HYDROLOGICAL PROCESSES
Hydrol. Process. **24**, 2111–2124 (2010)
Published online 24 March 2010 in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/hyp.7632

Copula-based uncertainty modelling: application to multisensor precipitation estimates

Amir AghaKouchak,[1]* András Bárdossy[1] and Emad Habib[2]

[1] *Institute of Hydraulic Engineering, University of Stuttgart, Pfaffenwaldring 61, 70569 Stuttgart, Germany*
[2] *Department of Civil Engineering, University of Louisiana at Lafayette, PO Box 42291, Lafayette, LA 70504, USA*

Abstract:

The multisensor precipitation estimates (MPE) data, available in hourly temporal and 4 km × 4 km spatial resolution, are produced by the National Weather Service and mosaicked as a national product known as Stage IV. The MPE products have a significant advantage over rain gauge measurements due to their ability to capture spatial variability of rainfall. However, the advantages are limited by complications related to the indirect nature of remotely sensed precipitation estimates. Previous studies confirm that efforts are required to determine the accuracy of MPE and their associated uncertainties for future use in hydrological and climate studies. So far, various approaches and extensive research have been undertaken to develop an uncertainty model. In this paper, an ensemble generator is presented for MPE products that can be used to evaluate the uncertainty of rainfall estimates. Two different elliptical copula families, namely, Gaussian and t-copula are used for simulations. The results indicate that using t-copula may have significant advantages over the well-known Gaussian copula particularly with respect to extremes. Overall, the model in which t-copula was used for simulation successfully generated rainfall ensembles with similar characteristics to those of the ground reference measurements. Copyright © 2010 John Wiley & Sons, Ltd.

KEY WORDS rainfall uncertainty; ensemble generation; multivariate simulation; multisensor precipitation estimates; t-copula; Gaussian copula

Received 7 October 2009; Accepted 25 January 2010

INTRODUCTION

With accurate information about surface rainfall and its spatial and temporal distributions, hydrologists and meteorologists have the potential to improve hydrological predictions and global climate studies. Hydrological and climate studies have long been relied on rain gauge measurements. Their applicability and usefulness, however, is limited due to its lack of spatial coverage and inability to accurately represent spatial variability of rainfall. Recent technological advances in the field of remote sensing have led to an increase in available rainfall data on a regional and global scale. Particularly, National Weather Service (NWS) Next Generation Weather Radar (NEXRAD) and several National Oceanic and Atmospheric Administration (NOAA) and National Aeronautics and Space Administration (NASA) sponsored satellite missions provide the hydrological community with amplitude of new rainfall data. The NEXRAD system was installed almost two decades ago across the entire continental United States (Crum and Alberty (1993); Young *et al.* (2000)). Since then, the River Forecast Center (RFC) of the National Weather Service (NWS) continuously provides multisensor precipitation estimates (MPE), also known as Stage IV data, in hourly temporal and 4 km × 4 km spatial resolution. Although MPE data provide much higher spatial resolution than was previously available from rain gauge networks, they suffer from different types of error. So far, various procedures have been developed and implemented to improve the quality of MPE products (e.g. merging of radar estimates and gage measurements, mean-field and local bias adjustments). It is worth mentioning that the uncertainties associated with gauge data (e.g. undercatch in windy situations, communication errors) and radar rainfall estimates (e.g. beam overshooting and blockage, measurement biases, melting effects, see Austin (1987), Young and Brunsell (2008), Westcott (2009) and Habib *et al.* (2009b) for details) will affect the quality of MPE data. Furthermore, unreliable gauge measurements, if used for radar adjustment and bias removal, may result in additional error in MPE data. Therefore, it is essential to evaluate the accuracy and uncertainty of MPE estimates, particularly due to the significance of MPE to current and future research and application in hydrology and climate studies.

The importance of the effect of uncertainty in temporal and spatial variability of rainfall in hydrological predictions and global climate studies have been highlighted in a number of studies (Seliga *et al.* (1992); Obled *et al.* (1994); Faures *et al.* (1995); Goodrich *et al.* (1995); Shah *et al.* (1996); Bell and Moore (2000); Arnaud *et al.* (2002); Syed *et al.* (2003)). In an effort to simulate radar data, Pegram and Clothier (2001) introduced the String of Bead model assuming log-normal marginal distribution for radar rainfall at the pixel scale. Seed and

* Correspondence to: Amir AghaKouchak, Department of Civil and Environmental Engineering, University of California—Irvine, Irvine, CA 92697-2175, USA. E-mail: amir.a@uci.edu

Srikanthan (2001) presented a space-time model for radar estimates using a multifractal (multiplicative cascade) approach where each level of the multiplicative cascade was linked in time with a different ARMA model. Using a reshuffling approach, Clark *et al*. (2004a) and Clark *et al*. (2004b) proposed a nonparametric rainfall simulation model that was able to capture spatial and temporal variabilities of rainfall by reshuffling the generated ensemble to match the characteristics of observed data. Germann *et al*. (2006) perturbed radar estimates with stochastically simulated radar error fields to obtain an ensemble of radar estimates assuming normal distribution for radar error. An operational approach based on empirical investigations of joint samples of radar and rain gauge data was proposed by Ciach *et al*. (2007) in which radar rainfall uncertainty consists of a systematic distortion function and a stochastic component that are estimated using a nonparametric scheme. AghaKouchak *et al*. (2010b) introduced a random error model whereby radar fields are perturbed using two error components: a proportion error field to account for errors that are proportional to the magnitude of rain rate and a purely random error field for the sum of errors from different sources. Villarini *et al*. (2009) presented a radar rainfall generator and a model to produce probability of exceeding maps of rainfall data conditioned on given radar rainfall estimates. In a recent work, AghaKouchak *et al*. (2010a) developed a model for simulation of remotely sensed rainfall estimates using v-transformed copulas. For a complete review on different rainfall simulation models, interested readers are referred to Mehrotra *et al*. (2006).

This study intends to develop an uncertainty model for MPE products using an ensemble of rainfall data which consists of a large number of stochastically generated rainfall realizations, each of which represents a rainfall event that can occur. It is worth pointing out that hydrological and meteorological phenomena are known to be multidimensional and thus require multivariate analysis as well as conditional probability distributions of variables (Genest and Favre (2007)). Classical families of multivariate distributions are commonly used for modelling joint probability distributions of several random variables. Classical multivariate distributions such as bivariate normal, log-normal and gamma are built with a number of model parameters that describe the behaviour of each random variable as well as the joint probability distribution. The main disadvantage of this approach is that modelling the dependence structure between variables is not independent of the choice of the marginal distributions (Dupuis (2007); Genest and Favre (2007)). The advent of copulas, however, allows hydrologists to avoid this restriction. The application of copulas in simulation of multivariate data, extreme value analysis and modelling dependence structure has become popular in hydrological analysis. In recent years, numerous copulas-based models have been introduced for different hydrological applications (Kelly and Krzysozto-fowicz (1997); De Michele and Salvadori (2002); Favre *et al*. (2004); Bárdossy (2006); Renard and Lang (2007); Bárdossy and Li (2008); Schölzel and Friederichs (2008); Serinaldi (2008a, b); Villarini *et al*. (2008); Zhang *et al*. (2008); AghaKouchak *et al*. (2010a)). In this paper, the well-known Gaussian copula and t-copula are used for simulation of rainfall error fields. The simulated fields are then imposed on MPE data to obtain an ensemble of precipitation estimates that can be used for uncertainty assessment. For example, the generated ensemble can then be used as input to hydrological models as possible rainfall events for different climate scenarios. Subsequent runs of a model using generated rainfall realizations would then allow an assessment of the model predictions uncertainty due to the input precipitation.

The paper is organized into seven sections. After the *Introduction* section, the study area and data resources are briefly introduced in Section *Study Area and Data Resources*. In the Section *Copulas*, the copula families used in this study are discussed. The Section *Methodolgy* is devoted to model description and simulation procedure. The details of the parameter estimation are explained in the Section *Parameter Estimation*. In Section *Rainfall Simulation*, the models are implemented for rainfall simulation and their performance are evaluated. The Section *Summary and Conclusions* summarizes the results and conclusions.

STUDY AREA AND DATA RESOURCES

The Little Washita research watershed located in Oklahoma was selected for this study. The area of the watershed is approximately 611 km^2 with the mean annual temperature of 16 °C and the average annual rainfall of approximately 760 mm (Elliott *et al*., (1993)). Stage IV rainfall estimates, provided by National Center for Atmospheric Research Earth Observing Laboratory (NCAR/EOL), are used in this study. Rainfall estimates are available with a spatial resolution of 4 km × 4 km and a temporal resolution of 1 h. The closest NEXRAD radar station is the Oklahoma City station located approximately 70 km from the centre of the watershed.

The Micronet network, operated by Agricultural Research Service (ARS), which includes 42 rain gauge

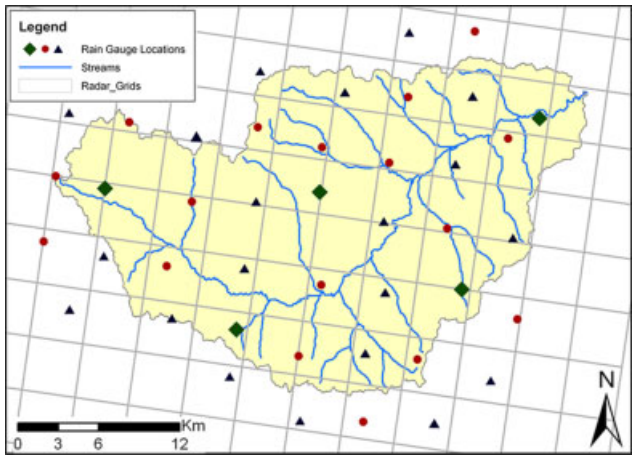

Figure 1. Little Washita watershed, Oklahoma, USA

Table I. Summary statistics of rainfall accumulations for each storm

Event ID	Date, time	Duration (h)	Mean (mm)	Standard deviation	Q10 (mm)	Q90 (mm)
Event 1	30/8/03, 07:00	12	4·3	5·5	0·8	8·8
Event 2	15/9/05, 06:00	14	3·8	4·3	1·0	10·7
Event 3	17/9/06, 15:00	10	10·1	15·5	1·0	22·7

stations, is located within and around the Little Washita watershed. The rain gauge stations are distributed almost uniformly across the watershed. The Micronet rain gauge data and MPE products can be considered independent because the gauges are not used for adjusting the MPE data. Figure 1 shows the watershed and the location of rain gauge stations and radar pixels. The Micronet network is equipped with tipping-bucket gauges that record rainfall data with a temporal resolution of 5 min and accuracy of 0·254 mm (Young *et al*. (2000)). The 5-min rain gauge accumulations are aggregated to hourly data to synchronize the rain gauge measurements with MPE data. The ground reference rainfall measurements are then used to obtain estimates of MPE error across the study area. Note that error is estimated after removing the overall bias of MPE data with respect to rain gauge measurements as described in Ciach *et al*. (2007). The presented models in this study are used for simulation of three rainfall events occurring in August 2003, September 2005 and September 2006. Table I displays summary statistics of the events including mean, standard deviation, 10 and 90% quantiles of the lumped rainfall accumulation for each storm.

COPULAS

Copulas are joint cumulative distribution functions that describe dependencies among variables independent of their marginals (Joe (1997) and Nelsen (2006)):

$$C^n(u_1, \ldots, u_n) = \Pr(U_1 \leq u_1, \ldots, U_n \leq u_n) \qquad (1)$$

where C^n is an n-dimensional joint cumulative distribution function of a multivariate random vector $U(U_1, \ldots, U_n)$ whose marginals are $u[0,1]$. Note that throughout this paper, a common statistical convention is used in which uppercase characters denote random variables and lowercase characters are their specified variables. Sklar (1959) showed that each continuous multivariate distribution F can be represented with a unique copula C:

$$F(x_1, \ldots, x_n) = C^n[F_1(x_1), \ldots, F_n(x_n)] \qquad (2)$$

The copula C^n is unique only if $F_1, \ldots, F_n$ are all continuous. Otherwise, the copula C^n is uniquely determined on $\mathrm{Ran}F_1, \ldots, \mathrm{Ran}F_n$, where Ran denotes range (Sklar (1996); Embrechts *et al*. (2001)). For proof and derivations, interested readers are referred to Sklar (1996). The Sklar theorem indicates that for multivariate distributions, the multivariate dependence structure and the univariate marginal distributions can be separated, and hence the dependence structure can be represented by a copula independent of the marginals. Having described the dependencies using a copula, a transformation function can be applied to each variable to transform the marginal distribution into the desired marginals (Nelsen (2006)):

$$C^n(u_1, \ldots, u_n) = F[F_1^{-1}(u_1), \ldots, F_n^{-1}(u_n)] \qquad (3)$$

where F is the multivariate cumulative distribution function (CDF) with marginals $F_1, \ldots, F_n$ belonging to different distribution families. In other words, using the Sklar theorem, one can simulate random variables with the same probability distribution as that of the input data while preserving the dependence structure of the variables.

It is important to remark that copulas are invariant to monotonic transformations of the variables. This is a great advantage in simulation as the variables may belong to different probability distributions. There are many families of copulas developed for different practical contexts. Each family of copulas has a number of parameters to describe the dependencies. The main difference associated with different copulas is in the detail of the dependence they represent. For instance, various copula families may differ in the part of their distributions (upper tail/lower tail) where the association is strongest/weakest. Tail dependence describes the significance of the dependence in lower left quantile or upper right quantile of a multivariate distribution function. The upper tail expresses the probability occurrence of positive large values (outliers) at multiple locations jointly. It is worth pointing out that the tail behaviour depends solely on the type of copula and not on the choice of marginal distribution. Thus, in copula-based simulation, the type of copula strongly affects the tail dependence of simulated realizations. In this work, two elliptical copulas, namely, normal copula and t-copula with two different tail behaviours are used for simulations. In the following, only the copula families used in this work are discussed. For additional information regarding different copula families, readers are referred to Joe (1997) and Nelsen (2006).

Gaussian copula

The Gaussian copula, derived from the multivariate normal distribution, is perhaps the most commonly used copula family mainly due to its simplicity. The

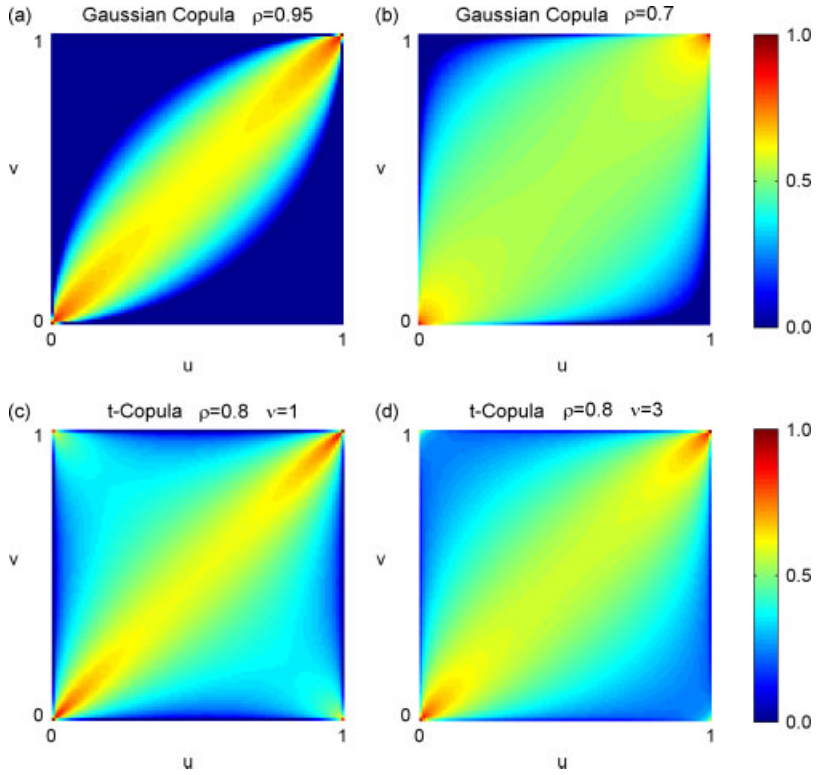

Figure 2. Bivariate copula density functions of the Gaussian copula and t-copula for different parameters

n-dimensional multivariate Gaussian copula with correlation matrix $\rho_{n\times n}$ can be expressed as (Nelsen (2006)):

$$C_\rho(u_1,\ldots,u_n) = F_\rho^n[F^{-1}(u_1),\ldots,F^{-1}(u_n)] \tag{4}$$

where F^n = multivariate Gaussian CDF whose density function is:

$$c(u_1,\ldots,u_n) = \frac{1}{\sqrt{\det\rho}}\exp\left[-\frac{1}{2}y(u)'(\rho^{-1}-I)y(u)\right] \tag{5}$$

where $y(u_i) = F^{-1}(u_i)$.

t-Copula

The t-copula, also known as Student copula, is an elliptical copula based on Student distribution that can be represented as:

$$C_{\nu,\rho}(u_1,\ldots,u_n) = t_{\nu,\rho}^n[t_\nu^{-1}(u_1),\ldots,t_\nu^{-1}(u_n)] \tag{6}$$

where t^n = multivariate Student CDF; ρ = shape matrix; ν = degrees of freedom
and

$$t_{\nu,\rho}^n(x) = \frac{1}{\sqrt{\det\rho}}\frac{\Gamma[(\nu+n)/2]}{\Gamma(\nu/2)(\pi\nu)^{n/2}} \times\int_{-\infty}^{x_1}\cdots\int_{-\infty}^{x_n}\frac{\mathrm{d}x}{\{1+[(x'\rho^{-1}x)/\nu]\}^{(\nu+n)/2}} \tag{7}$$

For $\nu > 2$, the shape matrix in Equation 6 is proportional to the correlation matrix (Malevergne and Sornette (2003)). The density function of the t-copula can be derived as (Malevergne and Sornette (2003)):

$$c(u_1,\ldots,u_n) = \frac{1}{\sqrt{\det\rho}}\frac{\Gamma[(\nu+n)/2][\Gamma(\nu/2)]^{n-1}}{\{\Gamma[(\nu+1)/2]\}^n} \times\frac{\Pi_{k=1}^n[1+(y_k^2/\nu)]^{(\nu+1)/2}}{\{1+[(y'\rho^{-1}y)/\nu]\}^{(\nu+n)/2}} \tag{8}$$

where $y_k = t_\nu^{-1}(u_k)$; t_ν = univariate Student distribution with ν degrees of freedom.

Figure 2 shows the bivariate copula density functions of the Gaussian copula and t-copula for different parameters. As shown, the density function is wider for lower correlations (compare Figure 2a and 2b). Additionally, note the difference between the density functions for different values of degrees of freedom (Figure 2c and 2d).

Both Gaussian and t-copulas are elliptical; however, they represent different tail dependence. For a multivariate distribution with n random variables $X(X_1,\ldots,X_n)$, the upper tail (λ_{up}) is described as (Melchiori (2003)):

$$\lambda_{\text{up}} = \lim_{u\longrightarrow 1}\Pr[X_1 \geq F_{X_1}^{-1}(u)|X_2 \geq F_{X_2}^{-1}(u)\ldots X_n \geq F_{X_n}^{-1}(u)] \tag{9}$$

where $F_1^{-1},\ldots,F_n^{-1}$ are the inverse cumulative distributions of the random variables $X_1,\ldots,X_n$ and $u[0,1]$. In other words, Equation 9 indicates the probability occurrence of outliers in X_1, conditioned on presence of outliers in $X_2,\ldots,X_n$. The conditional probability given in Equation 9 can be expressed as:

$$\Pr[X_1 \geq F_{X_1}^{-1}(u)|X_2 \geq F_{X_2}^{-1}(u)\ldots X_n \geq F_{X_n}^{-1}(u)]$$

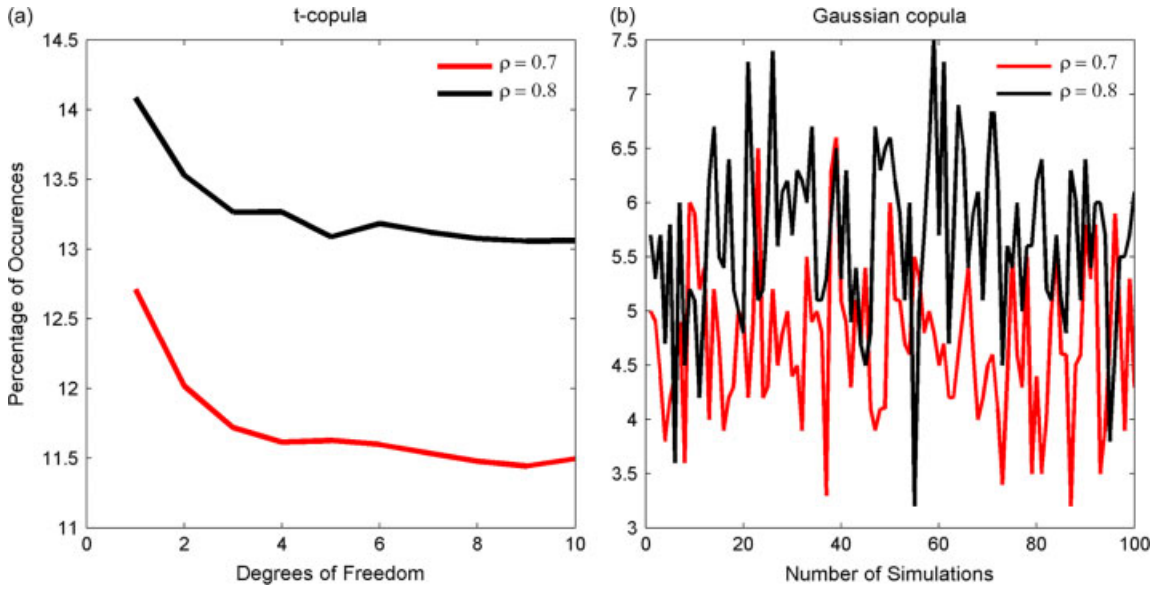

Figure 3. (a) Asymptotic dependent behaviour of the t-copula: the tail dependence becomes weaker as ν increases; (b) Tail behaviour of Gaussian copula

$$= \frac{\begin{array}{c}1 - \Pr[X_1 \le F_{X_1}^{-1}(u)] - \cdots - \Pr[X_n \le F_{X_n}^{-1}(u)] \\ +\Pr[X_1 \le F_{X_1}^{-1}(u), \ldots, X_n \le F_{X_n}^{-1}(u)]\end{array}}{1 - \Pr[X_2 \le F_{X_2}^{-1}(u)] - \cdots - \Pr[X_n \le F_{X_n}^{-1}(u)]} \tag{10}$$

It is worth remembering that $\Pr[X_1 \le F_{X_1}^{-1}(u)] = \cdots = \Pr[X_n \le F_{X_n}^{-1}(u)] = u$ (Melchiori (2003)). Having Equation 1 in mind, the second term in the numerator of Equation 10 can be described as (Embrechts *et al.* (2001)):

$$C^{(n)}(u, \ldots, u) = \Pr[X_1 \le F_{X_1}^{-1}(u), \ldots, X_n \le F_{X_n}^{-1}(u)] \tag{11}$$

where n is the dimension of the copula (number of random variables). Substituting Equation 11 into Equation 10 yields the following formulation for the upper tail:

$$\lambda_{\text{up}} = \lim_{u \longrightarrow 1} \frac{1 - nu + C^{(n)}(u, \ldots, u)}{1 - (n-1)u} \tag{12}$$

Similarly, one can derive the lower tail $\{\lambda_{\text{lo}} = \lim_{u \longrightarrow 1} \Pr[X_1 \le F_{X_1}^{-1}(u) | X_2 \le F_{X_2}^{-1}(u) \ldots X_n \le F_{X_n}^{-1}(u)]\}$ formulation as follows:

$$\lambda_{\text{lo}} = \lim_{u \longrightarrow 1} \frac{C^{(n)}(u, \ldots, u)}{(n-1)u} \tag{13}$$

The joint distribution is said to be asymptotically dependent if $0 < \lambda_{\text{up}} \le 1$ and asymptotically independent if $\lambda_{\text{up}} = 0$. The multivariate Gaussian copula shows asymptotic independence ($\lambda_{\text{up}} = 0$) regardless of the correlation between variables. That is, the extreme values in different variables occur independently even if there is a high correlation between variables. It is pointed out that for independent random variables, one could expect $\lambda_{\text{up}} = 0$; however, the converse does not hold (Malevergne and Sornette (2003)). That is, $\lambda_{\text{up}} = 0$ does not indicate that the variables are necessarily independent. While the Gaussian copula does not have upper or lower tail dependence, the t-copula shows a positive tail dependence, which indicates that when using the t-copula for simulation, one can expect larger extreme values to be generated simultaneously. The asymptotic dependent behaviour of t-copula is expected even when the variables are negatively correlated (Embrechts *et al.* (2001)). As ν increases, the tail dependence becomes weaker and the probability occurrence of joint extremes reduces. Figure 3(a) illustrates the tail behaviour of the bivariate t-copula with $\nu = 1$ to 10. The figure shows occurrences of $x > 0{\cdot}8$ (percentage) in both random vectors of the bivariate t-copula. It is noted that an increase in ν (degrees of freedom) results in less occurrences of extremes (values above a certain threshold). Figure 3(b), however, demonstrates the tail behaviour of Gaussian copula. For a constant threshold of 0·8, occurrence of extremes in Gaussian copulas is significantly less than t-copula. For additional information on elliptical copulas, the reader is referred to Fang *et al.* (2002) and Genest *et al.* (2007)).

METHODOLOGY

In the copula-based approach implemented here, multiple error fields (ε) are simulated using the Gaussian copula and t-copula and imposed on MPE fields (R_i). Equation 14 describes the general formulation of the model:

$$R_{\text{s}} = R_i + R_i \times \varepsilon \tag{14}$$

where R_{s} = simulated rainfall fields; ε = uncertainty component.

Previous studies indicate that radar rainfall error is proportional to the magnitude of rain rate (Ciach *et al.* (2007); Habib *et al.* (2008, 2009b); Villarini *et al.*

(2009)). The importance of error simulation conditioned on the magnitude of rain rate has been addressed in numerous studies (e.g. Habib *et al*. (2008), AghaKouchak *et al*. (2010b) and references therein). For example, while large rain rates are subject to large or small random error, small rainfall estimates are not subject to very large random error. This characteristics of rainfall error is included in the second term of Equation 14, where rainfall uncertainty is multiplied by the magnitude of rainfall rate ($R_i \times \varepsilon$). Using this formulation, the simulated error fields ($R_i \times \varepsilon$) will be proportional to the magnitude of rain rate. Application of a multiplicative error component for uncertainty of rating curves has been tested and verified in previous studies (e.g. Seber and Wild (1998) and Petersen-Ø verleir (2004)). For example, Petersen-Ø verleir (2004) shows that a multiplicative error term can be used to describe the uncertainty of measured rating curves. This approach is adopted here to avoid unrealistically large simulated errors for insignificant rain rates.

Previous research on properties of rainfall error shows that the spatial correlation of error is not negligible (Jordan *et al*. (2003); Ciach *et al*. (2007); Habib *et al*. (2008); Villarini *et al*. (2009)). Therefore, the error term (ε) in Equation 14 is assumed to be spatially correlated and copulas are used to describe the dependence structure. In a recent work, Habib *et al*. (2008) reported that the temporal autocorrelations of the rainfall error were rather low at the first time lag and close to zero for larger time lags. In this model, the temporal autocorrelation of rainfall error is not explicitly included in the presented models. However, AghaKouchak *et al*. (2010b) showed that when rainfall images are perturbed with error fields, the underlying temporal autocorrelation of rainfall data will be carried forward to the simulated fields.

Notice that the copula parameters are to be estimated based on the observations. In order to obtain estimates of the rainfall error across the study area, reference surface rainfall data are obtained from high-resolution rain gauge measurements over the study areas. Having estimated the parameters (see the next section), multiple univariate random fields are simulated using the selected copula C^n as described in Salvadori *et al*. (2007). The marginals of the observed error is then applied to the univariate random fields using the Sklar theorem [Equation (2)]. In this work, the empirical cumulative distribution function (CDF) of the observed error is numerically approximated using a stairs function and applied on the simulated uniform fields. It is worth remembering that copulas are invariant to monotonic transformations, and hence the simulated error fields will have the same spatial dependence structure after the transformation. For the details of multivariate simulation using copulas, interested readers are referred to Salvadori *et al*. (2007).

PARAMETER ESTIMATION

In order to simulate a random field using copulas, the copula parameters are to be estimated first. In this study, for each event, the parameters are estimated based on seasonal observations (e.g. summer 2003) in which the event occurs. The reason for this is explained further below. The first model introduced here is based on the multivariate Gaussian copula which is identified through a correlation matrix $\rho_{n\times n}$, where n is the dimension of the copula. It is worth remembering that the correlation matrix $\rho_{n\times n}$ is an $n \times n$ symmetrical matrix that can be expressed as a parameter dependent on different distance vectors of d: $\rho_{ij} = \xi(x_i - x_j)$ where the function $\xi(d)$ denotes a positive-definite correlation function such as a correlation function corresponding to the spherical variogram (Bárdossy and Li (2008)). This method is used to estimate the correlation matrix ($\rho_{n\times n}$) of the multivariate Gaussian copula. Assuming a spherical function for $\xi(d)$, the parameters of the function are estimated as described in Bárdossy and Li (2008). The function is then used to obtain the correlation matrix of the multivariate Gaussian copula. Table 2 gives the parameters of the spherical function used in this study. The parameters s_1, s_2, r_1 and r_2 represent the sills and ranges of a function consisting of two spherical components. The parameter s_0 denotes the sill of a nugget effect in the aforementioned function.

The parameters of the t-copula are the correlation matrix (similar to above) and degrees of freedom ν [Equation (6)]. For the t-copula, the degrees of freedom is estimated using the inference function for margins (IFM) method (Joe (1997); McLeish and Small (1988)) which is a special case of the generalized method of moments (GMM) with an identity weight matrix (Davidson and MascKinnon (1993)). The IFM methods can be used for estimation of univariate as well as dependence parameters from separate univariate and multivariate likelihood functions (Purwono (2005)). In this approach, the model parameters are estimated via a nonlinear system of equations, each of which being a partial derivative (score function) of a log-likelihood function from some marginal distribution of the multivariate model (Joe and Xu (1996)). Joe (1997) argues that the IFM method is highly efficient and reliable for parameter estimation and that this approach is computationally less extensive than using optimization techniques for parameter estimation. However, some statistical tests are required to make sure that the estimated parameters are reliable and that the available data is sufficient for parameter estimation. For this reason, the parameters are estimated based on four different types of data:

- Case 1: for each rainfall event separately.

Table II. The parameters of the spherical function

Event ID	s_0	s_1	s_2	r_1 (m)	r_2 (m)
Event 1	0·04	0·11	0·31	3100	14 500
Event 2	0·06	0·13	0·39	4000	13 000
Event 3	0·01	0·19	0·41	4300	14 500

- Case 2: seasonal data that each rainfall event occurred in (e.g. for Event 1 the entire summer 2003).
- Case 3: random subsets of the available data for each rainfall event (subsets are obtained from Case 1).
- Case 4: random subsets of the seasonal data for each event (subsets are obtained from Case 2).

Ideally, one may opt to estimate the parameters based on the available data for each rainfall event of interest separately. However, the data may not be sufficient for parameter estimation. Therefore, parameter estimation based on seasonal data is adopted to investigate this issue. In this approach, instead of using the data of one event, the entire season in which the event occurs will be used for parameter estimation. For example, in Event 1, which occurs in summer 2003, the entire summer (June, July and August) data is used for parameter estimation. The mean and confidence interval of 100 randomly selected subsets of the data in Cases 1 and 2 are also calculated and compared with the results obtained from Cases 1 and 2. Figure 4 plots the estimated degrees of freedom for each rainfall event using event-based (Case 1), and seasonal data (Case 2). Additionally, the graph shows the mean and confidence interval (± 3 times the standard error) of the estimated degrees of freedom for each rainfall event using random subsets of the data (Cases 3 and 4). The solid black lines represent the confidence intervals of the estimated parameters based on the random subsets of the seasonal data (Case 4), whereas the grey lines show the confidence intervals of the estimated parameters based on random subsets of the available data for each event (Case 3). Here, the standard error is used to provide an indication of amount of uncertainty. It is noted, however, that the estimation of the confidence intervals using the standard error is only acceptable when the sample size is *large* or at least *moderately large* (Bernardo and Smith (2000).

As seen in Figure 4, the estimated degrees of freedom (ν) based on the seasonal data (Case 2) fall within the confidence intervals of the estimated ν from the corresponding random subsets (Case 4). However, in rainfall Events 1 and 3, the event-based estimates (Case 1) did not fall within the range of the estimated ν from the random subsets. This may be due to the fact that the available data in each event were not sufficient to estimate the parameters reliably (Bouye *et al*. (2000); Purwono (2005)). Therefore, in this work, the model parameters are estimated based on the seasonal data (Case 2 in Figure 4).

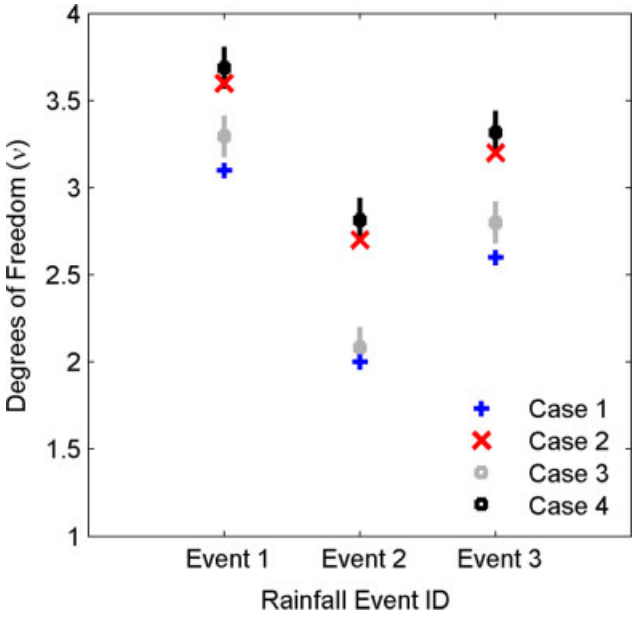

Figure 4. The estimated degrees of freedom for each rainfall event using event-based (Case 1) and seasonal data (Case 2). The solid black lines represents the confidence intervals of the random subsets of the seasonal data (Case 4), whereas the grey lines show the confidence intervals of the random subsets of the available data for each event (Case 3)

RAINFALL SIMULATION

For each rainfall event, multiple error fields are simulated using the Gaussian copula and t-copula based on the empirical CDF of observed error values. Figure 5(a) presents an MPE field occurred during Event 1, and Figure 5(b) shows the corresponding rain gauge measurements. Figure 5(c) and 5(d) displays two realizations of simulated rainfall fields using Gaussian copula and t-copula, respectively. Similar figures are presented for Events 2 and 3 (Figure 5(e)–5(l)). A visual comparison confirms that both models reproduce the rain gauge measurements (Figure 5(b), 5(f) and 5(j)) at their locations (Figure 5(c), 5(d), 5(g), 5(h), 5(k) and 5(l)). Note that in the first event (August 2003), 42 rain gauges were available whereas during the second (September 2005) and third events (September 2006), 22 gauges and 20 gauges, respectively, were in operation and available for analysis.

As mentioned previously, in this study an ensemble approach is used to describe the uncertainty associated with rainfall estimates. A statistical ensemble of a random process (here, rainfall error) is an idealization that consists of a large number of random realizations, each of which represents a possible true observation. For the rainfall events used in this study, rainfall ensembles are obtained by imposing 500 copula-based simulated error fields over MPE fields. Figure 6(a)–6(c) plots MPE data and simulated rainfall ensembles (500 realizations) using the Gaussian copula over one MPE pixel (the square-marked pixel shown in Figure 1). The solid black line represents MPE data and the grey lines (appear as shading) are simulated rainfall realizations over the length of the storm. Figure 6(d)–6(f) presents rainfall ensembles over the same pixel (the square-marked pixel shown in Figure 1) using t-copula. In both cases, the CDF of observed error is applied to the simulated fields using the Sklar Theorem [Equation (3)]. That is, the CDFs of the simulated error fields are expected to be similar to those of the observed regardless of the type of copula family.

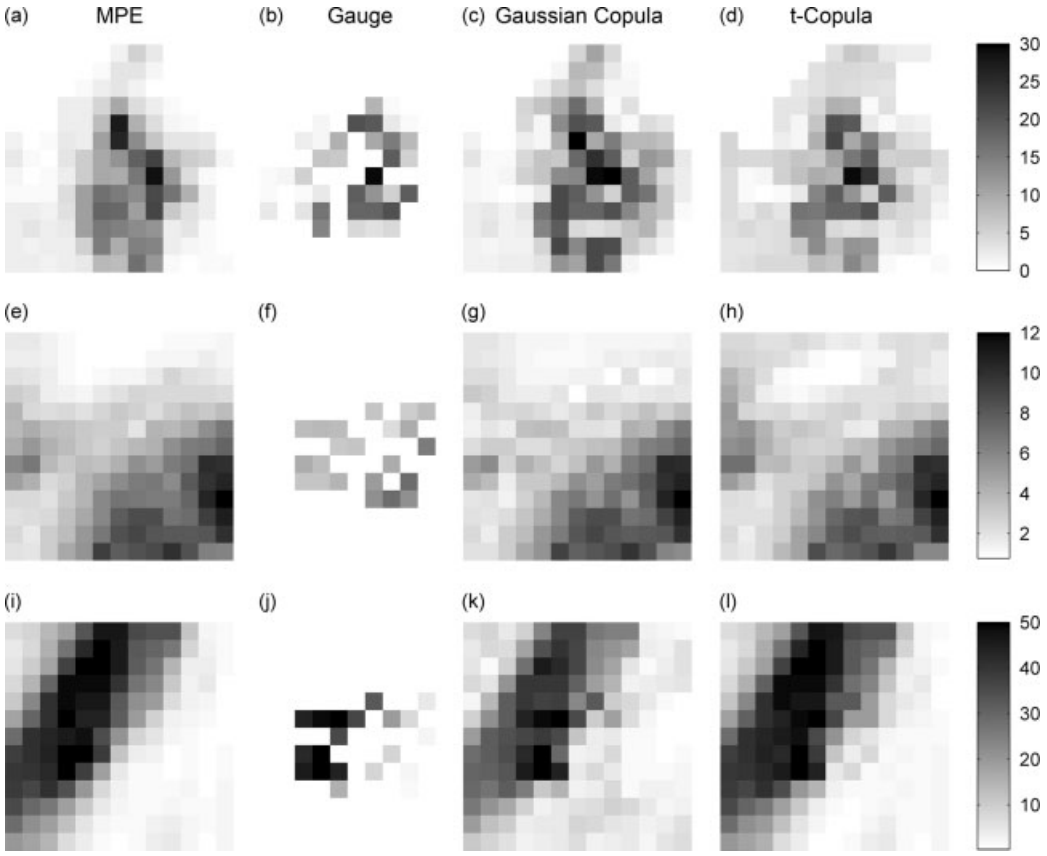

Figure 5. Event 1: (a) MPE data; (b) rain gauge measurements; (c) simulated rainfall using Gaussian copula; (d) simulated rainfall using t-copula. Figure (e)–(h) and (i)–(l) presents similar graphs for Events 2 and 3, respectively

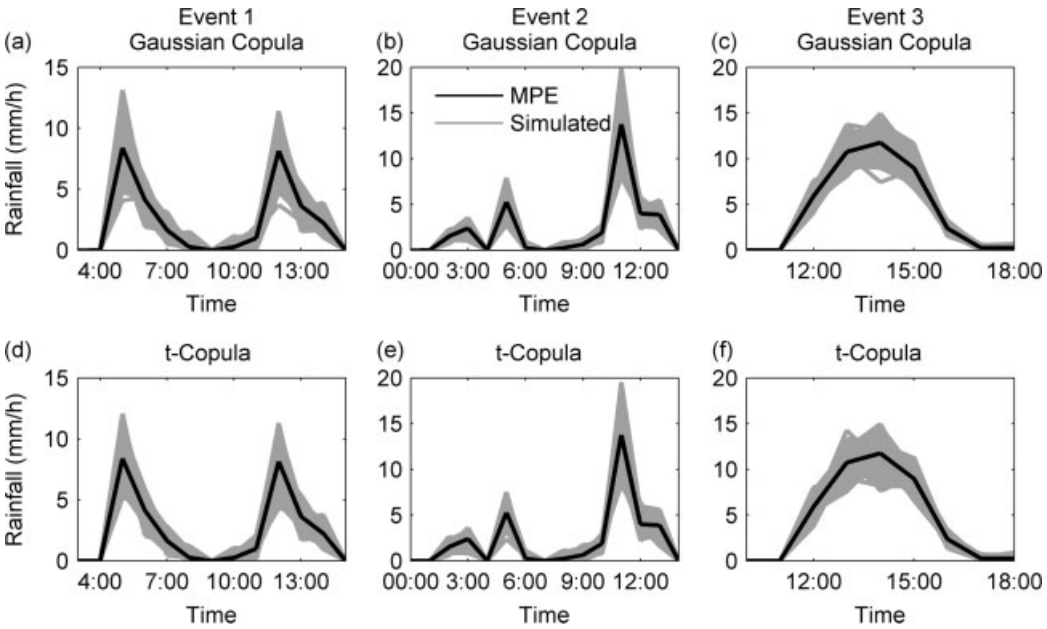

Figure 6. Rainfall ensembles (500 realizations, shaded areas) simulated using Gaussian copula (a–c) and t-copulas (d–f) over the square-marked pixel in Figure 1

In order to validate the presented models, multiple rainfall realizations were simulated with different numbers of gauges to investigate if the estimated uncertainty encompasses rain gauge measurements when less numbers of gauges were available. For the first event, Figure 7(a)–7(c) plots rainfall ensembles obtained using Gaussian copula over the *x*-marked pixel which contained one of the removed rain gauges. Figure 7(d)–7(f) presents simulated rainfall realizations obtained using t-copula over the same pixel. In Figure 7(a) and 7(d), all available gauges were included, whereas in Figure 7(b) and 7(e), 20 gauges (50% of the available

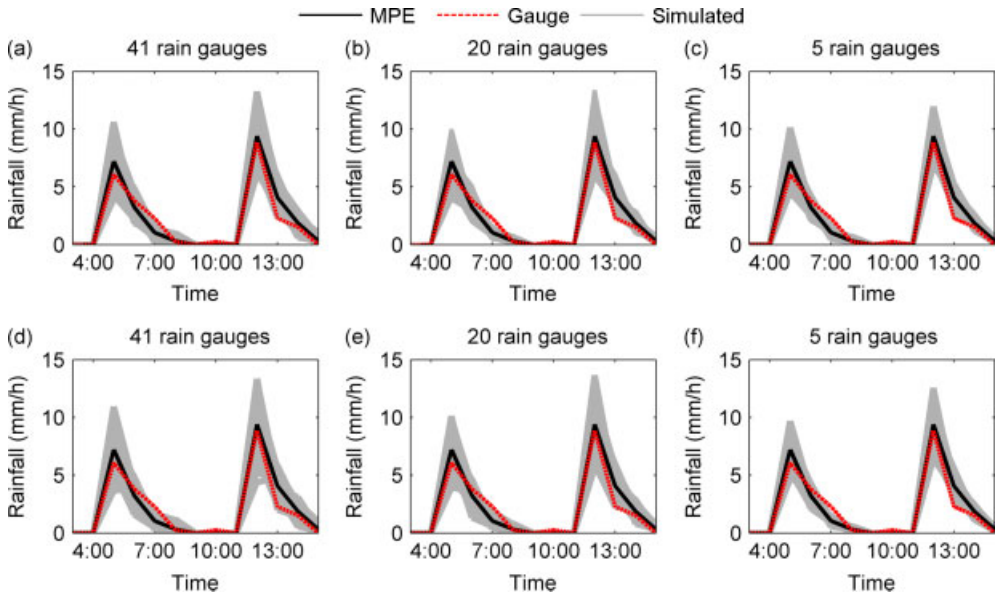

Figure 7. Simulated rainfall ensembles (shaded areas) over the *x*-marked pixel using Gaussian copula [(a, 41 gauges), (b, 20 gauges; marked with triangle and diamond in Figure 1) and (c, five gauges; marked with diamond in Figure 1)] and t-copulas [(d, 41 gauges), (e, 20 gauges) and (f, five gauges))

Table III. The number of time steps, n_{out} (%), that the estimated uncertainty did not enclose the rain gauge measurements

	All gauges		50% gauges		12% gauges	
Event ID	Gaussian	t-copula	Gaussian	t-copula	Gaussian	t-copula
Event 1	0·0	0·0	2·1	2·3	4·4	3·9
Event 2	0·0	0·0	1·3	0·9	3·8	4·3
Event 3	0·0	0·0	1·4	1·2	4·1	3·0

data, representing approximately one gauge in 31 km^2) are used for simulations. The selected rain gauges are marked with triangles and diamonds in Figure 1. In Figure 7(c) and 7(f), only five gauges (12% of the available data representing approximately one gauge in 122 km^2), marked with diamonds in Figure 1, are used for ensemble generation. The solid and dashed lines represent MPE and rain gauge measurements, whereas the grey lines (appear as shading) are simulated MPE realizations over the entire storm. One can see that with a few exceptions, the estimated uncertainties associated with the rainfall estimates reasonably enclosed the rain gauge measurements.

In order to quantitatively evaluate the rainfall ensembles, which are generated using different number of gauges, the number of time steps (n_{out}) where the rain gauge measurements fall outside the generated ensembles is counted and normalized with respect to the length of the storm. It is worth remembering that an ensemble of rainfall estimates is expected to encompass rainfall observations. Table III gives the number of time steps (n_{out} in percentage) that the estimated uncertainty did not enclose the rain gauge measurements. When all gauges are included, the estimated uncertainty for the rainfall events encompasses the rain gauge estimates entirely. When the number of rain gauges is reduced to 50%, n_{out} is increased but not significantly (most remain less than 2%). Further reduction of gauges to 12% resulted in an increase of approximately 4% in some instances. The results indicate that even with a few rain gauges, the simulated ensembles obtained from both Gaussian and t-copula models encompass approximately 96% of the rain gauge measurements. Overall, the t-copula model performs slightly better (compare columns 4 and 6 with columns 5 and 7, respectively).

The spatial dependence of rainfall is of particular interest. In th following, the Spearman rank correlation (ρ_s) matrix is used to assess the dependence structure of the simulated rainfall fields to that of the observed. The Spearman rank correlation is a nonparametric method of describing the dependencies in terms of ranks and is independent of the marginal distributions (Hollander and Wolfe (1973); Spearman (1904)):

$$\rho_{\text{s}} = \frac{n\sum \chi_i\xi_i - \sum \chi_i \sum \xi_i}{\sqrt{n\sum \chi_i^2 - \left(\sum \chi_i\right)^2}\sqrt{n\sum \xi_i^2 - \left(\sum \xi_i\right)^2}} \tag{15}$$

where χ_i = rank of x_i in X; ξ_i = rank of y_i in Y.

Figure 8(a)–8(c) presents the Spearman correlation matrices of rainfall estimates for ten MPE pixels, whereas Figure 8(d)–8(f) and 8(h)–8(i) shows the correlation

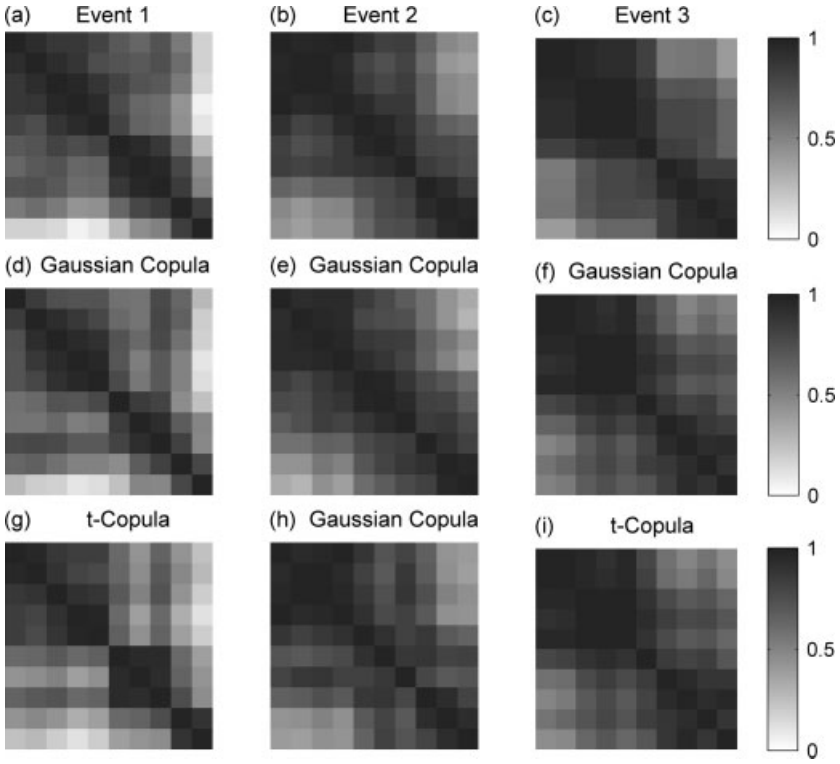

Figure 8. Rank correlation matrices. (a)–(c): Events 1–3 based on original MPE data; (d)–(f): one set of simulated data using Gaussian copula; (g)–(i): one set of simulated data using t-copula for Events 1–3

Table IV. The mean absolute error (MAE) of the correlation matrices of the simulated rainfall fields with respect to the observed

	All gauges		50% gauges		12% gauges	
Event ID	Gaussian	t-copula	Gaussian	t-copula	Gaussian	t-copula
Event 1	0·061	0·076	0·091	0·088	0·159	0·181
Event 2	0·040	0·039	0·105	0·093	0·183	0·169
Event 3	0·041	0·034	0·086	0·083	0·217	0·199

matrices of one set of simulated rainfall fields using Gaussian copula and t-copula, respectively (after imposing generated rainfall error on rainfall estimates). As shown in the provided examples, the overall spatial dependence structure is reasonably preserved. In order to analyse the correlation matrices of the simulated and observed MPE fields and to quantify the error, the mean absolute error (MAE) of the correlation matrices of the simulated MPE estimates is calculated with respect to the observed MPE data. The MAE, which is a quantity often used to measure how close simulations are to the observations, is defined as (Cox and Hinkley (1974); Bernardo and Smith (2000)):

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n} |S_i - O_i| \tag{16}$$

where S_i = simulated values; O_i = observed values; n = number of pairs (simulation, observation).

Table IV lists the MAE values for the correlation matrices of the simulated rainfall fields with respect to the correlation matrix of the observed MPE data. In this table, the MAE is also provided for the cases where the number of gauges is reduced (50% and 12%). As one may intuitively expect, the table shows that the MAE increases as the number of rain gauges decreases. Furthermore, the table indicates that even with few rain gauges, the MAE values are rather low for both Gaussian and t-copula models, meaning the presented models can be applied even with a few rain gauges available.

The model performance is also investigated with respect to the extremes. To investigate this issue, the 90th percentile of the rainfall estimates is assumed as the extreme value thresholds. Then, the number of occurrences of rainfall values above the thresholds in the simulated rainfall fields are compared with those of the initial rainfall estimates. Figure 9(a)–9(c) presents the total number of occurrences above the 90th percentile of MPE data (solid black lines), mean number of occurrences above the same threshold in 500 realizations simulated using the Gaussian copula (dashed lines) and the number of occurrences above the 90th percentile of rainfall estimates in each simulated realization. In the figures, the x-axis shows the number of simulated realizations (here 500) and the y-axis represents the number of occurrences

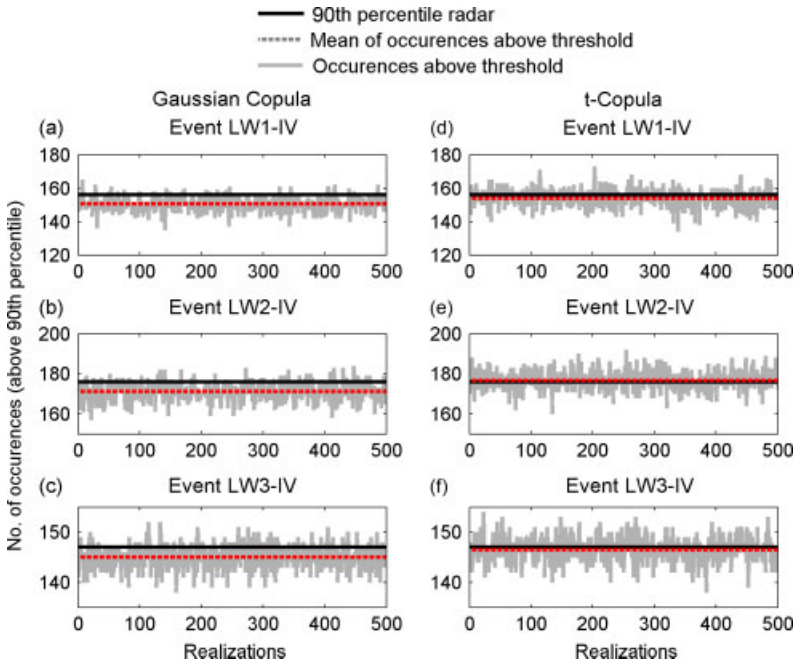

Figure 9. Total number of occurrences above the 90th percentile of MPE data (solid black lines), mean number of occurrences above the same threshold in 500 realizations (dashed lines) and the number of occurrences above the 90th percentile of MPE data in each simulated realizations

above the threshold (90th percentile) of the observations. Figure 9(d)–9(f) provides similar results for the case of t-copula simulations. One can observe that the Gaussian simulated fields slightly underestimate the number of extreme occurrences. Overall, the figures show that, in terms of the number of extreme occurrences, the t-copula is preferred. Table V summarizes the error (%) in the number of extreme occurrences in the simulated fields with respect to the observed rainfall data. It is remarked the asymptotic independence characteristics of the Gaussian copula results in less occurrence of extremes than the t-copula. As discussed earlier, the t-copula shows a positive tail dependence that indicates more extreme values are expected to be generated simultaneously using the t-copula than by the Gaussian copula (Embrechts *et al*. (2001)). This implies that for modelling extremes, Gaussian copula may not be an appropriate choice of copula family.

SUMMARY AND CONCLUSIONS

Multisensor precipitation estimates, unlike traditional gauge measurements, capture the details of spatial variability of rainfall. However, they are subject to various errors from different sources that may affect hydrological and meteorological applications. In this study, two models were presented for simulation of rainfall error estimates using the well-known Gaussian copula and t-copula. The rainfall uncertainty is assumed to be an additive term, as shown in Equation 14. A multiplicative error component was used to account for the proportionality of rainfall error ($R_i \times \varepsilon$) to rainfall rates reported in various studies (Ciach *et al*. (2007); Habib *et al*. (2008)). Multiple realizations of rainfall error are simulated based on observed error values, and proportional to the magnitude of rain rates for the corresponding pixel in rainfall estimates (MPE image). This approach guarantees that simulated rainfall error fields match observed error data conditioned on rainfall rates. That is, one can avoid unrealistically large errors while the magnitude of rain rate is not significant. To demonstrate the model performance, Stage IV radar rainfall data were used as the input to the models to generate ensembles of rainfall realizations.

Table V. The error (%) in the number of extreme occurrences in the simulated fields with respect to the observed rainfall data

	90th percentile	
Event ID	Gaussian	t-copula
Event 1	3·6	1·9
Event 2	2·8	0·9
Event 3	1·4	0·6

The copula parameters were estimated based on the available observations. In order to test how well the copula-based models fit the observations, cross-validation was used considering the MAE as the estimator. Cross-validation is a well-known approach that can be used to evaluate fitted models to observations and also to compare performances of different predictive modelling procedures (Efron and Tibshirani (1990); Picard and Cook (1997)). Table VI lists the results of the repeated random sub-sampling cross-validation (Picard and Cook (1997)) for the fitted Gaussian copula and t-copula to the observations. It is noted that for the case where only five rain gauges were included in simulations, a

Table VI. The results of cross-validation for the Gaussian copula and t-copula models (MAE estimator)

	All gauges		50% LW gauges		12% LW gauges	
Event ID	Gaussian	t-copula	Gaussian	t-copula	Gaussian	t-copula
Event 1	0·98	0·97	1·34	0·93	2·49	2·35
Event 2	0·85	0·84	1·31	1·23	2·50	2·28
Event 3	1·34	1·18	1·59	1·46	2·70	2·56

different cross-validation technique, leave-one-out cross-validation (LOOCV) was used due to the limited observations (see columns 6 and 7 in Table VI). In this approach, a single observation from the original sample is used for validation, and the remaining observations as the training data. This is repeated for the number of observations such that each observation is used once as the validation data. One can see that the t-copula model fits better to the observations. Furthermore, the table indicates that with less number of gauges available, the mean absolute error does not change dramatically.

In the presented models, instead of fitting a standard distribution function to the data, the empirical CDF of observed error is applied to the uniformly simulated error values so that simulated error values will have the same CDF as that of observed. This significant improvement helps avoiding the problem of overshooting (simulating unrealistically large values) which typically happens when a standard distribution function is fitted and used for simulations.

In terms of preserving spatial dependencies, both Gaussian and t-copula models were quite similar. However, for simulation of extremes (rainfall error above/below a certain threshold), using t-copula seemed to have significant advantages over Gaussian copula. The results showed that for a constant threshold, occurrence of extremes in Gaussian copula is less than the t-copula. This implies that for modelling extreme values, the t-copula is more appropriate, as the Gaussian copula may underestimate occurrence of extremes. However, one should be aware that the choice of copula family may be problem dependent and cannot be simply generalized with a few case studies.

This study intended to contribute to ongoing research regarding the uncertainty of MPE through ensemble simulation. Rainfall ensembles can be used for various analyses. For example, by subsequent runs of a hydrological model using an ensemble of rainfall realizations, one can investigate propagation of error in hydrological processes as well as the uncertainty of a given model due to errors in precipitation input. We need to emphasize that the presented model is meant to describe random error (stochastic sources of uncertainty) associated with initial MPE estimates and does not address systematic (bias) issues involved in quantification of rainfall uncertainty. For example, the data may be subject to physical biases and may require further consideration to remove/correct physical biases. Another hypothesis of the model is that error observations are representative for quantification of rainfall uncertainty and its statistical properties. It is worth reminding that the model is based on fitting a copula function to the observations and applying the empirical CDF of observations to the simulated fields. In ideal situations, more than one gauge per a single MPE pixel should be used to provide reliable estimates of the error. However, given the current gauge density, no more than one gauge is available. This limitation may alleviate the point-area effects by the hourly aggregation. We recognize and acknowledge that such effects may be significant especially during high-variability events (Habib *et al.* (2009b)). Additionally, one can argue that insufficient observations and sampling error, among others, may affect the quality of simulated ensembles. Quantifying/testing the effects of insufficient/erroneous observations requires extensive empirical analysis based on large data samples that are not typically available. An interesting issue for future research, that can potentially improve the presented models, is to investigate the dependence of rainfall uncertainty on rainfall regime (e.g. convective, stratiform) and different geographic regions. Unfortunately, at this point in time, long-term rainfall error observations for different rainfall regimes and climate regions are not available. Finally, the presented models are to be investigated to verify their robustness and transferability over different scales. Current efforts are underway by the authors to verify the statistical robustness of the presented models and their applicability in near real-time hydrological and meteorological applications.

ACKNOWLEDGEMENTS

Part of this work was carried out while the first author was a research associate with the University of Louisiana at Lafayette, USA. Financial support from the Louisiana Board of Regents Support Fund and the NASA EPSCoR/BoR is acknowledged.

REFERENCES

AghaKouchak A, Bárdossy A, Habib E. 2010a. Conditional simulation of remotely sensed rainfall data using a non-gaussian v-transformed copula. Submitted to *Advances in Water Resources* (in press).

AghaKouchak A, Habib E, Bárdossy A. 2010b. Modeling radar rainfall estimation uncertainties: a random error model. *Journal of Hydrologic Engineering* **15**(4): DOI:10.1061/(ASCE)HE.1943-5584.0000185 (in press).

Arnaud P, Bouvier C, Cisner L, Dominguez R. 2002. Influence of rainfall spatial variability on flood prediction. *Journal of Hydrology* **260**: 216–230.

Austin P. 1987. Relation between measured radar reflectivity and surface rainfall. *Monthly Weather Review* **115**: 1053–1070.

Bárdossy A. 2006. Copula-based geostatistical models for groundwater quality parameters. *Water Resources Research* **42**(11): W11416.1–W11416.12, DOI: 10.1029/2005WR004754.

Bárdossy A, Li J. 2008. Geostatistical interpolation using copulas. *Water Resources Research* **44**(7): W07412.1–W07412.15, DOI: 10.1029/2007WR006115.

Bell V, Moore R. 2000. The sensitivity of catchment runoff models to rainfall data at different spatial scales. *Hydrology and Earth System Sciences* **4**: 653–667.

Bernardo J, Smith A. 2000. *Bayesian Theory*. Wiley: New York.

Bouye E, Durrleman V, Nikeghbali A, Riboulet G, Roncalli T. 2000. Copulas for finance: a reading guide and some applications. Groupe de Recherche Operationnelle, Credit Lyonnais.

Ciach GJ, Krajewski WF, Villarini G. 2007. Product-error-driven uncertainty model for probabilistic quantitative precipitation estimation with nexrad data. *Journal of Hydrometeorology* **8**: 1325–1347.

Clark M, Gangopadhyay S, Brandon D, Werner K, Hay L, Rajagopalan B, Yates D. 2004a. A resampling procedure for generating conditioned daily weather sequences. *Water Resources Research* **40**(4): W04304.1–W04304.15, DOI: 10.1029/2003WR002747.

Clark M, Gangopadhyay S, Hay L, Rajagopalan B, Wilby R. 2004b. Schaake shuffle: a method for econstructing space-time variability in forecasted precipitation and temperature fields. *Journal of Hydrometeorology* **5**: 243–262.

Cox D, Hinkley D. 1974. *Theoretical Statistics*. Chapman & Hall: London; pp. 511.

Crum T, Alberty R. 1993. The wsr-88d and the wsr-88d operational support facility. *Bulletin of the American Meteorological Society* **74**: 1669–1687.

Davidson R, MascKinnon J. 1993. *Estimation and Inference in Econometrics*. Oxford University Press: Oxford.

De Michele C, Salvadori J. 2002. A generalized pareto intensity-duration model of storm rainfall exploiting 2-copulas. *Journal of Geophysical Research-Atmospheres* **108**(D2): 1–11.

Dupuis D. 2007. Using copulas in hydrology: benefits, cautions, and issues. *Journal of Hydrologic Engineering* **12**(4): 381–393.

Efron B, Tibshirani R. 1990. Improvements on cross-validation: the. 632 + bootstrap method. *Journal of the American Statistical Association* **92**(438): 548–560.

Elliott R, Schiebe F, Crawford K, Peter K, Puckett W. 1993. A unique data capability for natural resources studies. Paper No. 932529, International Winter Meeting; American Society of Agricultural Engineers. Chicago, IL, Dec 14–17.

Embrechts P, Lindskog F, McNeil A. 2001. Modelling dependence with copulas and applications to risk management. *Department of Mathematics*, ETHZ, Zurich, Switzerland, pp 50.

Fang H-B, Fang K-T, Kotz S. 2002. The meta-elliptica distribution with given marginals. *Journal of Multivariate Analysis* **82**: 1–16.

Faures J-M, Goodrich D, Woolhiser DSS. 1995. Impact of small-scale spatial rainfall variability on runoff modeling. *Journal of Hydrology* **173**: 309–326.

Favre A-C, El Adlouni S, Perreault L, Thiémonge N, Bobée B. 2004. Multivariate hydrological frequency analysis using copulas. *Water Resources Research* **40**: 1–12.

Genest C, Favre A, Béliveau J, Jacques C. 2007. Metaelliptical copulas and their use in frequency analysis of multivariate hydrological data. *Water Resources Research* **43**(9): W09401.1–W09401.12, DOI: 10.1029/2006WR005275.

Genest C, Favre A-C. 2007. Everything you always wanted to know about copula modeling but were afraid to ask. *Journal of Hydrologic Engineering* **12**(4): 347–368.

Germann U, Berenguer M, Sempere-Torres D, Salvade G. 2006. Ensemble radar precipitation - a new topic on the radar horizon. Proceeding of the 4th European Conference on Radar in Meteorology and Hydrology ERAD, Barcelona, Spain, 18–22 September 2006.

Goodrich D, Faures J, Woolhiser D, Lane L, Sorooshian S. 1995. Measurement and analysis of small-scale convective storm rainfall variability. *Journal of Hydrology* **173**: 283–308.

Habib E, Aduvala A, Meselhe E. 2008. Analysis of radar rainfall error characteristics and implications for streamflow simulations uncertainty. *Journal of Hydrologic Sciences* **53**(13): 568–587.

Habib E, Henschke A, Adler R. 2009a. Evaluation of TMPA satellite-based research and real-time rainfall estimates during six tropical-related heavy rainfall events over Louisiana, USA. *Atmospheric Research* **94**(3): 373–388, DOI: 10.1016/j.atmosres.2009.06.015.

Habib E, Larson B, Graschel J. 2009b. Validation of nexrad multisensor precipitation estimates using an experimental dense raingauge network in south louisiana. *Journal of Hydrology* **173**: 463–478.

Hollander M, Wolfe D. 1973. *Nonparametric Statistical Methods*. Wiley: New York.

Joe H. 1997. *Multivariate Models and Dependence Concepts*. Chapman Hall: London.

Joe H, Xu J. 1996. The estimation method of inference functions for margins for multivariate models. Technical Report 166. Department of Statistics, University of British Columbia: Vancouver, Canada.

Jordan P, Seed A, Weinmann P. 2003. A stochastic model of radar measurement errors in rainfall accumulations at catchment scale. *Journal of Hydrometeorology* **4**(5): 841–855.

Kelly K, Krzysoztofowicz R. 1997. A bivariate meta-gaussian density for use in hydrology. *Stochastic Environmental Research and Risk Assessment* **11**(1): 17–31.

Malevergne Y, Sornette D. 2003. Testing the gaussian copula hypothesis for financial assets dependences. *Quantitative Finance* **3**: 231–250.

McLeish D, Small C. 1988. *The Theory and Applications of Statistical Inference Functions, Lecture Notes in Statistics*, vol. 44. Springer-Verlag: New York.

Mehrotra R, Srikanthan R, Sharma A. 2006. A comparison of three stochastic multi-site precipitation occurrence generators. *Journal of Hydrology* **331**: 280–292.

Melchiori M. 2003. Which Archimedean Copula is the right one? Third version, YeldCurve e-Journal.

Nelsen R. 2006. *An Introduction to Copulas (Springer Series in Statistics)*. Springer Verlag: New York.

Obled C, Wendling J, Beven K. 1994. The sensitivity of hydrological models to spatial rainfall patterns: an evaluation using observed data. *Journal of Hydrology* **159**: 305–333.

Pegram G, Clothier A. 2001. Downscaling rainfields in space and time using the string of beads model in time series mode. *Hydrology and Earth System Sciences* **5**(2): 175–186.

Petersen-Ø verleir A. 2004. Accounting for heteroscedasticity in rating curve estimates. *Journal of Hydrology* **292**: 173–181.

Picard R, Cook D. 1997. Cross-validation of regression models. *Journal of the American Statistical Association* **79**(387): 575–583.

Purwono Y. 2005. Copula inference for multiple lives analysis - preliminaries. International Actuarial Association: 13th EAA Conference: Bali, Indonesia. 12–15 September 2005.

Renard B, Lang M. 2007. Use of a gaussian copula for multivariate extreme value analysis: Some case studies in hydrology. *Advances in Water Resources* **30**: 897–912.

Salvadori G, DeMichele C, Kottegoda N, Rosso R. 2007. *Extremes in Nature: An Approach Using Copulas*. Springers: Berlin.

Schölzel C, Friederichs P. 2008. Multivariate non-normally distributed random variables in climate research - introduction to the copula approach. *Nonlinear Processes in Geophysics* **15**(5): 761–772.

Seber G, Wild C. 1998. *Nonlinear Regression*. Wiley: Chichester.

Seed A, Srikanthan R. 2001. A space and time model for design storm rainfall. *Journal of Geophysical Research* **104**(24): 31623–31630.

Seliga T, Aron G, Aydin K, White E. 1992. Simulation using radar rainfall rates and a unit hydrograph model (syn-hyd) applied to greve watershed. American Meteorological Society, 25th International Conference on Radar Hydrology: Paris, France; 587–590.

Serinaldi F. 2008a. Analysis of inter-gauge dependence by kendall's τ, upper tail dependence coefficient, and 2-copulas with application to rainfall fields. *Stochastic Environmental Research and Risk Assessment* **22**: 671–688.

Serinaldi F. 2009. Copula-based mixed models for bivariate rainfall data: an empirical study in regression perspective. *Stochastic Environmental Research and Risk Assessment* **23**(5): 677–693.

Shah S, O'Connell P, Hosking J. 1996. Modeling the effects of spatial variability in rainfall on catchment response. 2. experiments with distributed and lumped models. *Journal of Hydrology* **175**: 89–111.

Sklar A. 1959. *Fonctions de Répartition à n Dimensions et Leurs Marges*, vol. 8. Publications de l'Institut de Statistique de L'Université de: Paris; 229–231.

Sklar A. 1996. *Random Variables, Distribution Functions, and Copulas - A Personal Look Backward and Forward, Distributions with Fixed Marginals and Related Topics*, Rüschendorff L, Schweizer B, Taylor MD, eds. Institute of Mathematical Statis: Hayward, CA; 1–14.

Spearman C. 1904. The proof and measurement of association between two things. *American Journal of Psychology* **15**: 72–101.

Syed K, Goodrich D, Myers D, Sorooshian S. 2003. Spatial characteristics of thunderstorm rainfall fields and their relation to runoff. *Journal of Hydrology* **271**: 1–21.

Villarini G, Krajewski W, Ciach G, Zimmerman D. 2009. Product-error-driven generator of probable rainfall conditioned on wsr-88d precipitation estimates. *Water Resources Research* **45**(1): W01404.1–W01404.11, DOI: 10.1029/2008WR006946.

Villarini G, Serinaldi F, Krajewski W. 2008. Modeling radar-rainfall estimation uncertainties using parametric and non-parametric approaches. *Advances in Water Resources* **31**(12): 1674–1686, DOI: 10.1016/j.advwatres.2008.08.002.

Westcott N. 2009. Differences in multi-sensor and rain-gauge precipitation amounts. Proceedings of the Institute of Civil Engineers, *Water Management* **162**(2): 73–81.

Young C, Bradley A, Krajewski W. 2000. Evaluating nexrad multisensor precipitation estimates for operational hydrologic forecasting. *Journal of Hydrometeorology* **1**(3): 241–254.

Young C, Brunsell N. 2008. Evaluating nexrad estimates for the missouri river basin: analysis using daily raingage data. *Journal of Hydraulic Engineering* **13**: 549–553.

Zhang LSR, Singh V. 2008. Bivariate rainfall frequency distributions using archimedean copulas. *Journal of Hydrology* **32**(1–2): 93–109.

