1 **Realistic sampling of anisotropic correlogram parameters for**
2 **conditional simulation of daily rainfields**

3
4 *Yeboah Gyasi-Agyei*

5 School of Engineering and Technology, Central Queensland University, Bruce Highway,
6 Rockhampton, QLD 4702, Australia.
7 y.gyasi-agyei@cqu.edu.au

8
9 **Abstract**

10 This paper has established a link between the spatial structure of radar rainfall, which
11 more robustly describes the spatial structure, and gauge rainfall for improved daily rainfield
12 simulation conditioned on the limited gauged data for regions with or without radar records.
13 A two-dimensional anisotropic exponential function that has parameters of major and minor
14 axes lengths, and direction, is used to describe the correlogram (spatial structure) of daily
15 rainfall in the Gaussian domain. The link is a copula-based joint distribution of the radar-
16 derived correlogram parameters that uses the gauge-derived correlogram parameters and
17 maximum daily temperature as covariates of the Box-Cox power exponential margins and
18 Gumbel copula. While the gauge-derived, radar-derived and the copula-derived correlogram
19 parameters reproduced the mean estimates similarly using leave-one-out cross-validation of
20 ordinary kriging, the gauge-derived parameters yielded higher standard deviation (SD) of the
21 Gaussian quantile which reflects uncertainty in over 90% of cases. However, the distribution
22 of the SD generated by the radar-derived and the copula-derived parameters could not be
23 distinguished. For the validation case, the percentage of cases of higher SD by the gauge-
24 derived parameter sets decreased to 81.2% and 86.6% for the non-calibration and the
25 calibration periods, respectively. It has been observed that 1% reduction in the Gaussian
26 quantile SD can cause over 39% reduction in the SD of the median rainfall estimate, actual

reduction being dependent on the distribution of rainfall of the day. Hence the main advantage of using the most correct radar correlogram parameters is to reduce the uncertainty associated with conditional simulations that rely on SD through kriging.

Keywords: rainfall; spatial; uncertainty; correlogram; kriging, radar

1. INTRODUCTION

The Great Barrier Reef (GBR) off the coast of Queensland, Australia, covers an area of 344,400 km² and is the largest coral reef system in the world. GBR has a significant economic importance to Australia as a major tourist attraction and efforts are being made to protect it from anthropogenic activities including agriculture and mining. One such effort includes modelling of the transport of sediments, nutrients and inhibiting herbicides from catchments to the reef system. A total area of 423,134 km² consisting of 35 major basins drain coastal Queensland to the GBR (Waters et al., 2014). Hydrological studies of such large catchments involving the use of spatially distributed hydrological models require adequate daily rainfall data. In order to make use of the long records of over 100 years daily rainfall, conditional simulation using kriging to estimate the mean and standard deviation (SD) of the Gaussian quantile is required to generate daily rainfields at 1 km² grid size over the catchments. Analysing radar images, Gyasi-Agyei and Pegram (2014) established that daily gauge network density should be better than one gauge per 16 km² to review the full spatial structure. Therefore, the daily rain gauge network density worldwide is not adequate to capture the spatial structure of rainfall. This shortfall calls for innovative ways of realistic spatial and temporal interpolation of the limited gauged data over the grid or functional units used in the hydrological models.

Several spatial interpolation techniques prevail in the literature, including inverse weighted distance (Teegavarapu and Chandramouli, 2005), nearest neighbours (Isaaks and Srivastava, 1990), thin plate splines (e.g., Jeffrey et al, 2001), kriging (e.g., Ly et al., 2011; Cressie, 1993) and spatial copula (e.g., Bárdossy and Pegram, 2009). However all the interpolation techniques share a common feature of weighted averages of the sampled data, and the differences are basically how the weights are calculated. A recent review on the spatial interpolation methods in environmental science can be found in Li and Heap (2014). With regards to rainfall, copula and kriging based methods are the most popular. Despite some problems, such as ground clutter and conversion methods of radar reflectivity to rainfall intensity (Villarini and Krajewski, 2010), interest in the use of radar sourced rainfall in hydrological modelling is increasing. Quirmbach and Schultz (2002) are among the few researchers that have directly used radar rainfall in water resources. Wang et al. (2013) and Rabiei and Haberlandt (2015) used mean bias reduction and error variance minimisation to adjust the radar data on limited gauged data. Verworn and Haberlandt (2011) and Velasco-Forero et al. (2009) used radar data in kriging with external drift. In the same vein, satellite-based high-resolution rainfall products (e.g. CMORTH, Joyce et al., 2004; TMPA 3B42 and TMPA 3B42RT, Huffman et al. 2007; PERSIANN, Hsu et al., 1997) are being looked into for regions without ground-based radar, despite problems associated with atmospheric effects and gaps in revisit times among others (e.g., Bitew and Gebremichael, 2011).

This paper further develops the work by Gyasi-Agyei and Pegram (2014) and Gyasi-Agyei (2016) on the use of rainfall radar to understand the spatial structure for possible transfer of information from radar regions, and time spans, to areas and time spans without radar records for maximum benefit in terms of uncertainty. In Gyasi-Agyei and Pegram (2014), daily accumulations of radar rainfall of a region within Free State, South Africa, were used to develop the spatial structure, and a weak link by way of a transition probability

matrix was established between the Gauge Wetness Ratio (GWR) and Radar Wetted Area Ratio (RWAR). GWR and RWAR are defined as the ratio of gauges/pixels that equalled or exceeded 2 mm rainfall and the total number of gauges/pixels, respectively. Application of the methodology to Queensland, Australia, data set required major structural changes as outlined in the ensuing sections. Of particular significance are the:

- use of gauge correlogram parameters and maximum temperature as covariates for the radar correlogram parameters through a generalised additive model;
- use of zero inflated standard distributions for the daily rainfall amounts;
- establishment of a copula-based joint distribution of the correlogram axes lengths;
- introduction of a new harmonic distribution for the correlogram direction parameter; and
- validation of the model for different catchments and time spans.

Gyasi-Agyei (2016) examined the effect of the locally varying anisotropy exhibited by radar rainfall within the same square region used in this paper. It was concluded that locally varying anisotropy did not significantly improve the estimates, and hence the use of the global anisotropy suffices.

The next section presents the data sets and processing used in this paper followed by the two-dimensional (2D) spatial correlogram development. What follows is the detailed analysis and modelling of the dependence structure (marginal and joint distributions) of the correlogram parameters. Application of the model to a test catchment and time span for validation is then presented before the concluding summary points. All analyses were done with R (R Core Team, 2015) statistical software.

2. STUDY AREA AND DATA PROCESSING

The study area includes the Brisbane river catchment (13,533 km²), and a square region (128 km x 128 km) within the range of the Mt. Stapylton weather radar station near Brisbane, Australia. As shown in Fig. 1, the two regions, which will be referred to as the catchment and the square region, have an intersection area of 4,982 km². The square region was chosen primarily because of the high density of daily rain gauges nearest to the radar station. The weather radar of a Meteor 1500 S-band Doppler type with a range of 256 km is located at latitude 27.718° S and longitude 153.240° E. In order to conform to the spatial unit of the radar data of 1 km² resolution, the latitude and longitude coordinate system of the daily rainfall gauged stations were converted to the Universal Transverse Mercator (UTM) easting and northing using the *spTransform* function of R package *rgdal* (Bivand et al., 2015) with the datum set to WGS84 for zone 56 (Fig. 1). This study area is located in a subtropical climate with an average temperature of 26.5°C, and an annual mean rainfall of about 990 mm with an average of 124 wet days per year. It experiences hot humid summers (December - February) and moderately dry winters (June - August), with torrential rain associated with thunderstorms being very common.

It needs to be mentioned that the square region is the same as used in Gyasi-Agyei (2016). However, the period considered in this paper is from 2000-01-01 to 2015-06-30. There were 234 daily rainfall gauges located in the square region (1 gauge per 70 km²), 115 within the catchment (1 gauge per 118 km²), and 238 within the 20 km buffer of the catchment (Fig. 1, dashed red line, 27,227 km²), for the period under investigation. Gauges outside the catchment boundary and within the buffer were included to minimise edge effects in the kriging process. Some of the gauging stations were closed and new ones opened during the study period. For each day, only the operational gauges were used as the days were considered independent. This caused the number of gauges to vary between 148 and 200 for

the square region, and between 67 and 100 within the catchment (130 to 196 including the buffer zone). Gyasi-Agyei (2013) found that the maximum daily temperature of the day has significant effects on the sub-daily rainfall characteristics, so it was considered as a covariate, notwithstanding that its correlation with the variables of interest was checked. Within the study area there were 19 temperature stations. For each day, the maximum daily temperature values of the operational stations were interpolated over the region by inverse weighted distance method and the average values within the catchment and the square region were used as representative. There were not enough temperature stations to develop a meaningful spatial structure required by kriging so inverse weighted distance method was considered appropriate. Also, compared with rainfall, temperature does not vary greatly with distance.

All data were obtained from the Australian Bureau of Meteorology (BOM) (<http://www.bom.gov.au/climate/data/>), and the daily rainfall data having a resolution of 0.1 mm was monitored over a 24-h period ending at 9 am. The data were split into two time periods, namely from 2000-01-01 to 2008-12-31 without radar records that were used for verification only, and from 2009-01-01 to 2015-06-30 with radar records for most days that were used to develop the model parameters. Quality checked reflectivity (Z) measurements received from the radar were classified into convective or stratiform rainfall. They were then converted into a rainfall intensity (R, mm/h) using a $Z \sim R$ power-law relationship with different constants for the two different types of rainfall. One minute accumulation rainfall was derived from the rainfall rate accounting for rainfall movement. Observed rainfall data sets within the range of the radar were used to adjust the estimated radar rainfall at an hourly timescale (http://www.bom.gov.au/australia/radar/about/calculating_rainfall_accumulations.shtml). The 6-min temporal resolution radar rainfall obtained from BOM were aggregated over the same sampling daily timescale.

Discontinuities in the data as result of the discretisation process used to sample and store rainfall data were eliminated before use. Following Gyasi-Agyei (2011, 2012a), the wet gauge values were made continuous by applying a random jitter (JT) sampled from a uniform distribution $JT \sim U(-0.05, 0.05)$. For example, a recorded rainfall amount of 20.1 mm is made continuous by adding any of $\{-0.04883066, 0.04764925, -0.01932201, 0.02861681, \text{etc}\}$ drawn from the uniform distribution. In the process duplicates in the data were removed making the whole data set continuous. Days where the number of wet gauges was greater than 5, 5% of the gauges recorded at least 1 mm rain, and the maximum wet gauge rain was at least 10 mm were selected for the analysis. Table 1 presents the summary characteristics of all the 918 wet days with radar records over the second time period. Collocated (radar data at gauged locations) data were also analysed. Clearly, the summary statistics are different for the 3 data sets. In Table 1 it is observed that the average wet gauge amount (the sum of all rainfall data divided by the number of wet gauges/pixels) and the wetted area ratio (number of gauges registering at or above 1mm rain / total number of operational gauges) values of the radar and collocated radar data were similar, in terms of their lower and upper limits, and also their mean. However, their maximum wet gauge amounts could be significantly different.

Because of the normality assumption of ordinary kriging, quantile-quantile (Q-Q) transformation was applied to the daily rainfall data (r) to obtain normal scores (w) which follows the standard Gaussian distribution. This was achieved by first fitting one of the seven 2-parameter distributions (Log-Logistic, Kappa, Generalised Pareto, Gumbel, Log-Normal, Weibull, and Gamma) to the wet gauge values on a given day. Log-Logistic, Kappa, and Generalised Pareto distributions were from the R package *FAdist* (Aucoin, 2015) and the rest were from the R package *fitdistrplus* (Delignette-Muller and Dutang, 2015). Based on the maximum p-value of the Anderson-Darling statistic, the best distribution for the wet day of interest was selected and used to generate the probability $F_R(r[s_k])$ from the rainfall amount

$r[s_k]$ at station k with coordinates s_k (with components of x-easting and y-northing, Fig. 1) which was then adjusted to reflect the zero values before transformation to normal quantile $w[s_k]$ as:

$$w[s_k] = \begin{cases} \Phi^{-1} \left[F_R(r[s_k])(1-p_0) + p_0 \right] & , \quad r[s_k] > 0 \\ \Phi^{-1} \left[p_0 \exp \left(-\frac{d[s_k]}{\bar{d}} \right) \right] & , \quad r[s_k] = 0 \end{cases} \quad (1)$$

where F_R is the selected cumulative distribution function, p_0 is the dry probability defined as the number of dry gauges/pixels over the total number of gauges/pixels, $d[s_k]$ is the distance of a dry gauge from the nearest wet gauge, $\bar{d}$ is the mean of $d[s_k]$ of the dry gauges, k is the station number varying from 1 to the total number of stations, and Φ is the cumulative normal distribution $N(0,1)$. In Gyasi-Agyei and Pegram (2014) and Gyasi-Agyei (2016), the empirical distribution was used and the zeros were treated differently. This new method of treating the zeros is a bit more realistic and improves data continuity. Also, the need to treat the quantiles of the simulated random fields exceeding the observed gauged values has been avoided by using standard distributions. Eq. (1) is inverted as:

$$r[s_k] = \begin{cases} F_R^{-1} \left[\left\{ \Phi(w[s_k]) - p_0 \right\} / (1-p_0) \right] & , \quad \Phi(w[s_k]) \geq p_0 \\ 0 & , \quad \Phi(w[s_k]) < p_0 \end{cases} \quad (2)$$

to obtain the rainfall amounts from the simulated Gaussian $N(0,1)$ values $w[s_k]$. Eqs. (1) and (2) are essentially zero inflated.

3. THE 2D SPATIAL CORRELOGRAM

Yao and Journel (1998) presented three steps for developing spatial correlograms for given values of the variable of interest at several locations. Their approach is based on Bochner's theorem that achieves a positive definite covariance lookup table through 'round-

trip' FFT (Fast Fourier Transform) with intermediate smoothing processes. Application of FFT requires the dimension of the empirical correlogram be a power of two to speed up the computations. In the first instance, the empirical covariance map, $R(i, j)$, was developed using the normal score $w[s_k]$ of the scattered gauged rainfall or the gridded radar data as:

$$R(i, j) = \frac{1}{N_h} \sum_{k=1}^{N_h} w[s_k] w[s_k + h] - m_0 m_{+h}, \quad \forall (s_k, s_k + h) : h \in [i \pm 1, j \pm 1] \quad (3)$$

where s_k (x-easting, y-northing) is the vector of spatial coordinates of gauge k , h (with components of δx -easting and δy -northing) is the lag vector representing the separation distance between two gauge locations, N_h is the number of contributing gauge pairs separated by lag h that falls within the bounded region of $[i \pm 1, j \pm 1]$, and m_0 and m_{+h} are the mean of the pair tail $w[s_k]$ and head $w[s_k + h]$ values. Note that the covariance is symmetric in h , $-h$. The covariance map is a square table with 256 pixels in each direction and centred in the middle, meaning i and j in Eq. (3) take on integer values from -127 to 128 noting that a pixel dimension is 1km x 1km. It is necessary, particularly for the scattered gauged data, to infill entries with zero or low N_h values by interpolation (the second step). There may be large fluctuations and missing entries that might be present after the first step, calling for a smoothing process involving a weighted average within a radial moving window. In the third step, the gridded sample covariance map is transformed to a quasi-density spectrum map in the frequency domain by FFT, with a subsequent smoothing to achieve the constraint of positivity and unit sum. The smoothed spectrum map is then inverse FFT to obtain the required correlogram lookup table which is positive definite and with all sample fluctuations, in particular for the scattered data, filtered out. Notable examples that have used this FFT approach in rainfall modelling are Velasco-Forero et al. (2009) who modelled anisotropic variograms of 1-hour radar accumulations, and Gyasi-Agyei and Pegram (2014) and Gyasi-Agyei (2016) for daily gauged and radar accumulations.

The primary goal of this paper is to fit a 2D exponential function to both the daily gauged and radar accumulation correlograms, and then link the gauged correlogram parameters to those of the radar for sampling ‘radar-like’ values at regions without radar information. Note that because we are dealing with normal scores the empirical correlogram is the same as the covariance table, $R(i, j)$. In essence, the assumption is that radar provides a better spatial structure which is to be reproduced through sampling methods. Eq. (4) describes the 2D anisotropic exponential function $R_{\Theta}(i, j)$ as:

$$R_{\Theta}(i, j) = R_{\Theta}(u, v) = a_n \exp \left\{ - \left[\left(\frac{u}{L_u} \right)^2 + \left(\frac{v}{L_v} \right)^2 \right]^{\frac{1}{2}} \right\}, \quad \begin{aligned} u &= j \sin(\theta) + i \cos(\theta) \\ v &= j \cos(\theta) - i \sin(\theta) \end{aligned} \quad (4)$$

In Eq. (4), u and v are the lag distances along the major and minor axes, respectively, of the elliptical contour making an angle of θ (between 0 and 180 degrees) measured anticlockwise from the +I (east) and +J (north) directions respectively. L_u and L_v are the correlation lengths of the major and minor axes which define the anisotropy ratio ($\eta = L_v/L_u$) that takes on values between 0 and 1. Parameter a_n is the correlogram value at lag 0 which should have a theoretical value of 1 but may be smaller due to the small-scale variability of the data, such as measurement errors and those associated with sample spacing and termed the nugget effect. In covariance terms the nugget effect may render the covariance value at lag 0 different from the expected maximum value called the sill that represents the variance. Because we are dealing with normal scores the sill has a value of 1.

In Fig. 2 is shown a 2D anisotropic exponential correlogram for $\theta = 120^\circ$, $L_u = 30$ km, $L_v = 18$ km, $a_n = 1$, and the effective range $L_R = -\log(0.05)L_u = 90$ km (where the correlation is 0.05). For each wet day, the four parameters $\Theta(a_n, L_u, L_v, \theta)$ of the correlogram were estimated for both gauged (gauge-derived) and radar (radar-derived) data using the global

optimisation technique of Duan et al. (1992) implemented in *R* function *SCEoptim* (Andrews and Guillaume, 2013) subject to the objective function Θ defined as:

$$\Theta = \arg \min_{\Theta} \sum_{i,j} [R_{\Theta}(i,j) - R(i,j)]^2 \quad (5)$$

where (i, j) is the grid cell location with respect to the centre of the sample correlogram table. An initial parameter estimation indicated that the nugget effect is close to zero and the parameter a_n was set to 1, leaving only the three parameters $\Theta(L_u, L_v, \theta)$ to be calibrated. Fig. 3 (left) shows the radar rainfall image for 2009-04-05 and Fig. 3 (right) compares the empirical correlogram with the fitted one. It is observed that the analytical function fits the empirical correlogram very well (correlation = 0.976), and anisotropy is very well pronounced.

4. DEPENDENCE STRUCTURE OF THE CORRELOGRAM PARAMETERS

The objective here is to develop relationships and marginal distributions that will allow sampling of radar correlogram parameters dependent on the observed counterparts, believing that the gauge network density is not high enough to capture the full spatial structure. Hereafter we use last subscripts to distinguish between parameters derived from radar data (radar-derived, last subscript r), gauge data (gauge-derived, last subscript g) and collocated data (collocated-derived, last subscript c). Table 2 presents the Pearson correlation between the correlogram parameters derived from the three data sources, as well as some basic characteristics of the gauged data and the maximum daily temperature. There is a very strong correlation (>80%) between the major (L_u) and minor (L_v) correlation lengths of all 3 data sets. However, the correlation between the radar-derived and gauge-derived parameters

are mild but significant, and of the same order as the maximum daily temperature. It appears there is little correlation between the direction and the other correlogram parameters. Also, the other gauged characteristics of mean amount of all gauges (MAG), maximum amount of gauges (MAX), number of wet gauges (NWG), wetted area ratio (WAR) and mean amount of wet gauges (MWG) are not correlated with correlogram parameters. From these observations, a bivariate joint distribution was sought between the radar-derived parameters L_{ur} and L_{vr} , and independently fit a distribution to the direction parameter θ_r . This approach was also adopted by Gyasi-Agyei and Pegram (2014), but with significant differences in the modelling of the marginal distribution functions in particular. Also, a bivariate distribution between L_{ur} and L_{vr} was sought instead of between L_{ur} and η_r as follows.

Fig. 4 displays the histogram of the 918 wet days of the radar-derived correlogram parameters. It is observed that the correlogram direction (Fig. 4, left) exhibits a very strong harmonic behaviour and should be treated as such. A second order harmonic density function $f_{\theta_r}(\theta)$ defined as:

$$f_{\theta_r}(\theta) = \frac{1}{180} \left[A + B \cos\left(2\pi \frac{\theta}{180}\right) + C \sin\left(2\pi \frac{\theta}{180}\right) + D \cos\left(4\pi \frac{\theta}{180}\right) + E \sin\left(4\pi \frac{\theta}{180}\right) \right] \quad (6)$$

where A, B, C, D and E are coefficients to be determined, was fitted. By integration of Eq. (6) and knowing that the cumulative distribution function $F_{\theta_r}(\theta)$ has a value of zero at $\theta=0$ and one at $\theta=180$, coefficient A should take a value of unity, yielding the harmonic distribution as:

$$F_{\theta_r}(\theta) = \frac{\theta}{180} + \frac{B}{2\pi} \sin\left(2\pi \frac{\theta}{180}\right) - \frac{C}{2\pi} \left[\cos\left(2\pi \frac{\theta}{180}\right) - 1 \right] + \frac{D}{4\pi} \sin\left(4\pi \frac{\theta}{180}\right) - \frac{E}{4\pi} \left[\cos\left(4\pi \frac{\theta}{180}\right) - 1 \right] \quad (7)$$

thus leaving coefficients B, C, D and E to be estimated. Fig. 4 (left panel) demonstrate a very good fit (correlation coefficient = 0.967) using the estimated parameters of B=-0.2361, C=-

0.5417, D=0.1131 and E=0.1124. This is a newly introduced distribution that can be used for data exhibiting multiple peaks and troughs in the density function. Harmonic functions have previously been fitted to parameters (e.g. Gyasi-Agyei, 1999, 2001) but not used as a distribution in the literature.

The copula methodology, presented later, used to model the strong dependence between the radar-derived correlogram parameters (L_{ur} and L_{vr}) requires their marginal distributions. It is observed that the marginal distributions of these parameters were highly positively skewed (Fig. 4, middle and right). We also explored the possibility of using the remaining variables in Table 2 as covariates to predict the marginal distribution parameters through link functions. The R package *gamlss* (Generalized Additive Models for Location, Scale and Shape; Rigby and Stasinopoulos, 2005) offers this opportunity. It is able to identify the appropriate marginal distributions which have up to four parameters: μ for location (median); σ for scale (coefficient of variation); v for skewness and τ for kurtosis being characterised as the shape parameters. Monotonic link functions are used to relate the distribution parameters to the explanatory variables (covariates) and random effects through an additive model. The coefficients of the linear functions of the explanatory variables are then estimated by maximum likelihood methods. The Generalised Akaike Information Criteria (GAIC) (Stasinopoulos and Rigby, 2007) is used for model selection and testing of the significance of an additional term in the linear functions. Thus the R package *gamlss* provides a suitable framework for fitting a suite of distributions and accounting for exogenous forcing variables by incorporating covariates in the location, scale and shape parameters. Linkage of the parameters to the explanatory variables provides a vehicle for improving the fit of the dispersion, skewness and kurtosis. Serinaldi and Kilsby (2014) used the *gamlss* framework to incorporate seasonality and sea surface pressure to the rainfall

distribution parameters. The interest here is in using gauged-derived correlogram parameters (L_{ug}, L_{vg}) and maximum temperature (T_m) as covariates.

A suite of 2-parameter (μ, σ), 3-parameter (μ, σ, ν) and 4-parameter (μ, σ, ν, τ) distributions in R package *gamlss* were tested, and the Box-Cox power exponential (BCPE, 4-parameter) distribution emerged as the best based on the GAIC statistic. This distribution was developed by Rigby and Stasinopoulos (2004) for random variables exhibiting both skewness and kurtosis (leptokurtosis or platykurtosis). It is defined through a transformed random variable that follows a truncated standard power exponential distribution with parameter τ , and the parameter ranges are $\mu > 0, \sigma > 0, -\infty < \nu < \infty$ and $\tau > 0$. A step-by-step comparison of the various linear combinations of the covariates to predict the parameters of the marginal distributions of the radar-derived parameters (L_{ur}, L_{vr}) was carried out. It was sufficient to link only the parameters μ and ν to the covariates as $\log(\mu) = m_1 + m_2 L_g + m_3 T_m$ and $\nu = n_1 + n_2 L_g$ considering the remaining parameters (σ and τ) as constants, and L_g being either L_{ug} or L_{vg} . For the chosen model, all p-values of the null hypothesis that the covariates do not contribute to the parameter variability were less than 0.001. Table 3 presents the estimated parameters with their standard errors of the linear link function coefficients. Using the logarithmic link function guarantees positivity of those parameters that should have positive values.

A copula modelling approach (see e.g., Nelsen, 2006) is adopted for the joint distribution between the radar-derived correlogram parameters of L_{ur} and L_{vr} . Denoting the continuous marginal distribution functions of L_{ur} and L_{vr} as F_U and F_V , respectively, their joint distribution H_{UV} by Sklar's theorem is written as:

$$\begin{aligned} H_{UV}(u, v) &= C[F_U(u), F_V(v)] \quad \text{or} \\ C(u, v) &= H_{UV}[F_U^{-1}(u), F_V^{-1}(v)] \end{aligned} \tag{8}$$

where $0 \leq u, v \leq 1$, F_U^{-1} and F_V^{-1} are the inverse distribution functions of F_U and F_V , and C is a copula. Gyasi-Agyei (2012b) used such a bivariate modelling approach for the joint distribution of the total daily rainfall amount and total wet periods duration.

It was observed that temperature has a significant effect on the copula parameter. Hence the data were ranked in order of temperature and binned into 11 overlapping classes of size 200 starting at every rank modulo 100 equals to 1. Because there are fewer days at the extremes of temperature, the 1st class is the first 100 and the last class constitute the last 100. That is, class 1 is from 1 to 100, class 2 from 1 to 200, class 3 from 101 to 300, and so on. For each class the mean temperature and the copula parameter were estimated. A pairwise scatterplots of the empirical rank distribution for classes 4 and 9 shown in Fig. 5 indicate a very strong positive correlation between L_{ur} and L_{vr} with a small upper tail dependence. Following the procedure presented in Gyasi-Agyei (2011), a single parameter bivariate copula was sought and the Gumbel copula emerged as the best. The Gumbel copula is defined as:

$$C(u, v) = \exp\left(-\left\{[-\log(u)]^\delta + [-\log(v)]^\delta\right\}^{1/\delta}\right) \quad (9)$$

where parameter $\delta \geq 1$. Generation of pair values (u, v) from the copula requires the h -function given as (Aas et al., 2009):

$$h(v/u) = \frac{\partial C(u, v)}{\partial u} = C(u, v) \frac{1}{u} [-\log(u)]^{\delta-1} \left\{ [-\log(v)]^\delta + [-\log(u)]^\delta \right\}^{1/\delta-1} \quad (10)$$

which should be numerically inverted to obtain the inverse of the h -function (h^{-1}). The generation process based on the probability integral transform (e.g., Salvadori et al. 2007) begins by drawing u and x from a uniform distribution $\sim U(0,1)$ and setting $v = h^{-1}(x/u)$.

This is followed by the estimation of the quantiles of L_{ur} and L_{vr} using their margins determined in the previous section. Fig. 5 (right) depicts the variation of the copula parameter with temperature (correlation = 0.985), indicating the dependence is stronger at the lower

temperatures. The `gofCopula` function of R package *copula* (Kojadinovic and Yan, 2010) based on the Cramer–von Mises S_n statistic was used to assess the goodness of fit of the Gumbel copula. p-values for all classes (0.416, 0.338, 0.380, 0.086, 0.014, 0.100, 0.063, 0.172, 0.489, 0.570, 0.305, in order of increasing temperature) indicate that all classes passed at the 5% significance level, with the exception of class 5 which failed at the 5% but passed at the 1%.

5. APPLICATION

It is important to critically examine the relationship between the radar-derived and the gauge-derived correlogram parameters of the major and minor axes. For easy depiction of trends, the data were grouped into 10 bins of equal size based on the quantiles of the gauge-derived correlogram parameters (L_{ug} and L_{vg}), maximum daily temperature (T_m) and the location parameter (μ) of the BCPE distribution. The means of the binned data are shown as blue solid circles in Fig. 6.

As observed in Fig. 6 (top left panel), only 6.5% of L_{ug} values exceeded their radar counterparts, and the percentage for L_{vg} (not shown) was slightly higher at 9.6%. Interestingly, for the collocated-derived parameters (L_{uc} and L_{vc}), the percentages were 0.2% for L_{uc} and 0.3% for L_{vc} , also exhibiting smaller variance with apparent stronger positive correlation (Fig. 6, bottom left panel). With respect to the daily maximum temperature (Fig. 6, middle panels), the correlogram axes parameters tend to decrease as the temperature increases. The perception that convective rainfall that normally occurs on hotter days towards the upper end of the temperature spectrum tends to be showery and have limited horizontal extent, and thus shorter correlation length, is supported by this observation. The combined effect of the gauge-derived parameters and temperature used to estimate the location

parameters of the BCPE distribution is shown in the right panels of Fig. 6 which indicate a strong positive correlation with the correlogram parameters. It is important to emphasise that the estimated coefficients of the covariates were all significant as the p-values of the null hypothesis that they do not contribute were less than 0.001.

Extremes are very important for flood studies, so the data were ranked according to the within-day maximum gauge reading and the top 20 extremes were examined for a possible trend. They are shown as red triangles in Fig. 6. There are no apparent specific features of the extreme wet days as they are scattered all over the plots, other than to say that they tend to be confined between 21°C and 31°C although there are fewer wet days below 21°C. Other features of the extreme events are presented in Table 4. Even the types of marginal distribution of the rainfall amounts reflect quite well the overall distribution shown in Table 5, Log-Normal (29%) and Log-Logistic (22%) dominating the gauged data and Log-Logistic accounting for over 75% of the radar data.

In order to understand the adequacy of the copula modelling component, L_{ug} , L_{vg} and T_m were used as inputs to sample the radar-derived counterparts from the copula-based joint distribution, as well as the radar-derived direction for the 918 wet days, labelled L_{us} , L_{vs} and θ_s and termed copula-derived parameters. This was repeated 100 times and, for each set, the Kolmogorov–Smirnov (K-S) test was used to compare the radar-derived and the copula-derived values. For the bivariate linkage between L_{ur} and L_{vr} , 2D K-S test (Press et al., 2007) was employed as discussed in Gyasi-Agyei and Pegram (2014). Table 6 shows the cumulative distribution of the p-values and the adequacy of the modelling effort is exemplified as all p values were in excess of 0.01, and only L_{vr} had 5% p-values lower than 0.05. In statistical terms, the assumptions that the copula-derived and the radar-derived variables come from the same underlying distribution cannot be rejected at the 1% significance level. With this confidence, the copula-derived correlogram parameters were

used to perform the ordinary kriging interpolation, which is a precursor for the desired application of conditional simulation of daily rainfields.

The mathematics behind the ordinary kriging interpolation method is well documented in the literature (e.g., Cressie, 1993), and requires no introduction here. R package *gstat* (Pebesma, 2004) was used to perform the ordinary kriging. It needs to be mentioned that *gstat* uses the effective length $L_R = -\log(0.05)L_u$, anisotropy ratio $\eta=L_v/L_u$, and the direction θ_{gstat} is measured from the north (+J) and increases clockwise, requiring the transformation:

$$\theta_{gstat} = \begin{cases} 90 - \theta & , \theta \leq 90 \\ 270 - \theta & , \theta > 90 \end{cases} \quad (11)$$

The parameter set $(L_R, \theta_{gstat}, \eta)$ were used to define the 2D anisotropic exponential variogram that was required by *gstat* for kriging. Since we are dealing with normal scores having a variance value of 1, the correlogram is the same as the covariance. Therefore, the variogram $V_\Theta(i, j)$ equals $1 - R_\Theta(i, j)$ where $R_\Theta(i, j)$ is defined in Eq. (4), the nugget being 0.

Cross-validation methodology (leave-one-out and 10-fold) was used to evaluate the performance of the copula-derived correlogram parameters in reproducing the gauged rainfall. The ordinary kriging was carried out for each day using three sets of correlogram parameters:

- gauge-derived: L_{ug}, L_{vg}, θ_g
- radar-derived: L_{ur}, L_{vr}, θ_r
- copula-derived: L_{us}, L_{vs}, θ_s

Only the results for the leave-one-out are presented as the 10-fold results were very close. Performance statistics of interest were the root-mean-square-error (RMSE), mean-bias (MB), and mean-absolute-bias (MAB) defined as:

$$427 \quad \text{RMSE} = \sqrt{\frac{1}{N} \sum_{k=1}^N [V_E(k) - V_O(k)]^2}$$

$$428 \quad \text{MB} = \frac{1}{N} \sum_{k=1}^N [V_E(k) - V_O(k)]$$

$$429 \quad \text{MAB} = \frac{1}{N} \sum_{k=1}^N |V_E(k) - V_O(k)|$$

430 where $V_O(k)$ and $V_E(k)$ are the observed and estimated mean values for gauge station k in mm.
 431 A very important statistic is the mean standard deviation (SD) in the Gaussian domain
 432 produced by ordinary kriging. SD defines the uncertainty associated with the mean estimates
 433 as the higher the SD the higher the errors at the unobserved locations.

434 Scatter plots of RMSE and SD are shown in Fig. 7 for the calibration period (2009-
 435 01-01 to 2015-06-30). With respect to the mean estimates, all correlogram parameter sources
 436 produced very similar results, hence statistics of RMSE, MB and MAB were equally
 437 predicted. For this reason, comparison of the performance statistics of MB and MAB are not
 438 shown. Kriging is generally quite robust to changes in the variogram, so the similarity of
 439 these mean dependent statistics is not surprising. What is very clear is that the gauge-derived
 440 correlogram parameters yielded higher SD, and hence greater uncertainty. In reality, only
 441 7.2% of the gauge-derived SDs were lower than those of the radar-derived values, and 6.6%
 442 were lower than the copula-derived correlogram parameters. Noting a K-S test p-value of
 443 0.094 (passing 5% significance level), the distribution of SD generated by the radar-derived
 444 parameters was indistinguishable from that generated by the copula-derived parameters,
 445 indicating that the copula-based model for sampling the correlogram parameters is working
 446 very well. Here too the scatter of the top 20 extreme wet days, shown as red triangles in Fig.
 447 7, do not appear to follow a specific trend, suggesting the methodology is good for all types
 448 of rainfall within the region.

For the validation phase, all rainfall data within the two time periods were examined for both the square region and the catchment. Using the same criteria for selecting the wet days, there were 1314 and 1061 wet days, respectively, during the earlier and later time periods for the square region, while the catchment had 1231 and 989 wet days, respectively. Leave-one-out cross-validation results presented in Fig. 8 show the same patterns as the calibration phase presented in Fig. 7. The mean estimate dependent performance statistics were produced similarly for both the gauge-derived and copula-derived correlogram parameters. With respect to SD, the gauge-derived parameters produced lower values in 8.4% and 5.7% of the cases during the time periods for the square region, and 18.8% and 13.5% for the catchment. Again, the top 20 extreme wet days are scattered all over the plots without a discernible pattern, and Table 7 shows some characteristics of the extreme wet days during the earlier time period.

In order to understand the implications of reduced SD on rainfall predictions we examined the estimates of the rainfall on 3 of the 20 extreme wet days. The distribution of the observed daily amounts were:

- 2011-01-11, Gamma (shape= 1.8897, rate= 0.03015), median rainfall = 51.4 mm
- 2013-01-28, Log-Logistic (shape= 0.2997, scale= 5.215), median rainfall = 184 mm
- 2014-03-28, Log-Logistic (shape= 0.2028, scale= 4.758), median rainfall = 116.6 mm

For a given normal distribution mean M from NM and SD from NSD, ten thousand normal quantiles were drawn from $\phi^{-1}(M=NM, SD=NSD)$, where NSD is a set from 0 to 0.5 at intervals of 0.05 and NM is a set of values between -3 and 3. These normal quantiles were converted into rainfall values in mm using Eq. (2). For each value of SD, the median rainfall standard deviation (RSD), 90% and 99% prediction limits (PL) were estimated. Fig. 9 shows the boxplots of the results as well as the variation of the rainfall statistics with SD. There is a

very strong nearly linear relationship between SD and RSD. The slopes of the plots of the 3 days, being 41.5, 94.0 and 39.0, can be interpreted as the percent reduction in RSD for 1% reduction in SD. It needs to be noted that the percentage reduction is dependent on the distribution of the day and the individual site value as shown in Fig. 10. Translation of the reduction in SD to PL is simply $M \pm SD\phi^{-1}[(1+PL)/2]$. What this analysis shows is that the use of the most correct correlogram of radar data, which tends to have longer correlation lengths, leads to smaller SD and in turn reduces the errors associated with the predicted values. In a nutshell, the use of radar correlogram parameters is to reduce the uncertainty associated with the conditional simulations that rely on SD through kriging. Example results of leave-one-out cross validation of one set of copula-derived correlogram parameters for one low and one extreme wet day within the catchment are shown in Fig. 11. The normal quantiles (Q_p) of the appropriate probability (p) were used to define the 90% prediction limits ($Q_{0.05}$, $Q_{0.95}$), and then converted into rainfall estimates in mm using Eq. (2).

In summary, the process of sampling the ‘radar-like’ correlogram parameters from gauge-derived estimates are the following.

- For a given day, gather all the observed daily rainfall and maximum temperature readings within the region of interest.
- Estimate the mean of maximum daily temperature of the day by inverse weighted distance method.
- Fit a 2-parameter marginal distribution to the rainfall values to convert rainfall to Gaussian values, with special treatment of the zeroes.
- Develop a correlogram table for the Gaussian values and fit the 2D anisotropic exponential function to determine the gauge-derived values of the major and minor axes correlation lengths.

- Sample the ‘radar-like’ correlogram direction from the second order harmonic distribution.
- Sample the ‘radar-like’ correlogram major and minor axes correlation lengths from the copula-based joint distribution that has Box-Cox power exponential margins and Gumbel copula, using the gauge-derived correlogram parameters and the mean of maximum daily temperature as covariates.

6. CONCLUDING REMARKS

This paper is the third in a series by the authors (Gyasi-Agyei and Pegram, 2014; Gyasi-Agyei, 2016) seeking to establish a link between the spatial structure of radar and gauge rainfall for improved simulation of daily rainfields conditioned on the limited gauged data for regions with or without radar records. While spatial gauged data records spans over 100 years, radar data are limited worldwide and only a few stations have data in excess of 10 years. Availability of such long rainfields will enhance catchment hydrological models for understanding the effects anthropogenic activities have on the transport of sediments, nutrients and inhibiting herbicides from catchments to the GBR for example.

918 wet days within the period of 2009-01-01 to 2015-06-30 were used to calibrate the model for a 128 km by 128 km region within the Mt. Stapylton, near Brisbane, weather radar station coverage area. The square region and the Brisbane river catchment (13,533 km²) were used for validation for two non-overlapping time periods (2000-01-01 to 2008-12-31 and 2009-01-01 to 2015-06-30). For the earlier time period, the number of wet days was 1314 and 1231 for the square region and catchment, respectively. Considering all data in the later period irrespective of radar information, the numbers were 1061 and 989. A 2D analytical exponential function was fitted to the gauge and radar data correlogram separately and their

parameters (major and minor axis lengths, and direction) compared. The established link between the gauge-derived and radar-derived correlogram parameters were evaluated through ordinary kriging, using leave-one-out cross-validation and performance statistics. The following are the concluding summary points.

- The introduced harmonic distribution fits the independent radar correlogram direction very well (correlation = 0.967), and can also be used for any data exhibiting multiple peaks and troughs in the density.
- The highly positively skewed distribution of the radar-derived correlogram axes parameters are best modelled by Box-Cox power exponential with the gauged-derived parameters and maximum daily temperature significantly influencing the location, while only the gauge-derived parameters have an effect on the skewness, the scale and kurtosis being constants.
- Gumbel copula with temperature dependent parameter fits the joint distribution of the radar-derived correlogram axes parameters very well, with an average p-value of 0.27.
- It was observed that the gauged-derived major axis length was lower than the radar-derived values in 93.5% of cases, and the minor axis values in 90.4% of cases during the calibration period for the square region. The percentages were 99.8% and 99.7% for the collocated-derived values. A shorter correlation length implies the correlation decays quickly and farther locations may not have direct influence. This leads to an increase in the errors associated with the estimates of the unobserved locations.
- The correlogram axes parameters tend to decrease with increasing temperature. This is evidence of convective rainfall that normally occurs on hotter days and has limited horizontal extent.

- 545 • With regards to the daily rainfall amounts distribution, the dominant ones were Log-
546 Normal (29%) and Log-Logistic (22%) for the gauged data, and over 76% of Log-Logistic
547 for the radar data.
- 548 • K-S test (1D for marginal distributions and 2D for the joint distribution) used to compare
549 100 sets of copula-derived and radar-derived correlogram parameters indicated that the
550 assumption of them coming from the same distribution could not be rejected at the 1%
551 significance level. In fact, only the minor axis length had 5% of p-values lower than 0.05.
- 552 • The gauged-derived, radar-derived and copula-derived correlogram parameters reproduced
553 the mean estimates very well, and thus the mean dependent performance statistics of
554 RMSE, MB and MAB were similarly predicted.
- 555 • The main differences in the gauged-derived, radar-derived and copula-derived
556 correlogram parameters were in the SD performance statistics that reflect the uncertainty
557 in the estimates. In 92.8% of cases, the radar-derived correlogram parameters produced
558 SDs lower than their gauge-derived counterparts, and with the copula-derived parameters
559 it increased marginally to 93.4%. It is interesting to note that the distribution of the SD
560 generated by the radar-derived and the copula-derived correlogram parameters could not
561 be distinguished with a K-S test p-value of 0.093, a testimony of the quality of the copula
562 model.
- 563 • Considering the gauge-derived and copula-derived correlogram parameters during the
564 validation phase, the same pattern in terms of the performance statistics were observed.
565 However, the percentage cases by which the gauge-derived sets produced higher SD were
566 91.6% and 94.3% for the first and second time periods for the square region, and 81.2%
567 and 86.5% in the case of the catchment.

- Extreme wet days, selected by within-day maximum gauge reading, do not seem to follow a discernible pattern as the top 20 wet days were found to be fairly widely distributed over the parameters' ranges for both the calibration and validation time periods.
- It has been demonstrated that 1% reduction in the Gaussian quantile SD can cause over 39% reduction in the uncertainty associated with the median rainfall estimate, and in general the percentage reduction depends on the distribution of the day and the actual amount at a given site.
- The main advantage of using the most correct radar correlogram parameters is thus to reduce the uncertainty associated with the conditional simulations that rely on SD through kriging.

The next stage of the research is to use Gómez-Hernández and Journel (1993) joint sequential simulation of multi-Gaussian fields methodology to simulate conditional random fields, and then convert to rainfall for the water quality modelling of the GBR. Effects of the gauge network density on the correlogram parameters is worth investigating. Other radar locations should be analysed to ascertain whether the developed relationships are region specific or not.

7. ACKNOWLEDGEMENT

Constructive comments by two anonymous reviewers are gratefully acknowledged.

8. REFERENCES

Aas, K., Czado, C., Frigessi, A., Bakken, H., 2009. Pair-copula constructions of multiple dependence. Insurance: Mathematics and Economics 44 (2), 182-198.

590 Andrews F., Guillaume J., 2013. hydromad: Hydrological Model Assessment and
 591 Development. R package version 0.9-18. URL <http://hydromad.catchment.org/>.
 592 Aucoin, F. (2015). FAdist: Distributions that are Sometimes Used in Hydrology. R package
 593 version 2.2. <https://CRAN.R-project.org/package=FAdist>.
 594 Bárdossy, A., Pegram, G.G.S., 2009. Copula based multisite model for daily precipitation
 595 simulation. Hydrol. Earth Syst. Sci. 13, 2299-2314. doi:10.5194/hess-13-2299-2009.
 596 Bivand, R., Keitt, T., Rowlingson, B., 2015. rgdal: Bindings for the Geospatial Data
 597 Abstraction Library. R package version 1.1-3. <https://CRAN.R-project.org/package=rgdal>.
 598 Bitew, M.M., Gebremichael, M., 2011. Assessment of satellite rainfall products for
 599 streamflow simulation in medium watersheds of the Ethiopian highlands. Hydrol. Earth
 600 Syst. Sci. 15, 1147–1155.
 601 Cressie, N., 1993. Statistics for Spatial Data. John Wiley and Sons, New York.
 602 Delignette-Muller, M.L., Dutang, C., 2015. fitdistrplus: An R Package for Fitting
 603 Distributions. Journal of Statistical Software 64(4), 1-34. doi:10.18637/jss.v064.i04.
 604 Duan, Q., Sorooshian, S., Gupta, V., 1992. Effective and efficient global optimization for
 605 conceptual rainfall- runoff models. Water Resour. Res., 28 (4), 1015-1031.
 606 Gómez-Hernández, J.J., Journel, A.G., 1993. Joint sequential simulation of multi-Gaussian
 607 fields. Geostat. Troia. 92 (1), 85-94.
 608 Gyasi-Agyei, Y., 1999. Identification of regional parameters of a stochastic model for rainfall
 609 disaggregation. J. Hydrol. 223 (3–4), 148–163.
 610 Gyasi-Agyei, Y., 2001. Modelling diurnal cycles in point rainfall properties, Hydrol. Process.
 611 15, 595–608. doi:10.1002/hyp.172.
 612 Gyasi-Agyei, Y., 2011. Copula-based daily rainfall disaggregation model, Water Resour.
 613 Res., 47 (7), W07535. doi:10.1029/2011WR010519.

614 Gyasi-Agyei, Y., 2012a. Modelling the dependence and internal structure of storm events for
615 continuous rainfall simulation. *J. Hydrol.* 464, 249-261.

616 Gyasi-Agyei, Y., 2012b. Use of observed scaled daily storm profiles in a copula based
617 rainfall disaggregation model. *Adv. Water Resour.* 45, 26-36.
618 doi:10.1016/j.advwatres.2011.11.003.

619 Gyasi-Agyei, Y., 2013. Evaluation of the effects of temperature changes on fine timescale
620 rainfall. *Water Resour. Res.* 49(7), 4379-4398.

621 Gyasi-Agyei, Y., Pegram, G., 2014. Interpolation of daily rainfall networks using simulated
622 radar fields for realistic hydrological modelling of spatial rain field ensembles. *J. Hydrol.*
623 519, 777-791, doi:10.1016/j.jhydrol.2014.08.006.

624 Gyasi-Agyei, Y., 2016. Assessment of radar based locally varying anisotropy on daily rainfall
625 interpolation, *Hydrolog. Sci. J.* 61(10), 1890-1902., [doi:10.1080/02626667.2015.1083652](https://doi.org/10.1080/02626667.2015.1083652).

626 Hsu, K. L., Gao, X., Sorooshian, S., Gupta, H. V., 1997. Precipitation estimation from
627 remotely sensed information using artificial neural networks. *J. Appl. Meteorol.* 36, 1176–
628 1190.

629 Huffman, G. J., et al., 2007. The TRMM multisatellite precipitation analysis: Quasi global,
630 multi-year, combined-sensor precipitation estimates at fine scale. *J. Hydrometeorol.* 8,
631 38–55.

632 Isaaks, E.H., Srivastava, R.M., 1990. *An Introduction to Applied Geostatistics*. Oxford
633 University Press, New York.

634 Jeffrey, S.J., Carter, J.O., Moodie, K.M., Beswick, A.R., 2001. Using spatial interpolation to
635 construct a comprehensive archive of Australian climate data. *Environ. Modell. Softw.* 16
636 (4), pp 309-330.

637 Joyce, R. J., Janowiak, J. E., Arkin, P. A., and Xie, P., 2004. CMORPH: A method that
 638 produces global precipitation estimates from passive microwave and infrared data at high
 639 spatial and temporal resolution, *J. Hydrometeorol.* 5(3), 487–503.

640 Kojadinovic, I, Yan, J., 2010. Modeling Multivariate Distributions with Continuous Margins
 641 Using the copula R Package. *J. Stat. Softw.* 34(9), 1-20. doi:10.18637/jss.v034.i09.

642 Li, J., Heap, A.D., 2014. Spatial interpolation methods applied in the environmental sciences:
 643 a review, *Environ. Modell. Softw.* 53, 173-198.

644 Ly, S., Charles, C., Degre, A., 2011. Geostatistical interpolation of daily rainfall at catchment
 645 scale: the use of several variogram models in the Ourthe and Ambleve catchments,
 646 Belgium. *Hydrol. Earth Syst. Sci.* 15, 2259-2274.

647 Nelsen, R.B., 2006. An Introduction to Copulas, second ed. Springer-Verlag, New York.

648 Pebesma, E.J., 2004. Multivariable geostatistics in S: The gstat package. *Comput. Geosci.* 30:
 649 683-691.

650 Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., 2007. Numerical Recipes:
 651 The Art of Scientific Computing, third ed., Cambridge University Press, New York p.
 652 1286.

653 Quirmbach, M., Schultz, G.A., 2002. Comparison of rain gauge and radar data as input to an
 654 urban rainfall-runoff model. *Water Sci. Technol.* 45 (2), 27-33.

655 R Core Team, 2015. R: A language and environment for statistical computing. R Foundation
 656 for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

657 Rabiei, E., Haberlandt, U., 2015. Applying bias correction for merging rain and radar data. *J.*
 658 *Hydrol.* 522, 544-557.

659 Rigby, R.A. Stasinopoulos, D.M., 2004. Smooth centile curves for skew and kurtotic data
 660 modelled using the Box-Cox power exponential distribution. *Stat. Med.* 23(19):3053-
 661 3076.

662 Rigby, B., Stasinopoulos, M., 2005. Generalized additive models for location, scale and
663 shape. *Appl. Stat.* 54, 507-554.

664 Salvadori, G., De Michele, C., Kottegoda, N., Rosso, R., 2007. *Extremes in Nature. An*
665 *Approach Using Copulas. Water Science and Technology Library (Vol. 56).* Springer,
666 Dordrecht.

667 Serinaldi, F. Kilsby, C.G., 2014. Simulating daily rainfall fields over large areas for collective
668 risk estimation. *J. Hydrol.*, 512, 285-302.

669 Stasinopoulos D.M., Rigby R.A., 2007. Generalized additive models for location scale and
670 shape (GAMLSS) in R. *J. Stat. Softw.* 23(7), doi:10.18637/jss.v023.i07.

671 Teegavarapu, R.S.V., Chandramouli, V., 2005. Improved weighting methods deterministic
672 and stochastic data-driven models for estimation of missing precipitation records. *J.*
673 *Hydrol.*, 312, 191-206.

674 Velasco-Forero, C.A., Sempere-Torres, D., Cassiraga, E.F., Gómez-Hernández, J.J., 2009. A
675 non-parametric automatic blending methodology to estimate rainfall fields from rain gauge
676 and radar data. *Adv. Water Resour.* 32, 986-1002.

677 Verworn, A., Haberlandt, U., 2011. Spatial interpolation of hourly rainfall - effect of
678 additional information, variogram inference and storm properties. *Hydrol. Earth Syst. Sci.*
679 15, 569-584.

680 Villarini, G., Krajewski, W.F., 2010. Review of the different sources of uncertainty in radar-
681 based estimates of rainfall, *Surv. Geophys.* 31, 107–129.

682 Wang, L.P., Ochoa-Rodríguez, S., Simões, N.E., Onof, C., Maksimovic, C., 2013. Radar-
683 raingauge data combination techniques: a revision and analysis of their suitability for
684 urban hydrology, *Water Sci. Technol.* 68 (4), 737-747. doi:10.2166/wst.2013.300.

685 Waters, D.K., Carroll, C., Ellis, R., Hateley, L., McCloskey, G.L., Packett, R., Dougall, C.,
686 Fentie, B., 2014. Modelling reductions of pollutant loads due to improved management

687 practices in the Great Barrier Reef catchments – Whole of GBR, Technical Report,
688 Volume 1, Queensland Department of Natural Resources and Mines, Toowoomba,
689 Queensland (ISBN: 978-1-7423-0999).
690 Yao, T., Journel, A.G., 1998. Automatic modeling of (cross) covariance tables using fast
691 Fourier transform. Math. Geol. 30 (6), 589-615.
692
693

Fig. 1. Study area: red solid irregular boundary defines the Brisbane river catchment; red dashed boundary defines the 20 km buffer; blue dots are the daily gauges inside the catchment boundary and the green dots are those within its 20 km buffer; violet square boundary defines the 128 km x 128 km region within the range of Mt. Stapylton (near Brisbane) weather radar station (red solid circle), Australia; black dots are gauges within the square region but not used as catchment data; red triangles are the temperature stations.

Fig. 2. Properties of the 2D anisotropic exponential model; the displayed contours are at 0.05, 0.1, 0.2 and 0.6 correlations; the parameters are; $\theta=120^\circ$, $L_u=30\text{km}$, $L_v=18\text{km}$, $a_n=1$, and $L_R=90\text{km}$ shown as blue dashed line.

Fig. 3. The left panel shows the radar rainfall (mm) of 2009-04-05, red dot indicates the location of Mt. Stapylton weather radar; and right panel compares the fitted correlogram (red) with the empirical correlogram (white); radar-derived parameters were $\theta=128.5^\circ$, $L_{ur}=38.2\text{km}$, $L_{vr}=19.2\text{km}$; correlation = 0.976.

Fig. 4. Comparison of radar-derived correlogram parameters and fitted (curve) probability density functions; left panel is for the Harmonic distribution fitted to the direction (θ_r); middle and right panels are for the Box-Cox power exponential (BCPE) distribution fitted to major (L_{ur}) and minor (L_{vr}) axis lengths; correlation: $\theta_r=0.967$, $L_{ur}=0.993$, $L_{vr}=0.988$.

Fig. 5. Bivariate rank distribution of the radar-derived correlogram major (L_{ur}) and minor (L_{vr}) axes lengths, and the variation of the Gumbel copula parameter with temperature (correlation = 0.985).

Fig. 6. Variation of the radar-derived correlogram parameters with the covariates; red triangles are the top 20 extreme wet days; green dashed lines are perfect agreement; blue solid circles are the mean of the binned data.

Fig. 7. Leave-one-out cross-validation for the square region between 2009-01-01 and 2015-06-30: comparison of RMSE (left panels) and mean standard deviation SD (right panels) of radar-derived, gauge-derived and copula-derived correlogram parameters; blue lines are perfect agreement; red triangles are the top 20 extreme wet days.

Fig. 8. Leave-one-out cross-validation for 2000-01-31 to 2008-12-31: comparison of the mean standard deviation (SD) using gauge-derived and copula-derived correlogram parameters; blue lines are perfect agreement; red triangles are the top 20 extreme wet days.

Fig. 9. Upper panels: sensitivity of simulated median rain to Gaussian quantile standard deviation (SD); lower panels: variation of simulated median rain SD to the Gaussian quantile SD and the effect on the prediction limits (PL) for 3 events of different median rainfall amounts.

Fig. 10. Variation of the slope of the Gaussian quantile standard deviation (SD) versus simulated rainfall standard deviation (RSD) and the rainfall amount; the median values are indicated in dotted lines.

Fig. 11. Leave-one-out cross-validation for Brisbane River catchment: comparison of the median estimates (solid blue line) and the observed values (red diamonds within the Brisbane river catchment and black triangles within its buffer zone) for one set of sampled correlogram parameters. The blue dashed lines represent the 90% prediction limits; gauge numbers are in increasing order of the median prediction for easy visualisation.

Fig. 1. Study area: red solid irregular boundary defines the Brisbane river catchment; red dashed boundary defines the 20 km buffer; blue dots are the daily gauges inside the catchment boundary and the green dots are those within its 20 km buffer; violet square boundary defines the 128 km x 128 km region within the range of Mt. Stapylton (near Brisbane) weather radar station (red solid circle), Australia; black dots are gauges within the square region but not used as catchment data; red triangles are the temperature stations.

Fig. 1. Study area: red solid irregular boundary defines the Brisbane river catchment; red dashed boundary defines the 20 km buffer; blue dots are the daily gauges inside the catchment boundary and the green dots are those within its 20 km buffer; violet square boundary defines the 128 km x 128 km region within the range of Mt. Stapylton (near Brisbane) weather radar station (red solid circle), Australia; black dots are gauges within the square region but not used as catchment data; red triangles are the temperature stations.

Fig. 2. Properties of the 2D anisotropic exponential model; the displayed contours are at 0.05, 0.1, 0.2 and 0.6 correlations; the parameters are; $\theta=120^\circ$, $L_u=30\text{km}$, $L_v=18\text{km}$, $a_n=1$, and $L_R=90\text{km}$ shown as blue dashed line.

Fig. 3. The left panel shows the radar rainfall (mm) of 2009-04-05, red dot indicates the location of Mt. Stapylton weather radar; and right panel compares the fitted correlogram (red) with the empirical correlogram (white); radar-derived parameters were $\theta=128.5^\circ$, $L_{ur}=38.2\text{km}$, $L_{vr}=19.2\text{km}$; correlation = 0.976.

Fig. 4. Comparison of radar-derived correlogram parameters and fitted (curve) probability density functions; left panel is for the Harmonic distribution fitted to the direction (θ_r); middle and right panels are for the Box-Cox power exponential (BCPE) distribution fitted to major (L_{ur}) and minor (L_{vr}) axis lengths; correlation: $\theta_r=0.967$, $L_{ur}=0.993$, $L_{vr}=0.988$.

Fig. 5. Bivariate rank distribution of the radar-derived correlogram major (L_{ur}) and minor (L_{vr}) axes lengths, and the variation of the Gumbel copula parameter with temperature (correlation = 0.985).

Fig. 6. Variation of the radar-derived correlogram parameters with the covariates; red triangles are the top 20 extreme wet days; green dashed lines are perfect agreement; blue solid circles are the mean of the binned data.

Fig. 7. Leave-one-out cross-validation for the square region between 2009-01-01 and 2015-06-30: comparison of RMSE (left panels) and mean standard deviation SD (right panels) of radar-derived, gauge-derived and copula-derived correlogram parameters; blue lines are perfect agreement; red triangles are the top 20 extreme wet days.

Fig. 8. Leave-one-out cross-validation for 2000-01-31 to 2008-12-31: comparison of the mean standard deviation (SD) using gauge-derived and copula-derived correlogram parameters; blue lines are perfect agreement; red triangles are the top 20 extreme wet days.

Fig. 9. Upper panels: sensitivity of simulated median rain to Gaussian quantile standard deviation (SD); lower panels: variation of simulated median rain SD to the Gaussian quantile SD and the effect on the prediction limits (PL) for 3 events of different median rainfall amounts.

Fig. 10. Variation of the slope of the Gaussian quantile standard deviation (SD) versus simulated rainfall standard deviation (RSD) and the rainfall amount; the median values are indicated in dotted lines.

Fig. 11. Leave-one-out cross-validation for Brisbane River catchment: comparison of the median estimates (solid blue line) and the observed values (red diamonds within the Brisbane river catchment and black triangles within its buffer zone) for one set of sampled correlogram parameters. The blue dashed lines represent the 90% prediction limits; gauge numbers are in increasing order of the median prediction for easy visualisation.

Table 1

Basic statistics of the 918 wet days analysed between 2009-01-01 and 2015-06-30 within the square region (GAU – collocated)

	total number of gauges	wetted area ratio (rain >= 1mm)			average wet gauge amount (mm) (rain >= 1mm)			maximum wet gauge amount (mm)			number rain >
		GAU	COL	RAD	GAU	COL	RAD	GAU	COL	RAD	GAU
min	147	0.05	0.00	0.01	1.76	0.00	1.08	10.00	0.40	2.30	9
mean	176	0.59	0.45	0.45	11.43	8.28	8.64	55.62	28.91	45.18	103
max	200	1.00	1.00	1.00	208	126	124	653	342	467	197

Table 2

Pearson correlation between the radar-derived correlogram parameters and other variables for the square region: 06-30

variable	L_{ur}	L_{vr}	θ_r
L_{ur}	1.000	0.902	-0.036
L_{vr}	0.902	1.000	-0.088
θ_r	-0.036	-0.088	1.000
L_{ug}	0.297	0.299	-0.065
L_{vg}	0.289	0.320	-0.103
θ_g	0.066	0.050	0.154
L_{uc}	0.867	0.835	-0.033
L_{vc}	0.839	0.903	-0.064
θ_c	0.039	-0.039	0.461
T_m	-0.210	-0.235	0.028
MAG	0.067	0.031	-0.034
MAX	0.028	-0.006	-0.062
NWG	-0.007	-0.029	0.024
WAR	-0.006	-0.029	0.018
MWG	0.070	0.031	-0.036

last subscript: r – radar, g – gauge, c – collocated; T_m – maximum temperature; L_u – major axis length; L_v – minor axis length; MAG – mean amount of rainfall of all gauges; MAX – maximum amount of rainfall of all gauges; NWG – wetted area ratio of all gauges; MWG – mean amount of rainfall of wet gauges

Table 3

Estimated linear link function coefficients of parameters of the Box-Cox power exponential (BCPE) distribution correlogram axes lengths for the square region

linear eqn component	L_{ur}		L_{vr}	
	Estimate	Std. Error	Estimate	Std. Error
$\log(\mu):m_1$	3.9705	0.0973	3.6556	0.0992
$\log(\mu):m_2$	0.0134	0.0014	0.0182	0.0017
$\log(\mu):m_3$	-0.0202	0.0033	-0.0224	0.0034
$\log(\sigma)$	-1.0058	0.0222	-0.9435	0.0203
$v:n_1$	-0.6348	0.1807	-0.7609	0.1476
$v:n_2$	0.0243	0.0075	0.0310	0.0079
$\log(\tau)$	0.7921	0.0758	0.9720	0.0794

L_{ur} – radar major axis length; L_{vr} – radar minor axis length; μ - location parameter; σ - scale parameter; v - skew parameter

Table 4 Top 20 extreme wet days (with good radar data) properties between 2009-01-01 and 2015-06-30 within

Date	MAX	MAG	MWG	WAR	Temp	L _{ur}	L _{vr}	θ _r	L _{ug}	L _{vg}	
2013-01-28	653.0	207.8	207.8	1.000	27.8	57.3	17.1	84.1	20.4	10.2	1
2010-02-07	372.0	103.2	103.2	1.000	27.1	26.3	20.3	109.3	25.8	16.9	1
2009-05-21	362.4	92.9	93.4	0.995	20.8	68.6	42.1	36.5	15.3	11.4	3
2013-01-31	350.0	9.0	15.8	0.567	28.1	29.4	20.9	132.0	30.4	27.3	1
2012-01-25	340.3	140.5	140.5	1.000	26.2	87.2	57.2	101.8	48.8	34.2	7
2012-01-26	338.8	30.1	31.2	0.965	30.4	35.3	26.3	133.0	31.5	29.5	1
2014-03-28	319.4	127.3	127.3	1.000	28.5	43.3	24.0	125.2	20.4	13.5	1
2011-01-12	308.8	52.3	52.6	0.994	29.2	28.8	24.2	149.6	51.1	33.4	8
2014-12-28	292.0	37.7	37.9	0.994	25.3	43.7	33.9	0.6	32.9	27.7	4
2015-01-24	291.0	70.2	73.4	0.956	30.1	82.5	71.2	105.2	14.5	13.6	1
2010-10-11	281.0	72.8	73.6	0.989	21.4	47.8	40.4	67.0	25.8	19.1	5
2011-01-11	280.0	61.9	62.7	0.989	27.6	74.9	58.0	91.9	26.5	17.6	7
2010-10-12	274.0	24.8	25.3	0.978	22.5	43.9	29.1	124.9	15.3	9.9	1
2015-01-27	270.0	7.6	17.3	0.433	30.6	42.7	29.0	135.2	17.4	13.8	1
2010-02-08	267.0	26.8	27.6	0.974	29.4	25.6	11.6	15.8	19.7	11.4	6
2015-02-21	258.0	63.3	63.8	0.994	26.2	84.4	72.4	71.8	18.7	16.3	1
2015-05-02	252.4	118.3	118.3	1.000	24.8	70.6	59.7	38.8	17.9	10.4	1
2009-05-22	249.0	19.2	25.0	0.767	23.5	23.3	15.7	8.1	50.7	40.8	1
2015-02-23	245.6	11.5	15.4	0.749	29.6	42.4	27.0	127.5	17.2	13.7	1
2014-03-31	236.0	5.3	6.2	0.855	28.8	22.8	16.3	100.2	17.1	17.1	1

last subscript: r – radar, g – gauge; L_u – major axis length; L_v – minor axis length; θ – correlogram direction; Ter – Terrestrial; DIST – marginal distribution; MAG – mean amount of rainfall of all gauges; MAX – maximum within-day rainfall; ratio of all gauges; MWG – mean amount of rainfall of wet gauges; LL – Log-Logistic; KA – Kappa; GP – Generalized Pareto; LN – Log-Normal; WE – Weibull; GA – Gamma

Table 5

Distribution of the marginal distributions over the 918 wet days for the period 2009-01-01 to 2015-06-30 for the

distribution	Gamma	Generalised Pareto	Gumbel	Kappa	Log-Logistic	Log
gauge	111	99	23	109	202	
radar	37	31	3	16	699	

Table 6

Distribution of Kolmogorov–Smirnov (K-S) test p-values comparing the radar-derived and copula-derived correlograms

cumulative probability	minimum	5%	25%	50%	75%	maximum
$L_{ur}-L_{vr}$	0.484	0.539	0.606	0.666	0.735	0.800
θ_r	0.118	0.265	0.517	0.787	0.885	0.950
L_{ur}	0.020	0.104	0.285	0.463	0.711	0.950
L_{vr}	0.011	0.045	0.146	0.334	0.480	0.650

Table 7 Top 20 extreme wet days' properties between 2000-01-01 and 2008-12-31 within the catchment

Date	MAX	MAG	MWG	WAR	Temp	L _{ug}	L _{vg}	θ_g
2008-11-20	250.0	88.6	88.6	1.000	29.2	29.5	12.9	168.6
2000-11-01	224.0	10.6	12.3	0.859	25.8	56.6	45.5	139.8
2001-03-10	218.0	31.6	40.7	0.775	30.0	37.3	20.2	108.0
2004-03-06	191.8	90.0	92.2	0.977	26.4	32.1	24.1	65.2
2001-02-02	190.8	87.6	88.5	0.989	25.2	29.1	19.6	135.4
2004-11-08	174.8	27.4	30.0	0.914	27.1	47.8	42.0	116.0
2007-11-24	174.4	33.6	34.8	0.966	25.0	41.7	19.7	65.0
2003-02-04	170.0	29.4	30.1	0.978	22.9	87.1	66.3	32.9
2003-02-25	170.0	24.1	24.9	0.966	26.3	61.4	53.9	93.5
2008-11-21	149.0	12.7	15.0	0.847	30.6	33.3	30.1	178.5
2000-10-31	139.4	20.7	21.0	0.988	22.5	70.6	56.1	171.5
2001-11-12	136.8	24.8	27.2	0.913	28.4	36.8	28.4	118.3
2001-03-11	133.0	14.2	17.3	0.818	27.8	40.1	29.6	94.2
2005-10-25	131.4	20.0	20.7	0.966	29.8	42.3	35.7	154.8
2004-01-15	131.2	36.5	36.5	1.000	27.1	20.9	14.1	117.4
2008-06-03	130.0	44.4	44.4	1.000	23.0	45.8	39.1	72.2
2004-01-31	126.0	6.6	9.4	0.703	33.1	27.4	14.9	162.6
2006-02-13	124.0	14.7	16.4	0.891	32.4	23.3	20.7	105.6
2008-02-04	122.0	27.3	27.3	1.000	26.5	17.7	13.2	84.3
2001-03-12	121.0	5.3	13.0	0.402	29.2	31.8	23.6	148.7

DIST_g – gauged distribution; LL – Log-Logistic; KA – Kappa; GP – Generalised Pareto; GU – Gumbel; LN – Log-Normal; WE – Weibull; axis length; L_{vg} – gauge minor axis length; θ_g – correlogram direction; Temp – maximum temperature; MAG – mean amount of rainfall within-day rainfall amount; WAR – wetted area ratio of all gauges; MWG – mean amount of rainfall of wet gauges; LL – Log-Logistic; Pareto; GU – Gumbel; LN – Log-Normal; WE – Weibull; GA – Gamma

ys analysed between 2009-01-01 and 2015-06-30 within the square region (GAU – gauged; RAD – radar; COL

area ratio (rain >=		average wet gauge amount (mm) (rain >= 1mm)			maximum wet gauge amount (mm)			number of gauges with rain >= 1 mm		
COL	RAD	GAU	COL	RAD	GAU	COL	RAD	GAU	COL	RAD
0.00	0.01	1.76	0.00	1.08	10.00	0.40	2.30	9	0	166
0.45	0.45	11.43	8.28	8.64	55.62	28.91	45.18	103	78	7450
1.00	1.00	208	126	124	653	342	467	197	198	16384

radar-derived correlogram parameters and other variables for the square region: data from 2009-01-01 to 2015-

variable	L_{ur}	L_{vr}	θ_r
L_{ur}	1.000	0.902	-0.036
L_{vr}	0.902	1.000	-0.088
θ_r	-0.036	-0.088	1.000
L_{ug}	0.297	0.299	-0.065
L_{vg}	0.289	0.320	-0.103
θ_g	0.066	0.050	0.154
L_{uc}	0.867	0.835	-0.033
L_{vc}	0.839	0.903	-0.064
θ_c	0.039	-0.039	0.461
T_m	-0.210	-0.235	0.028
MAG	0.067	0.031	-0.034
MAX	0.028	-0.006	-0.062
NWG	-0.007	-0.029	0.024
WAR	-0.006	-0.029	0.018
MWG	0.070	0.031	-0.036

e, c – collocated; T_m – maximum temperature; L_u – major axis length; L_v – minor axis length; θ – correlogram of rainfall of all gauges; MAX – maximum amount of rainfall of all gauges; NWG – number of wet gauges; gauges; MWG – mean amount of rainfall of wet gauges

efficients of parameters of the Box-Cox power exponential (BCPE) distribution fitted to the radar-derived square region

linear eqn component	L _{ur}		L _{vr}	
	Estimate	Std. Error	Estimate	Std. Error
log(μ):m ₁	3.9705	0.0973	3.6556	0.0992
log(μ):m ₂	0.0134	0.0014	0.0182	0.0017
log(μ):m ₃	-0.0202	0.0033	-0.0224	0.0034
log(σ)	-1.0058	0.0222	-0.9435	0.0203
v:n ₁	-0.6348	0.1807	-0.7609	0.1476
v:n ₂	0.0243	0.0075	0.0310	0.0079
log(τ)	0.7921	0.0758	0.9720	0.0794

– radar minor axis length; μ - location parameter; σ - scale parameter; v - skewness parameter; τ - kurtosis

ys (with good radar data) properties between 2009-01-01 and 2015-06-30 within the square region

AG	MWG	WAR	Temp	L _{ur}	L _{vr}	θ _r	L _{ug}	L _{vg}	θ _g	DIST _r	DIST _g
7.8	207.8	1.000	27.8	57.3	17.1	84.1	20.4	10.2	105.7	LL	LL
3.2	103.2	1.000	27.1	26.3	20.3	109.3	25.8	16.9	120.4	LL	LN
9.9	93.4	0.995	20.8	68.6	42.1	36.5	15.3	11.4	34.3	LL	LN
0	15.8	0.567	28.1	29.4	20.9	132.0	30.4	27.3	129.8	LL	GP
0.5	140.5	1.000	26.2	87.2	57.2	101.8	48.8	34.2	77.4	LL	WE
0.1	31.2	0.965	30.4	35.3	26.3	133.0	31.5	29.5	127.3	WE	LN
7.3	127.3	1.000	28.5	43.3	24.0	125.2	20.4	13.5	117.5	LL	LL
3.3	52.6	0.994	29.2	28.8	24.2	149.6	51.1	33.4	88.4	LL	LL
7.7	37.9	0.994	25.3	43.7	33.9	0.6	32.9	27.7	42.6	LL	LL
0.2	73.4	0.956	30.1	82.5	71.2	105.2	14.5	13.6	164.4	LL	LL
8.8	73.6	0.989	21.4	47.8	40.4	67.0	25.8	19.1	50.1	LL	WE
9.9	62.7	0.989	27.6	74.9	58.0	91.9	26.5	17.6	79.1	LL	GA
8.8	25.3	0.978	22.5	43.9	29.1	124.9	15.3	9.9	101.0	LL	KA
6	17.3	0.433	30.6	42.7	29.0	135.2	17.4	13.8	149.3	LL	GP
8.8	27.6	0.974	29.4	25.6	11.6	15.8	19.7	11.4	62.1	LL	LN
3.3	63.8	0.994	26.2	84.4	72.4	71.8	18.7	16.3	129.6	LL	GU
8.3	118.3	1.000	24.8	70.6	59.7	38.8	17.9	10.4	107.2	LL	GA
0.2	25.0	0.767	23.5	23.3	15.7	8.1	50.7	40.8	149.1	LL	GP
5	15.4	0.749	29.6	42.4	27.0	127.5	17.2	13.7	153.7	LL	LL
3	6.2	0.855	28.8	22.8	16.3	100.2	17.1	17.1	134.1	LL	LN

e; L_u – major axis length; L_v – minor axis length; θ – correlogram direction; Temp – maximum temperature;
AG – mean amount of rainfall of all gauges; MAX – maximum within-day rainfall amount; WAR – wetted area
n amount of rainfall of wet gauges; LL – Log-Logistic; KA – Kappa; GP – Generalised Pareto; GU – Gumbel;
ll; GA – Gamma

tributions over the 918 wet days for the period 2009-01-01 to 2015-06-30 for the square region

ma	Generalised Pareto	Gumbel	Kappa	Log-Logistic	Log-Normal	Weibull
1	99	23	109	202	263	111
7	31	3	16	699	52	80

irnov (K-S) test p-values comparing the radar-derived and copula-derived correlogram parameters, 100 repetitions.

cumulative bability	minimum	5%	25%	50%	75%	maximum
L _{vr}	0.484	0.539	0.606	0.666	0.735	0.836
	0.118	0.265	0.517	0.787	0.885	0.998
	0.020	0.104	0.285	0.463	0.711	0.999
	0.011	0.045	0.146	0.334	0.480	0.912

s' properties between 2000-01-01 and 2008-12-31 within the catchment

	MAX	MAG	MWG	WAR	Temp	L _{ug}	L _{vg}	θ _g	DIST _g
	250.0	88.6	88.6	1.000	29.2	29.5	12.9	168.6	LL
	224.0	10.6	12.3	0.859	25.8	56.6	45.5	139.8	KA
	218.0	31.6	40.7	0.775	30.0	37.3	20.2	108.0	LN
	191.8	90.0	92.2	0.977	26.4	32.1	24.1	65.2	GA
	190.8	87.6	88.5	0.989	25.2	29.1	19.6	135.4	GU
	174.8	27.4	30.0	0.914	27.1	47.8	42.0	116.0	LN
	174.4	33.6	34.8	0.966	25.0	41.7	19.7	65.0	WE
	170.0	29.4	30.1	0.978	22.9	87.1	66.3	32.9	LN
	170.0	24.1	24.9	0.966	26.3	61.4	53.9	93.5	KA
	149.0	12.7	15.0	0.847	30.6	33.3	30.1	178.5	LL
	139.4	20.7	21.0	0.988	22.5	70.6	56.1	171.5	LN
	136.8	24.8	27.2	0.913	28.4	36.8	28.4	118.3	GU
	133.0	14.2	17.3	0.818	27.8	40.1	29.6	94.2	WE
	131.4	20.0	20.7	0.966	29.8	42.3	35.7	154.8	KA
	131.2	36.5	36.5	1.000	27.1	20.9	14.1	117.4	LN
	130.0	44.4	44.4	1.000	23.0	45.8	39.1	72.2	WE
	126.0	6.6	9.4	0.703	33.1	27.4	14.9	162.6	KA
	124.0	14.7	16.4	0.891	32.4	23.3	20.7	105.6	WE
	122.0	27.3	27.3	1.000	26.5	17.7	13.2	84.3	GA
	121.0	5.3	13.0	0.402	29.2	31.8	23.6	148.7	LN

-Logistic; KA – Kappa; GP – Generalised Pareto; GU – Gumbel; LN – Log-Normal; WE – Weibull; GA – Gamma; L_{ug} – gauge major
length; θ_g – correlogram direction; Temp – maximum temperature; MAG – mean amount of rainfall of all gauges; MAX – maximum
ated area ratio of all gauges; MWG – mean amount of rainfall of wet gauges; LL – Log-Logistic; KA – Kappa; GP – Generalised
al; WE – Weibull; GA – Gamma

***Highlights (3 to 5 bullet points (maximum 85 characters including spaces per bullet point))**

- Radar correlogram better predicts the spatial structure than gauged correlogram.
- A second-order harmonic distribution is used to model radar correlogram direction.
- Copula joint distribution for radar correlogram axes length parameters.
- Gauged correlogram and temperature as covariates for predicting radar correlogram.
- Uncertainty is reduced in over 90% cases using radar 2D correlogram parameter sets.