

MODELLING AND SIMULATION OF VOLUMETRIC RAINFALL FOR A CATCHMENT IN THE MURRAY–DARLING BASIN

J. BOLAND¹, P. HOWLETT^{✉1}, J. PIANTADOSI¹ and R. ZAKARIA²

(Received 6 March, 2016; accepted 17 June, 2016; first published online 8 September 2016)

Abstract

We discuss modelling and simulation of volumetric rainfall in a catchment of the Murray–Darling Basin – an important food production region in Australia that was seriously affected by a recent prolonged drought. Consequently, there has been sustained interest in development of improved water management policies. In order to model accumulated volumetric catchment rainfall over a fixed time period, it is necessary to sum weighted rainfall depths at representative sites within each sub-catchment. Since sub-catchment rainfall may be highly correlated, the use of a Gamma distribution to model rainfall at each site means that catchment rainfall is expressed as a sum of correlated Gamma random variables. We compare four different models and conclude that a joint probability distribution for catchment rainfall constructed by using a copula of maximum entropy is the most effective.

2010 *Mathematics subject classification*: primary 62E17; secondary 62P12, 65C05, 65C20.

Keywords and phrases: rainfall modelling, correlated Gamma distributions, copulas of maximum entropy.

1. Introduction

We consider modelling and simulation of volumetric rainfall in a catchment of the Murray–Darling Basin in southeast Australia. Our methods elaborate and extend a previous work of Zakaria [30]. We compare the performance of four different models – a standard Gamma distribution fitted directly to the observed volumetric catchment rainfall by the method of maximum likelihood, the Alouini model [1] for the sum

¹Scheduling and Control Group (SCG), Centre for Industrial and Applied Mathematics (CIAM), School of Information Technology and Mathematical Sciences, University of South Australia, Mawson Lakes, 5095, South Australia; e-mail: john.boland@unisa.edu.au, phil.howlett@unisa.edu.au, julia.piantadosi@unisa.edu.au.

²Faculty of Industrial Sciences and Technology, Universiti Malaysia Pahang, Lebuhraya Tun Razak, 26300 Gambang, Kuantan Pahang, Malaysia; e-mail: roslinazairimah@ump.edu.my.

© Australian Mathematical Society 2016, Serial-fee code 1446-1811/2016 \$16.00

of known correlated Gamma distributions each having the same shape parameter, a modified Alouini model using the correlative coherence analysis (CCA) proposed by Getz [8] to simplify the correlation structure and a model which uses a checkerboard copula of maximum entropy [19] to construct a joint probability distribution for the volumetric catchment rainfall that incorporates the prescribed marginal sub-catchment rainfall distributions and matches the specified grade correlation coefficients.

We outline the structure of each model and discuss their relative merits. To demonstrate the models, we use official records from the Australian Bureau of Meteorology for seasonal rainfall in winter – the total rainfall in the months of June, July and August at stations 081013 (Dookie Agricultural College), 082039 (Rutherglen Research Station), 082001 (Beechworth Composite) and 072023 (Hume Reservoir) in the Murray–Darling Basin in southeast Australia. We use data for all years between 1922 and 2000. Longer term data is not available due to missing records in 2001 at one of the stations. The winter season was chosen, because winter rainfall makes a significant contribution to total volumetric catchment rainfall. The analysis could equally well be applied to other seasons or to other time periods. We show that all four models generate simulated catchment rainfalls that provide a good statistical match to the observed data. However, we emphasize that only one of the four models, the checkerboard copula of maximum entropy, can be used to simulate both the volumetric winter rainfall in the entire catchment and the corresponding components of volumetric winter rainfall in each of the four sub-catchments.

In this application, the Alouini model and the modified Alouini model are bedevilled by unwieldy numerical calculations. The probability density functions are represented as an infinite sum of weighted Gamma distributions and in each case some 3000 terms are required for adequate convergence. Moreover the numerical calculations require considerable care because individual terms may involve the products of very large and very small numbers. The modified Alouini model uses the CCA to replace the observed pairwise correlation coefficients by a single representative coefficient for all distinct pairs. Despite the simplified structure of the correlation matrix, the numerical difficulties are still present and the rate of convergence remains much the same.

Since each simulated outcome for the Alouini model generates an overall catchment rainfall but does not generate the corresponding components of sub-catchment rainfall, it seems that the main value of the model is to provide a *conceptual* solution to the problem of finding an exact distribution that represents a finite sum of known correlated Gamma random variables with a common shape parameter. Indeed, we will show that the probability density functions found by both the Alouini model and the modified Alouini model are almost identical to the density function defined by a standard two-parameter Gamma distribution fitted directly to the observed data by the method of maximum likelihood. Thus, the intensive numerical effort needed for the Alouini models is – for this application – a practical waste of time.

The remaining model – the checkerboard copula of maximum entropy – is much more useful than the others for simulation of catchment rainfall. In order to simulate

simultaneous winter rainfall in all four sub-catchments, it is necessary to construct a four-dimensional joint probability distribution with prescribed marginal distributions and specified grade correlation coefficients. For a given uniform subdivision of the unit four-dimensional hypercube, the checkerboard copula of maximum entropy [19] is the most disordered joint distribution with uniform marginals and constant density on each sub-hypercube that satisfies the constraints imposed by the specified grade correlation coefficients. The copula is defined by a multiply-stochastic hypermatrix $\mathbf{h} \in \mathbb{R}^\ell$, where $\ell = n^4$ is the number of equal subdivisions of the unit hypercube. By an elementary transformation on each marginal distribution, the copula can be used to find a corresponding joint probability distribution for the catchment rainfall with prescribed marginal Gamma distributions for each of the sub-catchment rainfalls and specified grade correlation coefficients.

Although some intricate numerical mathematics is required to determine the checkerboard copula of maximum entropy, the necessary calculations and the simulation itself can be performed accurately and efficiently on a standard laptop computer. Special MATLAB programs for the required calculations are available on request. The model assumes prescribed marginal distributions and specified pairwise grade correlation coefficients. In this case, we use marginal Gamma distributions fitted by the method of maximum likelihood to the observed sub-catchment rainfalls and grade correlation coefficients also estimated from the observed sub-catchment rainfall data using the Spearman rank correlation coefficients [21]. While there is no simple expression for the theoretical probability density defined by the checkerboard copula of maximum entropy, the empirical probability distribution for large samples generated by random simulations is stable and statistically consistent with the observed distribution.

1.1. Previous research on rainfall modelling One must distinguish between models, such as the ones discussed here, that are designed to simulate realistic rainfall statistics and those that also attempt to predict future rainfall. The latter task is immensely more difficult and must necessarily incorporate observations of indices which are known to be associated with short-term or long-term variations in climatological conditions.

The Gamma distribution has been widely used to model strictly positive rainfall accumulations at a single site [25] and also at multiple sites in the same geographic region [24]. Other authors [18, 22, 28] used a mixed Gamma distribution with an additional parameter to allow a positive probability for zero rainfall totals. A more sophisticated model [26] proposed random generation of wet and dry days by a Markov chain with independent generation of rainfall depths on wet days by a Gamma distribution. A similar model – with either a Markov chain or a binomial distribution to determine wet and dry days and an exponential distribution for rainfall depths on wet days – was used in [27] to generate monthly rainfall totals. More general chain-dependent processes [13] were devised to investigate *overdispersion* in rainfall models for different time scales.

The exponential dispersion model [12] is an important form of generalized linear model that has been used as a basis for modelling rainfall occurrence and depth. The Poisson–Gamma distribution [5, 9] is one such model where a Poisson process generates random rainfall events at a constant rate over a fixed time interval, and a Gamma distribution determines the rainfall depth during each event. These models belong to the *so-called* Tweedie family of generalized linear models [12] which allow the possibility of either a zero total, if there is no rainfall event, or a nonzero total. Special algorithms [6, 7] are required for calculation of probability density functions for the Tweedie distributions. A Tweedie distribution with different parameters for different times of the year [10] provides a good model for monthly rainfall, but recent work [11] suggests that more complex Tweedie models may be required to model the cyclical patterns that are sometimes evident in observed rainfall data. In particular, it was found that the addition of climatological variables improves the fit and makes substantial changes to the predicted monthly rainfall and the probability of a dry month. Climatological variables, such as NINO 3.4, are thought to have a significant impact on rainfall patterns in eastern Australia.

Synthetic rainfall data has often been used as a driver for simulation of water management systems where a full range of possible rainfall scenarios is desirable. The rGEN algorithm [20] uses a diagonal-band copula with marginal daily mixed Gamma distributions to generate synthetic rainfall totals on different time scales – daily, monthly and yearly.

Finally, we cite two papers [25, 29], which provide a more comprehensive review of weather generation models.

1.2. A probability distribution for the sum of Gamma random variables It is a common practice to model rainfall depth at a single site over a fixed time period as a random variable following a Gamma distribution. The volumetric rainfall in a particular region can then be modelled as the rainfall depth at a representative site multiplied by the area of the region. Thus, volumetric rainfall is also modelled by a Gamma distribution. Since catchment rainfall is calculated as the sum of the relevant sub-catchment rainfalls, we are now led to a model for catchment rainfall as a random variable defined by the sum of Gamma random variables.

Kotz and Adams [14] used moment generating functions to find the distribution of a sum of identically distributed Gamma random variables with a negative exponential correlation law. Their method has been the intuitive basis for much of the subsequent work. A later paper by Mathai [15] modelled storage capacity in a dam with Gamma-type distribution inputs and used zonal polynomials and confluent hypergeometric functions to find an exact expression for the probability density function of a sum of independent Gamma random variables. The Moschopoulos variation [16] on the Mathai method for inversion of the moment generating function allowed the probability density function for the sum of independent Gamma variables to be expressed as an infinite sum of weighted Gamma densities with a common scale parameter. A nice feature of the Moschopoulos series is a recursive formula for the weights. The Moschopoulos model was later exploited by Alouini et al. [1] to find a

probability density function for the sum of correlated Gamma variables, where each Gamma variable is defined by the same shape parameter.

Despite intensive research into multivariate Gamma distributions [4], there is no multidimensional classical distribution with prescribed marginal Gamma distributions that matches an arbitrary set of specified grade correlation coefficients. Nevertheless, the theory of copulas [17] can be used to construct joint probability densities that preserve the marginal distributions and match observed correlation coefficients. In particular, we note that checkerboard copulas of maximum entropy and checkerboard normal copulas have been used [2, 19, 20] to construct joint distributions for seasonal rainfall with correlated Gamma distributions for the marginal monthly rainfalls. Although these joint distributions were constructed specifically to model correlated temporal dependence at a single site for successive time periods, they can equally well be applied to model correlated spatial dependence at multiple sites over the same time period. The method is theoretically applicable to any number of marginal distributions, but becomes more complex numerically as the dimension increases. In this application, we consider a four-dimensional problem.

2. The observed data

We used official records from the Australian Bureau of Meteorology for stations 081013 (Dookie Agricultural College), 082039 (Rutherglen Research Station), 082001 (Beechworth Composite) and 072023 (Hume Reservoir) in the Murray–Darling Basin in southeast Australia. See the Australian Bureau of Meteorology website (<http://www.bom.gov.au/climate/data/>) for access to the detailed records. We considered all years between 1922 and 2000 – the longest time period with complete winter rainfall records at each site. The winter season was chosen as a suitable time period for the model. It is pertinent to note that the observed volumetric winter rainfall totals at each site are all strictly positive.

The areas in square kilometres for the sub-catchments in which each of the above stations is located were calculated approximately using the interactive map software from the Victorian Government Land Services and Spatial Information website [3]. The volumetric rainfall in a sub-catchment is given in cubic metres (m^3) by the formula

$$\text{volumetric rainfall} = [\text{sub-catchment area} \times 10^6] \times [\text{rainfall depth} \times 10^{-3}],$$

where the sub-catchment area is given in the usual units of square kilometres (km^2) and rainfall depth is measured in the standard units of millimetres (mm). The areas of the respective sub-catchments are Dookie 905 km^2 , Rutherglen 1538 km^2 , Beechworth 1847 km^2 and Hume 3390 km^2 . The mean winter rainfall depths are Dookie 169 mm, Rutherglen 182 mm, Beechworth 344 mm and Hume 223 mm. The mean sub-catchment volumetric winter rainfalls are Dookie $153 \times 10^6 \text{ m}^3$, Rutherglen $280 \times 10^6 \text{ m}^3$, Beechworth $636 \times 10^6 \text{ m}^3$ and Hume $756 \times 10^6 \text{ m}^3$. In order to work with reasonable numbers, henceforth, we use nonstandard units of 10^6 m^3 for volumetric rainfall. Histograms for the observed volumetric winter rainfall in

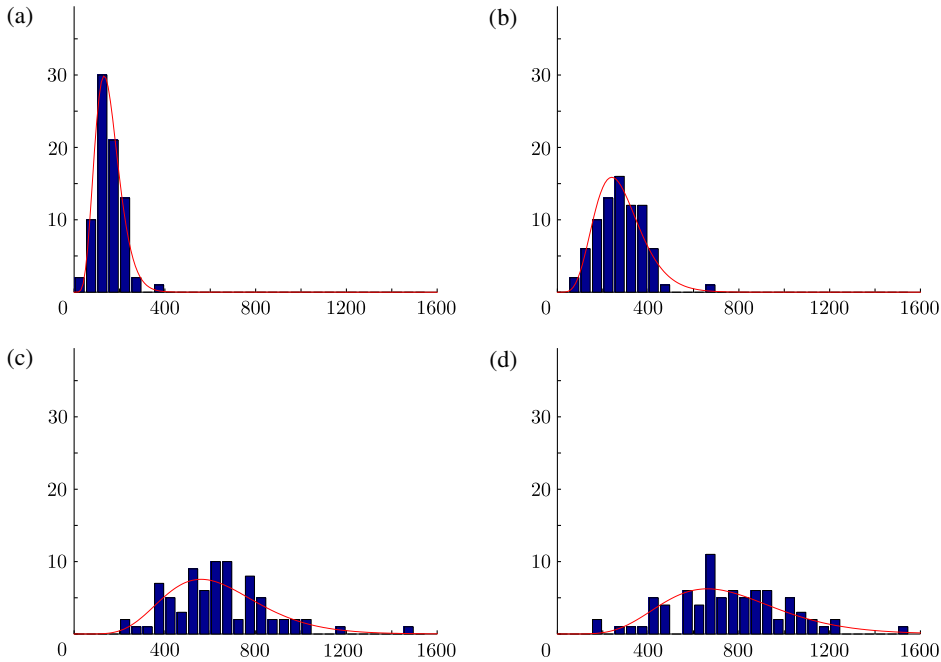


FIGURE 1. Histograms for observed volumetric winter rainfall in units of 10^6 m^3 at Dookie (a), Rutherglen (b), Beechworth (c) and Hume (d) with fitted maximum likelihood Gamma distributions.

the Dookie, Rutherglen, Beechworth and Hume sub-catchments with fitted Gamma distributions are shown in Figure 1 with the corresponding time series in Figure 2. We use a common scale for all histograms in Figure 1 and all graphs in Figure 2 to allow a realistic comparison of the relative contributions from each sub-catchment. In Figure 3, we show a histogram for the observed volumetric winter rainfall over the entire catchment with a fitted Gamma distribution and the corresponding time series.

3. A model for catchment rainfall using the Gamma distribution

When there is no probability of a zero total, an elementary model for volumetric catchment rainfall can be obtained by fitting a two-parameter Gamma distribution to the observed data by the method of maximum likelihood. The probability density function for the Gamma distribution $\Gamma(\alpha, \beta)$ with shape parameter $\alpha > 0$ and scale parameter $\beta > 0$ is defined for $x > 0$ by the formula

$$g_{\alpha, \beta}(x) = \frac{x^{\alpha-1} e^{-x/\beta}}{\beta^\alpha \Gamma(\alpha)}. \tag{3.1}$$

The coefficients α and β in the maximum likelihood estimate must satisfy the equations

$$\alpha\beta = \bar{x} \quad \text{and} \quad \psi(\alpha) + \log_e \beta = \overline{\log_e x},$$

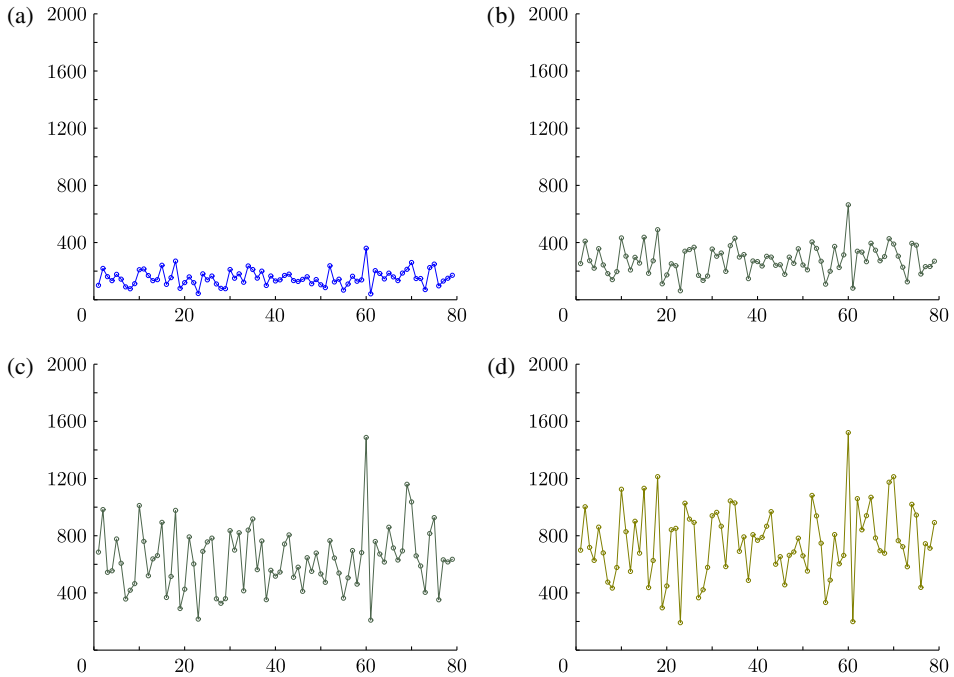


FIGURE 2. Time series for observed volumetric winter rainfall for Dookie (a), Rutherglen (b), Beechworth (c) and Hume (d).

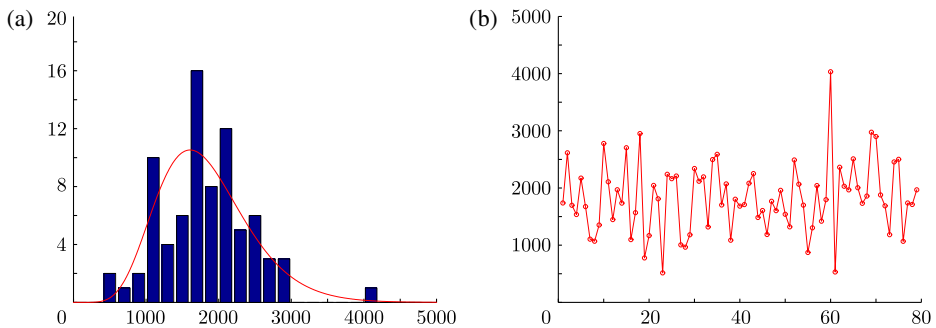


FIGURE 3. Histogram of observed volumetric catchment winter rainfall in units of 10^6 m^3 compared with maximum likelihood Gamma distribution (a) and time series for observed volumetric winter catchment rainfall in units of 10^6 m^3 (b).

where $\psi(\alpha) = \Gamma'(\alpha)/\Gamma(\alpha)$ is the Digamma function, $\bar{x} = (1/N) \sum_{k=1}^N x_k$ is the mean of the observed values $\{x_k\}_{k=1}^N$ and $\overline{\log_e x} = (1/N) \sum_{k=1}^N \log_e x_k$ is the mean of the logarithms of the observed values $\{\log_e x_k\}_{k=1}^N$. The observed catchment volumetric winter rainfall data gives $\bar{x} = 1824$ and $\overline{\log_e x} = 7.45$. Thus, we obtained $Y \sim \Gamma(\alpha, \beta)$,

where $\alpha = 8.36$ and $\beta = 218$. For later reference, note that the theoretical standard deviation for the fitted maximum likelihood Gamma distribution $\sigma = \sqrt{\alpha\beta^2} \approx 631$ is a little higher than the standard deviation for the observed data $s \approx 609$. Nevertheless, a Kolmogorov–Smirnov (KS) test [23] showed that the maximum likelihood Gamma distribution is a good fit to the observed data and should not be rejected at the 5% significance level.

4. The Alouini model for catchment rainfall

The Alouini model [1] is an extension of an earlier Moschopoulos model [16]. We consider a set of m correlated Gamma random variables $X_i \sim \Gamma(\alpha, \beta_i)$ with common shape parameter $\alpha > 0$ and with scale parameters $\beta_i > 0$ for each $i = 1, \dots, m$ and an associated random variable $Y = \sum_{i=1}^m X_i$. It is necessary to assume that the Pearson correlation coefficients $\rho_{i,j}$ are nonnegative for all $i, j = 1, \dots, m$. See [21] for more information. Define a diagonal matrix $B \in \mathbb{R}^{m \times m}$ and a symmetric matrix $C \in \mathbb{R}^{m \times m}$ by

$$B = \begin{bmatrix} \beta_1 & 0 & \cdots & 0 \\ 0 & \beta_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \beta_m \end{bmatrix}, \quad C = \begin{bmatrix} 1 & \sqrt{\rho_{1,2}} & \cdots & \sqrt{\rho_{1,m}} \\ \sqrt{\rho_{1,2}} & 1 & \cdots & \sqrt{\rho_{2,m}} \\ \vdots & \vdots & \ddots & \vdots \\ \sqrt{\rho_{1,m}} & \sqrt{\rho_{2,m}} & \cdots & 1 \end{bmatrix},$$

where $0 < \beta_1 \leq \dots \leq \beta_m$. Let $\lambda = (\lambda_1, \dots, \lambda_m)$ be the vector of eigenvalues for the matrix $A = BC$. We assume that A is nonnegative and that $0 \leq \lambda_1 \leq \dots \leq \lambda_m$. The probability density function for Y is given by

$$f_Y(y) = \sum_{k=0}^{\infty} \delta_k g_{m\alpha+k, \lambda_1}(y), \tag{4.1}$$

where $g_{m\alpha+k, \lambda_1}(\cdot)$ denotes the probability density function defined in (3.1) for the Gamma distribution $\Gamma(m\alpha + k, \lambda_1)$, and where the coefficients δ_k are defined recursively by the formula

$$\delta_{k+1} = \frac{\alpha}{k+1} \sum_{i=1}^{k+1} \left[\sum_{j=1}^m \left(1 - \frac{\lambda_1}{\lambda_j}\right)^i \right] \delta_{k+1-i}$$

for $k = 0, 1, \dots$ with $\delta_0 = \prod_{j=1}^m (\lambda_1/\lambda_j)^\alpha$. With this scaling, $\sum_{k=0}^{\infty} \delta_k = 1$ (see [1, Theorem 1] and the original paper by Moschopoulos [16] for more details). The Alouini method works well if $\delta_k \rightarrow 0$ rapidly. This requires $1 - \lambda_1/\lambda_m \ll 1$.

REMARK 4.1. Application of the Alouini model depends on knowing that the eigenvalues of the matrix C are nonnegative. In such cases, it is a common usage to say that C is nonnegative. Although C was originally described as nonnegative by Alouini et al. [1], we present a counter-example in Appendix B to show that C may fail to be nonnegative. If so, the Alouini method will fail.

TABLE 1. Maximum likelihood Gamma distributions for volumetric sub-catchment rainfall with $\mu_i = \alpha_i\beta_i$ and $\sigma_i = \sqrt{\alpha_i\beta_i^2}$ and key observed values \bar{x}_i , $\overline{\log_e x_i}$ and s_i .

	α_i	β_i	\bar{x}_i	σ_i	$\overline{\log_e x_i}$	s_i
Dookie	7.46	20.48	153	56	4.96	55
Rutherglen	7.00	39.99	280	106	5.56	101
Beechworth	8.36	76.06	636	220	6.39	220
Hume	8.02	94.19	756	267	6.56	250

TABLE 2. Spearman rank correlation coefficients.

	Dookie	Rutherglen	Beechworth	Hume
Dookie	1	0.87	0.87	0.83
Rutherglen	0.87	1	0.93	0.90
Beechworth	0.87	0.93	1	0.91
Hume	0.83	0.90	0.91	1

We begin by fitting a Gamma distribution to each of the four individual volumetric sub-catchment rainfall data sets using the method of maximum likelihood. The relevant data values and estimated parameters are shown in Table 1. A Kolmogorov–Smirnov test confirms that the maximum likelihood Gamma distribution is a good fit in all the four cases.

To estimate pairwise correlations, we used the Spearman rank correlation coefficients given by

$$\widehat{\rho}_{i,j} = 1 - \frac{6 \sum_{k=1}^N (r_{i,k} - r_{j,k})^2}{N(N^2 - 1)},$$

where $r_{i,k}$ and $r_{j,k}$ are the respective ranks for the observations $(x_{i,k}, x_{j,k})$ in the year k from the observed data sets $\{(x_{i,k}, x_{j,k})\}_{k=1}^N$ for $i, j = 1, \dots, m$. The Spearman coefficients are shown in Table 2.

The Alouini method uses an infinite series to represent the probability density function for the sum of a finite number of correlated Gamma random variables with a common shape parameter. Since the values of α_i in Table 1 are not too dissimilar, we decided to use a common shape parameter $\alpha_c = \sum_{i=1}^4 \alpha_i/4 \approx 7.71$ to calculate revised Gamma distributions to model $X_i \sim \Gamma(\alpha_c, \beta_{c,i})$ for each $i = 1, \dots, 4$ with a common shape parameter $\alpha = \alpha_c$ and distinct scale parameters $\{\beta_{c,i}\}_{i=1}^4$ using the method of maximum likelihood. The maximum likelihood equations, in this case, are reduced to the single equation $\alpha_c \beta_{c,i} = \bar{x}_i$. The revised values are shown in Table 3. Once again a Kolmogorov–Smirnov test confirms that the revised Gamma distributions provide a good fit to the observed data.

TABLE 3. Revised Gamma distributions for volumetric sub-catchment rainfall with $\mu_c = \alpha_c \beta_{c,i}$ and $\sigma_c = \sqrt{\alpha_c \beta_{c,i}^2}$.

	α_c	$\beta_{c,i}$	μ_c	σ_c
Dookie	7.71	19.82	153	55
Rutherglen	7.71	36.28	280	101
Beechworth	7.71	82.45	636	229
Hume	7.71	98.04	756	272

We used the revised parameter values to calculate the probability density function $f_Y(y)$ for $Y = \sum_{i=1}^4 X_i$ according to (4.1). The matrix $A = BC$ is given by

$$A \approx \begin{bmatrix} 19.8 & 18.4 & 18.5 & 18.1 \\ 33.8 & 36.3 & 34.9 & 34.5 \\ 77.1 & 79.4 & 82.4 & 78.7 \\ 89.5 & 93.1 & 93.6 & 98.0 \end{bmatrix}.$$

The eigenvalues of A are $\lambda \approx (1.66, 2.11, 4.50, 228)$. We used MATLAB to find the approximate Alouini probability density function from the series (4.1). Convergence is extremely slow, and it is necessary to use around 3000 terms to obtain a good approximation to $f_Y(y)$.

5. The modified Alouini model

The measured pairwise Spearman rank correlation coefficients for the volumetric monthly rainfall totals at the sub-catchment sites are all very similar (see Table 2). This suggests that we could apply the correlative coherence analysis (CCA) proposed by Getz [8] in order to construct a single overall correlation coefficient $r \in [0, 1]$ for the random vector $X = (X_1, \dots, X_m)$ with the matrix of Pearson correlation coefficients [21] $R(X) = [\rho_{i,j}(X)] \in \mathbb{R}^{m \times m}$. Getz suggested that the diversity of eigenvalues for the correlation matrix $R(X)$ could be described using a Shannon–Weaver measure [8]

$$H(X) = 1 - \frac{1}{\log_e(1/m)} \sum_{i=1}^m (\lambda_i/m) \log_e(\lambda_i/m) \in [0, 1].$$

Now consider a special correlation matrix $R(r) = [r_{i,j}] \in \mathbb{R}^{m \times m}$, where $r_{i,i} = 1$ and $r_{i,j} = r \in [0, 1]$ for $i \neq j$. In this case, the diversity measure $h(r) = H[R(r)]$ reduces to

$$h(r) = \frac{[1 + (m - 1)r] \log_e[1 + (m - 1)r] + (m - 1)(1 - r) \log_e(1 - r)}{m \log_e m}$$

and there is a unique correlative coherence value r for the random vector X defined by

$$h(r) = H(X) \iff r = h^{-1}[H(X)].$$

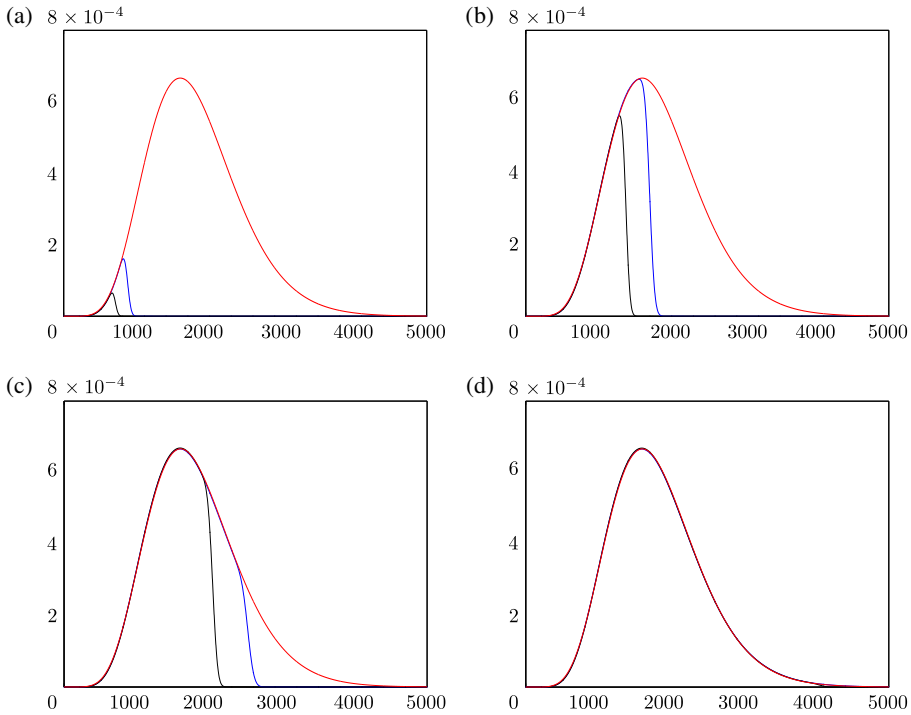


FIGURE 4. Density for the maximum likelihood Gamma distribution compared to partial sums for the Alouini model (converging more quickly left to right) and the modified Alouini model (converging more slowly left to right) showing 500 terms (a), 1000 terms (b), 1500 terms (c) and 3000 terms (d). Note that the convergence is visually complete with 3000 terms.

The Alouini method can now be applied with the observed correlation matrix $R = R(X)$ replaced by the special CCA matrix $R = R(r)$. Note that $C = R(\sqrt{r})$ is always nonnegative. We calculated $r \approx 0.89$ and hence $A = BC$ is given by

$$A \approx \begin{bmatrix} 19.8 & 18.7 & 18.7 & 18.7 \\ 34.2 & 36.3 & 34.2 & 34.2 \\ 77.8 & 77.8 & 82.4 & 77.8 \\ 92.4 & 92.4 & 92.4 & 98.0 \end{bmatrix}.$$

Although the Getz model is more robust, the convergence rate in this case is slightly slower. This is clearly shown by the graphs in Figure 4, where the partial sums are compared to the maximum likelihood Gamma density.

6. A comparison of the first three models

Each of the first three models – the maximum likelihood Gamma distribution, the Alouini model and the modified Alouini model – is designed to construct a density function for the total volumetric catchment winter rainfall. From a practical

viewpoint, at least in this application, the theoretical significance of the Alouini model is completely lost, because the resultant density is visually indistinguishable from the much simpler density generated by the maximum likelihood Gamma distribution. Our conclusion is that in purely practical terms, to simulate volumetric catchment winter rainfall, it is much better to use the elementary Gamma distribution.

The methods discussed so far simulate overall catchment rainfall, but do not simulate the corresponding components of sub-catchment rainfall. Therefore, the models replicate the overall risk of drought or flood, but do not simultaneously model the risk of drought or flood in the sub-catchments.

7. The model using a checkerboard copula of maximum entropy

The theory justifying the checkerboard copula of maximum entropy can be found elsewhere [19]. In this paper, we outline the basic ideas and present the results of the application. An m -dimensional *copula*, where $m \geq 2$, is a cumulative probability distribution $C(\mathbf{u}) \in [0, \infty)$, defined on the m -dimensional unit hypercube $\mathbf{u} = (u_1, \dots, u_m) \in [0, 1]^m$ for a vector-valued random variable $\mathbf{U} = (U_1, \dots, U_m)$ with uniform marginal distributions for each of U_1, \dots, U_m on $[0, 1]$. The Pearson correlation coefficients for the marginal distributions are defined by

$$\rho_{i,j} = \frac{E[(U_i - 1/2)(U_j - 1/2)]}{\sqrt{E[(U_i - 1/2)^2]E[(U_j - 1/2)^2]}} = 12E[U_i U_j] - 3$$

for each $i, j = 1, \dots, m$. In order to model the joint probability distribution for a vector-valued random variable $\mathbf{X} = (X_1, \dots, X_m) \in (0, \infty)^m$ with prescribed cumulative marginal distributions $F_1(x_1), \dots, F_m(x_m)$, we construct uniformly distributed random variables $U_i = F_i(X_i) \in (0, 1)$ for each $i = 1, 2, \dots, m$ and use an m -dimensional copula $C(\mathbf{u}) = C(\mathbf{F}(\mathbf{x})) = C(F_1(x_1), \dots, F_m(x_m))$. We say that the *grade correlation coefficients* for \mathbf{X} are simply the Pearson correlation coefficients for \mathbf{U} , that is,

$$\rho_{i,j} = 12E[F_i(X_i)F_j(X_j)] - 3 \tag{7.1}$$

for each $i, j = 1, \dots, m$. In theory, we distinguish between the Spearman rank correlation coefficients $\widehat{\rho}_{i,j}$ obtained from the observed data

$$\{(x_{1,k}, \dots, x_{m,k})\}_{k=1}^N$$

or equivalently from the transformed data $\{(u_{1,k}, \dots, u_{m,k})\}_{k=1}^N$, where $u_{i,k} = F_i(x_{i,k})$ and the grade correlation coefficients $\rho_{i,j}$ are defined by (7.1). In practice, it may be convenient to avoid explicitly transforming the observed data by simply using $\widehat{\rho}_{i,j}$ in place of $\rho_{i,j}$.

The m -dimensional multivariate *checkerboard* copula is defined by subdividing the unit hypercube $[0, 1]^m$ into n^m congruent sub-hypercubes

$$I_{(i_1, \dots, i_m)} = [(i_1 - 1)/n, i_1/n] \times \dots \times [(i_m - 1)/n, i_m/n]$$

for each $i_1, \dots, i_m = 1, \dots, n$ with constant density on each one. If the density on $I_{\mathbf{i}}$, where $\mathbf{i} = (i_1, i_2, \dots, i_m)$, is defined by $n^{m-1}h_{\mathbf{i}} \geq 0$, then the marginal distributions will be uniform if

$$\sum_{\mathbf{i} \in S(r,i)} h_{\mathbf{i}} = 1 \quad \text{for each } r \text{ and each } i,$$

where $S(r, i) = \{\mathbf{i} \mid i_r = i\}$ for each $r = 1, 2, \dots, m$ and $i = 1, 2, \dots, n$. In such cases we say that $\mathbf{h} = [h_{\mathbf{i}}] \in \mathbb{R}^{\ell}$, where $\ell = n^m$ is multiply stochastic. We wish to construct a joint density in this form with the desired grade correlation coefficients. For sufficiently large n , there are many ways that this can be done.

The *principle of maximum entropy* asserts that the best such distribution is the most disordered or least prescriptive solution – the multiply-stochastic hypermatrix $\mathbf{h} \in \mathbb{R}^{\ell}$ which has the most equal subdivision of probabilities but still allows the required correlations. This is the hypermatrix that satisfies the grade correlation constraints and has the highest possible entropy.

The mathematical problem to find the desired multiply-stochastic hypermatrix can be neatly and rigorously solved using the theory of *Fenchel duality*. A formal statement of the problem and the theoretical solution with details about numerical calculation procedures can be found in [19]. For a more detailed discussion of computational methods, see [20].

For our proposed application, we make the following remarks. The sub-catchment rainfalls $\{X_i\}_{i=1}^4$ are modelled by the maximum likelihood Gamma distributions $X_i \sim \Gamma(\alpha_i, \beta_i)$ with the parameters given in Table 1. The cumulative distribution functions $F_i(x) = G_{\alpha_i, \beta_i}(x) = \int_0^x g_{\alpha, \beta}(\xi) d\xi$ define transformed random variables $U_i = F_i(X_i)$, which are uniformly distributed on $(0, 1)$, and transformed observations $\{u_{i,k}\}_{k=1}^N$ defined by $u_{i,k} = F_i(x_{i,k})$ for each $i = 1, \dots, 4$. We use the Spearman rank correlation coefficients in place of the theoretical grade correlation coefficients, but, nevertheless, use the simpler notation $\rho_{i,j}$. This replacement is a convenience – not a necessity, and makes no discernible difference to the results. For highly correlated marginal distributions, it is important to realize that the correlation coefficients for the checkerboard copula must satisfy $\rho_{i,j} \in [-1 + 1/n^2, 1 - 1/n^2]$ for $i, j = 1, \dots, m$, where n is the number of uniform subdivisions of the unit interval [19]. Hence, in this application, we can match the observed rank correlation coefficients if we use $n = 6$. The multiply-stochastic hypermatrix contains $\ell = 6^4 = 1296$ elements.

MATLAB programs to calculate the multiply-stochastic hypermatrix $\mathbf{h} \in \mathbb{R}^{1296}$ that defines the checkerboard copula of maximum entropy and to run the simulations are available from the first-named author. It is convenient to define $\mathbf{h} = [\mathbf{h}_{pq}]$, where $\mathbf{h}_{pq} = [h_{pqrs}] \in \mathbb{R}^{6 \times 6}$ for each $p, q = 1, \dots, 6$. The elements \mathbf{h}_{pq} for $p = 1, 2, 3$ are shown in Appendix Appendix C to four decimal places accuracy.

The checkerboard copula of maximum entropy was used to simulate corresponding volumetric winter rainfalls for each sub-catchment over a period of $N = 79$ years. The sub-catchment totals were then added together to find the catchment rainfall. We used five successive trials to test the model. The sample statistics showed considerable

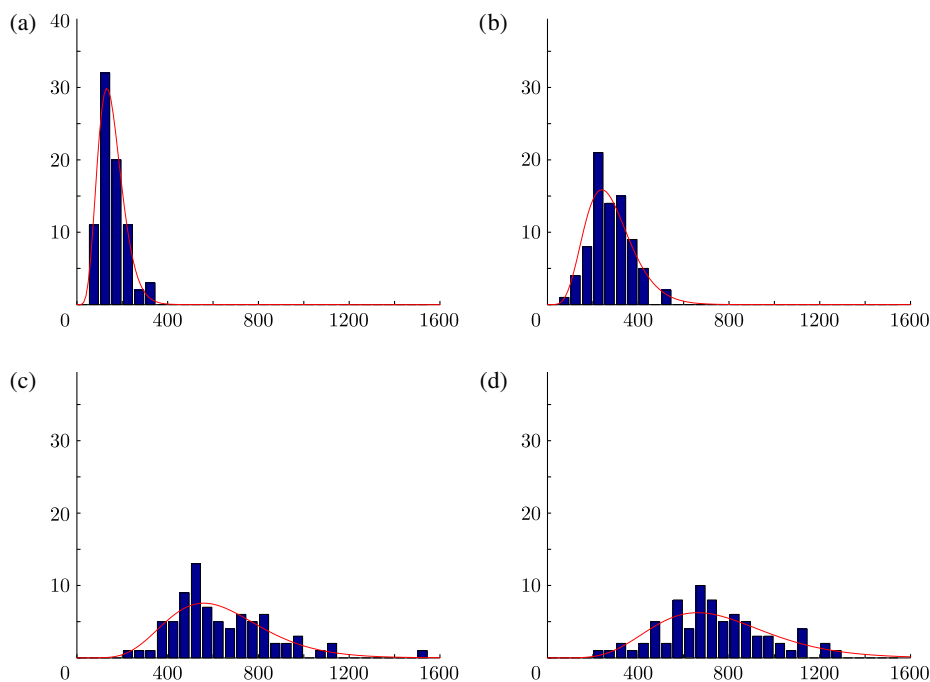


FIGURE 5. Histograms for typical simulated volumetric winter rainfall [Sim 79 #2] over a period of 79 years in units of 10^6 m^3 using a copula of maximum entropy compared to the relevant maximum likelihood Gamma distributions for Dookie (a), Rutherglen (b), Beechworth (c) and Hume (d). Note that there are three outliers in the range 1600–1850 not shown on the Hume histogram.

variation, but we begin by looking in detail at the results of an apparently typical trial – the second Monte Carlo simulation over 79 years denoted by [Sim 79 #2]. Figure 5 shows histograms for the volumetric sub-catchment winter rainfall generated in [Sim 79 #2] over a period of 79 years. The corresponding time series are shown in Figure 6. The histogram and time series for the volumetric catchment winter rainfall are shown in Figure 7. Repeated simulations confirm that the Kolmogorov–Smirnov test statistics for the distributions generated by the copula of maximum entropy and the observed distribution lie within the acceptable limits defined by the 0.05 significance level (see Appendix A for details of the Kolmogorov–Smirnov tests on the volumetric catchment winter rainfall). The overall mean and standard deviation in this trial were given by $(\bar{x}, s) = (1842, 611)$.

There is no simple expression for the theoretical probability density defined by the copula of maximum entropy, but an approximate density can be calculated by simulating very large samples. We ran three separate simulations over a time period of $N = 7900$ years. The mean and standard deviation for catchment rainfall in the first Monte Carlo simulation over 7900 years denoted by [Sim 7900 #1] were found to be

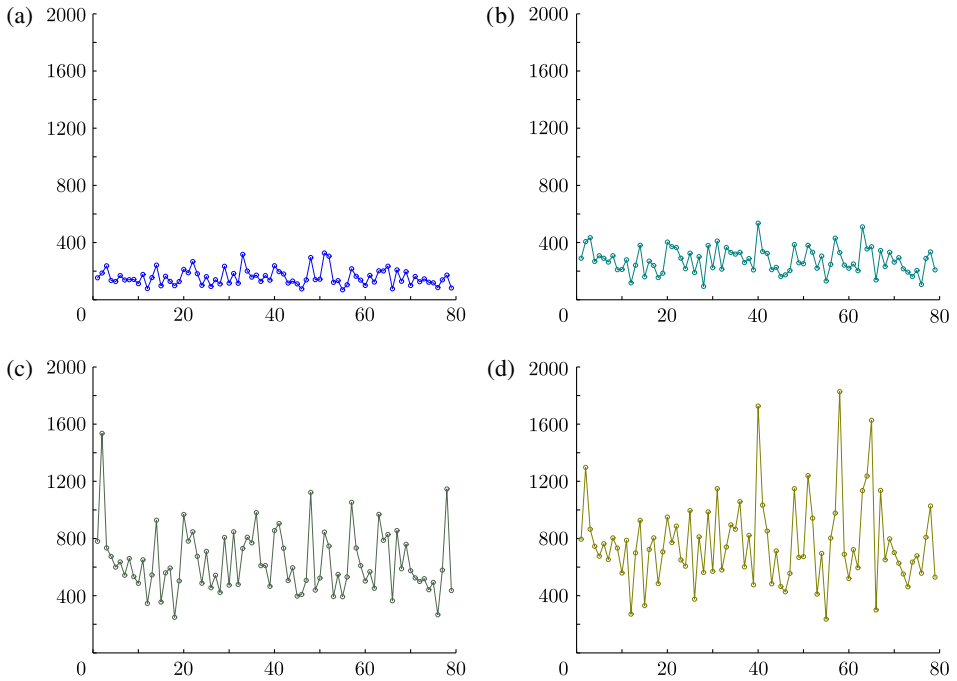


FIGURE 6. Time series for typical simulated volumetric winter rainfall [Sim 79 #2] in units of 10^6 m^3 using a copula of maximum entropy for Dookie (a), Rutherglen (b), Beechworth (c) and Hume (d).

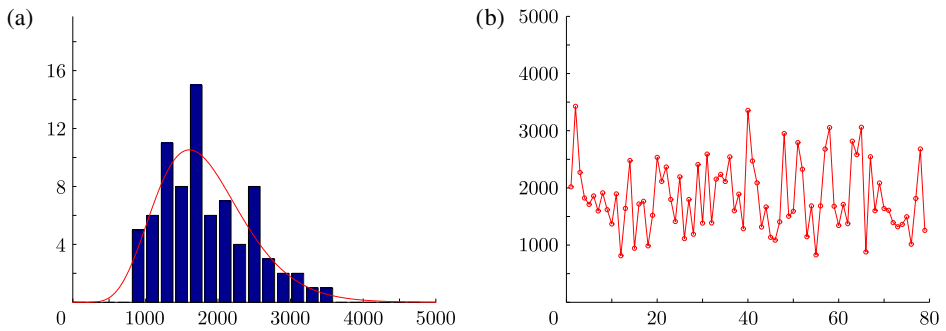


FIGURE 7. Histogram for typical simulated volumetric winter catchment rainfall [Sim 79 #2] in units of 10^6 m^3 using a copula of maximum entropy compared to the maximum likelihood Gamma distribution (a) and corresponding time series (b).

$(\bar{x}, s) = (1828, 596)$ with the Spearman rank correlation matrix given by

$$\rho_s = \begin{bmatrix} 1 & 0.86 & 0.87 & 0.83 \\ 0.86 & 1 & 0.93 & 0.90 \\ 0.87 & 0.93 & 1 & 0.91 \\ 0.83 & 0.90 & 0.91 & 1 \end{bmatrix}.$$

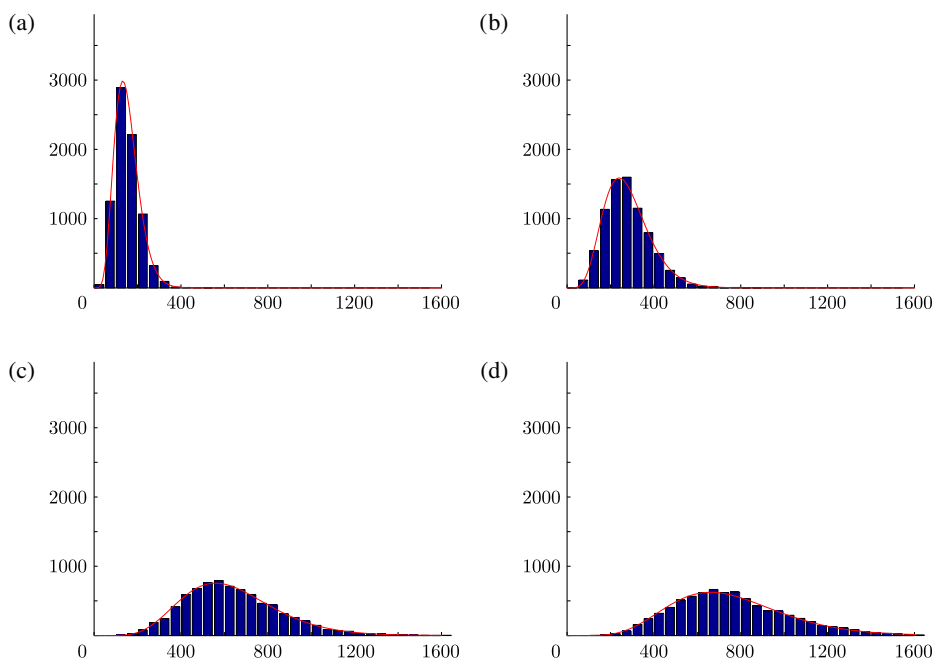


FIGURE 8. Histograms for simulated volumetric winter rainfall [Sim 7900 #1] over a period of $N = 7900$ years in units of 10^6 m^3 using the checkerboard copula of maximum entropy for Dookie (a), Rutherglen (b), Beechworth (c) and Hume (d) compared to the the relevant maximum likelihood Gamma distributions.

The histograms for the individual sub-catchment rainfalls in [Sim 7900 #1] are compared to the corresponding maximum likelihood Gamma distributions in Figure 8, while in Figure 9 we compare the simulated catchment rainfalls in [Sim 7900 #1] and [Sim 7900 #3] to the fitted maximum likelihood Gamma distribution. It is interesting to note that for the large samples of catchment rainfall generated by the copula of maximum entropy, the standard deviation is quite close to the observed standard deviation – much closer than it is for large samples generated by the maximum likelihood Gamma distribution.

8. Variability in catchment rainfall

The sub-catchment rainfalls are highly correlated, and so it is not possible to simulate realistic catchment rainfall by simply adding together independently generated sub-catchment rainfalls. Although summation of independent sub-catchment rainfalls produces a correct theoretical mean μ for the catchment rainfall, the corresponding variance σ_1^2 will grossly underestimate the true variance σ^2 . Indeed, the relevant estimates of variance using the fitted maximum likelihood marginal

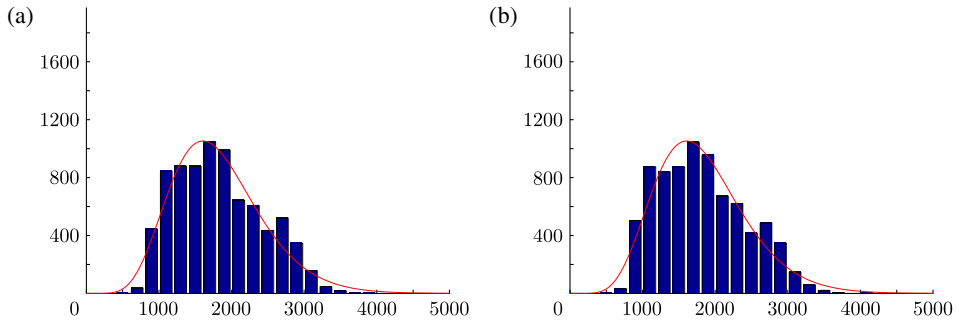


FIGURE 9. Histograms for simulated overall volumetric winter catchment rainfall using the checkerboard copula of maximum entropy showing [Sim 7900 #1] (a) and [Sim 7900 #3] (b) each over 7900 years in units of 10^6 m^3 both compared to the fitted maximum likelihood Gamma distribution for overall volumetric winter rainfall. Convergence to a smooth distribution is relatively slow with simulations using the copula of maximum entropy [2, 19, 20], but convergence is nevertheless consistent and there is no direct reason to expect that the distribution will converge to the maximum likelihood Gamma distribution.

Gamma distributions are

$$\sigma^2 \sim \sum_{i=1}^4 \alpha_i \beta_i^2 + \sum_{1 \leq i < j \leq 4} 2\rho_{i,j} \sqrt{\alpha_i \alpha_j} \beta_i \beta_j = \sigma_I^2 + \sum_{1 \leq i < j \leq 4} 2\rho_{i,j} \sqrt{\alpha_i \alpha_j} \beta_i \beta_j.$$

This leads to the estimates $\mu \sim 1824$, $\sigma \sim 626$ and $\sigma_I \sim 366$. One may argue that recent imprudent use of water resources in the Murray–Darling Basin is due in part to a failure to recognize the inherently high variance in catchment rainfall.

Finally, we sound a note of caution about the interpretation of apparent temporal trends in the observed rainfall data. In general terms, we believe that the 79 year time span of observations is too short to make definitive judgements about observed trends. Although we do not attempt a careful trend analysis in this study, other related studies [2] show that random samples generated by stationary distributions, such as those used here, frequently generate *phantom* trends of the same order of magnitude as *measured* trends in the observed data. This is simply a consequence of the inherent variability in small samples. In order to demonstrate the inherent variability in samples of this size, we used the checkerboard copula of maximum entropy to simulate five samples of catchment rainfall over a period of $N = 79$ years. The important point is that each sample is generated by the same underlying distribution and yet the sample statistics show a high degree of variation. Although there is no underlying theoretical trend in rainfall generated by this model, each sample time series shows a trend generated by chance alone. The detailed results for the selected *typical* simulation [Sim 79 #2] are shown in Figures 5–7. The histograms for total volumetric catchment rainfall for the remaining simulations [Sim 79 #1, #3, #4 and #5] are shown in Figure 10.

The sample means and standard deviations (\bar{x}, s) for the five simulated catchment winter rainfalls with $N = 79$ were (1926, 642), (1842, 611), (1705, 608), (1819, 663)

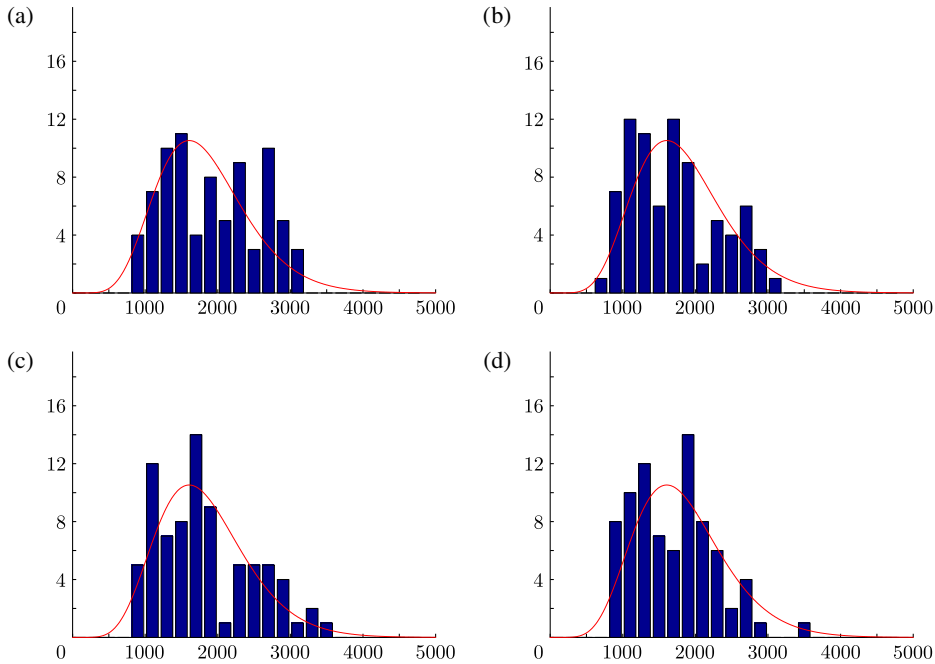


FIGURE 10. Histograms for simulated volumetric catchment winter rainfall [Sim 79 #1, #3, #4 and #5] using the checkerboard copula of maximum entropy in units of 10^6 m^3 over a period of 79 years compared to the maximum likelihood Gamma distribution.

and (1698, 559). This simple demonstration underlines the inherent variation in both the mean $\bar{x} \in [1698, 1926]$ and the standard deviation $s \in [559, 663]$ for small samples. The sample means and standard deviations (\bar{x}, s) for the three simulated catchment winter rainfalls with $N = 7900$ were (1828, 596), (1824, 614) and (1820, 600).

9. A note on extreme values

The Gamma distribution is the maximum-entropy distribution for situations where the means of both the observed values and the logarithms of the observed values are finite. Thus, it is no surprise that the Gamma distribution provides a good fit in general terms to each of the marginal distributions. However, there is one extreme value in each of the four observed sub-catchment rainfall records and in the observed total catchment rainfall record that lies – at the very least – on the outer margins of what could reasonably be expected from a simulation using the Gamma distribution. In this regard, our proposed model could possibly be improved by replacing the marginal Gamma distributions with appropriately chosen extreme-value distributions. It is important to note that the checkerboard copula of maximum entropy can be used just as effectively with any chosen marginal distribution.

10. Conclusions and future work

We have compared four models that generate volumetric monthly rainfall for a catchment in southeast Australia. The first three of these models – the maximum likelihood Gamma distribution, the Alouini model and the modified Alouini model – generate almost identical density functions, and each one generates catchment rainfall that is consistent with the observed data. The Gamma distribution is straightforward and easily implemented, but the Alouini models are more complicated and require much more careful management. Thus, for practical simulation, the maximum likelihood Gamma distribution is preferred. However, none of these models simulates corresponding sub-catchment rainfall totals for each simulated catchment rainfall.

The fourth model – the checkerboard copula of maximum entropy – is the model which we believe is the most effective. This model uses a joint probability distribution for the sub-catchment rainfalls that preserves the marginal distributions and matches the observed correlations. Each simulation produces separate sub-catchment rainfalls from which the total catchment rainfall is easily computed. Thus, the output from this model is much more informative. Although the required numerical calculations are quite demanding, we have written accurate and efficient MATLAB algorithms that compute the underlying multiply-stochastic hypermatrix and run the simulations on a standard laptop computer.

There is no theoretical bar to extending a model using the copula of maximum entropy to more than four dimensions, but it is clear that for m dimensions with n subdivisions the underlying hypermatrix with n^m elements may be extremely large. While this difficulty will be manifested in any method that requires a high-dimensional probability density, we nevertheless believe that some serious numerical work is needed to find efficient calculation and storage routines for the necessary numerical calculations.

Appendix A. The Kolmogorov–Smirnov (KS) test

The hypothesis that two samples of size n come from the same distribution should not be rejected at the 0.05 significance level if

$$D_n = \max_k |F_1^{(n)}(k) - F_2^{(n)}(k)| < 1.36 \sqrt{\frac{2}{n}},$$

where $F_i^{(n)}(k)$ denotes the cumulative distribution function for sample i at point k . All simulated samples using the checkerboard copula of maximum entropy with $n = 79$ were checked against the observed samples. In each case, the null hypothesis that the samples came from the same population could not be rejected. The difference between the cumulative frequency distribution functions for the observed and simulated volumetric catchment winter rainfall is shown for [Sim 79 #1], [Sim 79 #2], [Sim 79 #3] and [Sim 79 #4] in Figure 11. Note that in each case, the difference lies within the interval $[-0.216, 0.216]$ defined by the 0.05 significance level.

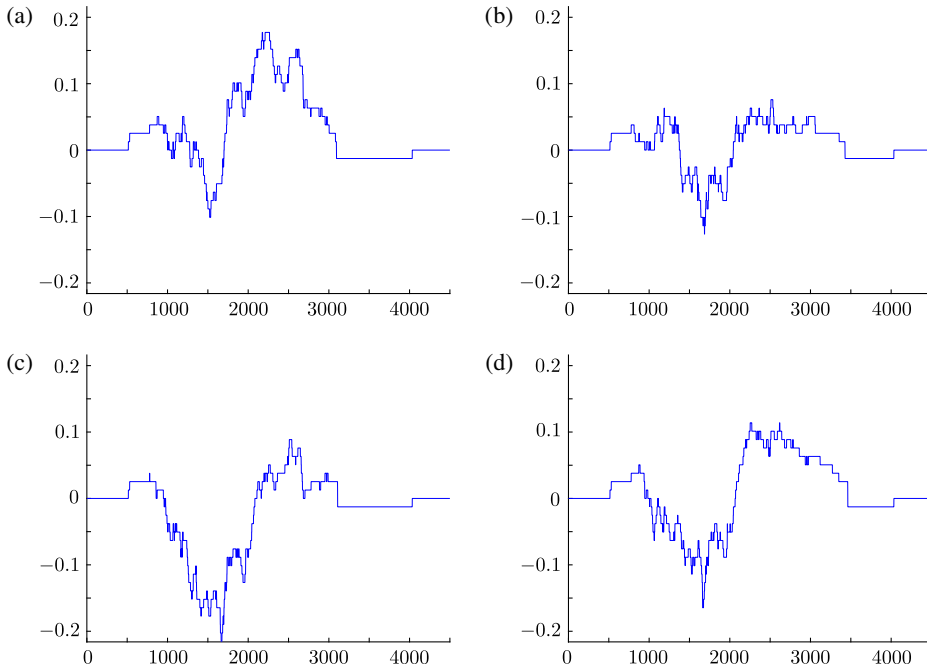


FIGURE 11. Difference between the observed and simulated cumulative frequency distribution functions for volumetric catchment winter rainfall for [Sim 79 #1], [Sim 79 #2], [Sim 79 #3] and [Sim 79 #4] showing that in each case the maximum modulus of the difference lies within the interval $[-0.216, 0.216]$ defined by the 0.05 significance level.

Appendix B. A counter-example

In this section, we show that the matrix C in the Alouini model may fail to be nonnegative, in which case the model cannot be directly applied. Let $\{Z_k\}_{k=1}^3$ be independent real-valued random variables with mean 0 and variance 1. Choose unit-length vectors $w_k = [w_{1k}, w_{2k}, w_{3k}]^T$ for each $k = 1, 2, 3$, and define new random variables $\{X_i\}_{i=1}^3$ by setting $X_i = \sum_{k=1}^3 w_{ik}Z_k$ for each $i = 1, 2, 3$. Now $E[X_iX_j] = \sum_{k=1}^3 w_{ik}w_{jk} = \langle w_i, w_j \rangle = \cos \theta_{ij}$. If we define a vector-valued random variable $X = [X_1, X_2, X_3]^T$, then $R = E[X_iX_j^T] = [\cos \theta_{ij}]$. If we assume that $|\theta_{ij}| \leq \pi/2$ for all i, j , then we can define $C = [\sqrt{\cos \theta_{ij}}]$. We will show with a simple counter-example that C is not necessarily nonnegative. If we let $\theta_{12} = 0.1$, $\theta_{13} = 0.3$ and $\theta_{23} = 0.2$, then

$$C = \begin{bmatrix} 1 & \sqrt{\cos 0.1} & \sqrt{\cos 0.3} \\ \sqrt{\cos 0.1} & 1 & \sqrt{\cos 0.2} \\ \sqrt{\cos 0.3} & \sqrt{\cos 0.2} & 1 \end{bmatrix}$$

and, according to MATLAB, the eigenvalues are approximately given by

$$\text{eig}(C) = \begin{bmatrix} -0.000033 \\ 0.023413 \\ 2.976620 \end{bmatrix}.$$

Hence, in this case, the matrix C fails to be nonnegative. The selected angles for the matrices are feasible. In fact, the sum of the two smaller angles equals the largest angle, so this represents an instance where the random variables are co-planar. The objective of the example is to show that there are feasible R matrices for which the corresponding C matrix fails to be nonnegative.

Appendix C. The hypermatrix specifying the checkerboard copula

The elements of the hypermatrix are shown to four decimal places. We recommend that all calculations should use machine accuracy. The very small elements correspond to the extreme events and truncation to (say) four decimal places in calculations would mean that rare events will be excluded from the simulations. We have $\mathbf{h} = [\mathbf{h}_{pq}]$, where $\mathbf{h}_{pq} = [h_{pqrs}] \in \mathbb{R}^{6 \times 6}$ for each $p, q = 1, \dots, 6$. The elements \mathbf{h}_{pq} for $p = 1, 2, 3$ are given by

$$\begin{aligned} \mathbf{h}_{11} &= \begin{bmatrix} 0.4736 & 0.1084 & 0.0022 & 0.0000 & 0.0000 & 0.0000 \\ 0.0354 & 0.0398 & 0.0039 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0002 & 0.0001 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \end{bmatrix}, \\ \mathbf{h}_{12} &= \begin{bmatrix} 0.0576 & 0.0367 & 0.0020 & 0.0000 & 0.0000 & 0.0000 \\ 0.0374 & 0.1171 & 0.0320 & 0.0006 & 0.0000 & 0.0000 \\ 0.0003 & 0.0046 & 0.0062 & 0.0006 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \end{bmatrix}, \\ \mathbf{h}_{13} &= \begin{bmatrix} 0.0002 & 0.0004 & 0.0001 & 0.0000 & 0.0000 & 0.0000 \\ 0.0011 & 0.0099 & 0.0076 & 0.0004 & 0.0000 & 0.0000 \\ 0.0001 & 0.0034 & 0.0126 & 0.0033 & 0.0001 & 0.0000 \\ 0.0000 & 0.0000 & 0.0002 & 0.0002 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \end{bmatrix}, \\ \mathbf{h}_{14} &= \begin{bmatrix} 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0001 & 0.0006 & 0.0004 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0001 & 0.0003 & 0.0001 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \end{bmatrix}, \end{aligned}$$

$$\begin{aligned}
 h_{21} &= \begin{bmatrix} 0.1758 & 0.0406 & 0.0008 & 0.0000 & 0.0000 & 0.0000 \\ 0.0334 & 0.0379 & 0.0038 & 0.0000 & 0.0000 & 0.0000 \\ 0.0001 & 0.0004 & 0.0002 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \end{bmatrix}, \\
 h_{22} &= \begin{bmatrix} 0.0391 & 0.0251 & 0.0014 & 0.0000 & 0.0000 & 0.0000 \\ 0.0645 & 0.2038 & 0.0562 & 0.0011 & 0.0000 & 0.0000 \\ 0.0013 & 0.0203 & 0.0275 & 0.0026 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0001 & 0.0001 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \end{bmatrix}, \\
 h_{23} &= \begin{bmatrix} 0.0003 & 0.0004 & 0.0001 & 0.0000 & 0.0000 & 0.0000 \\ 0.0036 & 0.0316 & 0.0243 & 0.0013 & 0.0000 & 0.0000 \\ 0.0006 & 0.0273 & 0.1030 & 0.0275 & 0.0005 & 0.0000 \\ 0.0000 & 0.0002 & 0.0039 & 0.0051 & 0.0005 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \end{bmatrix}, \\
 h_{24} &= \begin{bmatrix} 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0001 & 0.0002 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0008 & 0.0085 & 0.0063 & 0.0003 & 0.0000 \\ 0.0000 & 0.0001 & 0.0028 & 0.0102 & 0.0026 & 0.0001 \\ 0.0000 & 0.0000 & 0.0000 & 0.0001 & 0.0002 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \end{bmatrix}, \\
 h_{25} &= \begin{bmatrix} 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0004 & 0.0003 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0001 & 0.0002 & 0.0001 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \end{bmatrix}, \\
 h_{31} &= \begin{bmatrix} 0.0177 & 0.0041 & 0.0001 & 0.0000 & 0.0000 & 0.0000 \\ 0.0085 & 0.0098 & 0.0010 & 0.0000 & 0.0000 & 0.0000 \\ 0.0001 & 0.0003 & 0.0001 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \end{bmatrix}, \\
 h_{32} &= \begin{bmatrix} 0.0072 & 0.0047 & 0.0003 & 0.0000 & 0.0000 & 0.0000 \\ 0.0301 & 0.0962 & 0.0268 & 0.0005 & 0.0000 & 0.0000 \\ 0.0016 & 0.0243 & 0.0333 & 0.0032 & 0.0000 & 0.0000 \\ 0.0000 & 0.0001 & 0.0004 & 0.0002 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \end{bmatrix}.
 \end{aligned}$$

$$\begin{aligned}
 \mathbf{h}_{33} &= \begin{bmatrix} 0.0001 & 0.0002 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0031 & 0.0272 & 0.0211 & 0.0012 & 0.0000 & 0.0000 \\ 0.0014 & 0.0598 & 0.2278 & 0.0613 & 0.0012 & 0.0000 \\ 0.0000 & 0.0012 & 0.0219 & 0.0289 & 0.0027 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0001 & 0.0001 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \end{bmatrix}, \\
 \mathbf{h}_{34} &= \begin{bmatrix} 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0002 & 0.0004 & 0.0001 & 0.0000 & 0.0000 \\ 0.0000 & 0.0032 & 0.0343 & 0.0257 & 0.0014 & 0.0000 \\ 0.0000 & 0.0005 & 0.0286 & 0.1052 & 0.0274 & 0.0006 \\ 0.0000 & 0.0000 & 0.0002 & 0.0038 & 0.0049 & 0.0005 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \end{bmatrix}, \\
 \mathbf{h}_{35} &= \begin{bmatrix} 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0001 & 0.0002 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0008 & 0.0084 & 0.0061 & 0.0004 \\ 0.0000 & 0.0000 & 0.0001 & 0.0027 & 0.0095 & 0.0029 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0002 & 0.0003 \end{bmatrix}, \\
 \mathbf{h}_{36} &= \begin{bmatrix} 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0001 & 0.0005 & 0.0005 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0001 & 0.0004 \end{bmatrix}
 \end{aligned}$$

with $\mathbf{h}_{15} = \mathbf{h}_{16} = \mathbf{h}_{26} = [0.0000]$. For $p = 4, 5, 6$, the elements of \mathbf{h}_{pq} can be found from the symmetry condition, $h_{pqrs} = h_{tuvw}$, where $p + t = 7, q + u = 7, r + v = 7$ and $s + w = 7$.

References

- [1] M. S. Alouini, A. Abdi and M. Kaveh, "Sum of gamma variates and performance of wireless communication systems over Nakagami-fading channels", *IEEE Trans. Veh. Technol.* **50** (2001) 1471–1480; doi:10.1109/25.966578.
- [2] J. Borwein, P. Howlett and J. Piantadosi, "Modelling and simulation of seasonal rainfall using the principle of maximum entropy", *Entropy* **16** (2014) 747–769; doi:10.3390/e16020747.
- [3] L. Channel, "A government website that provides integrated access to information about land, resources and property services in Victoria", <http://www.land.vic.gov.au/>, 2014.
- [4] S. Das and D. K. Dey, "On Bayesian inference for generalized multivariate gamma distribution", *Statist. Probab. Lett.* **80** (2010) 1492–1499; doi:10.1016/j.spl.2010.05.018.
- [5] P. K. Dunn, "Occurrence and quantity of precipitation can be modelled simultaneously", *Int. J. Climatol.* **24** (2004) 1231–1239; doi:10.1002/joc.1063.
- [6] P. K. Dunn and G. K. Smyth, "Series evaluation of Tweedie exponential dispersion model densities", *Stat. Comput.* **15** (2005) 267–280; doi:10.1007/s11222-005-4070-y.
- [7] P. K. Dunn and G. K. Smyth, "Evaluation of Tweedie exponential dispersion model densities by Fourier inversion", *Stat. Comput.* **18** (2008) 73–86; doi:10.1007/s11222-007-9039-6.

- [8] W. M. Getz, “Correlative coherence analysis: variation from intrinsic and extrinsic sources in competing populations”, *Theor. Popul. Biol.* **64** (2003) 89–99; doi:10.1016/S0040-5809(03)00026-1.
- [9] M. M. Hasan and P. K. Dunn, “A simple Poisson–gamma model for modelling rainfall occurrence and amount simultaneously”, *Agric. Forest Meteorol.* **150** (2010) 1319–1330; doi:10.1016/j.agrformet.2010.06.002.
- [10] M. M. Hasan and P. K. Dunn, “Two Tweedie distributions that are near optimal for modelling monthly rainfall in Australia”, *Int. J. Climatol.* **31** (2011) 1389–1397; doi:10.1002/joc.2162.
- [11] M. M. Hasan and P. K. Dunn, “Understanding the effect of climatology on monthly rainfall amounts in Australia using Tweedie GLMs”, *Int. J. Climatol.* **32** (2012) 1006–1017; doi:10.1002/joc.2332.
- [12] B. Jørgensen, “Exponential dispersion models”, *J. Roy. Statist. Soc. Ser. B* **49** (1987) 127–162; jstor:2345415.
- [13] R. W. Katz and M. B. Parlange, “Overdispersion phenomenon in stochastic modelling of precipitation”, *J. Climate* **11** (1998) 591–601; doi:10.1175/1520-0442(1998)011<C0591:OPISMO>3E2.0.CO;2.
- [14] S. Kotz and J. W. Adams, “Distribution of sum of identically distributed exponentially correlated gamma variables”, *Ann. Math. Stat.* **35** (1964) 277–283; doi:10.1214/aoms/1177703611.
- [15] A. M. Mathai, “Storage capacity of a dam with gamma type inputs”, *Ann. Inst. Statist. Math. Part A* **34** (1982) 591–597; doi:10.1007/BF02481056.
- [16] P. G. Moschopoulos, “The distribution of the sum of independent gamma random variables”, *Ann. Inst. Statist. Math. Part A* **37** (1985) 541–544; doi:10.1007/BF02481123.
- [17] R. B. Nelsen, *An introduction to copulas*, Volume 139 of *Lect. Notes in Statist.* (Springer, New York, 1999); doi:10.1007/978-1-4757-3076-0.
- [18] J. Piantadosi, J. W. Boland and P. G. Howlett, “Generating synthetic rainfall on various time scales – daily, monthly and yearly”, *Environ. Model. Assess.* **14** (2009) 431–438; doi:10.1007/s10666-008-9157-3.
- [19] J. Piantadosi, P. G. Howlett and J. M. Borwein, “Copulas of maximum entropy”, *Opt. Lett.* **6** (2012) 99–125; doi:10.1007/s11590-010-0254-2.
- [20] J. Piantadosi, P. G. Howlett, J. M. Borwein and J. Henstridge, “Maximum entropy methods for generating simulated rainfall”, *Numer. Algebra Control Optim.* **2** (2012) 233–256 (Special Issue for Charles Pearce’s 70th birthday); doi:10.3934/naco.2012.2.233.
- [21] D. G. Rees, *Essential statistics*, 4th edn (Chapman and Hall, Boca Raton, FL, 2001); ISBN: 1584880074.
- [22] K. Rosenberg, J. W. Boland and P. G. Howlett, “Simulation of monthly rainfall totals”, *ANZIAM J.* **46** (2004) E85–E104; doi:10.0000/anziamj.v46i0.507.
- [23] N. Smirnov, “Table for estimating the goodness of fit of empirical distributions”, *Ann. Math. Stat.* **19** (1948) 279–281; doi:10.1214/aoms/1177730256.
- [24] R. Srikanthan, “Stochastic generation of daily rainfall at a number of sites”, Technical Report 05/7, Cooperative Research Centre for Catchment Hydrology, Melbourne, Australia, 2005; ISBN: 1 920813 26 8.
- [25] R. Srikanthan and T. A. McMahon, “Stochastic generation of annual, monthly and daily climate data: a review”, *Hydrol. Earth Syst. Sci.* **5** (2001) 653–670; doi:10.5194/hess-5-653-2001.
- [26] R. D. Stern and R. Coe, “Fitting models to daily rainfall data”, *J. Appl. Meteorol.* **21** (1982) 1024–1031; doi:10.1175/1520-0450(1982)021<C0420:CAPDFT>3E2.0.CO;2.
- [27] P. Todorovic and D. A. Woolhiser, “A stochastic model of n -day precipitation”, *J. Appl. Meteorol.* **14** (1975) 17–24; doi:10.1175/1520-0450(1975)014<C0017:ASMODP>3E2.0.CO;2.
- [28] D. S. Wilks, “Maximum likelihood estimation for the gamma distribution using data containing zeros”, *J. Climate* **3** (1990) 1495–1501; doi:10.1175/1520-0442(1990)003<C1495:MLEFTG>3E2.0.CO;2.
- [29] D. S. Wilks and R. L. Wilby, “The weather generation game: a review of stochastic weather models”, *Prog. Phys. Geog.* **23** (1999) 329–357; doi:10.1177/030913339902300302.
- [30] R. Zakaria, “Mathematical modelling of rainfall in the Murray–Darling Basin”, Ph.D. Thesis, University of South Australia, 2011; <http://researchoutputs.unisa.edu.au/1959.82f119866>.