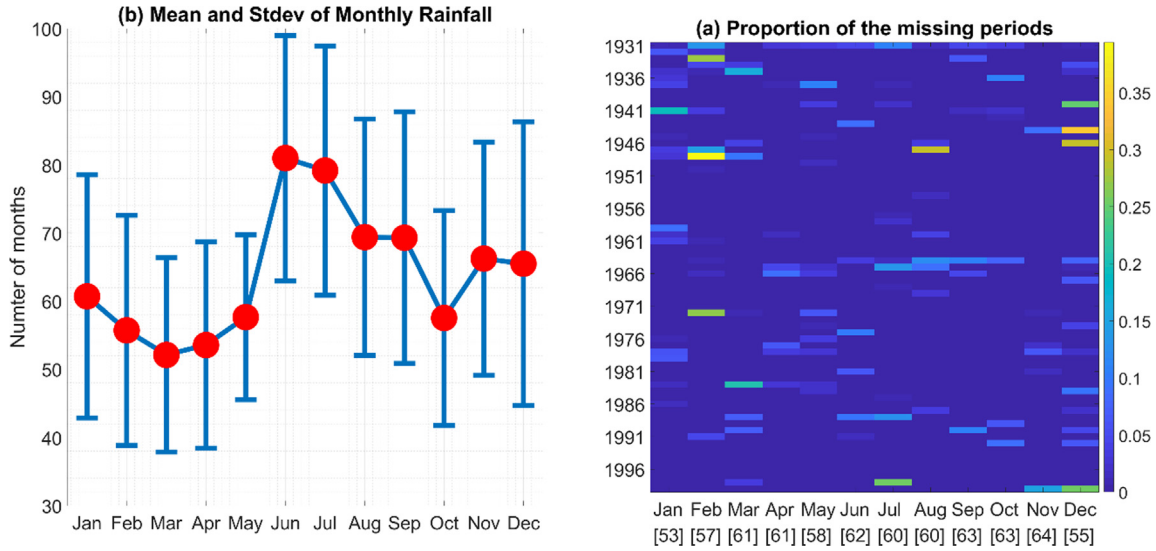

Fig. 2. (a) Seasonal variation of mean and standard deviation of monthly rainfall. (b) Proportion of the missing periods varying with calendar months and year of the 5-minute Bochum rainfall data. The number in the bracket represents the number of the months used to calculate the rainfall statistics.

December 31st, 1999. The mean monthly rainfall displays a clear seasonality and varies from 52 cm in March to 82 cm in June (Fig. 2a). The data have approximately 1 percent of missing periods that are distributed over the years and the calendar months (Fig. 2b). The months with the missing periods greater than 0.1 percent were excluded from the analysis.

2.2. Model description

The first module generates the fine-scale rainfall time series using a randomised Bartlett-Lewis rectangular pulse version of the Poisson cluster rainfall model. Then, the second module shuffles the sequence of the rainstorms to reflect the rainfall variability at time scales ranging from a couple of days to one month. The third module rearranges the adjusted sub-monthly rainfall so that it can reflect the observed rainfall statistics at time scales coarser than one month.

2.2.1. Module 1: Fine scale rainfall generation

For the generation of fine-scale (e.g. sub-hourly) rainfall, this study uses a recent version of the Randomised Bartlett-Lewis Rectangular Pulse (RBLRP) model of Kaczmarska et al. (2014) since it has been shown to outperform other types of Bartlett-Lewis models (ibid.; Onof and Wang, 2019). Unlike the non-randomised BLRP model (Rodriguez-Iturbe et al., 1987) and the traditional randomised RBLRP model (Rodriguez-Iturbe et al., 1988), the $RBLRP_x$ model introduces an the inverse correlation between rainfall cell duration and intensity, in line with the observed behaviour of intense convective rainfall lasting several minutes and milder frontal rainfall lasting for several days. The model generates rainfall based upon the following sequences:

(1) A series of rainstorms arrives in time according to a Poisson process. The parameter of the Poisson process is λ [1/T].
(2) The temporal scaling factor η [1/T] is a gamma distributed random variable with shape and rate parameters ν [−] and α [1/T] respectively. This scaling factor is used as parameter of the

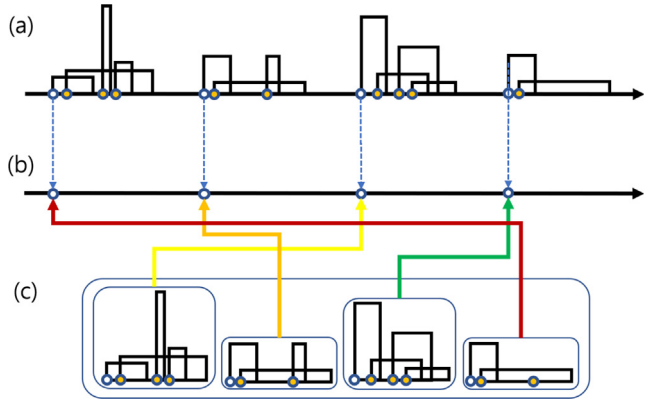

Fig. 3. The schematic of the storm shuffling algorithm of Module 2. (a) The original fine scale rainfall time series. (b) The rainstorms are removed from the original time series, but the times of the rainstorm occurrence in the time series are kept. (c) The rainstorms are randomly selected and placed back into the timeseries. Here, rainstorms with a depth resembling that of the previously selected rainstorm have greater probability of being selected, according to the probability defined in step (4).

exponential distribution of cell durations and to determine the distributions of rainstorm activity duration, rain cell depth, and rainfall cell arrival. All these random variables are mutually independent, which implies in particular that total storm rainfalls are mutually independent.

(3) For each rainstorm and conditionally upon η, the storm activity time is an exponentially distributed random variable with parameter $\eta\phi$, where ϕ [−] is a model parameter.

(4) For each rainstorm and conditionally upon η, rain cells arrive according to a truncated Poisson process with the parameter $\eta\kappa$ where κ [−] is a model parameter. The truncation is defined by the storm activity duration: rain cells can arrive only before the termination of the storm activity duration.

(5) Each rain cell is assigned a duration which is an exponentially distributed random variable with parameter η [1/T].

(6) Each rain cell is assigned an intensity that is a random variable whose distribution could e.g. be exponential, gamma or Pareto. Its mean is $\iota\eta$ where ι [L] is a model parameter. In the present study, we choose the Gamma distribution with shape parameter ω [−] and scale parameter $\iota\eta/\omega$.

The model is thus characterised by the following seven parameters: λ [1/T], ν [−], α [1/T], ι [L], ϕ [−],κ [−], and ω [−].

The parameters of the RBLRP$_x$ model (the "RBL" model hereafter for simplicity) are calibrated such that the statistics of the synthetically generated rainfall approximate those of the observed rainfall. Kaczmarska et al. (2014) derived the analytical expression of the first- to the third-order central moments of the synthetically generated rainfall, and Onof and Wang (2019) further developed the equations to extend the search domain of the parameter α ($\alpha < 1$), which improved reproduction of both extreme and standard statistics at fine timescales (e.g. hourly and sub-hourly). The analytical expression of the proportion of dry (or wet) periods derived by Rodriguez-Iturbe et al. (1988) was also included the calibration.

A variant of the particle swarm optimization algorithm (Cho et al., 2011) was used to identify the parameters which minimise the following objective function:

$$\mathrm{OF} = \sum_{i=1}^{n} w_i(\widehat{M}_i - M_i)^2 \tag{2}$$

where $\widehat{M}_i$ and M_i for $i = 1, \cdots, n$ respectively represent the n modelled and observed rainfall statistics selected for use in the calibration and w_i for $i = 1, \cdots, n$ represent the weight factors given to each statistic.

The type of statistics that were used for calibration (or M_is in Eq. (2)) are the mean, variance, covariance, skewness, and proportion of wet periods at 5-, 10-, 15-, 30-, 60-, 120-, 240-, 480-, 960-, 1440-minute aggregation, so n is 50 (5 different statistics × 10 aggregation intervals). The calibration was performed separately for each calendar month. The weight factors w_i may be determined in various manners depending on the purpose for which the synthetic rainfall is to be used (Kim and Olivera, 2011). In this study, each weight factor w_i is determined as the inverse of the variances of the corresponding observed property M_i: i.e.. using the following equation:

$$w_i = \frac{m}{\sum_{y=1}^{m} (M_i^y - \bar{M}_i)^2} \tag{3}$$

where M_i^y represents the ith statistic of a given calendar month of year y; $\bar{M}_i$ represents the mean of M_i^y over the years; and m represents the number of years of observation. This entails that statistics with greater inter-annual variability thus greater uncertainty have less weight. This choice has been shown to define an optimal generalized method of moments (Jesus and Chandler, 2011).

Note that the parameters of the RBL model is calibrated for entire period of rainfall time series, the diurnal cycle of rainfall cannot not be reproduced

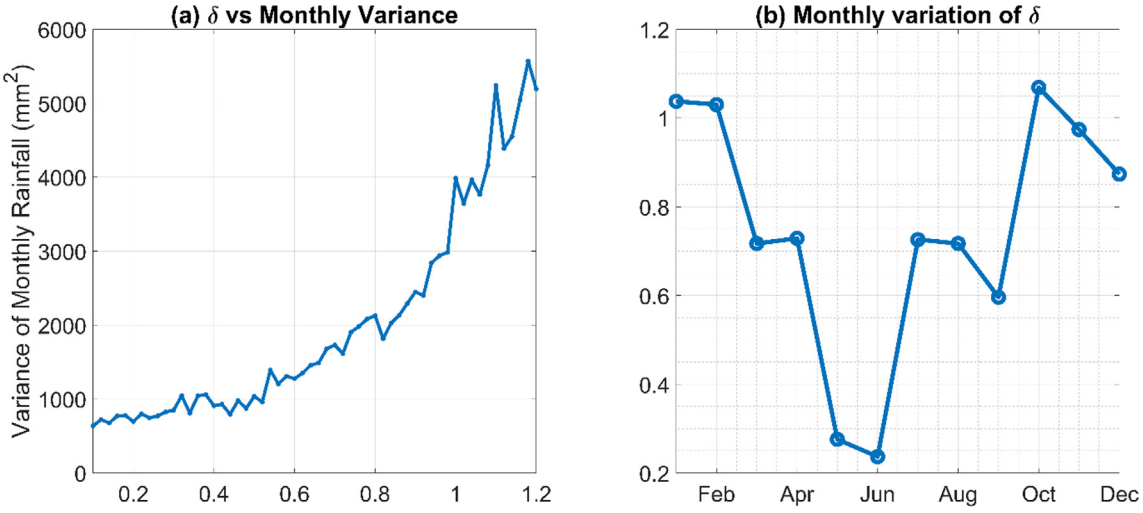

Fig. 4. Relationship between δ and the variance of the monthly rainfall depth.

2.2.2. Module 2: Rainstorm shuffling

This module shuffles the rainstorms generated by the RBL model. Fig. 3 describes the shuffling process. The rainstorms are shuffled based on the following sequence:

(1) While generating the fine-scale rainfall in the Module 1, the time of the rainstorm occurrence and the set of rain cells contained in each of the rainstorms are stored in the database.
(2) Empty the original time series except for the occurrence times of the rainstorms.
(3) Randomly select a rainstorm from the database and place it at the location of the first rainstorm occurrence. Here, each of the rainstorms has the same probability of being selected. The selected storm is then excluded from the database.
(4) Another storm is chosen from the database and placed at the next storm occurrence location. Here, the probability P_i of the rainstorm i being selected, is given by the following equation:

$$P_i = \frac{1}{\sum_{k=1}^{n_i} S_k} \bullet S_i \tag{4}$$

$$S_i = \left(\frac{1}{|\log(Q_i/Q_{prev})|} \right)^{\delta} \tag{5}$$

and Q_i and Q_{prev} represent the total depths of the ith rainstorm and that of the previously selected rainstorm respectively. Q_i is calculated as follows:

$$Q_i = \sum_{j=1}^{n_c} (I_{i,j} \bullet D_{i,j}) \tag{6}$$

where $I_{i,j}$ and $D_{i,j}$ represent the intensity and the duration of the rain cell, respectively, and the first and the second subscripts are the indices corresponding to the rainstorm and the rain cell respectively. For example, $D_{i,j}$ represents the duration of the j^{th} rain cell contained in the i^{th} rainstorm. n_c represents the total number of rain cells contained in the rainstorm. S_i represents the similarity between the i^{th} rainstorm and the previously selected rainstorm. n_i is the number of rainstorms remaining in the database at stage i of the process. δ is a model parameter to be calibrated. The selected storm is then excluded from the database.

(5) Step (4) is repeated until the entire storm occurrence places in the time series are filled with selected rainstorms.

Probability P_i (Eq. (5)) is designed to reflect the similarity in the depths of successive rainstorms characterising the observed rainfall. After this shuffling, it is therefore more likely that storms with relatively similar depths follow one another. This algorithm amounts to altering the RBL model by replacing the assumption of storm independence, with that of a dependence between consecutive total storm rainfalls defined by the conditional probabilitiesP_i. The other assumptions of the RBL model remain valid.

The rainstorms occur at times defined by the storm arrival Poisson process of the RBL model, and each of the selected rainstorms selected for relocation at each rainstorm occurrence time has of course already been generated in the simulation of the RBL model (see Section 2.2.1). Hence, if the time-scale is fine, aside from the two edges (start and end) of the each storm, time-steps are either completely internal or completely external to the storm, and that is the large majority of the time-steps. Therefore the statistical properties of the wet-dry process as well as of the marginal depth distribution (mean, variance, skewness) at such time scales are largely unaffected. When the scale increases, the proportion of unaffected time-steps decreases significantly and the sequence in which the storms arrive will start being relevant, thereby having a more noticeable effect on these statistics.

For larger scales, all but the mean depth will be altered. However, the alteration is small as shown in Fig. 5. The autocorrelation structure is also affected by the shuffling process, at least for scales and time-lags whose product exceeds the typical storm duration, but again the effect is small as seen in Fig. 6a.

Note that the probability P_i could be adjusted so as to reflect the similarity of not only depth but also duration of storms. This would, for instance, enable this reshuffling process to enable the generated rainfall to reflect the rainfall characteristics of the monsoon season, during which long-duration rainfall events successively occur.

Model parameter δ represents the impact of the degree of similarity between the successive rainstorms, as characterised by the modulus of the logarithm of the ratio of their total depths $|\log(Q_i/Q_{prev})|$. It is calibrated separately from the RBL parameters so that the monthly variance of the shuffled synthetic rainfall time series resembles that of the observed rainfall. An analytical approach to the calibration could not be implemented due to the absence of the equation representing the monthly variance of the shuffled synthetic rainfall. Therefore, the pattern search optimization algorithm (Audet and Dennis, 2002) was used to minimize the objective function that is numerically calculated in the following manner:

(1) 300 months of 5-minute rainfall time series are generated using Module 1.
(2) The original synthetic rainfall time series is shuffled based on a given value of δ.
(3) The shuffled 5-minute rainfall time series is aggregated to monthly rainfall, and the variance of the aggregated monthly rainfall is calculated.
(4) The objective function value is calculated as the absolute value of the difference between the observed monthly rainfall variance and the synthetic monthly rainfall variance.

Fig. 4a shows the relationship between δ and the variance of the shuffled synthetic rainfall aggregated to the monthly level. The parameter of July of the study area was used to generate the fine-scale rainfall. The figure a general increase of the variance as a function of δ. This is because, as δ increases, a greater value of P_i is assigned to the rainstorms with the depth similar to the previous rainstorm, so the greater δvalue, the more similar rainstorms flock together, which increases the occurrence of both wet and dry months, thereby increasing the monthly variance.

Note that the two variables do not have a smooth relationship because the variance shown in the y-value is calculated from the stochastically generated rainfall. For this reason, the pattern-search optimisation algorithm was employed, to identify the optimal parameters in the objective function surface with random sampling noise. Fig. 4b shows the calibrated δ for each of the calendar months. This exhibits a clear seasonal trend. Since greater δ values represent greater inter-storm correlations, this result reveals that consecutive summer rainfall events of the study area are less likely to resemble one another in terms of total storm depth, and vice versa for the winter rainfall events, which is to be expected.

2.2.3. Module 3: Monthly rainfall shuffling

Fig. 5 describes the process involved in Module 3. This module rearranges the stochastically generated rainfall time series so that it can account for the variability at timescales greater than 1 month (Tyralis et al., 2018) following the steps described below:

(1) The monthly rainfall time series is generated for the same length as the fine-scale rainfall time series using a separate coarse-scale rainfall model. Any coarse scale model can be used. Here, we call each of the monthly rainfalls generated by the coarse-scale model, a “coarse-scale rainfall block” (Fig. 5a).
(2) The shuffled synthetic fine-scale rainfall time series is segmented into different calendar months and years. Here, we call each segment as “fine-scale rainfall block” (Fig. 5b).
(3) The quantile matching between the fine-scale rainfall blocks and the coarse-scale rainfall blocks is performed based on the total

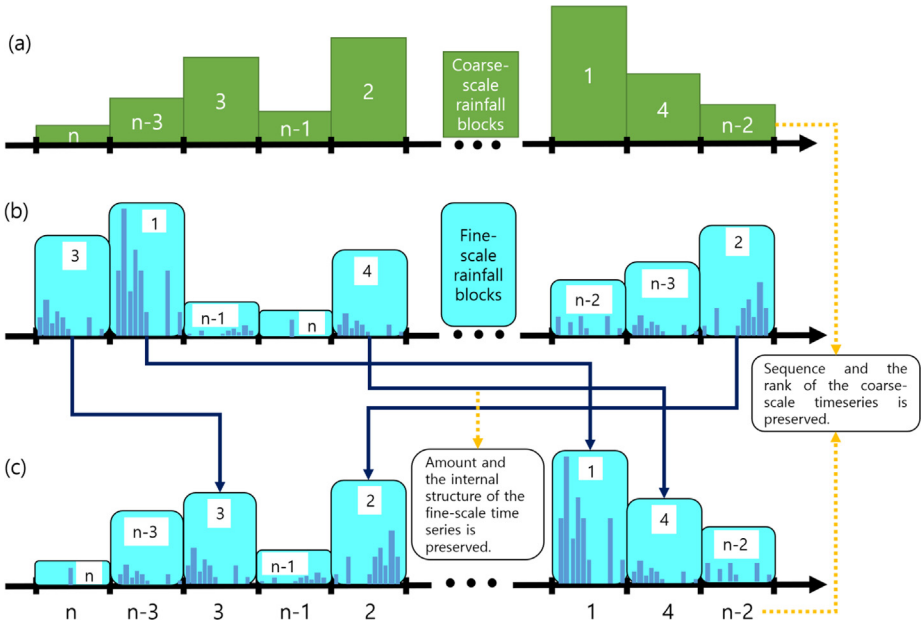

Fig. 5. The schematic of Module 3 of the model. (a) A monthly rainfall time series is generated using a coarse-scale model (e.g. the SARIMA model). (b) The Fine scale time series is segmented into monthly blocks. (c) The final time series is composed by adopting the sequence and rank of the coarse-scale time series and the amount and the internal structure of the fine-scale time series.

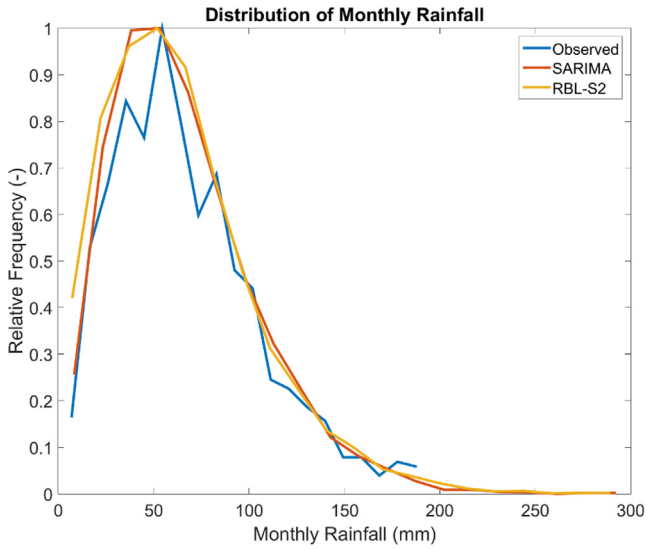

Fig. 6. Normalized monthly rainfall distribution of observation (blue), SARIMA model (red), and the model of this study (yellow). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

rainfall depth for each of the months of a given calendar month. For example, the fine-scale rainfall block with the n^{th} greatest depth is placed in the location of the coarse-scale rainfall block with the n^{th} greatest depth of the same calendar month. This process is repeated for all 12 calendar months (Fig. 5c).

This study uses the Seasonal Auto-Regressive Integrated Moving Average (SARIMA) model for monthly rainfall generation. The model structure is as follow:

$$y_t = \tau + \xi_1 y_{t-1} + \xi_{12} y_{t-12} + \xi_{24} y_{t-24} + \xi_{36} y_{t-36} + \xi_{48} y_{t-48} + \theta_{12} e_{t-12} \quad (7)$$

where y_t represents the monthly rainfall of the month t. ξ_1, ξ_{12}, ξ_{24}, ξ_{36}, ξ_{48} and θ_{12} are the parameters of the model that were estimated based on the method of maximum likelihood. The model structure was determined through an optimization process to minimize the Akaike Information Criterion (AIC) in the parameter space where non-seasonal terms (p,d,q) and the seasonal terms (P,D,Q) vary discretely in the ranges 0–2, 0–2, 0–2, 0–9, 0–1, and 0–9, respectively.

The shuffling algorithm suggested here adopts the sequence and the ranks from the blocks of the coarse-scale time series while it borrows the amount and the temporal structure from the fine scale rainfall time series internal to the blocks. Here, the key to a seamless coupling between the two models is whether the marginal distribution of the amounts of fine scale rainfall aggregated to the monthly level match the distribution of the amounts of monthly rainfall generated by the coarse scale model, at least up to the second-order. While the means of both distributions are identical because they are reproduced by the RBL model at all time-scales, Module 2 ensures a match at the second-order because the parameter δ of Module 2 is calibrated so as to preserve the variance of observed monthly rainfall, which the SARIMA model of Module 3 is also designed to reproduce. Fig. 6 compares the marginal distribution of the 69 years of monthly rainfall (blue) and the one based on the 500 years of synthetic rainfall (red and yellow). The red line corresponds to the SARIMA model and the yellow to the model of this study that went through both shuffling algorithms. The little difference between the red and yellow lines suggests that the monthly rainfall shuffling algorithm explained in this section successfully combined the fine- and coarse-scale rainfall models.

For convenience, we define the names of the models depending on the level of processing as follows:

(1) RBL: The Randomized Bartlett-Lewis model (Module 1 only)
(2) RBL-S: The RBL model with Module 2 (rainstorm shuffling algorithm)
(3) RBL-S2: The RBL model with Module 2 and Module 3 (both rainstorm shuffling and monthly rainfall shuffling algorithm)

2.3. Model application and validation

500 years of synthetic rainfall data were generated. Both the standard statistics and the extreme values were compared at timescales from 5-minutes to a decade.

3. Results and discussions

3.1. Reproduction of standard statistics.

Fig. 7 compares the statistics of the observed (x) and the synthetic (y) rainfall at timescales ranging from 5 min to 6 years. The coloured

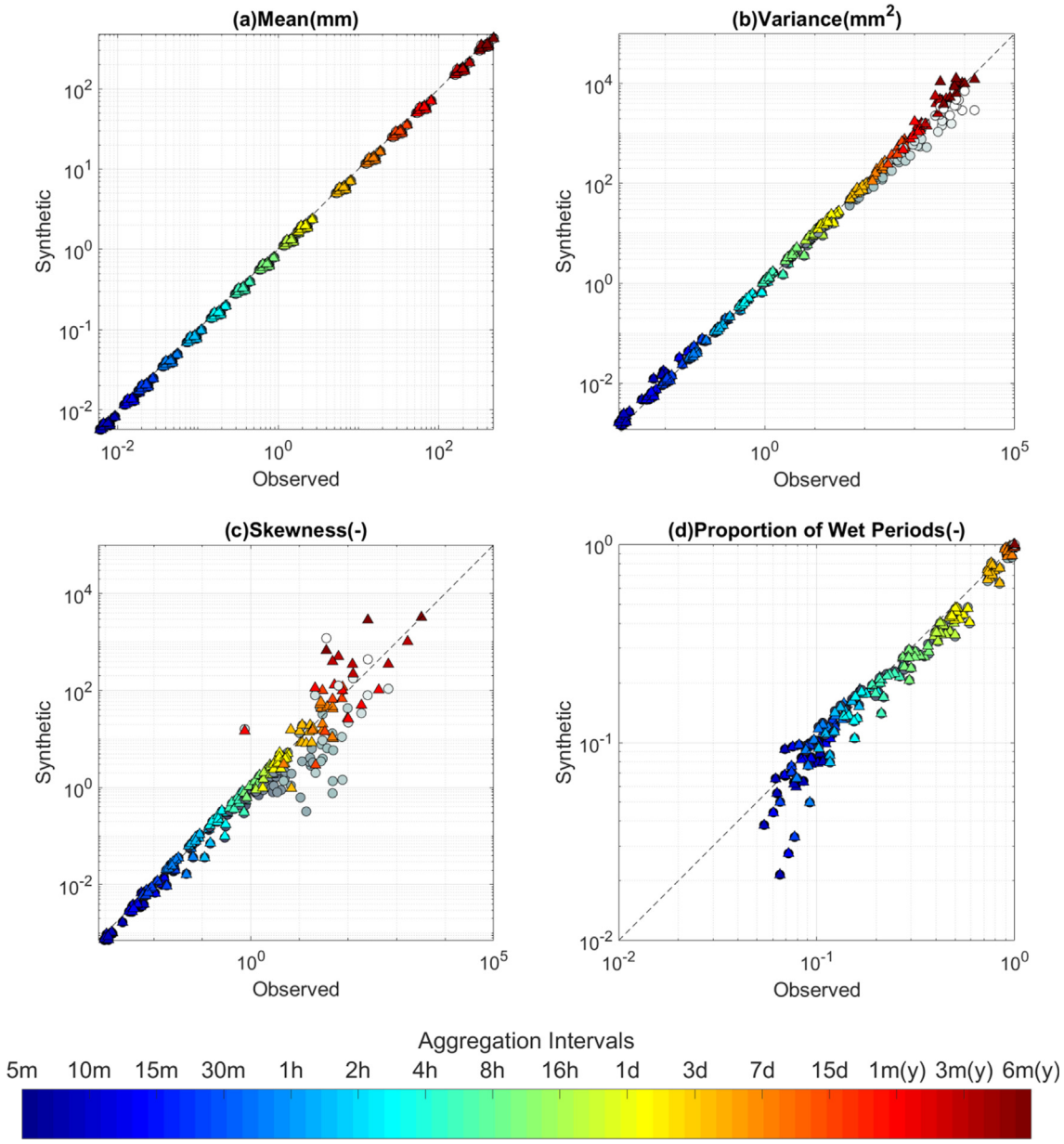

Fig. 7. Mean, variance, skewness, and proportion of wet periods of the observed (x) and synthetic (y) rainfall time series. The coloured triangles and grey circles represent the RBL-S2 and the RBL models respectively. The colours of the triangles and the brightness of the grey circles represent different aggregation intervals. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

triangles and grey discs represent the result of the RBL-S2 and the RBL models, respectively. The colours of the triangles and the brightness of the grey discs represent different aggregation intervals. Each colour has 12 triangles or discs representing each calendar month. For this, the time series of a given calendar month for consecutive years were constituted (e.g. January 1930, January 1931,…, January 1963 for the calibration period), then the time series were aggregated to a given timescale, from which statistics were calculated. For this reason, the timescales of 1, 3, and 6 months shown in this plot are 1, 3, and 6 years, respectively, which is denoted in the colour legend.

The mean rainfall is well reproduced regardless of the model type (and if this is true at one scale, it is true at all scales). The variance is well reproduced by both models at sub-hourly scales. The RBL model underestimates the variance for aggregation intervals exceeding approximately one hour and the degree of underestimation increases with the increase of the aggregation interval. The RBL-S2 model does not underestimate variances for any scales from 5 min to 6 years. This result suggests that the model also successfully reproduces the rainfall correlation structure across the timescales.

While the RBL model underestimates the skewness at time scales exceeding ~1 h, the RBL-S2 model significantly reduces the degree of underestimation. This is because the rainstorm shuffling algorithm of the RBL-S2 model makes the similar rainstorms flock together in the time series, which produces more of both smaller and greater rainfall depth values when the time series is aggregated to the coarser level. This not only increases the variance but also thickens both the head and tail part of the rainfall depth distribution increasing the skewness.

3.2. Correlation structure

Fig. 8 compares the correlation structure of the time series of February rainfall. The red, blue, and black lines are the Auto-Correlation Function (ACF) corresponding to the observed rainfall, the synthetic rainfall generated by the RBL model, and the synthetic rainfall

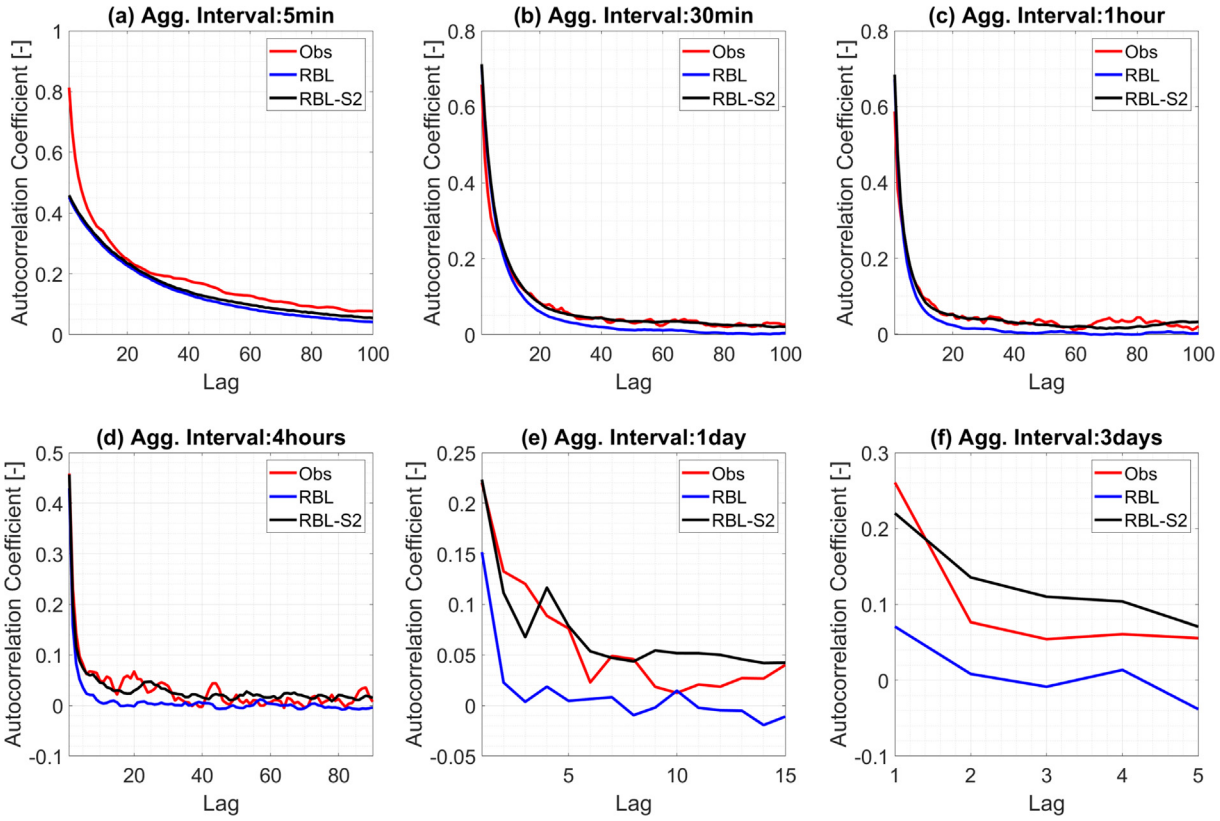

Fig. 8. Autocorrelation function (ACF) of the observed and the synthetic rainfall time series of February. ACF plots with the aggregation interval of 5 min, 30 min, 1 h, 4 h, 1 day, and 3 days are shown.

generated by the RBL-S2 model, respectively. The ACFs of the time series aggregated into 5 min, 30 min, 1 h, 4 h, 1 day, and 3 days are shown. The autocorrelation function (ACF) of the observed rainfall does not converge to 0 even at lag values corresponding to approximately 2 weeks (Fig. 8e and f). This gradual decaying trend of the ACF could not be reproduced by the RBL model, of which the ACF converges to 0 at the lag values corresponding to approximately 2 days. This value roughly coincides with the inter-storm arrival time (λ^{-1}) which varies between 1.3 days (November) and 2.8 days (February). This is because, as the lag of the ACF increases, the rainfall values from independent rainstorms are considered in the calculation of the autocorrelation coefficient (See Fig. 1b), which abruptly decreases the ACF value. Conversely, the RBL-S2 model successfully reproduces the gradual decaying tendency of the observed ACF. This is because the correlation between the rainfall values sampled from consecutive rainstorms tends to persist even though the lag of the ACF becomes longer than the inter-storm arrival time.

Fig. 8a reveals that, for small lags, the fine-scale correlation structure is underestimated by all models. This issue originates from the fundamental model assumption that the rain cells have rectangular shape so that the shape of the hyetograph resulting from their superposition is less smooth than that of the observed hyetograph (Kim et al., 2019). This issue may be resolved by applying a smoother shape of rain cell instead of rectangle, but this should be carefully addressed to maintain the mathematical framework of the rectangular pulse models and will be presented in a forthcoming paper.

3.3. Interannual variability

Fig. 9 compares the quantiles of observed (x) and synthetic (y) monthly rainfall of January, April, July, and October. For all months, the RBL model overestimates the low monthly rainfall values (dry period wetter than the observation) and underestimates the high monthly rainfall values (wet period drier than the observation). The RBL-S2 model resolves this problem. This is also because the rainstorm shuffling algorithm makes large rainstorms flock together with large rainstorms and small rainstorms with small rainstorms, so the months with extremely large and low rainfall occur in sequence more frequently than the case of the RBL model where a series of rainstorms have independent characteristics.

3.4. Moments across timescales

The primary purpose of this study is to develop a rainfall model that can reproduce the rainfall variability at all hydrologically relevant timescales so it can simultaneously be applied to all components of the modelling system. Fig. 10a compares the variances of the observed and the synthetic rainfall at aggregation intervals ranging between 5 min to a decade. While the RBL model underestimates the variance at timescales greater than approximately 1 day, the RBL-S model successfully reproduces the variances at time scales from 5 min to 6 months, but it also underestimates the variance at the timescale exceeding 6 months. The RBL-S2 model successfully reproduces the rainfall variability at timescales from 5 min to a decade. This is because the model reflects the rainfall variability at the large timescale (e.g. 1 to 10 years) that the SARIMA model of Module 3 reproduces.

The variation of higher-order moments with regard to varying timescales (Lovejoy and Schertzer, 1991; Deidda, 2000; Koutsoyiannis, 2006; Veneziano et al., 2006; Langousis et al., 2009; Dimitriadis and Koutsoyiannis, 2015; Lee et al., 2014; Iliopoulou and Koutsoyiannis, 2019) is closely associated with the extreme behaviour of rainfall (Langousis et al., 2009) as well as watershed response variables such as peak flow (Song et al., 2014). Fig. 8b and c shows the scaling behaviour of the third and the fourth central moments across the timescales. The result was similar to the case of the variance analysis: the RBL-S model reduces the underestimation of the moments caused by the RBL model

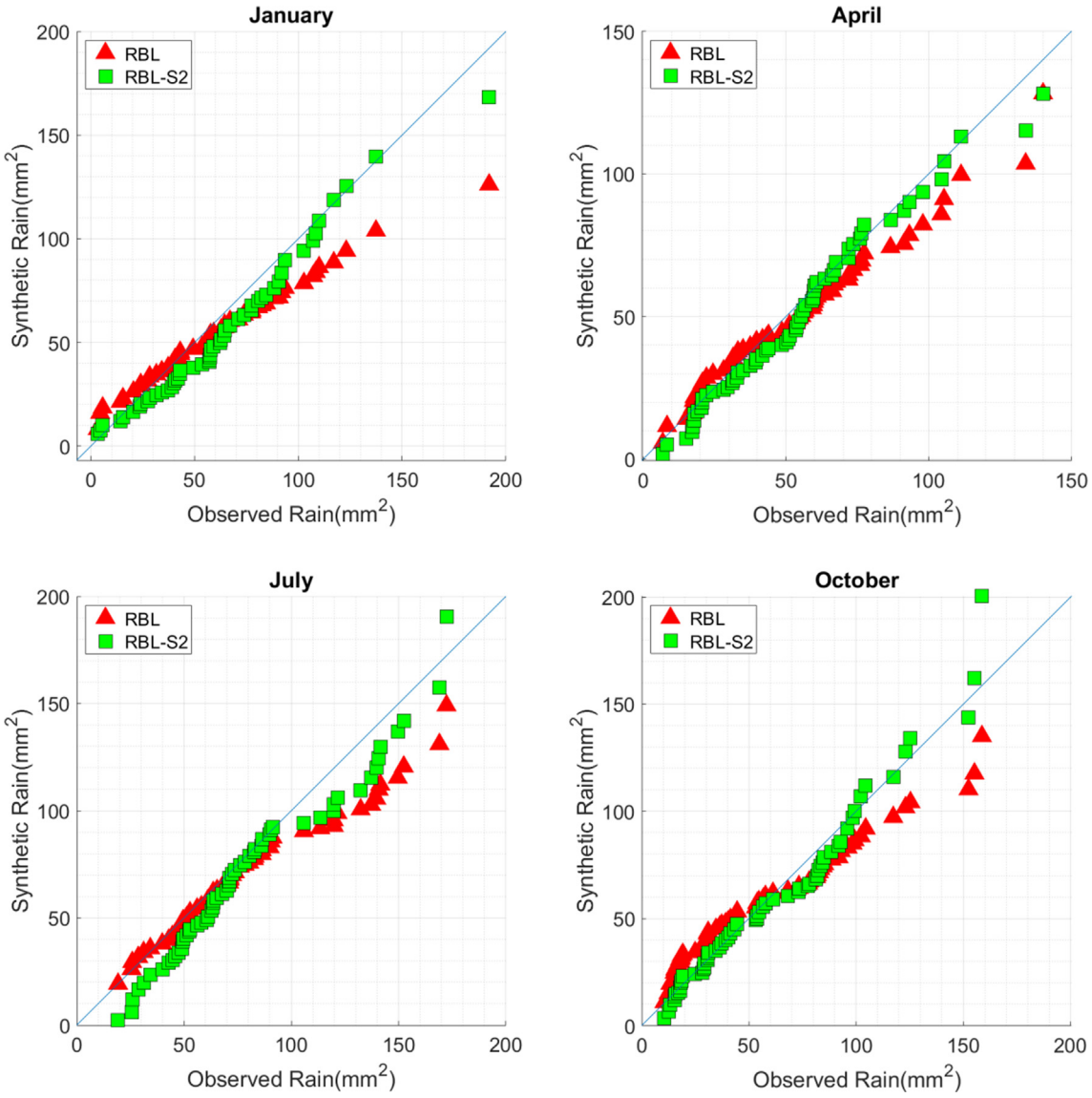

Fig. 9. Comparison of the observed rainfall quantiles (x) and the synthetic rainfall quantiles (y) for January, April, July, and October monthly rainfall.

at the timescale between 1 day and 6 months; and the RBL-S2 model further reduces the underestimation of the moments at the timescale greater than 6 months. Note that our model does not have an explicit algorithm to account for the third and the fourth moments at the time scale greater than 1 day, but these higher-order moments were well reproduced. This result implies that the mechanistic approach of incorporating rainfall memory such as the algorithms used in this study (rainstorm and monthly rainfall shuffling) can be a good alternative of the methods primarily focusing on statistical aspect of rainfall (Dimitriadis and Koutsoyiannis, 2018; Tsoukalas et al., 2018).

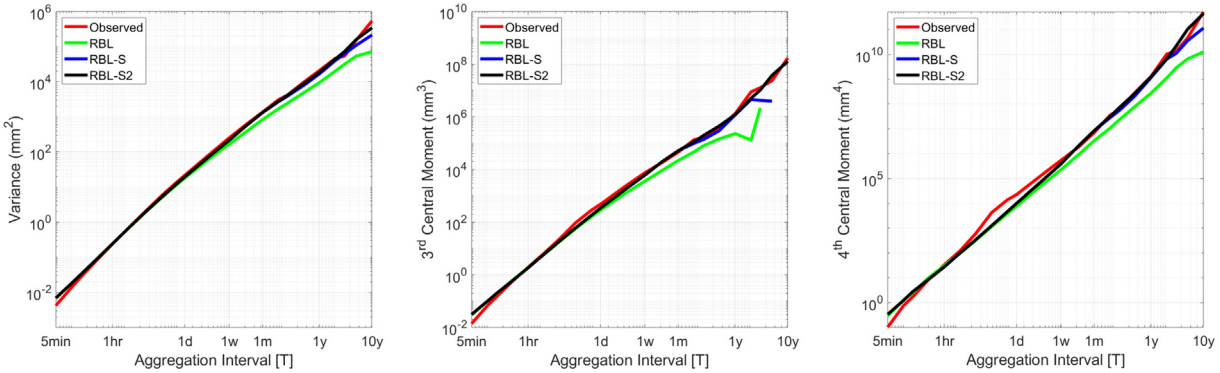

Fig. 10. Variance, third-, and fourth-central moment of observed and the synthetic rainfall across the timescales ranging from 5 min to 2 years. The results based on the RBL, RBL-S, and RBL-S2 models are shown.

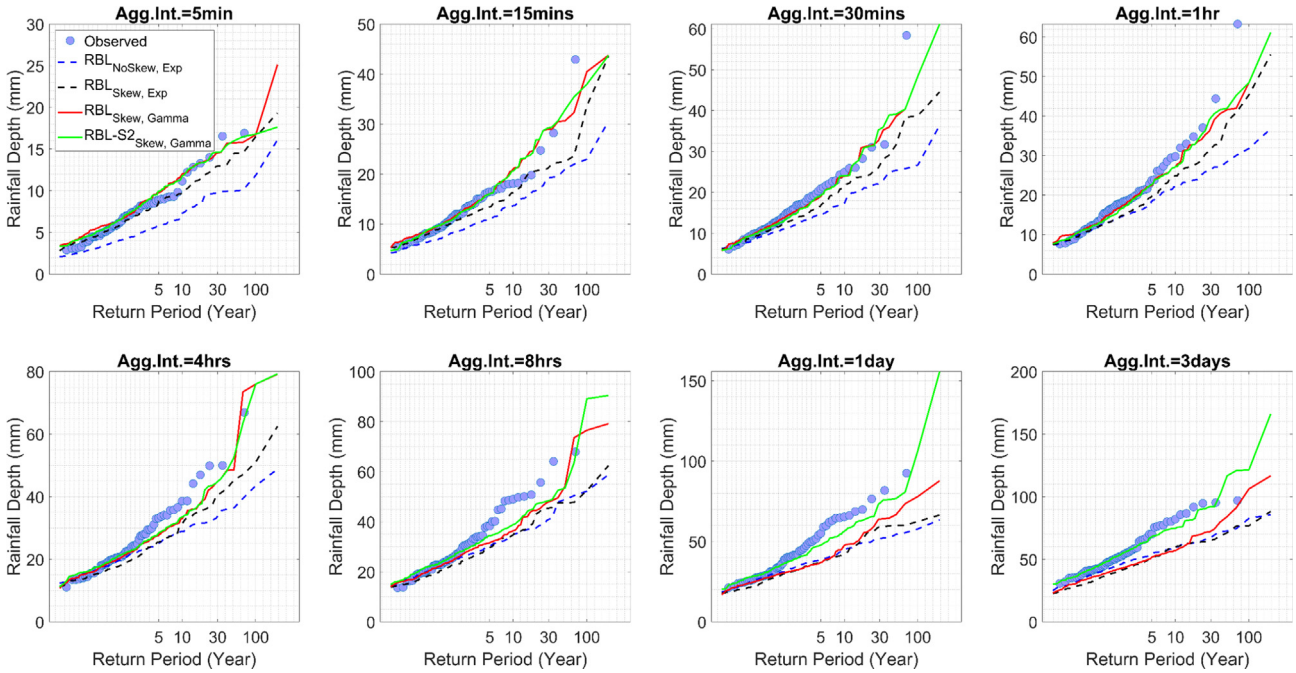

Fig. 11. The relationship between the annual maximum rainfall depth and the recurrence interval at the timescales between 5 min and 3 days. The blue dots represent the observed data for the whole observation period (1930–1999), The red and green solid lines represent the synthetic rainfall generated by the RBL model and the RBL-S2 model of this study. The black dashed line represents the synthetic rainfall generated by the RBL model based on the exponential rain cell intensity and with skewness included in the parameter calibration. The blue dashed line represents the synthetic rainfall based on the exponential rain cell intensity and no consideration of skewness in the parameter calibration. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3.5. Extreme values

Fig. 11 shows the relationship between annual maximum rainfall depths and recurrence intervals for both observed and synthetic rainfall. The x-axis was scaled based on the Gumbel transformation yielding the reduced variate. The blue dots represent the observed rainfall and the red and the green solid line represents the synthetic rainfall generated by the RBL and the RBL-S2 model. Both models successfully reproduce the observed extreme values without a significant trend of over- or underestimation at sub-hourly timescales. This is a significant improvement compared to the previous studies which found that the Poisson cluster rainfall models tend to systematically underestimate the extreme values. They attributed the causes to the parsimonious nature of the model (Kim et al., 2013, Park et al., 2019), the model calibration scheme in which skewness of the rainfall depth distribution is not considered (Cowpertwait, 1998; Kaczmarska et al., 2014; Onof and Wang, 2019), and the intrinsic limitation of the exponential distribution from which rain cell intensity values are drawn (Onof and Wang, 2019). The latter study found that the calibration scheme significantly affects the reproduction of extreme values, and suggested considering cell depth distributions other than the exponential that cannot reflect the heavy-tail of the observed rainfall distribution that is best modelled by sub-exponential distribution (Papalexiou et al., 2018; Nerantzaki and Papalexiou, 2019).

As opposed to the RBL model of this study that considered the skewness in the calibration process and the Gamma rain cell distribution, the model based on the exponential rain cell distribution with no consideration of skewness (blue dashed lines in Fig. 11) underestimated the 30-year rainfall by 38 percent and 41 percent at the 5 min and 1 h timescale, respectively. The one that considered the skewness but based on the exponential rain cell intensity (black dashed lines in Fig. 11) underestimated the same values by 18 percent and 25 percent, respectively.

The extreme values at timescale between 4 h and 1 day were underestimated by all models at the range of the recurrence interval between 5 years and 30 years. This trend of underestimation was reduced at the recurrence intervals exceeding 30 years. This is associated with the fundamental model structure of Poisson cluster models (See Fig. 1a). Indeed, first, note that the average duration of the rainstorms according to the model structure ranges between 1.9 h and 4.5 h according to the equations derived by Onof (2003) and the parameter values in Table 1.

For a rainstorm to reproduce extreme rainfall at timescales that are much finer than the rainstorm duration (e.g. 5 min through 1 h), it takes at most a couple of rain cells with very high intensity to overlap with each other. However, at timescales greater than this, it takes consecutive rainstorms to contain several rain cells with very high intensity, which happens with a low probability.

While the RBL model systematically underestimates the extreme rainfall at timescales of one day and more, the RBL-S2 model significantly eliminates this underestimation. The reason is as follows: First, note that the average inter-arrival time of rainstorms (λ^{-1}) ranges between 1.2 days and 2.7 days according to Table 1. Therefore, it is probable that time windows exceeding 1 day are likely to contain more than one rainstorm, so at this coarse timescale, the extreme rainfall depth is likely to be represented by more than a single storm. But the RBL model is less likely to have consecutive large rainstorms because the rainstorms are independent according to the model fundamental structure (See Section 2.2.1). On the contrary, the RBL-S2 model has an algorithm to induce the extreme rainstorms to gather together in the time series and fit in the time window yielding the extreme rainfall close to the observed one. This finding is also consistent with the results of Serinaldi and Kilsby (2016, 2018), which showed that the extreme rainfall is strongly associated with the spatio-temporal dependence structure of rainfall.

Fig. 12 compares the past-168 h (i.e. 7 days) rainfall of the annual maximum rainfall of the observed (x) and synthetic (y) rainfall. For convenience, we call this value the “P7 rainfall”. This value is important

Table 1
Parameters of the RBL model for the calibration period (1930–1964).

Month	λ	ν	α	ι	ϕ	κ	ω	Of (Eq. 1)
1	0.001294	0.004906	1.0650	0.0259	0.000142	0.02771	2.9804	3.9014
2	0.001129	0.005375	1.1289	0.0278	0.000109	0.02125	0.6163	2.2788
3	0.001482	0.013439	1.0958	0.0449	0.000425	0.03662	0.7988	3.0768
4	0.000960	0.003882	1.0510	0.0485	0.000081	0.01000	0.5993	3.8287
5	0.001343	0.005118	1.0390	0.0395	0.000324	0.03772	0.1980	3.8751
6	0.001819	0.065766	1.2424	0.0101	0.004000	1.84105	0.0100	5.9534
7	0.001693	0.031617	1.1690	0.0100	0.002943	1.39931	0.0129	5.9653
8	0.000717	1.000000	1.1447	1.0001	0.004000	0.03517	0.4096	3.2342
9	0.000585	0.452496	1.0485	1.0005	0.001000	0.01000	1.0179	6.9052
10	0.001153	0.009338	1.1275	0.0547	0.000221	0.02268	0.7556	3.0301
11	0.000814	0.001761	1.0459	0.0254	0.000029	0.01000	1.5975	9.1794
12	0.001359	0.010000	1.0788	0.0397	0.000241	0.03020	0.5082	4.4151

for the continuous hydrologic modelling studies in which antecedent moisture condition before the extreme events significantly influences the peak flow values. Several studies showed that the extreme rainfall does not always lead to the extreme flow discharge because of the varying antecedent soil moisture conditions (Briaud et al., 2009; Verhoest et al., 2010; Camici et al., 2011).

The RBL model systematically underestimated the P7 rainfall at the timescales between 5 min and 1 day. At the 3 day timescale, the value was well reproduced. The RBL-S2 model reduces the degree of underestimation of the P7 rainfall. This is also associated with the storm shuffling algorithm making similarly large storms gather together. It also suggests that the observed extreme rainfall events tend to occur during wet atmospheric and land surface conditions.

4. Conclusion

The rainfall we observe today is a result of an extremely complex set of interactions between a range of natural and anthropogenic causal factors that have occurred over a long time. Therefore, it is extremely difficult to precisely identify and quantify the rainfall persistence or "memory". For example, soil moistures do not necessarily lead to rainfall, but they are transferred into the atmosphere to make it unstable, which may or may not lead to rainfall (Seneviratne et al., 2010). In addition, the evapotranspired water vapor travels across continents (Ralph et al., 2006; van der Ent and Savenije, 2011) making the identification even more difficult when analyzed from an Eulerian viewpoint. Furthermore, an analysis based on thousands of stations of annual precipitation did not show any strong signs of long-term persistence (Iliopoulou et al., 2018; Tyralis et al., 2018).

For this reason, the implications of rainfall memory or persistence in practical applications have received relatively less attention. As Eq. (1) suggests, the rainfall memory causes the large rainstorms to cluster together with large rainstorms and the small rainstorms with small rainstorms. This entails the occurrence of very large or small rainfall depths at coarser timescales which govern the design of the hydrologic system and hydraulic structures. Therefore, a good rainfall model must correctly reproduce the rainfall memory and the corresponding temporal correlation structure at a wide range of timescales.

This study proposed a stochastic rainfall model with algorithms designed to reflect the rainfall memory existing at different timescales. In this approach, first, a series of rainstorms are generated based on the traditional Poisson cluster rainfall model. Second, the generated rainstorms are rearranged so that rainstorms with similar depth cluster together. Third, this rainfall time series is rearranged again at the monthly timescale to reflect the rainfall correlation at timescales equal to and coarser than a month. The suggested model was validated using 69 years of 5-minute rainfall data observed at Bochum, Germany. The model successfully reproduced the mean, variance, correlation

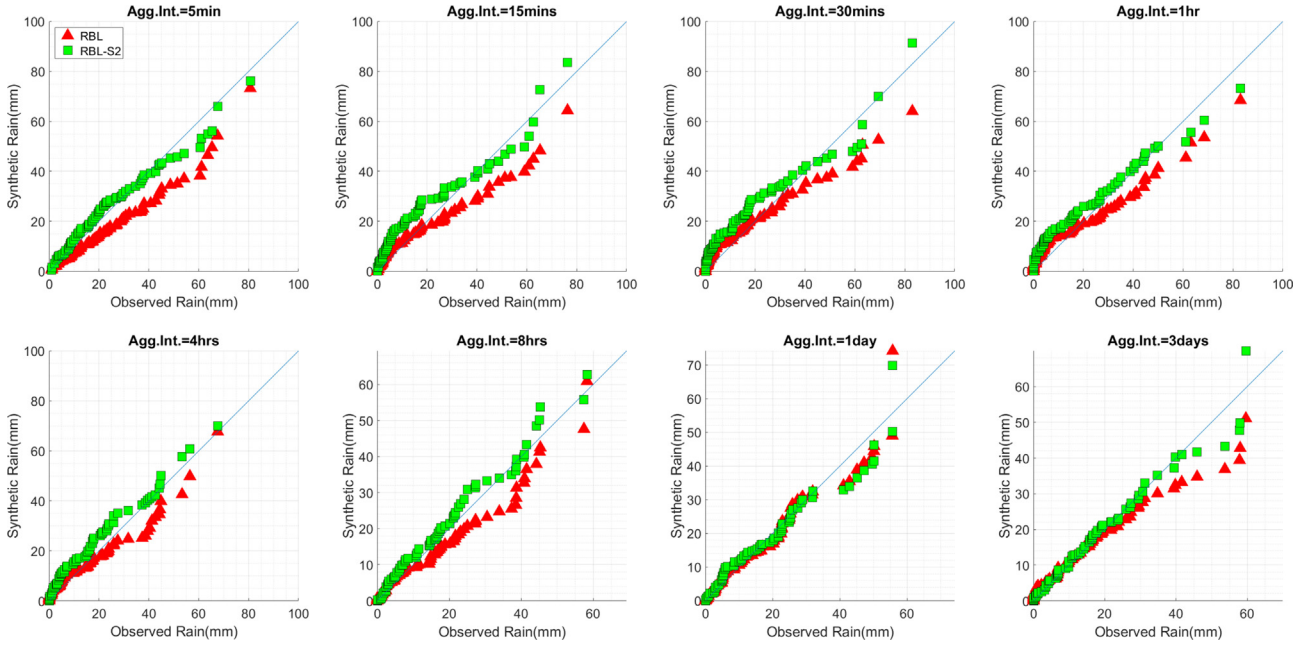

Fig. 12. The Quantile-Quantile plot of the past-168 h (7 days) rainfall of the annual maximum rainfall based on the RBL model and the RBL-S2 model.

structure and skewness of rainfall depths, the proportion of wet/dry periods, as well as the extreme values at timescales from 5 min to a decade. On the other hand, the traditional Poisson cluster rainfall model performed well in terms of all these statistics simultaneously only for timescales not exceeding the inter-storm arrival time (approximately 1 to 3 days). The suggested model reproduced well the past-7-day rainfall before an extreme rainfall event that the traditional model systematically underestimated. The difference in the performance of the two models shows the importance of designing stochastic rainfall models to include rainfall memory at a large range of timescales.

The strength of the suggested model from a practical viewpoint is that it can be applied to provide the input rainfall data not only to a wide range of modelling studies addressing, for example, urban flood, landslides, and droughts but also to the studies assessing the compound impacts of the disasters simultaneously occurring at different timescales (Chen et al., 2011). We expect that the model will gather more attention as the hydrologic societies started to recognize the hydrologic, human, and environmental systems from a holistic viewpoint and interpreting them based on continuous and dynamic simulations (Wagener et al., 2010, Kim et al., 2018).

Software Kit

The software kit that implements the methodology of this study can be downloaded at the following website: https://sites.google.com/site/hihydrology/projects

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This study was supported by the Basic Research Laboratory Program (NRF 2015-041523, 50 % grant) and the Basic Science Research Program (NRF 2018R1B5A2089503, 50% grant) through the NRF of Korea funded by the Ministry of Education. We truly appreciate the constructive reviews from Dr. Simon-Michael Papalexiou and Dr. Panayiotis Dimitriadis.

References

Afifi, T., Milan, A., Etzold, B., Schraven, B., Rademacher-Schulz, C., Sakdapolrak, P., Warner, K., 2016. Human mobility in response to rainfall variability: Opportunities for migration as a successful adaptation strategy in eight case studies. Migr. Develop. 5 (2), 254–274.

Anh, D., Kim, D., Kim, S., Park, J., 2020. Determination of flood-inducing rainfall and runoff for highly urbanized area based on high-resolution radar-gauge composite rainfall data and flooded area GIS Data. J. Hydrol. 584 (5). https://doi.org/10.1016/j.jhydrol.2020.124704.

Audet, C., Dennis Jr, J.E., 2002. Analysis of generalized pattern searches. SIAM J. Optim. 13 (3), 889–903.

Ayoub, A.T., 1999. Land degradation, rainfall variability and food production in the Sahelian zone of the Sudan. Land Degrad. Dev. 10 (5), 489–500.

Barbier, B., Yacouba, H., Karambiri, H., Zoromé, M., Somé, B., 2009. Human vulnerability to climate variability in the Sahel: farmers' adaptation strategies in northern Burkina Faso. Environ. Manage. 43 (5), 790–803.

Beecham, S., Rashid, M., Chowdhury, R.K., 2014. Statistical downscaling of multi-site daily rainfall in a South Australian catchment using a Generalized Linear Model. Int. J. Climatol. 34 (14), 3654–3670.

Berkelhammer, M., Sinha, A., Mudelsee, M., Cheng, H., Edwards, R.L., Cannariato, K., 2010. Persistent multidecadal power of the Indian Summer Monsoon. Earth Planet. Sci. Lett. 290 (1–2), 166–172.

Bernardara, P., De Michele, C., Rosso, R., 2007. A simple model of rain in time: An alternating renewal process of wet and dry states with a fractional (non-Gaussian) rain intensity. Atmos. Res. 84 (4), 291–301.

Burton, A., Kilsby, C.G., Fowler, H.J., Cowpertwait, P.S.P., O'connell, P.E., 2008. RainSim: A spatial–temporal stochastic rainfall modelling system. Environ. Modell. Software 23 (12), 1356–1369.

Briaud, J. L., Govindasamy, A. V., Kim, D., Gardoni, P., & Olivera, F. (2009). Simplified method for estimating scour at bridges (No. FHWA/TX-09/0-5505-1). Texas. Dept. of Transportation. Research and Technology Implementation Office.

Camici, S., Tarpanelli, A., Brocca, L., Melone, F., Moramarco, T., 2011. Design soil moisture estimation by comparing continuous and storm-based rainfall-runoff modeling. Water Resour. Res. 47 (5).

Carvalho, L.M., Jones, C., Liebmann, B., 2004. The South Atlantic convergence zone: Intensity, form, persistence, and relationships with intraseasonal to interannual activity and extreme rainfall. J. Clim. 17 (1), 88–108.

Chen, Y.S., Kuo, Y.S., Lai, W.C., Tsai, Y.J., Lee, S.P., Chen, K.T., Shieh, C.L., 2011. Reflection of typhoon morakot – The challenge of compound disaster simulation. J. Mountain Sci. 8 (4), 571–581.

Chandler, R.E., Wheater, H.S., 2002. Analysis of rainfall variability using generalized linear models: A case study from the west of Ireland. Water Resour. Res. 38 (10), 10–11.

Cho, H., Kim, D., Olivera, F., Guikema, S.D., 2011. Enhanced speciation in particle swarm optimization for multi-modal problems. Eur. J. Oper. Res. 213 (1), 15–23.

Coe, R., Stern, R.D., 1982. Fitting models to daily rainfall data. J. Appl. Meteorol. 21 (7), 1024–1031.

Cowpertwait, P.S., 1991. Further developments of the Neyman-Scott clustered point process for modeling rainfall. Water Resour. Res. 27 (7), 1431–1438.

Cowpertwait, P.S., 1995. A generalized spatial-temporal model of rainfall based on a clustered point process. Proc. R. Soc. Lon. Ser. A Mathemat. Phys. Sci. 450 (1938), 163–175.

Cowpertwait, P.S., 1998. A Poisson-cluster model of rainfall: some high-order moments and extreme values. Proc. R. Soc. Lon. Ser. A: Mathemat. Phys. Eng. Sci. 454 (1971), 885–898.

Cowpertwait, P., Isham, V., Onof, C., 2007. Point process models of rainfall: Developments for fine-scale structure. Proc. Royal Soc. A: Math. Phys. Eng. Sci. 463 (2086), 2569–2587.

David, J.Y., Sangwan, N., Sung, K., Chen, X., Merwade, V., 2017. Incorporating institutions and collective action into a sociohydrological model of flood resilience. Water Resour. Res. 53 (2), 1336–1353.

Deidda, R., 2000. Rainfall downscaling in a space - time multifractal framework. Water Resour. Res. 36 (7), 1779–1794.

De Lima, M.I.P., Grasman, J., 1999. Multifractal analysis of 15-min and daily rainfall from a semi-arid region in Portugal. J. Hydrol. 220 (1–2), 1–11.

Dimitriadis, P., Koutsoyiannis, D., 2015. Climacogram versus autocovariance and power spectrum in stochastic modelling for Markovian and Hurst-Kolmogorov processes. Stoch. Env. Res. Risk Assess. 29 (6), 1649–1669. https://doi.org/10.1007/s00477-015-1023-7.

Dimitriadis, P., Koutsoyiannis, D., 2018. Stochastic synthesis approximating any process dependence and distribution. Stochas. Environmental Res. Risk Assess. 32 (6), 1493–1515. https://doi.org/10.1007/s00477-018-1540-2.

Dodangeh, E., Shahedi, K., Solaimani, K., Kossieris, P., 2017. Usability of the BLRP model for hydrological applications in arid and semi-arid regions with limited precipitation data. Model. Earth Syst. Environ. 3 (2), 539–555.

Gommes, R., Petrassi, F., 1996. Rainfall variability and drought in sub-Saharan Africa. SD dimensions, FAO.

Eltahir, E.A., 1998. A soil moisture–rainfall feedback mechanism: 1. Theory and observations. Water Resourc. Res. 34 (4), 765–776.

Entekhabi, D., Rodriguez-Iturbe, I., Castelli, F., 1996. Mutual interaction of soil moisture state and atmospheric processes. J. Hydrol. 184 (1–2), 3–17.

Fatichi, S., Ivanov, V.Y., Caporali, E., 2011. Simulation of future climate scenarios with a weather generator. Adv. Water Res. 34 (4), 448–467.

Gyasi-Agyei, Y., Willgoose, G.R., 1997. A hybrid model for point rainfall modeling. Water Resour. Res. 33 (7), 1699–1706.

Haan, C.T., Allen, D.M., Street, J.O., 1976. A Markov chain model of daily rainfall. Water Resour. Res. 12 (3), 443–449.

Iliopoulou, T., Papalexiou, S.M., Markonis, Y., Koutsoyiannis, D., 2018. Revisiting long-range dependence in annual precipitation. J. Hydrol. 556, 891–900. https://doi.org/10.1016/j.jhydrol.2016.04.015.

Iliopoulou, T., Koutsoyiannis, D., 2019. Revealing hidden persistence in maximum rainfall records. Hydrol. Sci. J. 64 (14), 1673–1689. https://doi.org/10.1080/02626667.2019.1657578.

Jesus, J., Chandler, R.E., 2011. Estimating functions and the generalized method of moments. Interface Focus 1 (6), 871–885.

Kaczmarska, J., Isham, V., Onof, C., 2014. Point process models for fine-resolution rainfall. Hydrol. Sci. J. 59 (11), 1972–1991.

Kim, D., Olivera, F., Cho, H., Socolofsky, S.A., 2013. Regionalization of the modified bartlett-lewis rectangular pulse stochastic rainfall model. Terrestrial, Atmos. Oceanic Sci. 24 (3).

Kim, D., Kwon, H.H., Lee, S.O., Kim, S., 2016. Regionalization of the Modified Bartlett-Lewis rectangular pulse stochastic rainfall model across the Korean Peninsula. J. Hydro-environ. Res. 11, 123–137.

Kim, D., Kwon, H., Giustolisi, O., Savic, D., 2018. Current water challenges require holistic and global solutions. J. Hydroinf. 20 (3), 533–534.

Kim, J., Lee, J., Kim, D., Kang, B., 2019. The role of rainfall spatial variability in estimating areal reduction factors. J. Hydrol. 568, 416–426.

Kim, Y., Wang, G., 2007. Impact of initial soil moisture anomalies on subsequent precipitation over North America in the coupled land–atmosphere model CAM3–CLM3. J. Hydrometeorol. 8 (3), 513–533.

Kim, D., Olivera, F., 2011. Relative importance of the different rainfall statistics in the calibration of stochastic rainfall generation models. J. Hydrol. Eng. 17 (3), 368–376.

Khaliq, M.N., Cunnane, C., 1996. Modelling point rainfall occurrences with the modified Bartlett-Lewis rectangular pulses model. J. Hydrol. 180 (1–4), 109–138.

Kossieris, P., Efstratiadis, A., Tsoukalas, I., & Koutsoyiannis, D. (2015). Assessing the performance of Bartlett-Lewis model on the simulation of Athens rainfall. European

Geosciences Union General Assembly.
Koutsoyiannis, D., 2006. Nonstationarity versus scaling in hydrology. J. Hydrol. 324 (1–4), 239–254.
Koutsoyiannis, D., Onof, C., 2001. Rainfall disaggregation using adjusting procedures on a Poisson cluster model. J. Hydrol. 246 (1–4), 109–122.
Kovats, R.S., Bouma, M.J., Hajat, S., Worrall, E., Haines, A., 2003. El Niño and health. Lancet 362 (9394), 1481–1489.
Kossieris, P., Makropoulos, C., Onof, C., Koutsoyiannis, D., 2018. A rainfall disaggregation scheme for sub-hourly time scales: Coupling a Bartlett-Lewis based model with adjusting procedures. J. Hydrol. 556, 980–992.
Kwon, H.H., Lall, U., Obeysekera, J., 2009. Simulation of daily rainfall scenarios with interannual and multidecadal climate cycles for South Florida. Stoch. Env. Res. Risk Assess. 23 (7), 879–896.
Langousis, A., Veneziano, D., Furcolo, P., Lepore, C., 2009. Multifractal rainfall extremes: theoretical analysis and practical estimation. Chaos Solitons Fractals 39 (3), 1182–1194.
Lee, D., Lee, J., Kim, D., 2014. Applicability of a space-time rainfall downscaling algorithm based on multifractal framework in modeling heavy rainfall events in Korean Peninsula. J. Korea Water Resourc. Assoc. 47 (9), 839–852.
Lovejoy and Schertzer, D. (1991). Multifractal analysis techniques and the rain and cloud fields from 10 − 3 to 10 6 m. In Non-Linear Variability in Geophysics (pp. 111-144). Springer, Dordrecht.
Marani, M., 2005). Non-power-law-scale properties of rainfall in space and time. Water Resourc. Res. 41 (8).
Marshall, J.D., Shimada, B.W., Jaffe, P.R., 2000. Effect of temporal variability in infiltration on contaminant transport in the unsaturated zone. J. Contam. Hydrol. 46 (1–2), 151–161.
Mooley, D.A., Parthasarathy, B., 1984. Fluctuations in all-India summer monsoon rainfall during 1871–1978. Clim. Change 6 (3), 287–301.
Menabde, M., Sivapalan, M., 2000. Modeling of rainfall time series and extremes using bounded random cascades and levy-stable distributions. Water Resour. Res. 36 (11), 3293–3300.
Milan, A., Ruano, S., 2014. Rainfall variability, food insecurity and migration in Cabricán, Guatemala. Climate Develop. 6 (1), 61–68.
Mishra, A., Desai, V., 2005. Drought forecasting using stochastic models. Stoch. Env. Res. Risk A. 19, 326–339.
Modarres, R., Ouarda, T.B., 2014. Modeling the relationship between climate oscillations and drought by a multivariate GARCH model. Water Resour. Res. 50, 601–618.
Nerantzaki, S.D., Papalexiou, S.M., 2019. Tails of extremes: Advancing a graphical method and harnessing big data to assess precipitation extremes. Adv. Water Resour. 134, 103448. https://doi.org/10.1016/j.advwatres.2019.103448.
Oh, M., Lee, D., Kwon, H., Kim, D., 2016. Development of flood inundation area GIS database for Samsung-1 drainage sector, Seoul, Korea. J. Korea Water Resourc. Assoc. 49 (12), 981–993.
Olsson, Jonas, Burlando, Paolo, 2002. Reproduction of temporal scaling by a rectangular pulses rainfall model. Hydrol. Process. 16 (3), 611–630.
Onof, C. (2003). DEFRA Project: Improved methods for national spatial-temporal rainfall and evaporation modelling for BSM, Internal Report, No. 8, Mathematical expressions of generalized moments used in single-site rainfall models, December 16, 2003.
Onof, C., Chandler, R.E., Kakou, A., Northrop, P., Wheater, H.S., Isham, V., 2000. Rainfall modelling using Poisson-cluster processes: A review of developments. Stoch. Env. Res. Risk Assess. 14 (6), 384–411.
Onof, C., Wheater, H.S., 1993. Modelling of British rainfall using a random parameter Bartlett-Lewis rectangular pulse model. J. Hydrol. 149 (1–4), 67–95.
Onof, C., Wang, L.-P., 2019. Modelling rainfall with a Bartlett-Lewis process: New developments. Hydrol. Earth Syst. Sci. Discuss. https://doi.org/10.5194/hess-2019-406, in review.
Papalexiou, S.M., 2018. Unified theory for stochastic modelling of hydroclimatic processes: Preserving marginal distributions, correlation structures, and intermittency. Adv. Water Resour. 115, 234–252. https://doi.org/10.1016/j.advwatres.2018.02.013.
Papalexiou, S.M., AghaKouchak, A., Foufoula-Georgiou, E., 2018. A diagnostic framework for understanding climatology of tails of hourly precipitation extremes in the United States. Water Resour. Res. https://doi.org/10.1029/2018WR022732.
Park, J., Onof, C., Kim, D., 2019. A hybrid stochastic rainfall model that reproduces some important rainfall characteristics at hourly to yearly timescales. Hydrol. Earth Syst. Sci. 23, 989–1014. https://doi.org/10.5194/hess-23-989-2019.
Patz, J.A., Campbell-Lendrum, D., Holloway, T., Foley, J.A., 2005. Impact of regional climate change on human health. Nature 438, 310. https://doi.org/10.1038/nature04188.
Ralph, F.M., Neiman, P.J., Wick, G.A., Gutman, S.I., Dettinger, M.D., Cayan, D.R., White, A.B., 2006. Flooding on California's Russian River: Role of atmospheric rivers. Geophys. Res. Lett. 33 (13).
Rodriguez-Iturbe, I., Cox, D.R., Isham, V., 1987. Some models for rainfall based on stochastic point processes. Proceedings of the Royal Society of London. A. Math. Phys. Sci. 410 (1839), 269–288.
Rodriguez-Iturbe, I., Cox, D.R., Isham, V., 1988. A point process model for rainfall: Further developments. Proceedings of the Royal Society of London. A. Math. Phys. Sci. 417 (1853), 283–298.
Serinaldi, F., Kilsby, C.G., 2016. Understanding persistence to avoid underestimation of collective flood risk. Water 8, 152. https://doi.org/10.3390/w8040152.
Serinaldi, F., Kilsby, C.G., 2018. Unsurprising surprises: The frequency of record-breaking and overthreshold hydrological extremes under spatial and temporal dependence. Water Resour. Res. 54, 6460–6487. https://doi.org/10.1029/2018WR023055.
Seneviratne, S.I., Corti, T., Davin, E.L., Hirschi, M., Jaeger, E.B., Lehner, I., Teuling, A.J., 2010. Investigating soil moisture–climate interactions in a changing climate: A review. Earth Sci. Rev. 99 (3–4), 125–161.
Singh, V.P., 1997. Effect of spatial and temporal variability in rainfall and watershed characteristics on stream flow hydrograph. Hydrol. Process. 11 (12), 1649–1669.
Singh, S.V., Kripalani, R.H., Shaha, P., Ismail, P.M.M., Dahale, S.D., 1981. Persistence in daily and 5-day summer monsoon rainfall over India. Arch. Meteorol. Geophys. Bioclimatol. Series A 30 (3), 261–277.
Song, H., Kim, D., Kim, B., Hwang, S., Kim, T., 2014. Hydrological assessment of multifractal space-time rainfall downscaling model: Focusing on application to the upstream watershed of Chungju dam. J. Korea Water Resourc. Assoc. 47 (10), 959–972.
Tucker, G.E., Bras, R.L., 2000. A stochastic approach to modeling the role of rainfall variability in drainage basin evolution. Water Resour. Res. 36 (7), 1953–1964.
Tsoukalas, I., Makropoulos, C., Koutsoyiannis, D., 2018. Simulation of stochastic processes exhibiting any-range dependence and arbitrary marginal distributions. Water Resour. Res. 54 (11), 9484–9513. https://doi.org/10.1029/2017WR022462.
Tyralis, H., Dimitriadis, P., Koutsoyiannis, D., O'Connell, P.E., Tzouka, K., Iliopoulou, T., 2018. On the long-range dependence properties of annual precipitation using a global network of instrumental measurements. Adv. Water Resour. 111, 301–318. https://doi.org/10.1016/j.advwatres.2017.11.010.
Van der Ent, R.J., Savenije, H.H.G., 2011. Length and time scales of atmospheric moisture recycling. Atmos. Chem. Phys. 11 (5), 1853–1863.
Verhoest, N., Troch, P.A., De Troch, F.P., 1997. On the applicability of Bartlett-Lewis rectangular pulses models in the modeling of design storms at a point. J. Hydrol. 202 (1–4), 108–120.
Veneziano, D., Langousis, A., Furcolo, P., 2006. Multifractality and rainfall extremes: a review. Water Resour. Res. 42 (6).
Verhoest, N.E., Vandenberghe, S., Cabus, P., Onof, C., Meca-Figueras, T., Jameleddine, S., 2010. Are stochastic point rainfall models able to preserve extreme flood statistics? Hydrol. Process. 24 (23), 3439–3445.
Wagener, T., Sivapalan, M., Troch, P.A., McGlynn, B.L., Harman, C.J., Gupta, H.V., Wilson, J.S., 2010. The future of hydrology: An evolving science for a changing world. Water Resour. Res. 46 (5).
Yoo, J., Kim, D., Kim, H., Kim, T., 2016. Application of copula functions to construct confidence intervals of bivariate drought frequency curve. J. Hydro-Environ. Res. 11, 113–122.