Stoch Environ Res Risk Assess (2011) 25:107–122
DOI 10.1007/s00477-010-0427-7

ORIGINAL PAPER

Stochastic simulation of bivariate gamma distribution: a frequency-factor based approach

Ke-Sheng Cheng · Ju-Chen Hou · Jun-Jih Liou · Yii-Chen Wu · Jie-Lun Chiang

Published online: 9 July 2010

Abstract A frequency-factor based approach for stochastic simulation of bivariate gamma distribution is proposed. The approach involves generation of bivariate normal samples with a correlation coefficient consistent with the correlation coefficient of the corresponding bivariate gamma samples. Then the bivariate normal samples are transformed to bivariate gamma samples using the well-known general equation of hydrological frequency analysis. We demonstrate that the proposed bivariate gamma simulation approach is capable of generating random sample pairs which not only have the desired marginal densities of component random variables but also their correlation coefficient. Scatter plots of simulated bivariate sample pairs also exhibit appropriate linear patterns (dependence structure) that are commonly observed in environmental and hydrological applications. Caution should also be exercised when specifying combinations of coefficients of skewness and the correlation coefficient for bivariate gamma simulation.

Keywords Bivariate gamma distribution · Stochastic simulation · Frequency factor · Frequency analysis

K.-S. Cheng (✉) · J.-C. Hou · J.-J. Liou · Y.-C. Wu
Department of Bioenvironmental Systems Engineering, National Taiwan University, Taipei, Taiwan, ROC
e-mail: rslab@ntu.edu.tw

K.-S. Cheng
Hydrotech Research Institute, National Taiwan University, Taipei, Taiwan, ROC

J.-L. Chiang
Department of Soil and Water Conservation, National Pingtung University of Science and Technology, Pingtung, Taiwan, ROC

1 Introduction

In the field of hydrology and water resources stochastic simulation has becoming more and more important and widely applied, mainly because issues related to hydrological forecasting and water resources planning often involve different sources of uncertainty. Recent developments in hydrological modeling, flood risk analysis, environmental impact assessment, etc. have demonstrated the usefulness of stochastic simulation (National Research Council 2000; Kim and Lee 2007).

Many applications of stochastic simulation in hydrology and water resources deal with univariate simulation (Cowpertwait et al. 1996; Scholz 1997; Werner and Kadlec 2000; Cheng et al., 2007; Benke and Hamilton 2008). However, due to the dependent nature of hydrological variables, multivariate stochastic simulation is gaining more attention (Clark 1980; Loganathan et al. 1987; Singh and Singh 1991; Bacchi et al. 1994; Kelly and Krzysztofowicz 1997; Ashkar et al. 1998; Goel et al. 2000; Yue 1999, 2000, 2001a, b; Loaiciga and Leipnik 2005; Khalili et al. 2009; Lee et al. 2010). For tractability, simulation modelers sometimes ignore the correlation between variables or assume marginal distributions such as the normal, for which correlated variates can be easily generated (Schmeiser and Lal 1982). However, the multivariate normal distribution is inappropriate for applications in many real world data which are essentially non-negative and asymmetric. In contrast, the bivariate gamma distribution has been found useful for modeling such data. For example, flood volume and storm duration were found to be correlated and could be characterized by a bivariate gamma distribution (Yue et al. 2001). In addition, the exponential distribution is a special case of the gamma distribution, and the sum of a set of n independent and identically distributed exponential random variables has a gamma distribution. As a result, the waiting time until the nth occurrence in a Poisson process can be characterized by the gamma distribution. Hydrologists often consider the duration and average intensity of storm events as two correlated gamma random variables. Nadarajah and Gupta (2006a) showed that the product of two correlated gamma random variables also has a gamma distribution, and that the duration and the total depth (which is the product of duration and intensity) of storm events form a bivariate gamma distribution. Thus, stochastic simulation of bivariate gamma distributions may prove to be very useful for flood risk analysis, assessment of climate change effect on hydrological regimes, and environmental impact assessment.

There have been a few models of bivariate gamma distribution in the literature (Cherian 1941; Kibble 1941; Izawa 1953; Moran 1969; D'este 1981; Schmeiser and Lal 1982; Loaiciga and Leipnik 2005; Nadarajah and Gupta 2006b). Yue et al. (2001) reviewed several models of bivariate gamma distributions and demonstrated that these models may be useful for hydrological engineers to analyze joint statistical behavior of multivariate hydrological events such as floods and storms. In their study, a number of random sample-pairs (X, Y) from marginal distributions of pre-specified parameters were generated. The simulation was conducted with various combination scenarios and yielded sample-pairs with different correlation coefficients, instead of a pre-specified or desired correlation coefficient. However, for practical applications, one needs to simulate bivariate gamma random samples with the desired or pre-specified correlation coefficient. Recently stochastic simulations using copula functions which are more widely found in financial literature have also been applied for flood frequency analysis (Grimaldia and Serinaldi 2006), modeling of rainfall intensity and duration (De Michele and Salvadori 2003) and flow peak and volume (Favre et al. 2004), bivariate rainfall frequency analysis (Zhang and Singh 2007), and regional risk analysis (Renard and Langa 2007).

Many bivariate gamma distributions are difficult to implement in practical situations, and seldom succeeded in gaining popularity among practitioners in the field of hydrological frequency analysis (Yue et al. 2001). Additionally, there is no agreement about what the multivariate gamma distribution should be, and in practical applications we often only need to specify the marginal gamma distributions and the correlations between the component random variables (Law 2007). Thus, the objective of this study is to propose a new bivariate gamma simulation approach based on the frequency factor which is well-known to scientists and engineers in water resources field. The proposed approach aims to yield random vectors which have not only the desired marginal distributions but also a pre-specified correlation coefficient between the component variates.

In the following sections we first describe the methodology of the proposed bivariate gamma simulation approach. Detailed derivations of the correlation coefficient between two component variates as a function of the shape factors of marginal densities are also given in Sect. 2 and Appendix 1. Section 3 presents properties of the sample estimates of distribution parameters which are in accordance with real storm rainfall data in central Taiwan. Simulation results are evaluated by considering the bias and mean squared error (MSE) of the parameter estimators. Section 4 further discusses properties of the correlation coefficient with respect to ranges of coefficient of skewness. The proposed approach is also shown to yield random samples of the bivariate gamma distribution proposed by Moran (1969). The final section constitutes conclusions drawn from this study.

2 Frequency-factor based bivariate gamma simulation

Cheng et al. (2007) proposed a frequency-factor based method for random number generation of five distributions commonly used in hydrological frequency analysis. The method stems from the following general equation of hydrological frequency (Chow 1951; Chow et al. 1988):

$$X_T = \mu_X + K_T \sigma_X, \tag{1}$$

where μ_X and σ_X are, respectively, the mean and standard deviation of a random variable X, and X_T represents the magnitude of X with exceedence probability $1/T$. The frequency factor K_T is a distribution-specific function of T.

From the view point of random number generation, the frequency factor can be considered as a random variable K, and K_T is a value of K with exceedence probability $1/T$. For example, frequency factor of the Pearson type III distribution can be approximated by (Kite 1988)

$$K_T \approx z + (z^2 - 1)\frac{\gamma_X}{6} + \frac{1}{3}(z^3 - 6z)\left(\frac{\gamma_X}{6}\right)^2 - (z^2 - 1)\left(\frac{\gamma_X}{6}\right)^3 + z\left(\frac{\gamma_X}{6}\right)^4 - \frac{1}{3}\left(\frac{\gamma_X}{6}\right)^5, \tag{2}$$

where z is the standard normal deviate and γ_X is the coefficient of skewness of X. Given μ_X, σ_X and γ_X, if we can generate a set of random numbers of K, say $k_1, k_2, \ldots, k_n$, then a random sample of X, say $x_1, x_2, \ldots, x_n$, can be obtained by

$$x_i = \mu_X + k_i \sigma_X. \tag{3}$$

Note that each k_i, $i = 1, 2, \ldots, n$, corresponds to its own exceedence probability $1/T_i$.

The gamma distribution is a special case of the Pearson type III distribution with a zero location parameter. Therefore, it seems plausible to generate random samples of a bivariate gamma distribution based on two jointly distributed frequency factors of Eq. 2. Details of such frequency-factor based bivariate gamma simulation are described below.

Assume two random variables X and Y are jointly distributed, each with the following marginal gamma density:

$$f_X(x;\alpha,\beta,\varepsilon)=\frac{1}{\alpha\Gamma(\beta)}\left(\frac{x}{\alpha}\right)^{\beta-1}e^{-(x/\alpha)},\quad 0\le x<+\infty \tag{4}$$

$$\alpha=\frac{\sigma}{\sqrt{\beta}}>0, \tag{5}$$

$$\beta=\left(\frac{2}{\gamma}\right)^2>0, \tag{6}$$

$$\mu=\alpha\beta=\sigma\sqrt{\beta}>0, \tag{7}$$

where μ, σ, and γ are the mean, standard deviation, and coefficient of skewness of X (or Y), respectively, and α and β are, respectively, the scale and shape parameters of the gamma distribution. From Eq. 3, random variables X and Y are, respectively, associated with their frequency factors K_X and K_Y.

Equation 2 indicates that frequency factor K_X of a gamma random variable X can be approximated by a function of the standard normal deviate and the coefficient of skewness γ_X of the gamma distribution. Therefore, simulation of the frequency factor K_X can be achieved by generating a random sample of the standard normal deviate U, say $u_1, u_2, \ldots, u_n$, and then utilizing Eq. 2 to obtain $k_{x_1}, k_{x_2}, \ldots, k_{x_n}$ from $u_1, u_2, \ldots, u_n$. However, for a bivariate gamma density $f_{XY}(x, y)$, the two frequency factors K_X and K_Y are correlated through two standard normal deviates U and V which have a correlation coefficient ρ_{UV}. Thus, random number generation of the second frequency factor K_Y must take into consideration the correlation between K_X and K_Y which in turn stems from the correlation between U and V.

Given a random number of U, say u, the conditional density of V is expressed by the following conditional normal density

$$\varphi_{V|U}(v|U=u)=\frac{1}{\sqrt{2\pi(1-\rho_{UV}^2)}}\cdot\exp\left\{-\frac{1}{2}\left[\frac{v-\rho_{UV}u}{\sqrt{1-\rho_{UV}^2}}\right]^2\right\}, \tag{8}$$

with mean $\rho_{UV}u$ and variance $1-\rho_{UV}^2$. Thus, based on a random sample $u_1, u_2, \ldots, u_n$ of U, a random sample of V, say $v_1, v_2, \ldots, v_n$, can be generated by a normal random number generator with means $\rho_{UV}u_i$ ($i = 1, 2, \ldots, n$) and variance $1-\rho_{UV}^2$. Finally, from the two sets of random samples $u_1, u_2, \ldots, u_n$ and $v_1, v_2, \ldots, v_n$, Eq. 2 can be used to obtain random samples of the two frequency factors K_X and K_Y, i.e. $k_{x_1}, k_{x_2}, \ldots, k_{x_n}$ and $k_{y_1}, k_{y_2}, \ldots, k_{y_n}$.

Given the expected values (μ_X and μ_Y) and standard deviations (σ_X and σ_Y) of random variables X and Y, random samples of the bivariate gamma distribution can thus be obtained by transferring $k_{x_1}, k_{x_2}, \ldots, k_{x_n}$ and $k_{y_1}, k_{y_2}, \ldots, k_{y_n}$ to $x_1, x_2, \ldots, x_n$ and $y_1, y_2, \ldots, y_n$ using Eq. 3. The above frequency-factor based procedures for generation of random samples of a bivariate gamma distribution is referred to as the *BVG-FF algorithm* and is illustrated in Fig. 1. It is worthy to note that the mean, standard deviation and coefficient of skewness of a gamma distribution are related by

$$\sigma=\frac{\mu\gamma}{2}. \tag{9}$$

In practice, a set of pre-specified mean, standard deviation, and coefficient of skewness [i.e. (μ_X, σ_X, γ_X) and (μ_Y, σ_Y, γ_Y)], and correlation coefficient ρ_{XY} are given for stochastic simulation from a bivariate gamma distribution. To ensure the generated pair of random samples ($x_1, x_2, \ldots, x_n$ and $y_1, y_2, \ldots, y_n$) having the desired correlation relationship, correlation coefficient ρ_{UV} must be determined from

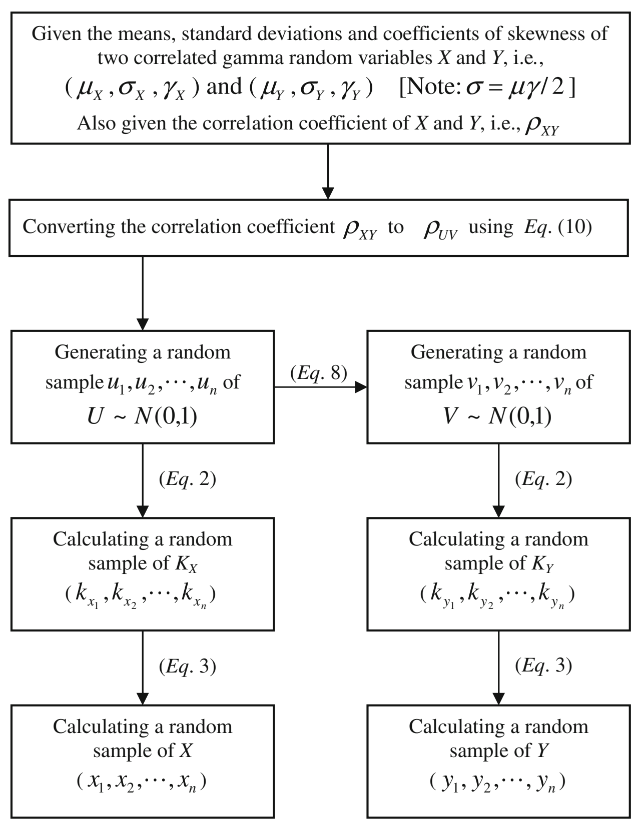


Fig. 1 The BVG-FF algorithm for stochastic simulation of the bivariate gamma distribution

the pre-specified γ_X, γ_Y and ρ_{XY} through the following equation:

$$\rho_{XY} \approx (A_X A_Y - 3A_X C_Y - 3C_X A_Y + 9C_X C_Y)\rho_{UV} + 2B_X B_Y \rho_{UV}^2 + 6C_X C_Y \rho_{UV}^3, \quad (10)$$

where

$$A_X = 1 + \left(\frac{\gamma_X}{6}\right)^4, \quad B_X = \frac{\gamma_X}{6} - \left(\frac{\gamma_X}{6}\right)^3, \quad C_X = \frac{1}{3}\left(\frac{\gamma_X}{6}\right)^2,$$

$$A_Y = 1 + \left(\frac{\gamma_Y}{6}\right)^4, \quad B_Y = \frac{\gamma_Y}{6} - \left(\frac{\gamma_Y}{6}\right)^3, \quad C_Y = \frac{1}{3}\left(\frac{\gamma_Y}{6}\right)^2.$$

Proof of Eq. 10 is given in Appendix 1. It is interesting to note that the correlation coefficient ρ_{XY} does not depend on the scale parameters α_X and α_Y.

Equation 10 shows that ρ_{XY} can be expressed as a third order polynomial of ρ_{UV}. It is therefore of practical concern whether there exits a unique ρ_{UV} for a given set of (γ_X, γ_Y, ρ_{XY}). Or equivalently, given a set of (γ_X, γ_Y, ρ_{XY}), does Eq. 10 return a single-value of ρ_{UV}? Figure 2 demonstrates the relationships between ρ_{XY} and ρ_{UV} for $\gamma_X = 1.0$ and 3.0, with γ_Y varying from 0.1 to 3.0. We proved that Eq. 10 is indeed a single-value function of ρ_{UV}, and details of the proof are given in Appendix 2.

3 Simulation and verification

In order to assess the performance of the proposed BVG-FF simulation algorithm and to demonstrate how this approach can be implemented, we chose to base our simulation on real rainfall data observed at two raingauge stations (C1I020 and C1G690) in central Taiwan using Monte Carlo random number generation techniques. Fourteen years (1993–2006) of hourly rainfall records are available at both stations (http://www.rslabntu.net/). In general, rainfalls in Taiwan are brought about by four season-specific dominant storm types: winter frontal rainfall (from November to April of the next year), Mei-Yu (from late April to end of June), convective rainfall and typhoons (both from July to October). Among these seasonal storms, typhoon rainfalls account for the highest percentage of annual total rainfall. Storms of different types exhibit different characteristics. For example, typhoons tend to have longer storm durations and larger amount of total rainfall than convective storms. In this study, we focus on the joint distribution of the duration and total rainfall depth of typhoon events.

From historical rainfall records, a minimum *inter-event time* (the time span from the end of one storm event to the beginning of the next storm event) of 8 h and a minimum storm duration of 12 h were adopted for extraction of typhoon events. Table 1 summarizes statistical properties of

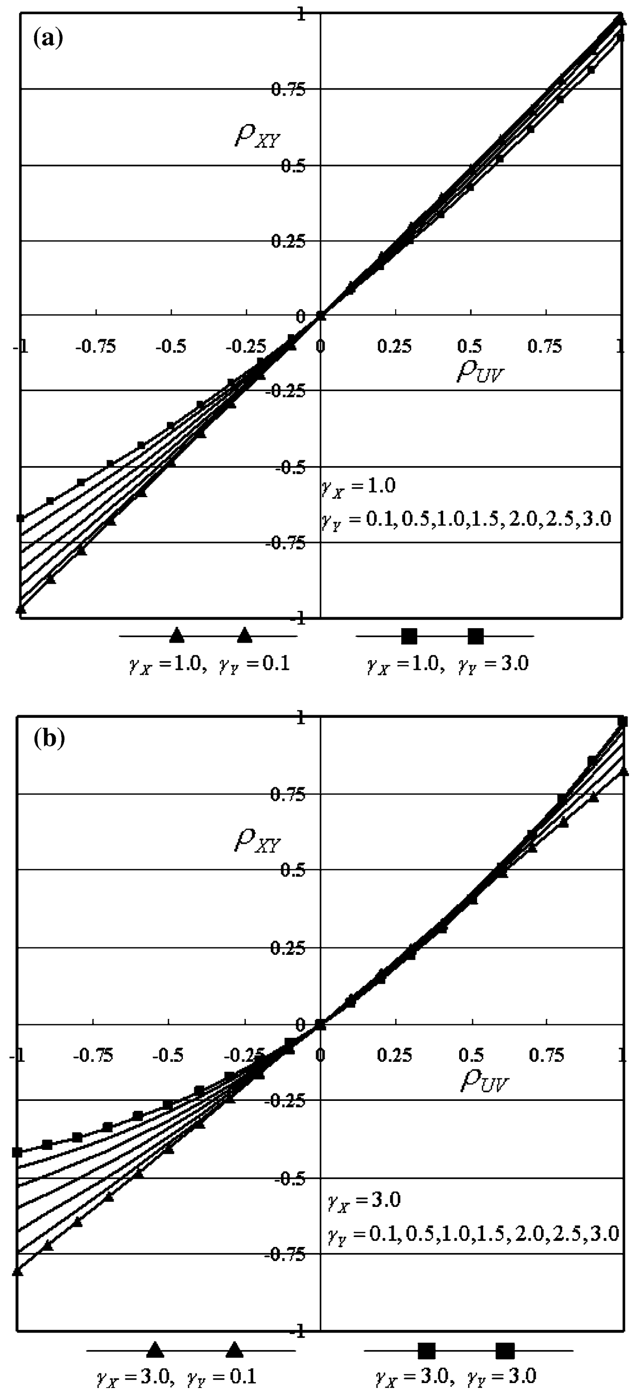


Fig. 2 The single-value relationship between ρ_{XY} and ρ_{UV}. **a** $\gamma_X = 1.0$, **b** $\gamma_X = 3.0$

the total rainfall depth and duration of typhoon events at the two exemplar raingauge sites. Total rainfall depth and duration of typhoon events were considered having marginal gamma distributions based on the results of the Kolmogorov–Smirnov goodness-of-fit tests (see Table 2), although other methods (Kottegoda 1980; Tsukatani and Shighemitsu 1980; D'Agostino and Stephens 1986; Hosking and Wallis 1997; Leiva et al. 2008; Liou et al. 2008) could also be applied. Additionally, methods of model selection based on loss of information, such as Akaike information criteria (AIC; Akaike 1974), Schwarz's Bayesian information criteria (BIC; Schwarz 1978) and Hannan–Quinn information criteria (HQIC; Hannan and Quinn 1979), were also applied for distribution model selection among three candidate distributions, including the gamma, log-normal and extreme value type I (Gumbel). The results showed that the duration could be best modeled by the gamma distribution whereas lognormal distribution was the best model for the total depth. Gamma modeling of the total depth was associated with a slightly higher information loss than the lognormal distribution. To demonstrate the BVG-FF simulation algorithm, we further assume the total depth and duration of typhoon events are bivariate-gamma distributed, although validity of such assumption requires further test.

Let X and Y represent the duration and total rainfall depth of typhoon events, respectively. Based on the statistical properties of historical rainfall data shown in Table 1, we assume $\mu_X = 26.83$ h, $\mu_Y = 105.19$ mm, $\gamma_X = 1.45$, $\gamma_Y = 1.63$, and $\rho_{XY} = 0.71$ for Station C1I020. For Station C1G690, corresponding parameters are $\mu_X = 22.78$ h, $\mu_Y = 95.65$ mm, $\gamma_X = 1.66$, $\gamma_Y = 2.27$, and $\rho_{XY} = 0.76$. Then, for a specific sample size $n (n = 20, 40, 60, 100, 150, 250, 500)$, a total of 10,000 random samples were generated using the proposed BVG-FF simulation algorithm.

From each of the 10,000 generated random samples, estimates of the mean and coefficient of skewness of the methods of moments and maximum likelihood were calculated. Details of these estimators are given in Appendix 3. For the gamma distribution, both the methods of moments and maximum likelihood have their estimators of the population mean the same as the sample mean. In addition to the method of moment estimator ($\hat{\gamma}_{\mathrm{MOM}}$) and the maximum likelihood estimator ($\hat{\gamma}_{\mathrm{MLE}}$), a commonly used estimator of the coefficient of skewness is as follows (Chow et al. 1988):

$$\hat{\gamma} = \frac{n \sum_{i=1}^{n} (x_i - \bar{x}_n)^3}{(n-1)(n-2)S_n^3}. \tag{11}$$

Estimator of the above coefficient of skewness is very sensitive to sample size n and Bobee and Robitaille (1977) suggested using the following sample-size-adjusted coefficient of skewness for the gamma distribution:

$$\hat{\gamma}' = \hat{\gamma}_{\mathrm{MOM}} \frac{[n(n-1)]^{1/2}}{n-2}\left(1 + \frac{8.5}{n}\right) = \hat{\gamma}\left(1 + \frac{8.5}{n}\right). \tag{12}$$

Hereafter, the coefficient of skewness $\hat{\gamma}$ and $\hat{\gamma}'$ are, respectively, referred to as the sample coefficient of skewness and adjusted sample coefficient of skewness. Sample correlation coefficient is calculated as

$$r_{XY} = \frac{\sum_{i=1}^{n} (x_i - \bar{x}_n)(y_i - \bar{y}_n)}{(n-1)S_X S_Y}. \tag{13}$$

Estimates of distribution parameters (μ_X, μ_Y, γ_X, γ_Y, and ρ_{XY}) vary with samples and uncertainties of these estimates are affected by the sample size n. It is therefore necessary to investigate the estimation uncertainty with respect to sample size when assessing the simulation results. Sample means and sample standard deviations of the above parameter estimators calculated from 10,000 simulated random samples are listed in Tables 3 and 4, with respect to sample size n ranging from 20 to 500. In general, estimates of the population means of the duration and total depth and the correlation coefficient of the two random variables are very close to their corresponding population parameters, with negligible estimation errors even for sample size $n = 20$. However, the four estimators of coefficient of skewness ($\hat{\gamma}_{\mathrm{MOM}}$, $\hat{\gamma}_{\mathrm{MLE}}$, $\hat{\gamma}$ and $\hat{\gamma}'$) do not perform equally well. This is illustrated by Figs. 3 and 4 using bias and MSE of individual estimators.

The bias and MSE of individual estimators can be, respectively, expressed by

$$b(\hat{\theta}) = E(\hat{\theta}) - \theta, \tag{14}$$

and

$$\mathrm{MSE}(\hat{\theta}) = \mathrm{Var}(\hat{\theta}) + [b(\hat{\theta})]^2, \tag{15}$$

where θ and $\hat{\theta}$, respectively, represent a parameter and its estimator, and $b(\hat{\theta})$ is the bias of $\hat{\theta}$. From the sample means and sample standard deviations of parameter estimators (see Tables 3, 4), the bias and MSE of individual estimators can be estimated. Generally speaking, both the bias and MSE decrease with increasing sample size for all parameter estimators. At sample size 500, all biases are

Table 1 Statistical properties of typhoon events at two raingauge stations based on rainfall data recorded in a 14-year period (1993–2006)

Station ID	Duration, X (h)			Total depth, Y (mm)			Correlation coefficient r_{XY}
	Mean	Standard deviation	Coefficient of skewness	Mean	Standard deviation	Coefficient of skewness	
C1I020	26.83	13.87	1.45	105.19	118.77	1.63	0.71
C1G690	22.78	11.18	1.66	95.65	106.44	2.27	0.76

Table 2 Results of the Kolmogorov–Smirnov goodness-of-fit tests for duration and total depth of stations C1I020 and C1G690

Station	Variable	Value of the test statistic	p value	Critical value ($\alpha = 0.05$)
C11020 ($n = 48$)	Duration	0.086	0.840	0.192
	Total depth	0.135	0.317	0.192
C1G690 ($n = 23$)	Duration	0.116	0.879	0.275
	Total depth	0.164	0.510	0.275

n and α represent the sample size and level of significance, respectively

Table 3 Summary statistics of the estimators of distribution properties (mean, coefficient of skewness, and correlation coefficient) with respect to sample size n ranging from 20 to 500 (Station C1I020, $\mu_X = 26.83$, $\mu_Y = 105.19$, $\gamma_X = 1.45$, $\gamma_Y = 1.63$, $\rho_{XY} = 0.71$, $\alpha_X = 7.744$, $\beta_X = 1.915$, $\alpha_Y = 70.251$, $\beta_Y = 1.497$)

Parameter estimator	Summary statistics	Sample size (n)						
		20	40	60	100	150	250	500
$\bar{x}_n$	Mean	26.826	26.827	26.822	26.816	26.840	26.824	26.826
	SD	2.398	1.709	1.396	1.067	0.890	0.677	0.481
$\hat{\gamma}_{X,\mathrm{MOM}}$	Mean	0.986	1.175	1.262	1.339	1.379	1.413	1.451
	SD	0.589	0.538	0.510	0.456	0.407	0.342	0.271
$\hat{\gamma}_{X,\mathrm{MLE}}$	Mean	1.411	1.441	1.449	1.456	1.460	1.462	1.465
	SD	0.231	0.164	0.135	0.103	0.086	0.066	0.048
$\hat{\gamma}_X$	Mean	1.068	1.221	1.294	1.359	1.393	1.421	1.455
	SD	0.638	0.560	0.523	0.462	0.411	0.344	0.272
$\hat{\gamma}'_X$	Mean	1.521	1.481	1.478	1.475	1.472	1.470	1.480
	SD	0.908	0.679	0.597	0.502	0.434	0.356	0.276
$\bar{y}_n$	Mean	105.128	105.025	105.133	105.066	105.274	105.104	105.127
	SD	19.280	13.461	11.174	8.549	7.094	5.421	3.807
$\hat{\gamma}_{Y,\mathrm{MOM}}$	Mean	1.093	1.309	1.408	1.496	1.547	1.588	1.627
	SD	0.602	0.580	0.555	0.499	0.446	0.384	0.302
$\hat{\gamma}_{Y,\mathrm{MLE}}$	Mean	1.603	1.634	1.643	1.650	1.655	1.658	1.661
	SD	0.266	0.190	0.158	0.121	0.100	0.078	0.055
$\hat{\gamma}_Y$	Mean	1.184	1.361	1.444	1.519	1.563	1.598	1.632
	SD	0.652	0.602	0.570	0.506	0.450	0.386	0.302
$\hat{\gamma}'_Y$	Mean	1.687	1.650	1.649	1.649	1.651	1.652	1.660
	SD	0.929	0.730	0.650	0.549	0.476	0.399	0.308
r_{XY}	Mean	0.703	0.706	0.707	0.710	0.710	0.711	0.711
	SD	0.139	0.097	0.081	0.062	0.051	0.040	0.028

very small and can be practically ignored. It is worthy to note that among different estimators of coefficient of skewness ($\hat{\gamma}_{\mathrm{MOM}}, \hat{\gamma}_{\mathrm{MLE}}, \hat{\gamma}$ and $\hat{\gamma}'$) the maximum likelihood estimator is most superior with significantly lower MSE than other estimators. The method of moment estimator $\hat{\gamma}_{\mathrm{MOM}}$ and the sample coefficient of skewness $\hat{\gamma}$ seem to have compatible performance in terms of bias and MSE. It can also be observed that although the adjusted sample coefficient of skewness $\hat{\gamma}'$ can effectively reduce the estimation bias for random samples of small to moderate sample size ($n \leq 60$), its estimates are associated with higher standard deviations. As a result, the adjusted sample coefficient of skewness has the highest MSE among the four estimators of coefficient of skewness.

In addition to verifying the marginal distributions and correlation, Schmeiser and Lal (1982) pointed out that the scatter plot of generated (X, Y) pairs must also be investigated. Schmeiser and Lal (1982) proposed a bivariate gamma simulation algorithm using a trivariate reduction technique which yields two dependent random variables from three independent random variables. Given the shape and scale parameters of two dependent gamma random variables and their correlation coefficient, the algorithm uses five variables to satisfy a set of three equality

Table 4 Summary statistics of the estimators of distribution properties (mean, coefficient of skewness, and correlation coefficient) with respect to sample size n ranging from 20 to 500 (Station C1G690, $\mu_X = 22.78$, $\mu_Y = 95.65$, $\gamma_X = 1.66$, $\gamma_Y = 2.27$, $\rho_{XY} = 0.76$, $\alpha_X = 7.391$, $\beta_X = 1.459$, $\alpha_Y = 122.869$, $\beta_Y = 0.778$)

Parameter estimator	Summary statistics	Sample size (n)						
		20	40	60	100	150	250	500
$\bar{x}_n$	Mean	22.718	22.774	22.792	22.756	22.768	22.772	22.774
	SD	2.005	1.418	1.163	0.884	0.729	0.570	0.400
$\hat{\gamma}_{X,\text{MOM}}$	Mean	1.117	1.326	1.426	1.526	1.576	1.620	1.651
	SD	0.618	0.582	0.553	0.506	0.463	0.387	0.307
$\hat{\gamma}_{X,\text{MLE}}$	Mean	1.635	1.668	1.673	1.682	1.688	1.691	1.694
	SD	0.275	0.194	0.162	0.124	0.102	0.080	0.056
$\hat{\gamma}_X$	Mean	1.210	1.378	1.462	1.549	1.592	1.630	1.656
	SD	0.669	0.605	0.567	0.514	0.467	0.390	0.308
$\hat{\gamma}'_X$	Mean	1.724	1.671	1.669	1.681	1.682	1.686	1.685
	SD	0.953	0.734	0.647	0.558	0.494	0.403	0.314
$\bar{y}_n$	Mean	95.134	95.459	95.609	95.101	95.311	95.409	95.385
	SD	24.281	17.097	14.094	10.851	8.776	6.810	4.861
$\hat{\gamma}_{Y,\text{MOM}}$	Mean	1.459	1.760	1.891	2.039	2.143	2.217	2.275
	SD	0.668	0.701	0.664	0.650	0.638	0.561	0.457
$\hat{\gamma}_{Y,\text{MLE}}$	Mean	2.306	2.346	2.355	2.368	2.371	2.375	2.379
	SD	0.386	0.270	0.226	0.173	0.140	0.109	0.077
$\hat{\gamma}_Y$	Mean	1.580	1.829	1.940	2.070	2.165	2.231	2.282
	SD	0.723	0.729	0.682	0.660	0.645	0.564	0.458
$\hat{\gamma}'_Y$	Mean	2.252	2.218	2.215	2.246	2.288	2.307	2.321
	SD	1.030	0.884	0.778	0.716	0.681	0.583	0.466
r_{XY}	Mean	0.758	0.762	0.762	0.763	0.762	0.763	0.763
	SD	0.126	0.089	0.073	0.058	0.048	0.037	0.026

conditions, and thus the solutions are feasible solutions, rather than an optimal solution (Schmeiser and Lal 1982). Substantial computation is required to determine whether or not the selected parameter values satisfy the required equality conditions. Such computation involves another numerical root finding algorithm and a numerical integration for calculation of expected value. Using the bivariate gamma simulation algorithm, Schmeiser and Lal (1982) demonstrated a case that generated random samples may have the desired marginal distributions and correlation; however, the scatter pattern of (X, Y) pairs is inappropriate in most applications. The pattern is that of many independent variates superimposed over many other variates lying on a well-behaved curve. Figure 5 illustrates an example of such inappropriate scatter plot of (X, Y) pairs.

From the results of bivariate gamma simulation using the BVG-FF algorithm, two sets of random sample pairs of sample size 40 and 500, respectively, were arbitrarily chosen for examination of their scatter plots. As can be observed in Figs. 6 and 7, irrespective of the sample sizes, scatter plots of (X, Y) pairs exhibit appropriate linear patterns that are commonly observed in real world environmental or hydrological studies.

Comparing to the algorithm proposed by Schmeiser and Lal (1982), the proposed BVG-FF algorithm is much less intensive in computation. It only requires simulation of bivariate standard normal random samples with a specific correlation coefficient, and bivariate gamma samples can be easily transformed from standard normal samples using the general equation of frequency analysis. Another important advantage of the BVG-FF algorithm is that it generates random samples of known bivariate gamma density, i.e. the Moran bivariate gamma model (see details in Sect. 4) whereas the joint density of the Schmeiser and Lal's approach was not given.

4 Further discussions on the feasible region of ρ_{XY} and joint PDF

In this section we further examine two important issues—feasible range of the correlation coefficient ρ_{XY} and joint

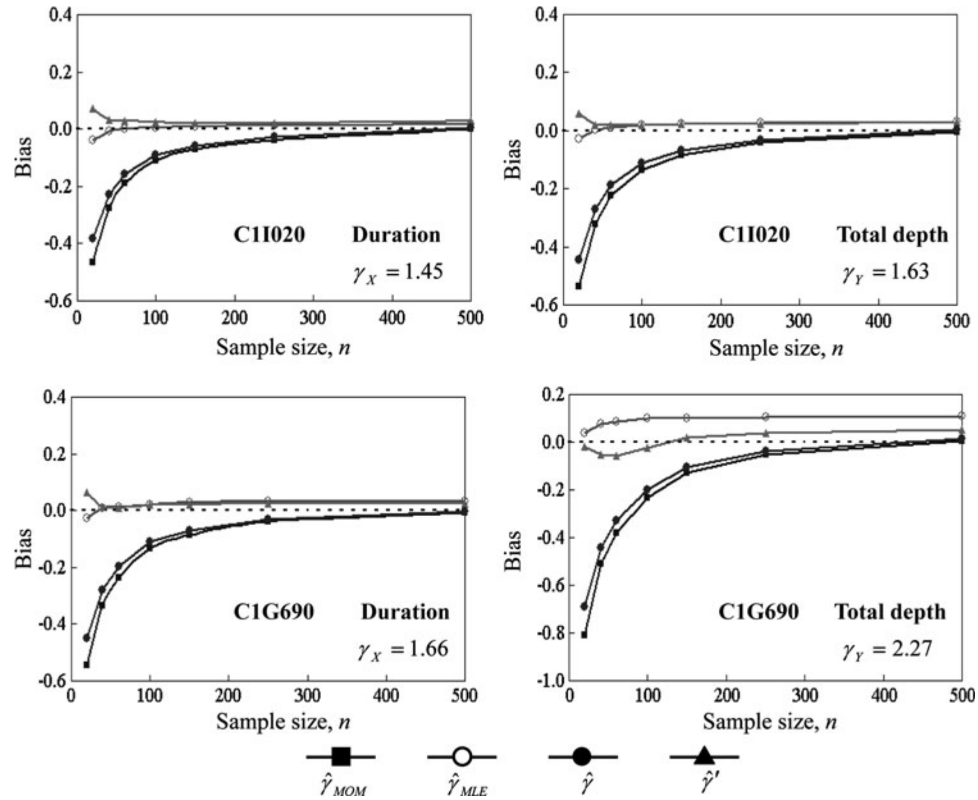


Fig. 3 Biases of different estimators of coefficient of skewness of duration (*X*) and total depth (*Y*) of stations C1I020 and C1G690

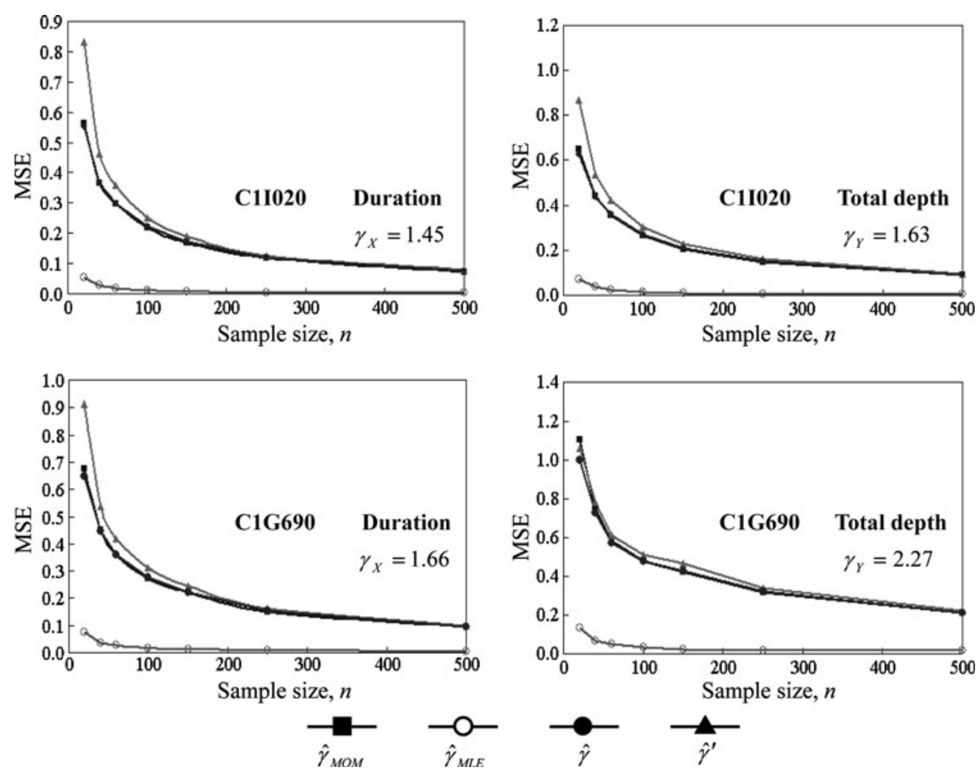


Fig. 4 Mean squared errors of different estimators of coefficient of skewness of duration (*X*) and total depth (*Y*) of stations C1I020 and C1G690

probability density function (PDF) of the bivariate gamma distribution, which are of important concerns to stochastic simulation of the proposed approach.

A key element in stochastic simulation using the proposed approach is the conversion between the correlation coefficient ρ_{XY} of two gamma random variables X and Y

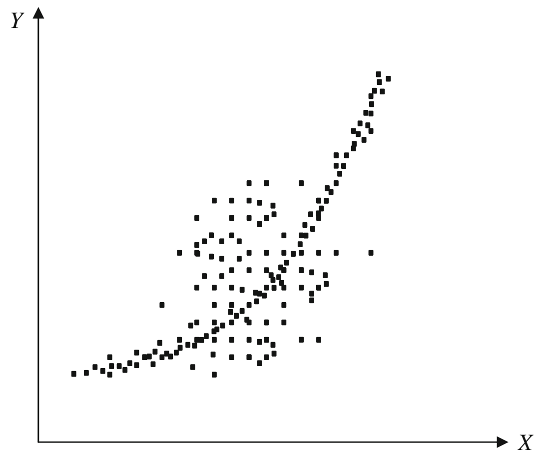
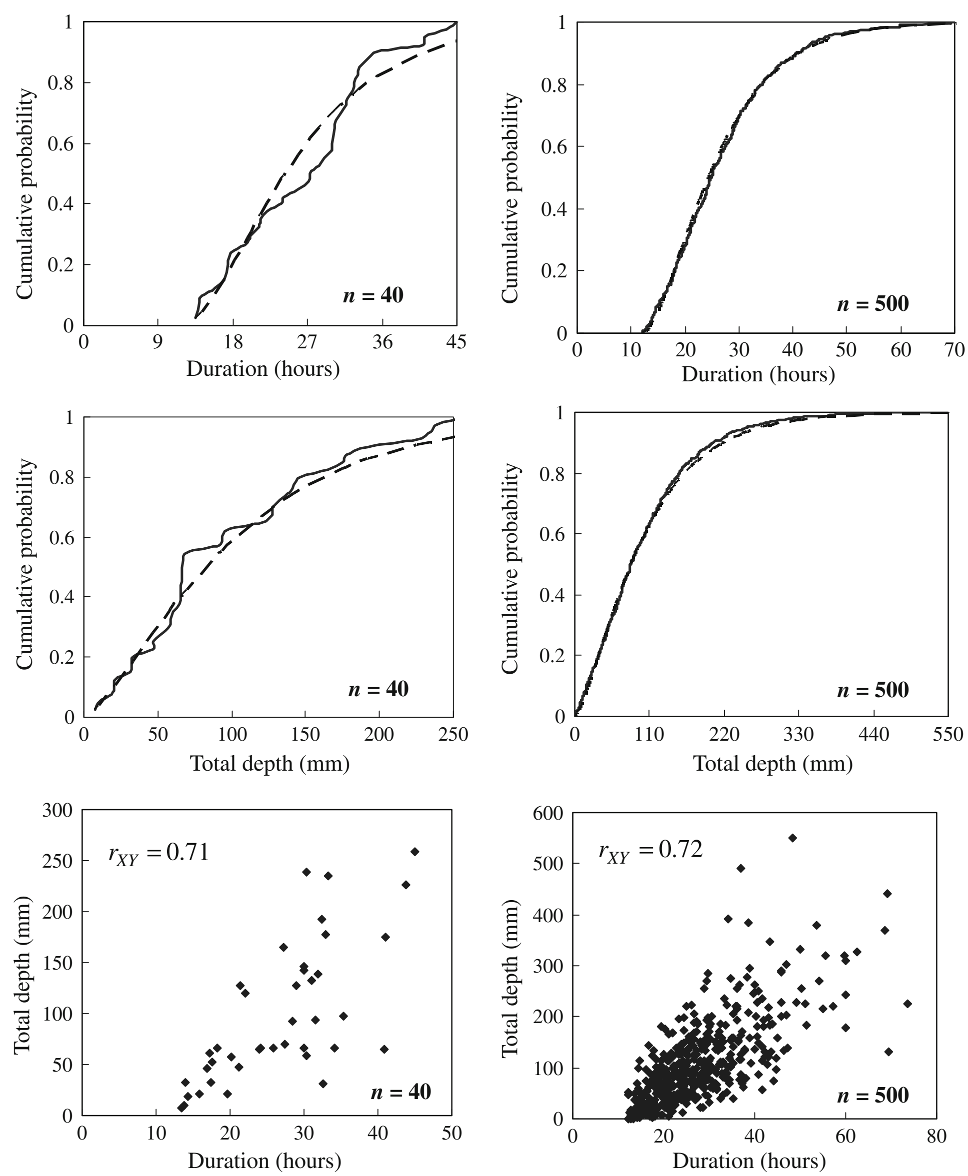


Fig. 5 Illustration of a scatter plot of simulated random samples (X, Y) with an inappropriate pattern

and the corresponding correlation coefficient ρ_{UV} of two standard normal random variables U and V using Eq. 10. Such conversion is a function of the coefficients of skewness γ_X and γ_Y (or equivalently, shape factors β_X and β_Y). It is worthy to note that, for given ranges of γ_X and γ_Y, the correlation coefficient ρ_{XY} cannot assume all values between -1 and 1 since the property of $-1 \leq \rho_{UV} \leq 1$ must be satisfied. With $0 < \gamma_X, \gamma_Y \leq 3$, Figs. 8 and 9, respectively, demonstrate the feasible regions of ρ_{XY} for $\rho_{UV} = 0.6$ and $\rho_{UV} = -0.6$. Clearly, under certain circumstances the values of ρ_{XY} must fall within a restricted region in order for X and Y to form a bivariate gamma distribution. Thus, one should exercise extra caution when specifying coefficients of skewness (γ_X, γ_Y) and the correlation coefficient ρ_{XY} for bivariate gamma simulation. For instance, two random variables X and Y with $\gamma_X = 1.5$,

Fig. 6 ECDF (*solid line*), CDF (*dashed line*) and scatter plots of simulated bivariate gamma random samples (Station C1I020). Distribution parameters are shown in Table 3



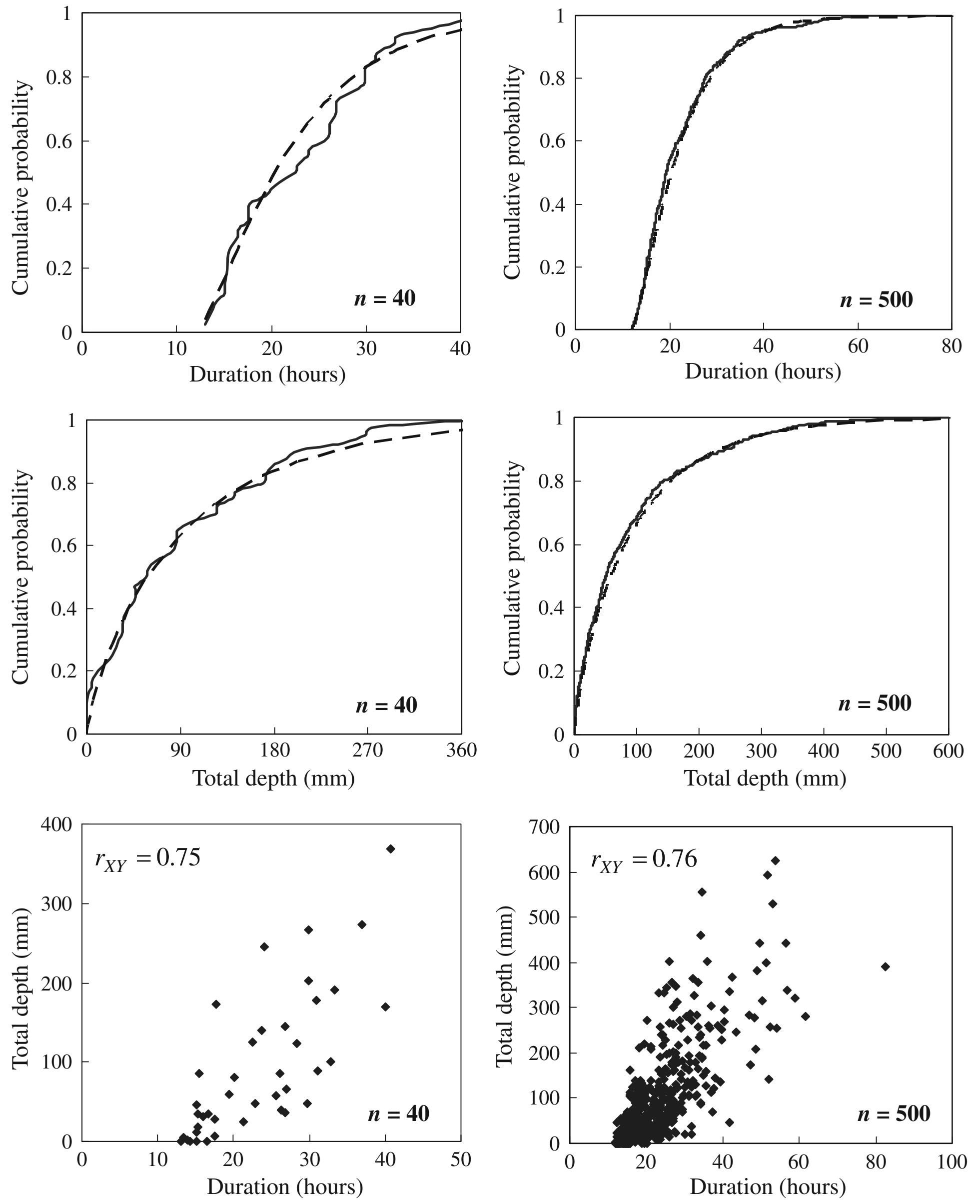


Fig. 7 ECDF (*solid line*), CDF (*dashed line*) and scatter plots of simulated bivariate gamma random samples (Station C1G690). Distribution parameters are shown in Table 4

$\gamma_Y = 2$, and $\rho_{XY} = -0.75$ cannot constitute a bivariate gamma distribution.

Until now we have not explicitly mentioned the joint PDF of the generated random samples using the proposed approach. However, a careful comparison of the Moran bivariate gamma distribution (Moran 1969; Yue et al. 2001) and the proposed approach reveals that random samples generated by the proposed approach are indeed distributed with the following joint PDF of the Moran bivariate gamma model:

$$f_{XY}(x,y) = \frac{1}{\sqrt{1-\rho_{UV}^2}} f_X(x;\alpha_x,\beta_x) f_Y(y;\alpha_y,\beta_y) \cdot \exp\left[-\frac{(\rho_{UV}u)^2 - 2\rho_{UV}uv + (\rho_{UV}v)^2}{2(1-\rho_{UV}^2)}\right], \quad (16a)$$

$$u = \Phi^{-1}[F_X(x;\alpha_x,\beta_x)], \quad (16b)$$

$$v = \Phi^{-1}[F_Y(y;\alpha_y,\beta_y)], \quad (16c)$$

where U and V are jointly distributed standard normal random variables with correlation coefficient ρ_{UV}, and X and Y constitute a bivariate gamma distribution with marginal densities $f_X(x;\ \alpha_x,\ \beta_x)$ and $f_Y(y;\ \alpha_y,\ \beta_y)$, respectively. Details of the Moran bivariate gamma distribution are also included in Appendix 4.

As a final note we like to point out that the proposed approach can also be applied for stochastic simulation of the more general bivariate Pearson type III distribution since the frequency factor equation (Eq. 2) remains the same, although a location parameter will be introduced. R-code simulation program of the proposed approach is

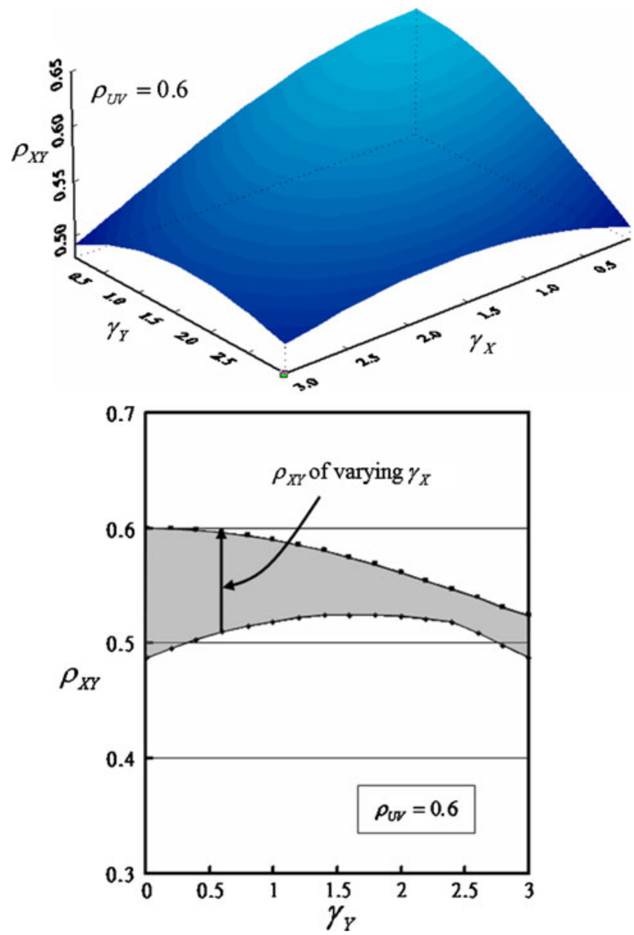


Fig. 8 Feasible regions of ρ_{XY} for two gamma random variables X and Y with $0 < \gamma_X, \gamma_Y \leq 3$ and $\rho_{UV} = 0.6$

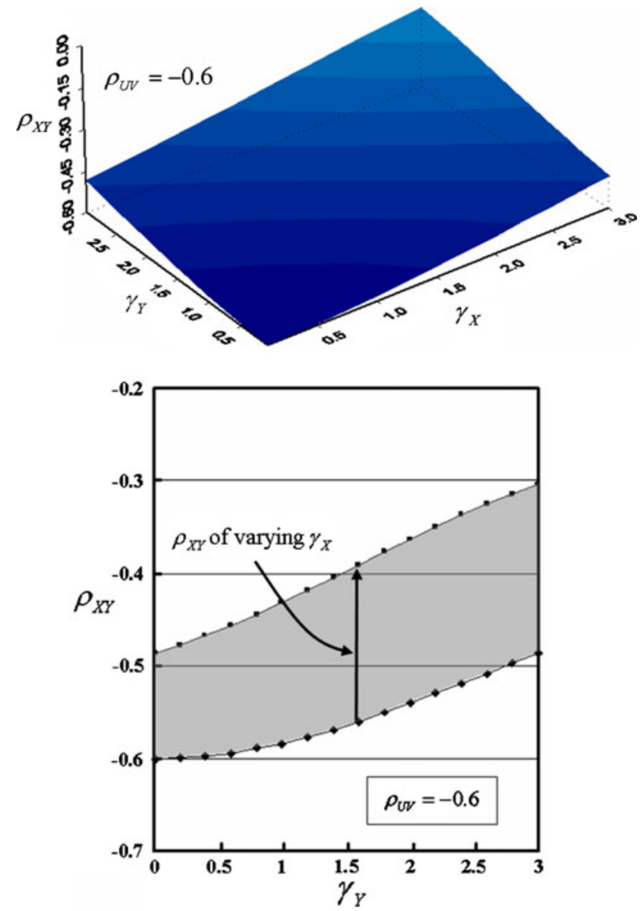


Fig. 9 Feasible regions of ρ_{XY} for two gamma random variables X and Y with $0 < \gamma_X, \gamma_Y \leq 3$ and $\rho_{UV} = -0.6$

available at http://www.rslabntu.net or by an email request to the corresponding author.

5 Conclusions

We have demonstrated that the proposed frequency-factor based bivariate gamma simulation approach is capable of generating random sample pairs which not only have the desired marginal densities of component random variables but also their correlation coefficient. The approach involves generation of bivariate normal samples with a correlation coefficient consistent with the correlation coefficient of the corresponding bivariate gamma samples. Then the bivariate normal samples are transformed to bivariate gamma samples using the well-known general equation of hydrological frequency analysis. A key element in the proposed approach lies on the single-value relationship between the correlation coefficients of the bivariate normal distribution and bivariate gamma distribution. Other concluding remarks are as follows:

1. Among the four estimators of coefficient of skewness, the maximum likelihood estimator is most superior. Although the adjusted sample coefficient of skewness can effectively reduce estimation bias for random samples with small to moderate sample sizes, its estimates are associated with higher standard deviations.
2. Scatter plots of simulated bivariate sample pairs exhibit appropriate linear patterns (dependence structure) that are commonly observed in environmental or hydrological applications.
3. The single-value relationship of ρ_{XY} and ρ_{UV} is dependent only on the coefficients of skewness of the two component random variables. For given ranges of γ_X andγ_Y (or equivalently, shape factors β_X and β_Y) the correlation coefficient ρ_{XY} will fall within a corresponding feasible range.
4. Random samples generated using the proposed approach are distributed with a joint probability density function of the Moran (1969) model.

5. The proposed approach can also be applied for stochastic simulation of the more general bivariate Pearson type III distribution.

Acknowledgments This research was originally initiated from a project funded by the Council of Agriculture, ROC. We also thank the National Science Council for subsequent funding support. We are grateful to the Sinotech Engineering Consultants, Inc. for providing scholarship for the second author during this study. We also appreciate the constructive and insightful review comments of two anonymous referees which have made this article more informative and in depth.

Appendix 1: Derivation of the relationship between ρ_{XY} and ρ_{UV} (Eq. 10)

Suppose that random variables X and Y form a bivariate gamma distribution. According to what described in Sect. 2, given the means (μ_X and μ_Y) and standard deviations (σ_X and σ_Y), X and Y can be, respectively, expressed in terms of their corresponding frequency factors K_X and K_Y, i.e.

$$X = \mu_X + K_X\sigma_X \tag{17}$$

and

$$Y = \mu_Y + K_Y\sigma_Y. \tag{18}$$

Note that, with given means μ_X and μ_Y and standard deviations σ_X and σ_Y, the coefficients of skewness γ_X and γ_Y can be readily determined by Eq. 9.

From the above equations, it can be easily shown that frequency factors K_X and K_Y are distributed with zero mean and unit standard deviation, and correlation coefficient of X and Y is equivalent to correlation coefficient of K_X and K_Y, i.e.

$$E[K_X] = E[K_Y] = 0, \tag{19}$$

$$\mathrm{Var}[K_X] = \mathrm{Var}[K_Y] = 1, \tag{20}$$

and

$$\rho_{XY} = \rho_{K_XK_Y}. \tag{21}$$

According to Eq. 2, the frequency factors K_X and K_Y can be, respectively, approximated by

$$K_X \approx U + (U^2-1)\frac{\gamma_X}{6} + \frac{1}{3}(U^3-6U)\left(\frac{\gamma_X}{6}\right)^2 - (U^2-1)\left(\frac{\gamma_X}{6}\right)^3 + U\left(\frac{\gamma_X}{6}\right)^4 - \frac{1}{3}\left(\frac{\gamma_X}{6}\right)^5 \tag{22}$$

and

$$K_Y \approx V + (V^2-1)\frac{\gamma_Y}{6} + \frac{1}{3}(V^3-6V)\left(\frac{\gamma_Y}{6}\right)^2 - (V^2-1)\left(\frac{\gamma_Y}{6}\right)^3 + V\left(\frac{\gamma_Y}{6}\right)^4 - \frac{1}{3}\left(\frac{\gamma_Y}{6}\right)^5, \tag{23}$$

where U and V both are random variables with standard normal density and are correlated with correlation coefficient ρ_{UV}.

From Eqs. 19 and 20, correlation coefficient of K_X and K_Y can be derived as follows:

$$\begin{aligned}\rho_{K_XK_Y} &= \mathrm{Cov}(K_X,K_Y) = E[K_XK_Y]\\ &\approx E\left\{\begin{array}{l}\left[U + (U^2-1)\frac{\gamma_X}{6} + \frac{1}{3}(U^3-6U)\left(\frac{\gamma_X}{6}\right)^2 - (U^2-1)\left(\frac{\gamma_X}{6}\right)^3 + U\left(\frac{\gamma_X}{6}\right)^4 - \frac{1}{3}\left(\frac{\gamma_X}{6}\right)^5\right]\cdot\\ \left[V + (V^2-1)\frac{\gamma_Y}{6} + \frac{1}{3}(V^3-6V)\left(\frac{\gamma_Y}{6}\right)^2 - (V^2-1)\left(\frac{\gamma_Y}{6}\right)^3 + V\left(\frac{\gamma_Y}{6}\right)^4 - \frac{1}{3}\left(\frac{\gamma_Y}{6}\right)^5\right]\end{array}\right\},\end{aligned} \tag{24}$$

$$\begin{aligned}K_X &\approx U\left[1+\left(\frac{\gamma_X}{6}\right)^4\right] + (U^2-1)\left[\frac{\gamma_X}{6} - \left(\frac{\gamma_X}{6}\right)^3\right]\\ &\quad + \frac{1}{3}(U^3-6U)\left(\frac{\gamma_X}{6}\right)^2 - \frac{1}{3}\left(\frac{\gamma_X}{6}\right)^5 = A_XU + B_X(U^2-1)\\ &\quad + C_X(U^3-6U) + D_X,\end{aligned} \tag{25}$$

$$\begin{aligned}K_Y &\approx V\left[1+\left(\frac{\gamma_Y}{6}\right)^4\right] + (V^2-1)\left[\frac{\gamma_Y}{6} - \left(\frac{\gamma_Y}{6}\right)^3\right]\\ &\quad + \frac{1}{3}(V^3-6V)\left(\frac{\gamma_Y}{6}\right)^2 - \frac{1}{3}\left(\frac{\gamma_Y}{6}\right)^5 = A_YV + B_Y(V^2-1)\\ &\quad + C_Y(V^3-6V) + D_Y\end{aligned} \tag{26}$$

where

$$A_X = 1+\left(\frac{\gamma_X}{6}\right)^4, B_X = \frac{\gamma_X}{6} - \left(\frac{\gamma_X}{6}\right)^3, C_X = \frac{1}{3}\left(\frac{\gamma_X}{6}\right)^2, D_X = -\frac{1}{3}\left(\frac{\gamma_X}{6}\right)^5, \tag{27}$$

$$A_Y = 1+\left(\frac{\gamma_Y}{6}\right)^4, B_Y = \frac{\gamma_Y}{6} - \left(\frac{\gamma_Y}{6}\right)^3, C_Y = \frac{1}{3}\left(\frac{\gamma_Y}{6}\right)^2, D_Y = -\frac{1}{3}\left(\frac{\gamma_Y}{6}\right)^5, \tag{28}$$

In the above equation, all expectations of the terms marked by arrow signs are zero and the above equation reduces to

$$\approx E\begin{bmatrix} A_X A_Y UV + \cancel{A_X B_Y U(V^2-1)} + A_X C_Y U(V^3-6V) + \cancel{A_X D_Y U} \\ + \cancel{B_X A_Y V(U^2-1)} + B_X B_Y (U^2-1)(V^2-1) + \cancel{B_X C_Y (U^2-1)(V^3-6V)} + \cancel{B_X D_Y (U^2-1)} \\ + C_X A_Y (U^3-6U)V + \cancel{C_X B_Y (U^3-6U)(V^2-1)} + C_X C_Y (U^3-6U)(V^3-6V) + \cancel{C_X D_Y (U^3-6U)} \\ + \cancel{D_X A_Y V} + \cancel{D_X B_Y (V^2-1)} + \cancel{D_X C_Y (V^3-6V)} + D_X D_Y \end{bmatrix} \tag{29}$$

$$E[K_X K_Y] \approx E\begin{bmatrix} A_X A_Y UV + A_X C_Y U(V^3-6V) + B_X B_Y (U^2-1)(V^2-1) \\ + C_X A_Y (U^3-6U)V + C_X C_Y (U^3-6U)(V^3-6V) + D_X D_Y \end{bmatrix}. \tag{30}$$

From Eq. 25,

$$E[K_X] \approx E\left[A_X U + B_X(U^2-1) + C_X(U^3-6U) + D_X\right] = D_X. \tag{31}$$

Since K_X and K_Y are distributed with zero means, it follows that

$$D_X = D_Y \approx 0. \tag{32}$$

Therefore, the correlation coefficient of K_X and K_Y can be approximated by

$$\begin{aligned} \rho_{K_X K_Y} &= E[K_X K_Y] \\ &\approx E\begin{bmatrix} A_X A_Y UV + A_X C_Y U(V^3-6V) + B_X B_Y (U^2-1)(V^2-1) \\ + C_X A_Y (U^3-6U)V + C_X C_Y (U^3-6U)(V^3-6V) \end{bmatrix} \\ &= A_X A_Y \rho_{UV} + A_X C_Y\left[E(UV^3) - 6\rho_{UV}\right] + B_X B_Y\left[E(U^2V^2) - 1\right] \\ &\quad + C_X A_Y\left[E(U^3V) - 6\rho_{UV}\right] \\ &\quad + C_X C_Y\left[E(U^3V^3) - 6E(UV^3) - 6E(U^3V) + 36\rho_{UV}\right]. \end{aligned} \tag{33}$$

Note that in the above derivations we have used the following properties

$$E(UV) = \rho_{UV}, \tag{34}$$

$$E(U^2) = E(V^2) = 1. \tag{35}$$

It can also be shown that (Kendall and Stuart 1977; Kan 2008)

$$E(U^2V^2) = 2\rho_{UV}^2 + 1, \tag{36}$$

$$E(U^3V^3) = 6\rho_{UV}^3 + 9\rho_{UV}, \tag{37}$$

$$E(U^3V) = E(UV^3) = 3\rho_{UV}. \tag{38}$$

Thus,

$$\begin{aligned} \rho_{XY} &= \rho_{K_X K_Y} \\ &\approx (A_X A_Y - 3A_X C_Y - 3C_X A_Y + 9C_X C_Y)\rho_{UV} \\ &\quad + 2B_X B_Y \rho_{UV}^2 + 6C_X C_Y \rho_{UV}^3. \end{aligned} \tag{39}$$

Appendix 2: Proof of Eq. 10 as a single-value function

Let $f(\rho_{UV}) = \partial\rho_{XY}/\partial\rho_{UV}$. From Eq. 10 we have

$$f(\rho_{UV}) = (A_X A_Y - 3A_X C_Y - 3C_X A_Y + 9C_X C_Y) + 4B_X B_Y \rho_{UV} + 18C_X C_Y \rho_{UV}^2, \tag{40}$$

$$A_X A_Y - 3A_X C_Y - 3C_X A_Y + 9C_X C_Y = (A_X - 3C_X)(A_Y - 3C_Y), \tag{41}$$

$$A_X - 3C_X = 1 + \left(\frac{\gamma_X}{6}\right)^4 - \left(\frac{\gamma_X}{6}\right)^2 = \left[\left(\frac{\gamma_X}{6}\right)^2 - 1\right]^2 + \left(\frac{\gamma_X}{6}\right)^2 > 0. \tag{42}$$

Similarly,

$$A_Y - 3C_Y = 1 + \left(\frac{\gamma_Y}{6}\right)^4 - \left(\frac{\gamma_Y}{6}\right)^2 = \left[\left(\frac{\gamma_Y}{6}\right)^2 - 1\right]^2 + \left(\frac{\gamma_Y}{6}\right)^2 > 0. \tag{43}$$

Therefore,

$$A_X A_Y - 3A_X C_Y - 3C_X A_Y + 9C_X C_Y = \left\{\left[\left(\frac{\gamma_X}{6}\right)^2 - 1\right]^2 + \left(\frac{\gamma_X}{6}\right)^2\right\}\left\{\left[\left(\frac{\gamma_Y}{6}\right)^2 - 1\right]^2 + \left(\frac{\gamma_Y}{6}\right)^2\right\}. \tag{44}$$

Let

$$\begin{aligned} g(\rho_{UV}) &= 4B_X B_Y \rho_{UV} + 18C_X C_Y \rho_{UV}^2 \\ &= \frac{\gamma_X\gamma_Y}{9}\left[\left(\frac{\gamma_X}{6}\right)^2 - 1\right]\left[\left(\frac{\gamma_Y}{6}\right)^2 - 1\right]\rho_{UV} \\ &\quad + \frac{18}{9}\left(\frac{\gamma_X}{6}\right)^2\left(\frac{\gamma_Y}{6}\right)^2\rho_{UV}^2. \end{aligned} \tag{45}$$

Also, let $G_X = \frac{\gamma_X}{6}$, $G_Y = \frac{\gamma_Y}{6}$. We then have

$$g(\rho_{UV}) = 4G_X G_Y (G_X^2 - 1)(G_Y^2 - 1)\rho_{UV} + 2G_X^2 G_Y^2 \rho_{UV}^2, \tag{46}$$

$$\begin{aligned} f(\rho_{UV}) &= \left[(G_X^2 - 1)^2 + G_X^2\right]\left[(G_Y^2 - 1)^2 + G_Y^2\right] \\ &\quad + 4G_X G_Y (G_X^2 - 1)(G_Y^2 - 1)\rho_{UV} + 2G_X^2 G_Y^2 \rho_{UV}^2, \end{aligned} \tag{47}$$

$$\frac{\partial f}{\partial \rho_{UV}} = 4G_X G_Y \left(G_X^2 - 1\right)\left(G_Y^2 - 1\right) + 4G_X^2 G_Y^2 \rho_{UV}. \qquad (48)$$

Let $\frac{\partial f}{\partial \rho_{UV}} = 0$, it yields an extreme value of f at

$$\rho_{UV}^* = -\frac{\left(G_X^2 - 1\right)\left(G_Y^2 - 1\right)}{G_X G_Y}. \qquad (49)$$

Substituting ρ_{UV}^* into Eq. 47 gives the extreme value of f as

$$\begin{aligned} f(\rho_{UV}^*) &= \left[\left(G_X^2 - 1\right)^2 + G_X^2\right]\left[\left(G_Y^2 - 1\right)^2 + G_Y^2\right] \\ &\quad - 4\left(G_X^2 - 1\right)^2\left(G_Y^2 - 1\right)^2 + 2\left(G_X^2 - 1\right)^2\left(G_Y^2 - 1\right)^2 \\ &= \left[\left(G_X^2 - 1\right)^2 + G_X^2\right]\left[\left(G_Y^2 - 1\right)^2 + G_Y^2\right] \\ &\quad - 2\left(G_X^2 - 1\right)^2\left(G_Y^2 - 1\right)^2 \\ &= G_X^2\left(G_Y^2 - 1\right)^2 + \left(G_X^2 - 1\right)^2 G_Y^2 + G_X^2 G_Y^2 \\ &\quad - \left(G_X^2 - 1\right)^2\left(G_Y^2 - 1\right)^2. \end{aligned} \qquad (50)$$

Since $-1 \leq \rho_{UV} \leq 1$ (or equivalently, $\rho_{UV}^2 \leq 1$), Eq. 49 yields

$$\left(G_X^2 - 1\right)^2\left(G_Y^2 - 1\right)^2 \leq G_X^2 G_Y^2. \qquad (51)$$

Thus,

$$f\left(\rho_{UV}^*\right) \geq G_X^2\left(G_Y^2 - 1\right)^2 + \left(G_X^2 - 1\right)^2 G_Y^2 > 0. \qquad (52)$$

We now check the second derivative of $f(\rho_{UV})$, i.e.

$$\frac{\partial^2 f}{\partial (\rho_{UV})^2} = 4G_X^2 G_Y^2 > 0. \qquad (53)$$

Therefore, $f\left(\rho_{UV}^*\right) > 0$ is the minimum of the function $f(\rho_{UV}) = \partial\, \rho_{XY}/\partial\, \rho_{UV}$. It follows that

$$f(\rho_{UV}) = \partial \rho_{XY}/\partial \rho_{UV} > 0 \qquad (54)$$

for all possible values of ρ_{UV}. The above equation indicates ρ_{XY} increases with increasing ρ_{UV}, and thus Eq. 10 is a single-value function.

Appendix 3: Method of moments and maximum likelihood estimators of the mean and coefficient of skewness of the gamma distribution

Equations 4–7 give the density function of a gamma distribution and expressions of the distribution parameters (α, β) as functions of the mean, standard deviation and coefficient of skewness. Given a random sample of size n, i.e. (x_1, …, x_n), we adopt the following conventional expressions of sample moments and central moments:

$$m_r' = \sum_{i=1}^{n} x_i^r \Big/ n, \quad m_r = \sum_{i=1}^{n} (x_i - \bar{x}_n)^r \Big/ n.$$

The method of moments estimator of the mean, variance and coefficient of skewness are, respectively, expressed as $\hat{\mu}_{\text{MOM}}$, $\hat{\sigma}^2_{\text{MOM}}$ and $\hat{\gamma}_{\text{MOM}}$. Thus,

$$\hat{\mu}_{\text{MOM}} = m_1' = \sum_{i=1}^{n} x_i \Big/ n = \bar{x}_n, \qquad (55)$$

$$\hat{\sigma}^2_{MOM} = m_2 = m_2' - (m_1')^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x}_n)^2 = \frac{(n-1)}{n} S_n^2, \qquad (56)$$

$$\left[S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x}_n)^2\right]$$

$$\hat{\gamma}_{\text{MOM}} = \frac{m_3}{(m_2)^{3/2}} = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x}_n)^3}{\left[(n-1)S_n^2/n\right]^{3/2}} = \frac{n\sum_{i=1}^{n}(x_i - \bar{x}_n)^3}{(n-1)\sqrt{n(n-1)}S_n^3}. \qquad (57)$$

The maximum likelihood estimators of the distribution parameters (α, β) have been documented by Kite (1988) and Rao and Hamed (2000):

$$\hat{\beta}_{MLE} = \begin{cases} \frac{1}{U}(0.500876 + 0.1648852U - 0.054427U^2) & 0 \leq U \leq 0.5772 \\ \frac{8.898919 + 9.05995U + 0.9775373U^2}{U(17.7928 + 11.968477U + U^2)} & 0.5772 \leq U \leq 17 \end{cases}, \qquad (58)$$

$$\hat{\alpha}_{\text{MLE}} = \bar{x}_n \Big/ \hat{\beta}_{\text{MLE}}, \qquad (59)$$

where $U = \ln \bar{x}_n - \ln G$ and $G = (x_1 x_2 \ldots x_n)^{1/n}$. From the invariance property of maximum likelihood estimators, we have

$$\hat{\mu}_{\text{MLE}} = \hat{\alpha}_{\text{MLE}}\hat{\beta}_{\text{MLE}} = \bar{x}_n, \qquad (60)$$

$$\hat{\gamma}_{\text{MLE}} = 2/\hat{\beta}_{\text{MLE}}. \qquad (61)$$

Appendix 4: The Moran bivariate gamma distribution (adapted from Moran 1969)

Suppose that U and V are two jointly distributed standard normal random variables. The cumulative distribution function of U and V are, respectively, $\Phi(u) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{u} e^{-\frac{1}{2}t^2}dt$ and $\Phi(v) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{v} e^{-\frac{1}{2}t^2}dt$, and the joint normal density is

$$\varphi_{UV}(u, v) = \frac{1}{2\pi\sqrt{1 - \rho_{UV}^2}}\exp\left[-\frac{u^2 - 2\rho uv + v^2}{2(1 - \rho_{UV}^2)}\right]. \qquad (62)$$

From the theory of probability integral transformation, we have

$$\Phi(U) = U^*, \qquad (63)$$

$$\Phi(V) = V^*, \qquad (64)$$

where U^* and V^* are jointly distributed uniform random variables with data interval (0, 1).

Now define two random variables X and Y by the equations

$$U^* = \frac{1}{\alpha_x \Gamma(\beta_x)} \int_0^X \left(\frac{t}{\alpha_x}\right)^{\beta_x - 1} e^{-t/\alpha_x} dt = F(X; \alpha_x, \beta_x) \quad (65)$$

and

$$V^* = F(Y; \alpha_y, \beta_y). \quad (66)$$

Then the joint probability density of X and Y can be written as

$$\begin{aligned} & f_{XY}(x, y; \alpha_x, \beta_x, \alpha_y, \beta_y, \rho_{UV}) \\ & = \frac{1}{2\pi\sqrt{1-\rho_{UV}^2}} \exp\left[-\frac{u^2 - 2\rho_{UV}uv + v^2}{2(1-\rho_{UV}^2)}\right] \frac{du\, du^*\, dv\, dv^*}{du^*\, dx\, dv^*\, dy} \\ & = \frac{1}{\sqrt{1-\rho_{UV}^2}} \exp\left\{-\frac{1}{2(1-\rho_{UV}^2)}\left[\rho_{UV}^2 \left(\Phi^{-1}(F_X(x; \alpha_x, \beta_x))\right)^2,\right.\right. \\ & \quad - 2\rho_{UV}\Phi^{-1}(F_X(x; \alpha_x, \beta_x))\Phi^{-1}(F_Y(y; \alpha_y, \beta_y)) \\ & \quad \left.\left. + \rho_{UV}^2 \left(\Phi^{-1}(F_Y(y; \alpha_y, \beta_y))\right)^2\right]\right\} \\ & \quad \times \frac{1}{\alpha_x \alpha_y \Gamma(\beta_x)\Gamma(\beta_y)} \left(\frac{x}{\alpha_x}\right)^{\beta_x - 1} \left(\frac{y}{\alpha_y}\right)^{\beta_y - 1} e^{-x/\alpha_x} e^{-y/\alpha_y}. \end{aligned} \quad (67)$$

Or equivalently,

$$\begin{aligned} & f_{XY}(x, y; \alpha_x, \beta_x, \alpha_y, \beta_y, \rho_{UV}) \\ & = \frac{1}{\sqrt{1-\rho_{UV}^2}} f_X(x; \alpha_x, \beta_x) f_Y(y; \alpha_y, \beta_y) \\ & \quad \times \exp\left[-\frac{(\rho_{UV}u)^2 - 2\rho_{UV}uv + (\rho_{UV}v)^2}{2(1-\rho_{UV}^2)}\right], \end{aligned} \quad (68)$$

where $u = \Phi^{-1}[F_X(x;\, \alpha_x,\, \beta_x)]$ and $v = \Phi^{-1}[F_Y(y;\, \alpha_y,\, \beta_y)]$.

Notice that the correlation coefficient between X and Y, i.e. ρ_{XY}, does not explicitly appear in the above equation of bivariate gamma density. In this study we have derived the relationship between ρ_{XY} and ρ_{UV} (Eqs. 10, 39) with which ρ_{XY} is implicitly included in the bivariate density function.

References

Akaike H (1974) A new look at statistical model identification. IEEE Trans Autom Control AC-19:716–723

Ashkar F, El-Jabi N, Issa M (1998) A bivariate analysis of the volume and duration of low-flow events. Stoch Hydrol Hydraul 12:97–116

Bacchi B, Becciu G, Kottegoda NT (1994) Bivariate exponential model applied to intensities and durations of extreme rainfall. J Hydrol 155:225–236

Benke KK, Hamilton AJ (2008) Quantitative microbial risk assessment: uncertainty and measures of central tendency for skewed distributions. Stoch Env Res Risk Assess 22:533–539

Bobee B, Robitaille R (1977) The use of the Pearson type III and log Pearson type III distribution revisited. Water Resour Res 13(2): 427–443

Cheng KS, Chiang JL, Hsu CW (2007) Simulation of probability distributions commonly used in hydrological frequency analysis. Hydrol Process 21:51–60

Cherian KC (1941) A bi-variate correlated gamma-type distribution function. J Indian Math Soc 5:133–144

Chow VT (1951) A general formula for hydrologic frequency analysis. Trans Am Geophys Union 32:231–237

Chow VT, Maidment DR, Mays LW (1988) Applied hydrology. McGraw-Hill, New York

Clark RT (1980) Bivariate gamma distribution for extending annual streamflow records from precipitation: Some large sample results. Water Resour Res 16:863–870

Cowpertwait PSP, O'Connell PE, Metcalfe AV, Mawdsley JA (1996) Stochastic point process modelling of rainfall I. Single-site fitting and validation. J Hydrol 175:17–46

D'Agostino RB, Stephens MA (1986) Goodness-of-fit techniques. Marcel Dekker, New York

D'este GM (1981) A Morgenstern-type bivariate gamma distribution. Biometrika 68(1):339–340

De Michele C, Salvadori G (2003) A generalized pareto intensity duration model of storm rainfall exploiting 2-copulas. J Geophys Res 108(D2):4067

Favre AC, El Adlouni S, Perreault L, Thiémonge N, Bobée B (2004) Multivariate hydrological frequency analysis using copulas. Water Resour Res 40:W01101

Goel NK, Kurothe RS, Mathur BS, Vogel RM (2000) A derived flood frequency distribution for correlated rainfall intensity and duration. J Hydrol 228:56–67

Grimaldia S, Serinaldi F (2006) Asymmetric copula in multivariate flood frequency analysis. Adv Water Resour 29(8):1155–1167

Hannan EJ, Quinn BG (1979) The determination of the order of an autoregression. J R Stat Soc B 41(2):190–195

Hosking JRM, Wallis JR (1997) Regional Frequency analysis: an approach based on L-moments. Cambridge University Press, Cambridge

Izawa T (1953) The bivariate gamma distribution. Clim Stat 4(1): 9–15 (in Japanese)

Kan R (2008) From moments of sum to moments of products. J Multivar Anal 99(3):542–554

Kelly KS, Krzysztofowicz R (1997) A bivariate meta-Gaussian density for use in hydrology. Stoch Hydrol Hydraul 11:17–31

Kendall MG, Stuart A (1977) The advanced theory of statistics, vol 1, Distribution theory, 4th edn. Charles Griffin, London

Khalili M, Brissette F, Leconte R (2009) Stochastic multi-site generation of daily weather data. Stoch Env Res Risk Assess 23:837–849

Kibble WF (1941) A two-variate gamma type distribution. Sankhya 5:137–150

Kim KH, Lee KK (2007) Optimization of groundwater-monitoring networks for identification of the distribution of a contaminant plume. Stoch Env Res Risk Assess 21:785–794

Kite GW (1988) Frequency and risk analysis in hydrology. Water Resources Publications, Fort Colins

Kottegoda NT (1980) Stochastic water resources technology. Macmillan, London

Law AM (2007) Simulation modeling and analysis. McGraw Hill, Singapore

Lee CH, Kim TW, Chung G, Choi M, Yoo C (2010) Application of bivariate frequency analysis to the derivation of rainfall–frequency curves. Stoch Env Res Risk Assess 24:389–397

Leiva V, Barros M, Paula GA, Sanhueza A (2008) Generalized Birnbaum–Saunders distributions applied to air pollutant concentration. Environmetrics 19:235–249

Liou JJ, Wu YC, Cheng KS (2008) Establishing acceptance regions for *L*-moments-based goodness-of-fit test by stochastic simulation. J Hydrol 355:49–62

Loaiciga HA, Leipnik RB (2005) Correlated gamma variables in the analysis of microbial densities in water. Adv Water Resour 28: 329–335

Loganathan GV, Kuo CY, Yannaccone J (1987) Joint probability distribution of streamflows and tides in estuaries. Nordic Hydrol 18:237–246

Moran PAP (1969) Statistical inference with bivariate gamma distributions. Biometrika 56(3):627–634

Nadarajah S, Gupta AK (2006a) Some bivariate gamma distributions. Appl Math Lett 19:767–774

Nadarajah S, Gupta AK (2006b) Intensity-duration models based on bivariate gamma distributions. Hiroshima Math J 36:387–395

National Research Council (2000) Risk analysis and uncertainty in flood damage reduction studies. National Academy Press, Washington

Rao AR, Hamed KH (2000) Flood frequency analysis. CRC Press, New York

Renard B, Langa M (2007) Use of a Gaussian copula for multivariate extreme value analysis: Some case studies in hydrology. Adv Water Resour 30(4):897–912

Schmeiser BW, Lal R (1982) Bivariate gamma random vectors. Oper Res 30(2):355–374

Scholz K (1997) Stochastic simulation of urban hydrological processes. Water Sci Technol 36:25–31

Schwarz G (1978) Estimating the dimension of a model. Ann Stat 6(2):461–464

Singh K, Singh VP (1991) Derivation of bivariate probability density functions with exponential marginals. J Stoch Hydrol Hydraul 5:55–68

Tsukatani T, Shighemitsu K (1980) Simplified Pearson distributions applied to air pollutant concentration. Atmos Environ 14:245–253

Werner TM, Kadlec RH (2000) Stochastic simulation of partially-mixed, event-driven treatment wetlands. Ecol Eng 14(3): 253–267

Yue S (1999) Applying bivariate normal distribution to flood frequency analysis. Water Int 24(3):248–254

Yue S (2000) Joint probability distribution of annual maximum storm peaks and amounts as represented by daily rainfalls. Hydrol Sci J 45(2):315–326

Yue S (2001a) A statistical measure of severity of El Niño events. Stoch Env Res Risk Assess 15:153–172

Yue S (2001b) A bivariate gamma distribution for use in multivariate flood frequency analysis. Hydrol Process 15:1033–1045

Yue S, Ouarda TBMJ, Bobee B (2001) A review of bivariate gamma distributions for hydrological application. J Hydrol 246:1–18

Zhang L, Singh VP (2007) Bivariate rainfall frequency distributions using Archimedean copulas. J Hydrol 332(1–2):93–109