Application of the stochastic model for temporal rainfall disaggregation for hydrological studies in north western England

M. Abdellatif, W. Atherton and R. Alkhaddar

ABSTRACT

Assessment of climate change on any hydrological system requires higher temporal resolution at hourly or less in terms of time scale. This paper implements the Bartlett–Lewis Rectangular Pulses (BLRP) model coupled with a proportional adjusting procedure to disaggregate daily rainfall to hourly rainfall in order to demonstrate the reliability of this method. Three stations in northwestern England have been selected that represent different climates in the region. Parameters estimation of the BLRP model has been performed under different levels of hourly rainfall aggregation for a combination of rainfall statistics. The HYETOS model, which applies BLRP, reproduced standard statistics such as mean, variance, Lag-1, autocorrelation as well as dry proportions. Moreover, the model was proven to have the capability to disaggregate the rainfall extremes. The fitted BLRP model could then be used to disaggregate future daily rainfall in order to investigate the climate change impact of different rainfall intensities.

Key words | Bartlett–Lewis rectangular pulses model, daily rainfall, disaggregation, extremes, HYETOS

M. Abdellatif (corresponding author)
W. Atherton
R. Alkhaddar
School of the Built Environment,
Liverpool John Moores University,
The Peter Jost Centre,
Byrom Street,
Liverpool, L3 3AF,
UK
E-mail: *m.e.abdellatif@2010.ljmu.ac.uk*

INTRODUCTION

Stochastic models have a wide range of application in fields such as flood risk estimation, river flow forecasting and water resources engineering. Many studies use the stochastic model for rainfall disaggregation from daily to sub-daily level for the purpose of flood design, although some of these methods are limited in their applications to specific rainfall conditions (Debele *et al.* 2007).

Some of the hourly rainfall modelling in the literature makes use of the dry–wet structure of rainfall. In these approaches the rainfall occurrence and depth process are described separately and then both are superimposed to form the overall rainfall model (e.g. Eagleson 1978; Istok & Boersma 1989; Acreman 1990; Koutsoyiannis & Xanthopoulos 1990). In recent approaches, both occurrence and depth rainfall processes are combined together and parameters estimation is performed from the hourly and aggregated hourly rainfall data (e.g. Khaliq & Cunnane 1996). Examples of these approaches are the Random Cascade Models, the Bartlett–Lewis Rectangular Pulses (BLRP) or Neyman–Scott rectangular pulse models based on point process theory (Rodriguez-Iturbe *et al.* 1987a), as well as three-state continuous Markov model of Hutchinson (1990).

The BLRP model in its original form was applied strictly as a rainfall simulator (Pui *et al.* 2009). It has since been modified with an appropriate adjusting procedure to be applied in rainfall disaggregation.

Many studies have examined the ability of the BLRP and how realistically it simulates rainfall variability and extremes. Rodriguez-Iturbe *et al.* (1987b) found that the BLRP model is able to reproduce some of the rainfall depth statistics and perform relatively well in regard to the extreme values of rainfall for different periods of aggregation, but they are less able to preserve the proportion dry at the level of aggregation of 1 hour when applied to Denver rainfall as stated by Khaliq & Cunnane (1996). When examining the method for two case studies in the

doi: 10.2166/hydro.2012.090

UK and the USA, Koutsoyiannis & Onof (2001) found that the BLRP model performed well in preserving the most important statistical properties of the rainfall process. Pui *et al.* (2009) found that the BLRP model performed better on average than the cascade models for Sydney rainfall with a slightly inflated reproduction of dry proportions at an hourly scale. However, complications were encountered during the parameter estimation stage and the choice of statistics remains subjective. Hanaish *et al.* (2011) tested the modified BLRP model for a station in Malaysia and compared the disaggregated synthetic rainfall data with the observed and expected data. They found that the mean values for the three rainfall data are quite close; however, the synthetic data are quite different from the observed and expected depths when comparison was based on autocorrelation and standard deviation. Moreover, Hanaish *et al.* (2011) compared the extreme values series and a poor fit (i.e. underestimation) was found at 1, 6 and 12 hour levels of aggregation.

This paper addresses the issue of temporal rainfall disaggregation by applying combined hydroinformatics data-driven tools of the BLRP stochastic model and HYETOS model (Koutsoyiannis & Onof 2000) on three selected case studies in northwestern England. The stations are located in one region in the UK but have different climates (i.e. different rainfall intensities). The objective is to explore the effects of rainfall depth on the ability of the model to reproduce hourly statistics and rainfall extremes. The resulting model can then be used to examine the design storm of any hydrological system under current and future climate change conditions. It demonstrates an example of how hydroinformatics modelling technology can be used to resolve some problems in water and environmental fields.

CASE STUDY AND DATA COLLECTION

Three stations have been selected in northwestern England (NW) to represent three catchments or drainage areas with various climatic conditions: Tower Wood in the north (TW), Worthington in the middle (WN), and Worleston (WR) in the south (see Figure 1). The exposure of the NW region to westerly maritime air masses and the presence of extensive areas of high ground, especially in the TW area,

Figure 1 | Location of studied stations.

mean that the region is considered as one of the wettest places in the UK.

Hourly rainfall data for the selected stations were obtained from the Environment Agency for England and Wales for the period 2000–2008 in TW, for period 1999–2010 in WN and for the period 1991–2010 in WR. The hourly data obtained (which is measured) was extracted from a data series that contains missing data sets. As calibration of the temporal disaggregation model requires only a small amount of hourly rainfall data at the same locations as the daily data being disaggregated, the small missing data sets are not going to affect the model calibration. Daily rainfall data were obtained from the daily recording rainfall gauges in the three locations.

BLRP MODEL STRUCTURE

The BLRP model assumes that the storm origins (T) occur according to the Poisson Process with rate λ and the cell origins (t) arrive following Poisson Process with rate β, as depicted in Figure 2. The process of considering new cell origins terminates after a time (S), which is exponentially

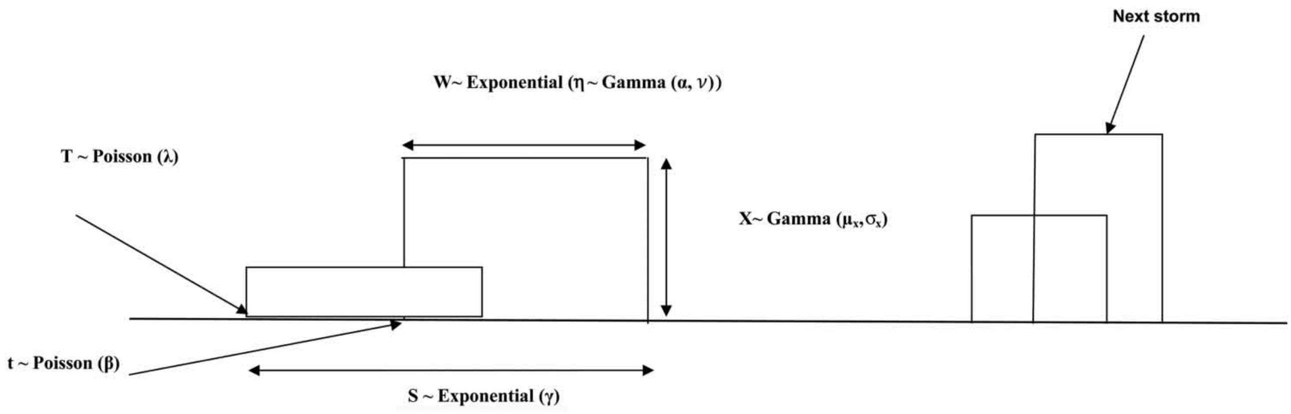

Figure 2 | Sketch of the Modified Bartlett–Lewis Rectangular Pulses Model (reproduced from Hanaish *et al*. (2011)).

distributed with parameter γ. The distribution of the duration of the cells (W) is an exponential distribution with parameter η. The cell has constant intensity with a specified distribution (Koutsoyiannis & Onof 2001).

The above description for the original BLRP model assumes all parameters are constant, however in the modified model; the parameter η varies randomly from one storm to another with a gamma distribution of shape parameter, α and scale parameter v. Rodriguez-Iturbe *et al.* (1987a, b) applied the old version of BLRP model to a single month of Denver data (Onof & Wheater 1993), where he found a major deficiency to reproduce the proportion of dry periods but Rodriguez-Iturbe *et al.* (1988) and Entekhabi *et al.* (1989) overcome this problem by this randomisation. Parameters β and γ are re-parameterised so that $K = \beta/\eta$ and $\phi = \gamma/\eta$. Moreover it is desirable as it will be more consistent with the nature of the storm, which should be varying in the characteristics from storm to storm rather than consider them constant. Essentially, the effect of these parameters is that all storms have a common structure, but distinct storms occur on different (random) timescales.

The constant cell depth (X) is exponentially distributed with parameter $1/\mu$. Alternatively, it can be chosen as a two parameter gamma with mean, μ_x and standard deviation, σ_x.

The number of cells per storm has a geometric distribution of mean, $\mu_c = 1 + k/\phi$ (Khaliq & Cunnane 1996; Koutsoyiannis & Onof 2001). In this paper, seven parameters have been used, namely λ, K, ϕ, α, v, μ_x, σ_x, as shown in Figure 2.

ESTIMATION OF BLRP MODEL PARAMETERS

The parameters for the BLRP were estimated on a monthly basis, assuming local stationary within the month according to procedures set out in Koutsoyiannis & Onof (2000). The combination of all months cannot be considered in one model (Khaliq & Cunnane 1996). Each month has different characteristics and rainfall pattern, so it was found to be better to model each month separately.

In general, the main fitting techniques used are moment, likelihood and Bayesian. The latter two are based on likelihood function, which cannot be obtained for models based on the Poisson cluster (Hanaish *et al.* 2011). Therefore the method of generalised moments is the best choice to fit the BLRP model. The equations of the BLRP model are solved by equating the statistical feature of the historical rainfall with the theoretical one according to minimising sum of weighted squared errors criterion. That results in a set of non-linear equations that, in the current study, has been solved by employing the Newton optimisation algorithm. So the theory behind the method of generalised moment is to find the parameters (θ) which minimise the objective function given by Equation (1): minimising sum of weighted squared errors:

$$S(\theta) = \sum_i^n w_i[T_i - \tau_i(\theta)]^2 \quad (1)$$

where $\theta = (\theta_1, \theta_2 \ldots \theta_n)$ is a vector of unknown parameters of the rainfall statistics used for calibration, w_i r represents the weights, T_i the historical statistics obtained for the data and $\tau_i(\theta)$ the theoretical model statistics as a function of θ.

The sensitivity of the statistical moments used to calibrate the model is still under investigation, however, for the seven parameters at least seven equations are needed, which should be obtained from historical data. The constant cell depth has been selected to be a two parameter gamma distributed with mean, μ_x and standard deviation, σ_x, which have been chosen to be equally in this study for the modified version of BLRP model ($\mu_x = \sigma_x$).

The equation of the modified BLRP model relates the statistical properties of the rainfall to the seven BLRP parameters as given in Equations (2)–(5) below.

The first and the second order properties of the aggregated process, Y_i^h, are reproduced here from Rodriguez-Iturbe *et al.* (1988), where Y is the cumulative amount of rainfall in the ith arbitrary interval of length, h, hours.

Average rainfall depth at time scale h:

$$E\left(Y_i^h\right) = \frac{\lambda h v \mu_x \mu_c}{\alpha - 1} \quad (2)$$

where: $\mu_c - 1 + \dfrac{k}{\phi}$

Variance of rainfall depth at time scale h:

$$\begin{aligned}\mathrm{Var}\left(Y_i^h\right) = {} & 2A_1\left[(\alpha-3)hv^{2-\alpha} - v^{3-\alpha} + (v+h)^{3-\alpha}\right] \\ & - 2A_2\left[\varnothing(\alpha-3)hv^{2-\alpha} - v^{3-\alpha} + (v+\varnothing h)^{3-\alpha}\right]\end{aligned} \quad (3)$$

and for $k \geq 1$

Covariance of rainfall depth at time scale h:

$$\begin{aligned}&\mathrm{cov}\left(Y_i^n, Y_{i+k}^n\right) = \\ &A_1\{[v+(k+1)h]^{3-\alpha} - 2(v+kh)^{3-\alpha} + [v+(k-1)h]^{3-\alpha}\} \\ &\quad -A_2\{[v+(K+1)\phi h]^{3-\alpha} - 2(v+k\varnothing h)^{3-\alpha} \\ &\quad +[v+(k-1)\phi h]^{3-\alpha}\}\end{aligned} \quad (4)$$

where,

$$A_1 = \frac{\lambda \mu_c \ v^\alpha}{(\alpha-1)(\alpha-2)(\alpha-3)}\left[E(X^2) + \frac{K\varnothing\mu_x^2}{\varnothing - 1}\right]$$

and

$$A_2 = \frac{\lambda \mu_c k \mu_X^2 \ v^\alpha}{\varnothing^2(\varnothing^2-1)(\alpha-1)(\alpha-2)(\alpha-3)}$$

The probability that a period of length h is dry is given by:

$$P\left(Y_i^h = 0\right) = \exp\left\{-\lambda h - \lambda\mu_T + \lambda G_P^*(0,0)\frac{\varnothing + k\left[\dfrac{v}{v+(k+\varnothing)h}\right]^{\alpha-1}}{\varnothing + k}\right\} \quad (5)$$

where μ_T is the average mean storm duration whose exact expression is given in Rodriguez-Iturbe *et al.* (1987a) and can be approximated to a third degree in k and ϕ by:

$$\begin{aligned}\mu_T \approx {} & \frac{v}{\phi(\alpha-1)}\left[1 + \varnothing(k+\varnothing) - \frac{1}{4}\varnothing(k+\varnothing)(k+4\varnothing)\right. \\ & \left. + \frac{1}{72}\varnothing(k+\varnothing)(4k^2 + 27k\varnothing + 72\varnothing^2\right]\end{aligned}$$

The above approximation for μ_T has not been considered accurate enough and was corrected by Onof & Wheater (1993) as follows:

$$\begin{aligned}\mu_T \approx {} & \frac{v}{\phi(\alpha-1)}\left[1 + \varnothing\left(k + \frac{\varnothing}{2}\right) - \frac{1}{4}\varnothing\left(5\,\varnothing\,k + K^2 + 2\,\varnothing^2\right)\right. \\ & \left. + \frac{1}{72}\varnothing(4\,K^3 + 31\,K^2\,\varnothing + 99k\varnothing^2 + 36\varnothing^2)\right]\end{aligned}$$

The expression for $G_P^*(0,0)$ by Rodriguez-Iturbe *et al.* (1987a):

$$G_P^*(0,0) \approx \frac{v}{\varnothing(\alpha-1)}\left(1 - k - \varnothing + \frac{3}{2}k\varnothing + \varnothing^2 + \frac{1}{2}k^2\right)$$

SINGLE SITE DISAGGREGATION USING HYETOS

An application was developed by Koutsoyiannis & Onof (2000, 2001) and resulted in a computer program called

HYETOS, which can be easily applied at any location as long as it is provided with a minimal amount of data that can support the parameter estimation. The higher and lower levels of time scales in this application are daily and hourly, respectively. These levels are found to be the most suitable ones for typical hydrological applications. The HYETOS model itself does not include a module to estimate model parameters; therefore, another module (Solver in Excel) has been used in this study to fit the model. Once the BLRP parameters are obtained, they are subsequently used with the HYETOS model to obtain a single site disaggregated rainfall series (Koutsoyiannis & Onof 2000).

The procedure that HYETOS follows to disaggregate the daily rainfall at a single site into hourly data using the BLRP model as a background stochastic model can be described as follows. Four levels of repetition scheme used so as to optimise computer time. Level 0, BLRP model run several times ($t > L + 1$) until a sequence of exactly L wet days is generated (which is selected in the current study based on the maximum observed wet spell and should not exceed 12). Different sequences separated by at least one dry day can be assumed independent. Then (level 1), the intensities of all cells and storms are generated and the resulting daily depths are calculated. For each cluster of the wet days the generated synthetic daily depths (which simulated at hourly scale) should matched the sequence of original daily totals with a tolerance distance, d defined as:

$$d = \left[\sum_{i=1}^{L} \ln^2 \left(\frac{N_i + C}{\tilde{N}_i + C} \right) \right] \tag{6}$$

where N_i and $\tilde{N}_i$ are, the original and simulated daily totals at the rain gauge station, with L as the length of the sequence of wet days and c small constant (0.1 mm). If d is greater than acceptable limit, which is selected in this study to be 0.1 as suitable value (various values around 0.1 were tested), regenerate the intestines of cells (level 1 repetitions) without modifying the time location of the storms and their cells. If however after a large number of level 1 repetitions, the distance still higher than 0.1, in this case discard the arrangement of the storm and cells, which is not consistent with the original one and generate new one which is level 2 repetitions. In the case of the very long sequences where it is impossible to obtain the criteria of d less than 0.1, the sequence should divided into subsequent randomly and each treated separately from others (level 3 repetitions).

For all levels, the study used a total number of repetitions is 5,000 as maximum. A correction procedure, referred to as proportional adjustment, to make the generated hourly series fully consistent with given daily totals, is then applied based on Koutsoyiannis & Onof (2001). The proportional adjusting procedure modifies the initially generated values to get the modified values according to:

$$X_s = \tilde{X}_s \left(\frac{N}{\sum_1^{24} \tilde{X}_s} \right) \qquad s = 1, 2 \ldots 24 \tag{7}$$

where N is the daily depth to be disaggregated.

RESULTS AND DISCUSSION

Performance of the temporal disaggregation model

The sensitivity of the number of observations used to calibrate the model affects the final performance of the model (Segond *et al.* 2006). This suggests that a model needs to be calibrated with a large number of data. Wheater *et al.* (2005) suggested that at least 15 years of rainfall data should be used in order to obtain a model that would have good overall performance. However, that amount of data was not available for all stations in the current study. Instead, the disaggregating model calibrated for each station was found to have good performance with 9, 12, and 18 years in TW, WN, and WR stations, respectively. Although the model for TW station has been calibrated with only 9 years of data, but resulted in a good fit model compared with the other two sites which have more than 9 years used in the fitting. This could be attributed to the increased number of sequences of wet day.

This reflects that the BLRP model does not need only more data to have a well calibrated model, but rather it also depends on the statistics and type of algorithm used to estimate the parameter. Different combinations of rainfall statistic for 1 hour and multiples thereof have been tried in order to obtain the best fit across the three catchments. The suitable moments (statistics) found are: 1-hour mean, 1-hour variance, 1-hour lag-1 autocovariance, 1-hour proportion dry, 6-hour variance and 24-hour proportion dry for TW and WR stations. For WN station, all the above statistics were selected additional to 24-hour variance, which was found to give good results when added. The difference between the two sets of statistics used in fitting the models as explained earlier due to equality of parameters μ_x and σ_x (leads to six parameters, so six statistics needed as in TW and WR but in WN seven statistics employed to improve the fitting). However, the assumption of randomisation still exists for the three stations which associated only with parameter η.

The values for these parameters were chosen for a certain tolerance limit in order for the solution to converge (see Table 1).

Table 1 | Boundary constraints of parameters used at the three sites

Parameter	Unit	Lower constraint	Upper constraint
λ	d^{-1}	0.000001	99
$\kappa = \beta/\eta$	(–)	0.000001	99
$\varphi = \gamma/\eta$	(–)	0.000001	99
α	(–)	0.000001	99
v	d	0.000001	99
μ_X	mm d^{-1}	0.000001	99
σ_X	mm d^{-1}	0.000001	99

The weights used in these models have been taken as having a value of 1. This decision came after carrying out numerous trials for different statistics computed from the historical data using different values for the weights which were found to be unsatisfactory for the model fit. A summary of the modified BLRP model parameters obtained is presented in Tables 2–4 using Equations (2)–(5) with the original approximation of the mean storm duration by Rodriguez-Iturbe *et al.* (1987a). Parameter estimates vary considerably over the months, indicating that separate monthly simulations are required to accurately reproduce actual rainfall distributions in the area.

Moreover another fit was performed using the same set of Equations (2)–(5) with the approximation of the mean storm duration by Onof & Wheater (1993), which corrected an error in the original one. Table 5 shows an example of this fit for TW employing the corrected approximation, however no significant difference to the results was found. There was difference in the optimisation method as conjugate gradient yield better results than the Newton method which was used in the original approximation. So it is recommended for any future application of BLRP model to consider the correct approximation as it is not obvious that the difference between the two approximations will always have negligible effects.

Plots of the model fit (BLRP) and that produced from HYETOS against the observed data for each month in

Table 2 | Estimated BLRP parameter *s* for the master gauge of TW station using Rodriguez-Iturbe *et al.* (1987a) approximation of the mean storm duration

Month	λ (mm d^{-1})	$\kappa = \beta/\eta$	$\varphi = \gamma/\eta$	α	*v* (d)	μ_X (mm d^{-1})	σ_X (mm d^{-1})	OF
1	0.48063	0.34870	0.06212	4.25885	0.21034	33.10716	33.10716	4.29×10^{-5}
2	0.58331	0.25759	0.10290	3.47918	0.19618	34.46851	34.46851	2.35×10^{-5}
3	0.71708	1.04539	0.11131	2.07677	0.02885	20.60905	20.60905	0.000131
4	0.68744	0.93569	0.08705	2.26942	0.02486	21.94154	21.94154	7.31×10^{-5}
5	0.69741	0.72757	0.07488	2.65820	0.02469	27.61123	27.61123	0.000286
6	0.65496	0.59934	0.13289	5.05571	0.12527	30.67456	30.67456	4.01×10^{-5}
7	0.67621	0.17180	0.01646	3.29331	0.01107	99.00000	99.00000	0.001053
8	0.77587	0.40874	0.06009	4.10360	0.04429	50.56184	50.56184	0.000163
9	0.65940	0.21706	0.07396	6.68564	0.24459	45.35329	45.35329	4.71×10^{-5}
10	0.75736	0.50493	0.07860	2.44029	0.03932	34.97688	34.97688	4.16×10^{-5}
11	1.06377	0.49538	0.06715	2.33371	0.02535	35.06887	35.06887	5.34×10^{-5}
12	0.82024	0.58854	0.07954	2.31739	0.04747	32.97733	32.97733	2.86×10^{-5}

Table 3 | Estimated BLRP parameters for the master gauge of WN station using Rodriguez-Iturbe *et al.* (1987a) approximation of the mean storm duration

Month	λ (mm d^{-1})	$\kappa = \beta/\eta$	$\varphi = \gamma/\eta$	α	v (d)	μ_X (mm d^{-1})	σ_X (mm d^{-1})	OF
1	0.94353	0.95068	0.11449	6.67558	0.08308	21.96963	21.96963	0.01678
2	1.02358	1.66562	0.10615	2.91308	0.02001	14.84352	14.84352	0.03928
3	0.68053	0.50863	0.05772	3.60237	0.03445	21.04118	21.04118	0.04827
4	0.73608	0.83627	0.07895	6.51238	0.05872	24.75399	24.75399	0.00355
5	0.66516	0.53682	0.03441	2.90789	0.00997	40.92153	40.92153	0.03622
6	0.42464	0.11639	0.02584	3.58376	0.04359	60.99989	60.99989	0.00942
7	0.66348	0.26703	0.05439	5.55417	0.05306	59.36534	59.36534	0.00244
8	0.89540	0.24281	0.05346	4.88165	0.02709	90.00382	90.00382	0.02361
9	0.72438	0.49474	0.06884	3.91303	0.03618	45.24926	45.24926	0.00585
10	1.07305	0.63846	0.05074	2.86708	0.01305	38.94103	38.94103	0.00438
11	1.31967	0.78842	0.08485	4.02959	0.03092	26.32452	26.32452	0.01104
12	1.14823	0.92793	0.08472	4.50687	0.03944	21.73045	21.73045	0.01430

Table 4 | Estimated BLRP parameters for the master gauge of WR station using Rodriguez-Iturbe *et al.* (1987a) approximation of the mean storm duration

Month	λ (mm d^{-1})	$\kappa = \beta/\eta$	$\varphi = \gamma/\eta$	α	v (d)	μ_X (mm d^{-1})	σ_X (mm d^{-1})	OF
1	0.94782	1.07250	0.09040	2.99318	0.01432	19.80680	19.80680	0.00284
2	0.79134	1.29547	0.11121	4.45990	0.03587	13.81512	13.81512	0.00159
3	0.62701	1.16025	0.09904	8.15786	0.07493	15.68464	15.68464	0.00131
4	0.60466	0.35884	0.06343	9.24289	0.13078	26.09495	26.09495	0.02655
5	0.56557	0.43836	0.03196	2.61501	0.00729	41.16541	41.16541	0.01098
6	0.58699	0.62010	0.09791	6.20092	0.06204	32.46399	32.46399	0.00029
7	0.54831	0.09508	0.03254	4.23765	0.03596	77.97608	77.97608	0.00052
8	0.59787	0.25854	0.05471	3.26056	0.01899	68.59553	68.59553	0.00045
9	0.55469	0.27013	0.07864	9.60892	0.18119	39.54048	39.54048	0.00941
10	0.82557	0.72706	0.15378	7.00000	0.10420	29.41798	29.41798	0.06440
11	1.02375	0.90255	0.20520	7.00000	0.12395	17.77258	17.77258	0.04228
12	0.86177	1.05417	0.12521	7.00000	0.07604	19.21809	19.21809	0.04636

respect of mean, standard deviation and proportion dry day, are shown in Figures 3–5.

The HYETOS model uses the full test mode with hourly historical data as input (the data are read from a file). This mode is appropriate for testing the entire model performance including the appropriateness of the Bartlett–Lewis model and its parameters and disaggregation model by comparing the original and disaggregated statistics.

All the models generally perform well, as they reproduce the mean exactly and fit the standard deviation well in the three stations of the study. The proportion of dry days is a very important property for hydrological applications and also useful to reserve as an important feature for subsequent model validation. The plots show that the simulated rainfall obtained by HYETOS slightly overestimates the proportion of dry in all stations. However, the expected values in respect to the fit perfectly match the observed ones in TW but are underestimated in the WN and WR stations.

Another important demonstration for the ability of the HYETOS model to disaggregate the daily rainfall to hourly under the full test mode employing the parameters estimated in Tables 2–4 is shown in terms of the obtained statistics and

Table 5 | Estimated BLRP parameter s for the master gauge of TW station using Onof & Wheater (1993) approximation of the mean storm duration

Month	λ (mm d^{-1})	$\kappa = \beta/\eta$	$\varphi = \gamma/\eta$	α	v (d)	μ_X (mm d^{-1})	σ_X (mm d^{-1})	OF
1	0.46725	0.33610	0.05852	4.18179	0.20615	33.26702	33.26702	0.00010
2	0.55207	0.23104	0.08775	3.41445	0.19284	34.79662	34.79662	0.00006
3	0.71651	1.02863	0.11134	2.07536	0.02915	20.69567	20.69567	0.00006
4	0.68729	0.92960	0.08723	2.26919	0.02504	21.96087	21.96087	0.00032
5	0.69733	0.72499	0.07531	2.66502	0.02502	27.58908	27.58908	0.00026
6	0.65300	0.57822	0.13006	4.95667	0.12323	30.87397	30.87397	0.00017
7	0.67624	0.17180	0.01646	3.29331	0.01107	99.00000	99.00000	0.00107
8	0.77571	0.40750	0.05998	4.09599	0.04425	50.56000	50.56000	4.94×10^{-5}
9	0.64193	0.20421	0.06785	6.39133	0.23257	45.59212	45.59212	1.46×10^{-5}
10	0.75579	0.49795	0.07782	2.43323	0.03926	35.05539	35.05539	0.000126
11	1.06267	0.49115	0.06684	2.33055	0.02538	35.10419	35.10419	0.000126
12	0.81671	0.57610	0.07818	2.30507	0.04717	33.13377	33.13377	0.000238

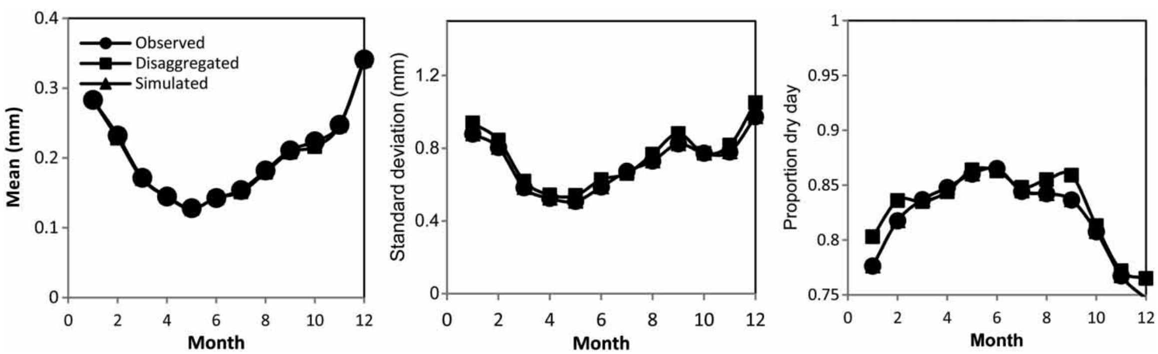

Figure 3 | Properties of hourly rainfall for TW single site. Circle represents the observed, square the disaggregated using HYETOS and triangle is simulated using BLRP.

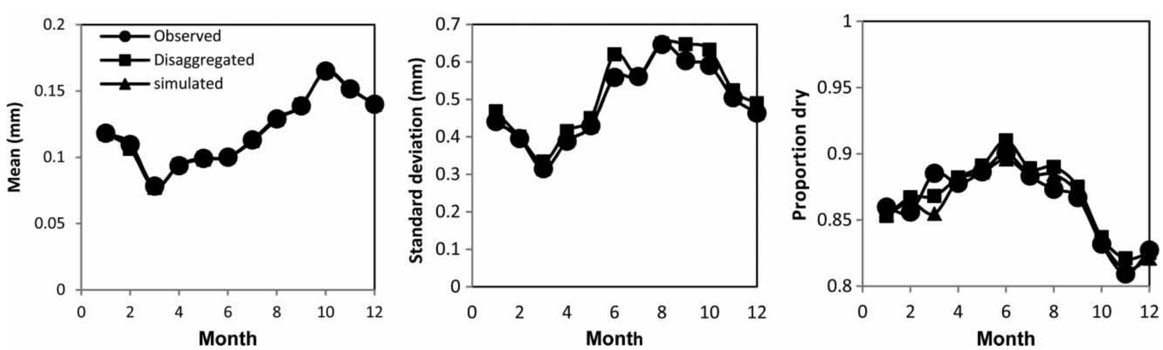

Figure 4 | Properties of hourly rainfall for WN single site. Circle represents the observed, square the disaggregated using HYETOS and triangle is the simulated using BLRP.

presented in Figures 6–8. These statistics can be used to assess the model's adequacy. The figures depict the goodness of fit test represented here by the computed lag-1 autocorrelation function and skewness between the theoretical and disaggregated hourly rainfall data at the three stations. The plots of the two statistics in these figures

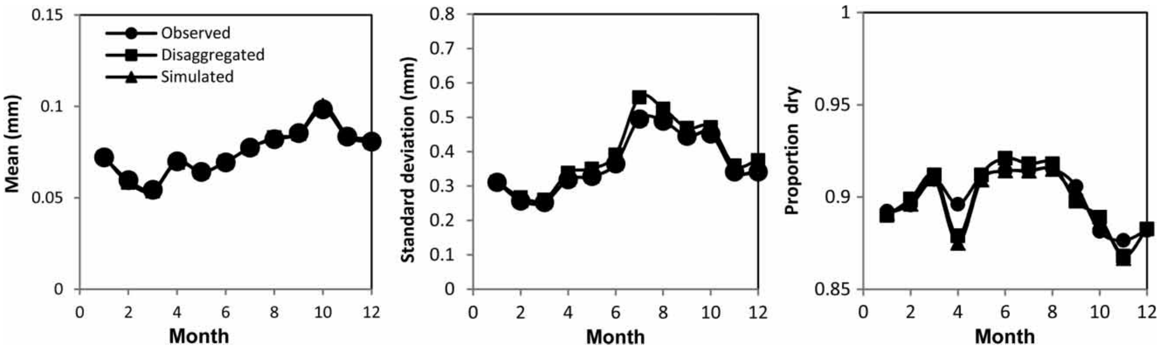

Figure 5 | Properties of hourly rainfall for WR single site. Circle represents the observed, square the disaggregated using HYETOS and triangle is the simulated using BLRP.

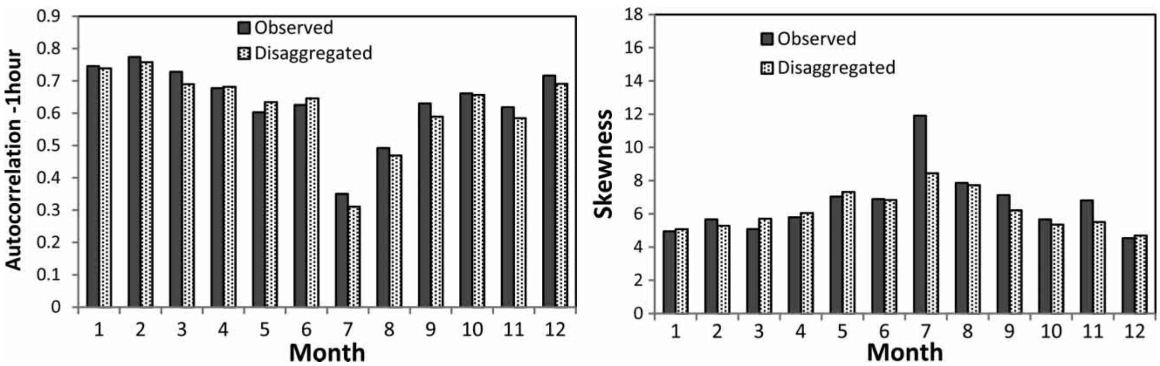

Figure 6 | Comparison of lag-1 autocorrelation (left) and skewness (right) of observed and disaggregated for the case study of TW station at hourly scale.

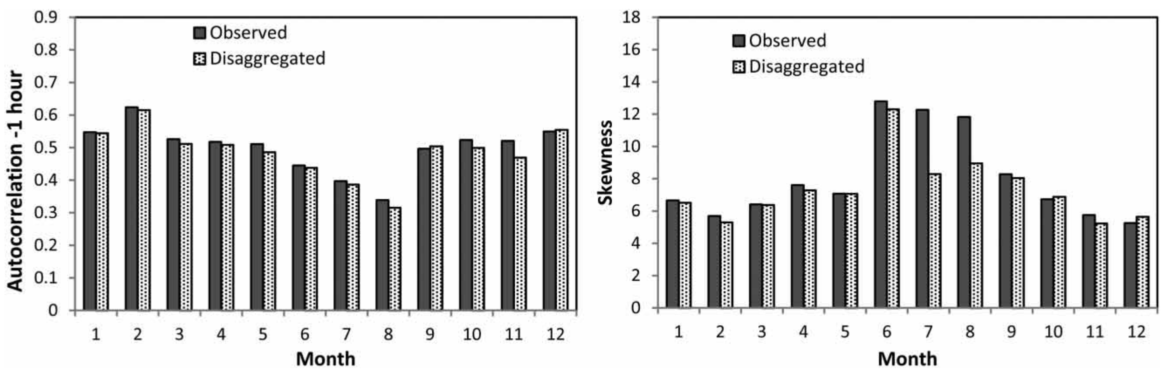

Figure 7 | Comparison of lag-1 autocorrelation (left) and skewness (right) of observed and disaggregated for the case study of WN station at hourly scale.

confirm that the HYETOS model performed very satisfactorily and produced acceptable results, especially for the lag-1 autocorrelation.

Nevertheless, looking at the model results in terms of skewness, there is a significant underestimation for the hourly rainfall skewness in the months of July and August at TW and WN stations, respectively. This can be attributed to the fact that summer months normally have the highest rainfall variability and skewness. Results for the WR station tend to be best in reproducing the skewness property, which

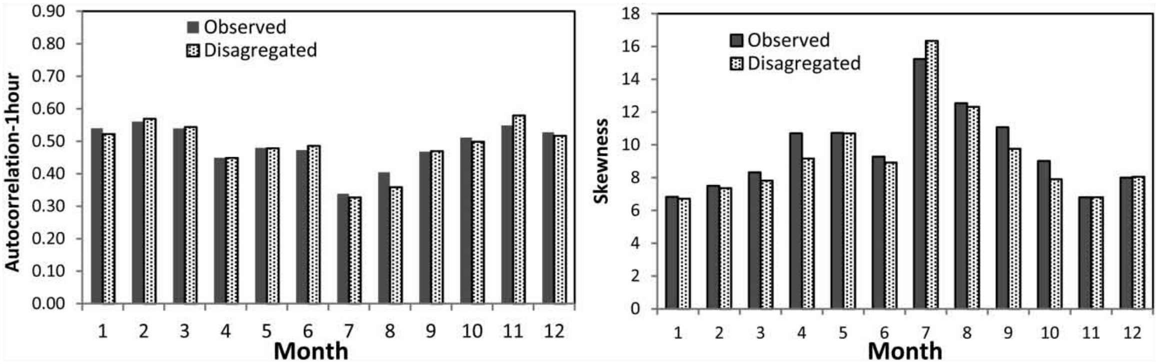

Figure 8 | Comparison of lag-1 autocorrelation (left) and skewness (right) of observed and disaggregated for the case study of WR station at hourly scale.

can be attributed to the nature of rainfall in the area, which is characterised by having lower intensity rainfall and shorter continuous wet days. WR station results are also a reflection of the effect of rainfall intensity on the disaggregation scheme.

The disaggregated results obtained through the full test mode of HYETOS when the hourly input aggregated to daily, which serve as an original series and then disaggregated producing another synthetic hourly series using BLRP model parameters. However, the operational mode of HYETOS can be used with daily input for model application (e.g. climate change studies) no means for testing and can also produce disaggregated rainfall.

Different levels of aggregation have been compared for the rainfall statistics simulated by BLRP model and disaggregated by HYETOS. The mean statistic for 6- and 24-hourly rainfall was preserved well for TW, WN and WR for both disaggregated and simulated rainfall (see Figures 9–11). Moreover, the randomised BLRP model brought an achievement for proportion dry at hourly and even for 24-hourly rainfall and slightly overestimated the 6-hourly aggregations for the three sites. Although the 6-hourly rainfall was not

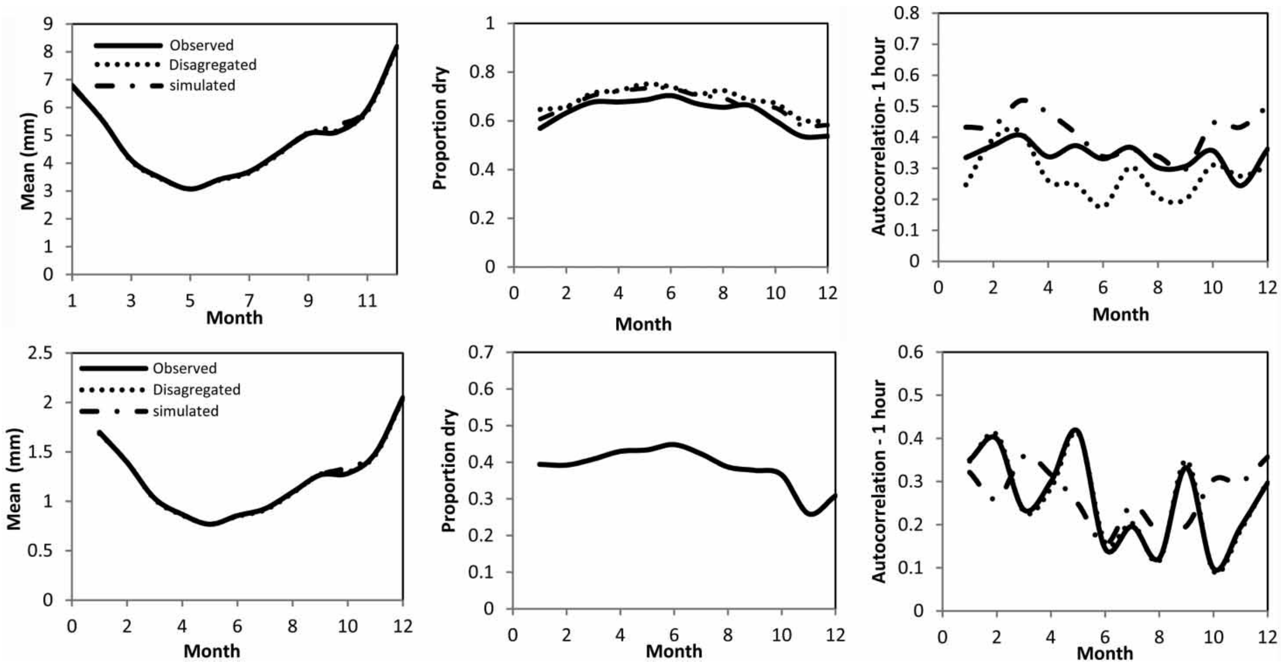

Figure 9 | Statistical properties of 6-hourly (top) and 24-hourly (bottom) rainfalls for TW single site for the observed, simulated by BLRP and disaggregated using HYETOS.

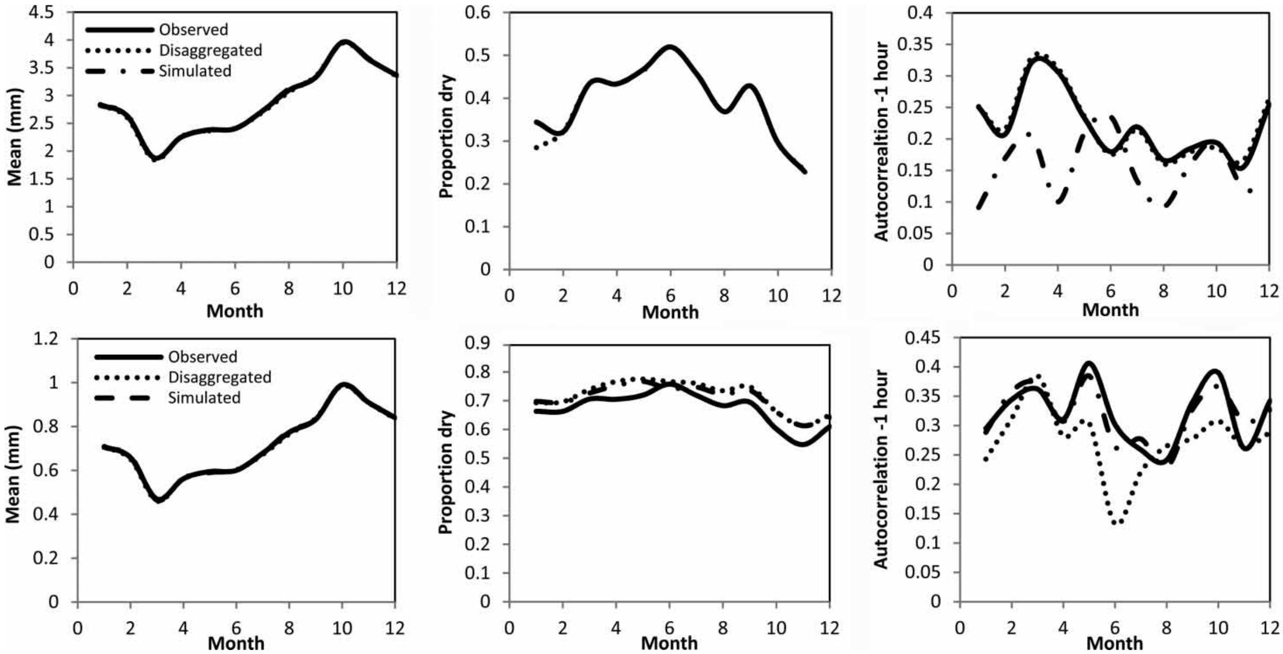

Figure 10 | Statistical properties of 6-hourly (top) and 24-hourly (bottom) rainfalls for WN single site for the observed, simulated by BLRP and disaggregated using HYETOS.

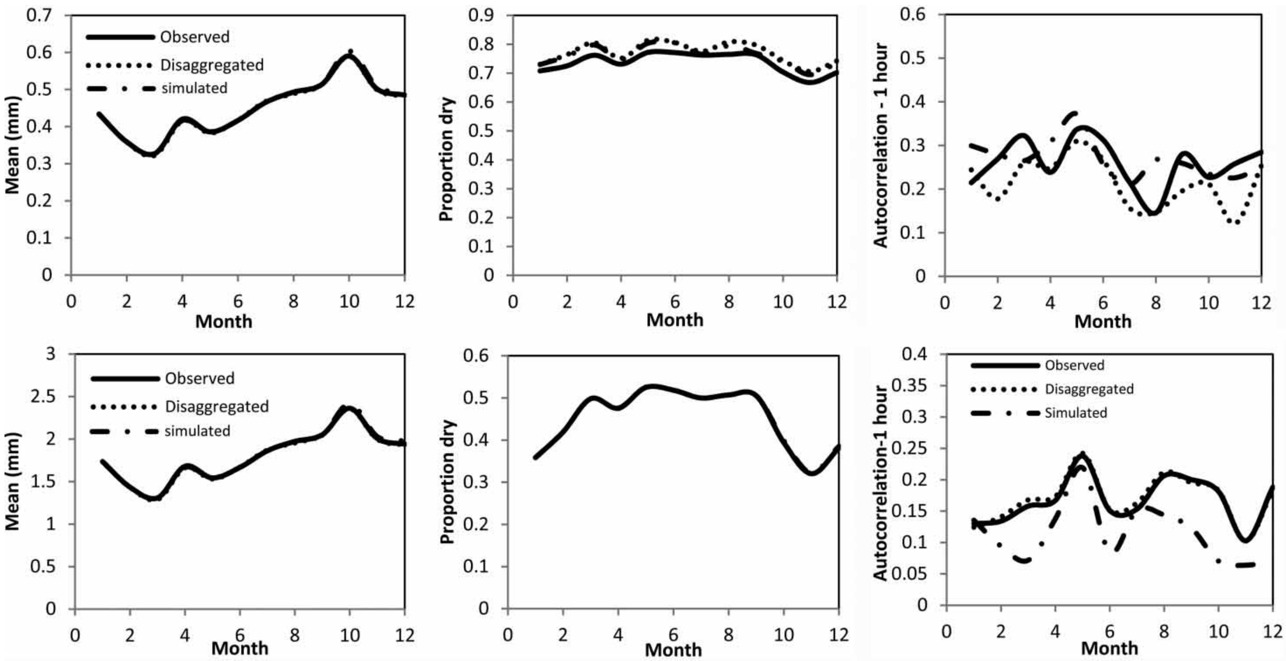

Figure 11 | Statistical properties of 6-hourly (top) and 24-hourly (bottom) rainfalls for WR single site for the observed, simulated by BLRP and disaggregated using HYETOS.

used in the fitting, but has shown a good agreement with the observed data as depicted by Figures 9–11.

For the disaggregated rainfall BLRP model has tendency to underestimate autocorrelation lag-1 at scale of 6 hour for all sites, but reproduced it well for 24-hour scale (Figures 9–11). Whereas for the simulated rainfall at 6-hourly aggregations, the autocorrelation lag-1 was overestimated in TW and WR but not in WN. This may be attributed to the different statistics used in fitting the model. However for 24-hourly scale rainfall BLRP model has found to underestimate the autocorrelation lag-1.

Effects of disaggregation on extreme rainfall

Extreme rainfall is considered one of the most important parameters used in the design of any hydrological system. So the ability of the HYETOS model to reproduce extreme values of rainfall has also been assessed in this study using a combined approach of Peak Over Threshold (POT) and Generalised Pareto Distribution (Gilleland *et al.* 2005). The combined approach, which uses the method of maximum likelihood to estimate the parameters of the Generalised Pareto Distribution, is programmed R Language and contained in software called extRemes. Figures 12–14 show plots of specified return period and corresponding computed hourly return level obtained by extRemes in the three studied stations. The plots demonstrate that HYETOS slightly overestimated the values of computed return levels; however, it was able to capture some features of the observed extremes at hourly scale. The thresholds selected to fit the observed and disaggregated models are not significantly different. For TW, the observed and disaggregated thresholds used were 1.3 and 2 mm, for WN they were 1.5 and 2 mm, and for WR they were 1.2 and 2 mm, respectively.

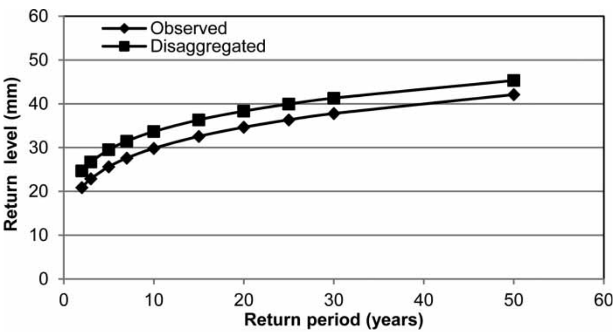

Figure 12 | Impact of the disaggregation scheme on extreme rainfall at TW.

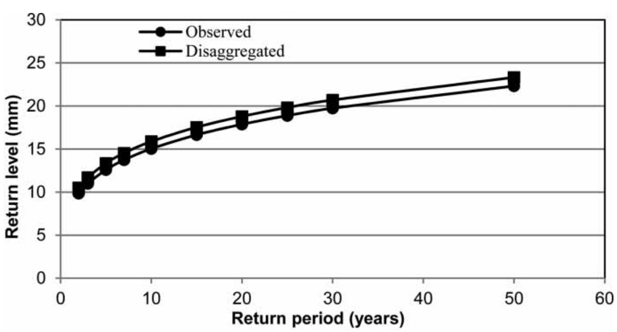

Figure 13 | Impact of the disaggregation scheme on extreme rainfall at WN.

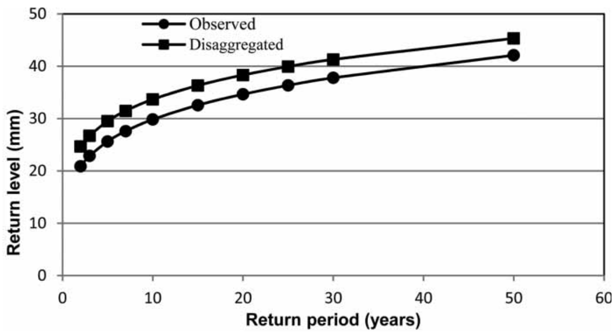

Figure 14 | Impact of the disaggregation schemes on extreme rainfall at WR.

CONCLUSIONS

In this paper, the performances of the combined stochastic rainfall model (BLRP) and the daily rainfall disaggregator (HYETOS) for reproducing hourly rainfall are evaluated. Historical rainfall data from three stations were used to assess the effect of different rainfall intensities on these models. Fitting of the BLRP model has been achieved by using a combination of different moments generated from the statistical properties of the historical rainfall data. Then, fitted parameters obtained from the BLRP model were used in the HYETOS software to disaggregate the daily rainfall to test the performance of the HYETOS model. The disaggregated rainfall series obtained by HYETOS was then used to reproduce the standard statistics and extreme values in each station. The results obtained have shown reasonable estimates for the standard statistics and extreme values under different climates and consequently, encourage use of the BLRP and HYETOS models in studies involving modelling of hydrological systems, especially for climate change studies. Moreover, the results also demonstrated that hydroinformatics modelling technology is very useful in resolving complex issues such as temporal disaggregation of rainfall.

REFERENCES

Acreman, M. C. 1990 A simple stochastic model of hourly rainfall for Farnborough, England. *Hydrol. Sci. J.* **35**, 119–148.

Eagleson, P. S. P. 1978 Climate, soil and vegetation. 2. The distribution of annual precipitation derived from observed storm sequences. *Water Resour. Res.* **27** (7), 1431–1438.

Entekhabi, D., Rodriguez-Iturbe, I. & Eagleson, P. S. 1989 Probabilistic representation of the temporal rainfall process by modified Neyman-Scott rectangular pulses model: parameter estimation and validation. *Water Resour. Res.* **25** (2), 295–302.

Debele, B., Srinivasan, R. & Parlange, J. Y. 2007 Accuracy evaluation of weather data generation and disaggregation methods at finer timescales. *Adv. Water Res* **30** (5), 1286–1300.

Gilleland, E., Katz, R. & Young, G. 2005 Extremes toolkit (extRemes): Weather and climate applications of extreme value statistics. http://www.assessment.ucar.edu/toolkit/.

Hanaish, I. S., Ibrahim, K. & Jemain, A. 2011 Daily rainfall disaggregation using HYETOS model for peninsular Malaysia, recent researches in applied mathematics, simulation and modelling. *Fifth International Conference on Applied Mathematics*, Simulation, Modelling (ASM '11), ISBN: 978-1-61804-016-9, Corfu Island, Greece, pp. 146–150.

Hutchinson, M. F. 1990 A point rainfall model based on a three-state continuous Markov occurrence process. *J. Hydrol.* **114**, 125–148.

Istok, J. D. & Boersma, L. 1989 A stochastic cluster model for hourly precipitation data. *J. Hydrol.* **106**, 257–285.

Khaliq, M. N. & Cunnane, C. 1996 Modelling point rainfall occurrences with the Modified Bartlett-Lewis Rectangular Pulses Model. *J. Hydrol.* **180**, 109–138.

Koutsoyiannis, D. & Onof, C. 2000 HYETOS a computer program for stochastic disaggregation of fine scale rainfall. Available from: http://www.itia.ntua.gr/e/softinfo/3/>

Koutsoyiannis, D. & Onof, C. 2001 Rainfall disaggregation using adjusting procedures on a Poisson cluster model. *J. Hydrol.* **246**, 109–122.

Koutsoyiannis, D. & Xanthopoulos, T. 1990 A dynamical model for short scale rainfall disaggregation. *Hydrol. Sci. J.* **35**, 303–322.

Onof, C. & Wheater, H. S. 1993 Modelling of British rainfall using a random parameter Bartlett-Lewis rectangular pulses model. *J. Hydrol.* **149**, 67–95.

Pui, A., Sharma, A. & Mehrotra, R. 2009 A comparison of alternatives for daily to sub-daily rainfall disaggregation. Eighteenth World IMACS / MODSIM Congress, Cairns, Australia, 13–17 July.

Rodriguez-Iturbe, I., Cox, D. R. & Isham, V. 1987a Some models for rainfall based on stochastic point processes. *Proc. R. Soc. Lond. A* **410**, 269–288.

Rodriguez-Iturbe, I., Febres, D. & Valdes, J. B. 1987b Rectangular pulses point process models for rainfall: analysis of empirical data. *J. Geophys. Res.* **92**, 9645–9656.

Rodriguez-Iturbe, I., Cox, D. R. & Isham, V. 1988 A point process model for rainfall: further developments. *Proc. R. Soc. Lond. A.* **417**, 283–298.

Segond, M. L., Onof, C. & Wheater, H. S. 2006 Spatial-temporal disaggregation of daily rainfall from a generalized linear model. *J. Hydrol.* **331**, 674–689.

Wheater, H., Isham, V., Chandler, R., Onof, C., Stewart, E., Bellone, E., Yang, C., Lekkas, D., Lekkas, D., Lourmas, G., Segond, M.-L., Forst, A., Prudhomme, C. & Crooks, S. 2005 Improved methods for national spatial-temporal rainfall and evaporation modelling for BSM.R7D Technical Report F2105/TR. Imperial College, London and University College, London.

First received 8 May 2012; accepted in revised form 23 October 2012. Available online 27 December 2012