STUDENT COPULA METHOD IN RAINFALL DISTRIBUTION

Sharainie binti Sahrin, PM.Dr.Fadhilah Yusof

1.1 INTRODUCTION

Copulas are tools for modelling dependence of several random variables. The term copula was first used in the work of Sklar (1959) and is derived from the latin word copulare, to connect or to join. The main purpose of copulas is to describe the interrelation of several random variables. (Thorsten Schmidt, 2006).

Copula is a function that joins the two distributions and known as dependence functions. Copula connect multivariate distribution function to its univariate marginal distribution. When we have two models having the problems relating to dependence, we can join that models becoming one model using marginal function. So, the dependency is taken care. It means that copula played an important role to join multivariate distributions to their one dimensional marginal distribution function.

1.2 PROBLEM STATEMENT

This study will focus on expressing the bivariate distribution function from two univariate distribution functions. Considering these two distribution functions are independent, the multiplication of these two distributions can be done. However, in this case, the dependency between these two distributions must be considered. One of the way allowing the dependency of these distributions by using copula method.

1.3 LITERATURE REVIEW

The Word Copula is a Latin noun that means "A link, tie, bond". In 1959, the word Copula appeared for the first time.

Copula played an important role to join multivariate distributions to their one dimensional marginal distribution function. When we are having the problems relating to dependence, we can join that models becoming one model using marginal function as we want dependency is taken care.

Copula method has been used for flood frequency analysis. This analysis was based on t-copula for Johor River, Malaysia. Student Copula was used to model the joint dependence of peak flow-volume, volume-duration and peak flow-duration. (Salarpour *et al.*, 2013).

Copula also has been used to model bivariate rainfall distribution. This model has been simulated at two sites in the Murray-Darling Basin, Australia. The selected sites are Hume and Beechworth. In this study, the asymmetric t-copula, also known as skew-t is used to analyse the monthly rainfall data.(Zakaria*et al.*, 2010)

1.4 METHODOLOGY

Joint Distribution Function

If $Y_1$and Y_2 are jointly continuous random variables with a joint density function given by $f(y_1,y_2)$, then

1. $f(y_1,y_2) \geq 0$ for all y_1, y_2
2. $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(y_1, y_2) dy_1 dy_2 = 1$

Joint distribution function is:

$H(x, y) = P[X \leq x, Y \leq y]$ (1.1)

Marginal distribution function:

$F(x) = P[X \leq x], G(y) = P[Y \leq y]$ (1.2)

Gamma Distribution

The PDF an d CDF of the gamma distribution with continuous shape parameter (α), continuous scale parameter (β) and continuous location parameter (γ) are as depicted below:

$$f(x) = \frac{(x-\gamma)^{\alpha-1}}{\beta^{\alpha}\Gamma(\alpha)} \exp(-(x-\gamma)/\beta) \quad (1.3)$$

$$F(x) = \frac{\Gamma_{(x-\gamma)/\beta}(\alpha)}{\Gamma(\alpha)} \quad (1.4)$$

Where: $\gamma \leq x \leq +\infty$

Parameter Estimation:

Likelihood function

$$L = \frac{\sum_{i=1}^{n}(x_i - \gamma)^{\alpha-1}}{\beta^{n\alpha}\sum_{i=1}^{n}\Gamma(\alpha)} \exp \sum_{i=1}^{n}\left(\frac{-x_i + \gamma}{\beta}\right) \quad (1.5)$$

The log-likelihood function iS $(\alpha - 1)\ln \sum_{i=1}^{n}(x_i - \gamma) - n\alpha \ln \beta - n \ln \Gamma(\alpha) + \sum_{i=1}^{n}\left(\frac{-x_i + \gamma}{\beta}\right)$

(1.6)

$$\frac{\partial l}{\partial \alpha} = \ln \sum_{i=1}^{n}(x_i - \gamma) - n \ln \beta - n\psi(\alpha) = 0$$ where $\psi(\alpha)$ is digamma function (1.7)

$$\frac{\partial l}{\partial \beta} = \frac{-n\alpha}{\beta} - \frac{\sum_{i=1}^{n}(-x_i + \gamma)}{\beta^2} = 0 \quad (1.8)$$

$$\frac{\partial l}{\partial \gamma} = -(\alpha - 1)\sum_{i=1}^{n}(x_i - \gamma)^{-1} + \frac{n}{\beta} = 0 \quad (1.9)$$

Use Newton-Raphson method.

$$\begin{bmatrix} \alpha_{k+1} \\ \beta_{k+1} \\ \gamma_{k+1} \end{bmatrix} = \begin{bmatrix} \alpha_k \\ \beta_k \\ \gamma_k \end{bmatrix} - J_{k_i}^{-1} \begin{bmatrix} f_1(\alpha) \\ f_2(\beta) \\ f_3(\gamma) \end{bmatrix} \quad (1.10)$$

$$J_{k_i} = \begin{pmatrix} \frac{\partial f_1(\alpha)}{\partial \alpha} & \frac{\partial f_1(\alpha)}{\partial \beta} & \frac{\partial f_1(\alpha)}{\partial \gamma} \\ \frac{\partial f_2(\beta)}{\partial \alpha} & \frac{\partial f_2(\beta)}{\partial \beta} & \frac{\partial f_2(\beta)}{\partial \gamma} \\ \frac{f_3(\gamma)}{\partial \alpha} & \frac{f_3(\gamma)}{\partial \beta} & \frac{f_3(\gamma)}{\partial \gamma} \end{pmatrix} = \begin{pmatrix} \frac{\partial^2 \ln L}{\partial \alpha^2} & \frac{\partial^2 \ln L}{\partial \alpha \partial \beta} & \frac{\partial^2 \ln L}{\partial \alpha \partial \gamma} \\ \frac{\partial^2 \ln L}{\partial \beta \partial \alpha} & \frac{\partial^2 \ln L}{\partial \beta^2} & \frac{\partial^2 \ln L}{\partial \beta \partial \gamma} \\ \frac{\partial^2 \ln L}{\partial \gamma \partial \alpha} & \frac{\partial^2 \ln L}{\partial \gamma \partial \beta} & \frac{\partial^2 \ln L}{\partial \gamma^2} \end{pmatrix} \quad (1.11)$$

Weibull Distribution

The PDF and CDF of the weibull distribution with continuous shape parameter (α), continuous scale parameter (γ) and continuous location parameter (β) are as depicted below:

$$f(x) = \frac{\alpha}{\gamma}\left(\frac{x-\beta}{\gamma}\right)^{\alpha-1} \exp\left(-\left(\frac{x-\beta}{\gamma}\right)^{\alpha}\right) \quad (1.12)$$

$$F(x) = 1 - \exp\left(-\left(\frac{x-\beta}{\gamma}\right)^{\alpha}\right) \quad (1.13)$$

Where: $\gamma \le x \le +\infty$

Parameter Estimation:

Likelihood function

$$L = \left(\frac{\alpha}{\gamma}\right)\prod_{i=1}^{n}\left[\left(\frac{x_i-\beta}{\gamma}\right)^{\alpha-1}\right]\exp\left[-\sum_{i=1}^{n}\left(\frac{x_i-\beta}{\gamma}\right)^{\alpha}\right] \quad (1.14)$$

Taking the logarithm for the likelihood function will get the log-likelihood function:

$$l(\alpha,\beta,\gamma;X)=n\ln\alpha-n\ln\gamma+\alpha\sum_{i=1}^{n}\ln(x_i-\beta)-\sum_{i=1}^{n}\ln(x_i-\beta)-\alpha\sum_{i=1}^{n}\ln\gamma+\sum_{i=1}^{n}\ln\gamma-\sum_{i=1}^{n}\left(\frac{x_i-\beta}{\gamma}\right)^{\alpha}$$

(1.15)

The partial derivatives for the log-likelihood function with respect to α,β and γ are:

$$\frac{\partial l}{\partial\alpha}=\frac{n}{\alpha}+\sum_{i=1}^{n}\ln\left(\frac{x_i-\beta}{\gamma}\right)-\sum_{i=1}^{n}\left(\frac{x_i-\beta}{\gamma}\right)^{\alpha}\ln\left(\frac{x_i-\beta}{\gamma}\right)=0 \qquad (1.16)$$

$$\frac{\partial l}{\partial\gamma}=-\frac{n}{\gamma}-\alpha\sum_{i=1}^{n}\frac{1}{\gamma}+\frac{\alpha}{\gamma}\sum_{i=1}^{n}\left(\frac{x_i-\beta}{\gamma}\right)^{\alpha}=0 \qquad (1.17)$$

$$\frac{\partial l}{\partial\beta}=-(\alpha-1)\sum_{i=1}^{n}\left[\frac{1}{x_i-\beta}\right]+\frac{\alpha}{\gamma}\sum_{i=1}^{n}\left(\frac{x_i-\beta}{\gamma}\right)^{\alpha-1}=0 \qquad (1.18)$$

$$\begin{bmatrix}\alpha_{k+1}\\ \beta_{k+1}\\ \gamma_{k+1}\end{bmatrix}=\begin{bmatrix}\alpha_{k}\\ \beta_{k}\\ \gamma_{k}\end{bmatrix}-J_{k_i}^{-1}\begin{bmatrix}f_1(\alpha)\\ f_2(\beta)\\ f_3(\gamma)\end{bmatrix} \qquad (1.19)$$

$$J_{k_i}=\begin{pmatrix}\frac{\partial f_1(\alpha)}{\partial\alpha} & \frac{\partial f_1(\alpha)}{\partial\beta} & \frac{\partial f_1(\alpha)}{\partial\gamma}\\ \frac{\partial f_2(\beta)}{\partial\alpha} & \frac{\partial f_2(\beta)}{\partial\beta} & \frac{\partial f_2(\beta)}{\partial\gamma}\\ \frac{f_3(\gamma)}{\partial\alpha} & \frac{f_3(\gamma)}{\partial\beta} & \frac{f_3(\gamma)}{\partial\gamma}\end{pmatrix}=\begin{pmatrix}\frac{\partial^2\ln L}{\partial\alpha^2} & \frac{\partial^2\ln L}{\partial\alpha\partial\beta} & \frac{\partial^2\ln L}{\partial\alpha\partial\gamma}\\ \frac{\partial^2\ln L}{\partial\beta\partial\alpha} & \frac{\partial^2\ln L}{\partial\beta^2} & \frac{\partial^2\ln L}{\partial\beta\partial\gamma}\\ \frac{\partial^2\ln L}{\partial\gamma\partial\alpha} & \frac{\partial^2\ln L}{\partial\gamma\partial\beta} & \frac{\partial^2\ln L}{\partial\gamma^2}\end{pmatrix}$$

(1.20)

Log-normal Distribution

The PDF and CDF of the log-normal distribution with shape parameter (σ), scale parameter (μ) and location parameter (γ) are as depicted below

$$f(x)=\frac{1}{(x-\gamma)\sigma\sqrt{2\pi}}\exp\left(-\frac{(\ln(x-\gamma)-\mu)^2}{2\sigma^2}\right) \qquad (1.21)$$

$$F(x)=\Phi\left(\frac{\ln(x-\gamma)-\mu}{\sigma}\right) \qquad (1.22)$$

Where: $0 \le \gamma < x$

Parameter Estimation:

Likelihood function $L(\mu,\sigma,\gamma \mid X) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \prod_{i=1}^{n}(x_i-\gamma)^{-1} \exp\left(-\frac{\sum_{i=1}^{n}(\ln(x_i-\gamma)-\mu)^2}{2\sigma^2}\right)$

(1.23)

The log-likelihood function is

$$\ln L(\mu,\sigma,\gamma \mid X) = -n\ln\sigma - n\ln\sqrt{2\pi} - \sum_{i=1}^{n}\ln(x_i-\gamma) - \frac{\sum_{i=1}^{n}(\ln(x_i-\gamma)-\mu)^2}{2\sigma^2}$$

(1.24)

Differentiating the log-likelihood function and equating to zero will obtain the maximum likelihood estimating equations.

$$\frac{\partial \ln L}{\partial \mu} = \frac{1}{\mu\sigma^2}\sum_{i=1}^{n}[\ln(x_i-\gamma)-\mu] = 0 \qquad (1.25)$$

$$\hat{\mu} = \frac{\sum_{i=1}^{n}\ln(x_i-\hat{\gamma})}{n} \qquad (1.26)$$

$$\frac{\partial \ln L}{\partial \sigma} = -\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^{n}[\ln(x_i-\gamma)-\mu]^2 = 0 \qquad (1.27)$$

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(\ln(x_i-\hat{\gamma}-\hat{\mu})^2 \qquad (1.28)$$

$$\frac{\partial \ln L}{\partial \gamma} = \sum_{i=1}^{n}\frac{1}{x_i-\gamma} + \frac{1}{\sigma^2}\sum_{i=1}^{n}\frac{\ln(x_i-\gamma)-\mu}{x_i-\gamma} = 0 \qquad (1.29)$$

Maximum likelihood estimate for γ by replacing eq.1.26and eq.1.28in the last equation, eq.1.39 to obtain an equation in γ, eq.1.30.

$$\left[\sum_{i=1}^{n}\frac{1}{x_i-\gamma}\right]\left[\sum_{i=1}^{n}\ln(x_i-\gamma) - \sum_{i=1}^{n}(\ln(x_i-\gamma))^2 + \frac{1}{n}\left(\sum_{i=1}^{n}\ln(x_i-\gamma)\right)^2\right] - n\sum_{i=1}^{n}\frac{\ln(x_i-\gamma)}{x_i-\gamma} = 0 \qquad (1.30)$$

Kolmogorov-Smirnov (K-S) test

The K-S Test statistic is defined as:

$$D = \max_{1<i<N}\left(\left|F(x_i) - \frac{i-1}{N}\right|, \left|\frac{i}{N} - F(x_i)\right|\right) \qquad (1.31)$$

where x_i is sorted from smallest to largest, F the theoretical cumulative distribution and N is the number of sample size.The best fit distribution is selected based on the test statistic, D, where the distribution that gives the smallest value of D is known as the best fit distribution.

Evaluating the dependency

The dependency can be evaluated using Kendall's tau (τ) or Kendall's tau coefficient as below:

$$\tau = \frac{\text{(number of concordant pairs) - (number of discordant pairs)}}{\frac{1}{2}n(n-1)}$$

(1.32)

Where (x_1, y_1) and (x_2, y_2) are taken as concordant or dependent when:

$(x_1 - x_2)\ (y_1 - y_2) > 0$

and discordant or independent when:

$(x_1 - x_2)\ (y_1 - y_2) < 0$

Finding the Parameters.

After getting the value of τ, the value of θ can be obtained using the formula below.Kendall's tau for t copula as written below:

$$\tau = \frac{2}{\pi}\arcsin(\theta)$$

$-1 < \theta < 1$

Then, make θ as the subject,

$$\theta = \sin\left(\frac{\pi}{2}\tau\right) \qquad (1.33)$$

In dependent t-test for paired samples, in which two samples are matched or paired, the degree of freedom used is n-1

t (student) copula

$$= \int_{-\infty}^{t_v^{-1}(u)} \int_{-\infty}^{t_v^{-1}(v)} \frac{1}{2\pi\sqrt{(1-\theta^2)}} \times \exp\left(1 + \frac{(x^2 - 2\theta xy + y^2)}{v(1-\theta^2)}\right)^{-1-\frac{v}{2}} dydx \qquad (1.34)$$

$-1 < \theta < 1$

where $t_v^{-1}(.)$ denotes the inverse function of the CDF $t_v(.)$ with v as the degree of freedom , and θ is between $t_v^{-1}(u)$ and $t_v^{-1}(v)$ and controls the strength of the tails. x and y are the dependent variables.

Simulation of data

i. Generate multivariate t-distributed random numbers which can be between $-\infty$ to $+\infty$
ii. Transform the multivariate data to uniform data using marginals of univariate t distribution.
iii. Find t-copula values using the uniform data, correlation and degrees of freedom

Comparing The Observed Data And Simulated Data

Afterward, the copula is compared by using Kolmogorov-Smirnov goodness of fit test at 5% significance level. If P-value is greater than 0.05 significance level, conclusion can be made that the two distributions are not significantly different at 5% significance level and vice versa.

1.5 RESULT AND DISCUSSION

Finding the Marginal Distribution $F_x(x)$ and $F_y(y)$

Table 4.1: Values of Parameters and Test Statistics

	Malacca (x)			Tangkak (y)		
	Parameters	Test Statistics	Rank	Parameters	Test Statistics	Rank
Lognormal	$\sigma = 0.20143$ $\mu = 7.1412$ $\gamma = 707.96$	0.14009	1	$\sigma = 0.4$ $\mu = 6.4373$ $\gamma = 1222.1$	0.1125	3
Gamma	$\alpha = 7.415$ $\beta = 97.529$ $\gamma = 1273.5$	0.14495	2	$\alpha = 3.2383$ $\beta = 154.32$ $\gamma = 1398.4$	0.10967	2
Weibull	$\alpha = 1.8411$ $\beta = 526.59$ $\gamma = 1527.5$	0.1454	3	$\alpha = 1.6896$ $\beta = 501.71$ $\gamma = 1450$	0.09936	1

i) Kendall's tau, τ

Table 4.2: The Kendall's tau Correlation between x = Station in Malacca and y=Station in Tangkak in 33 years

Year	**Kendall's tau, τ**
1975	-0.0606
1976	0.5152
1977	0.3333

1978	0.3698
1979	0.3333
1980	0.1212
1981	0.5758
1982	0.3636
1983	0.4848
1984	0.5152
1985	0.5152
1986	0.4242
1987	0.2901
1988	0.5152
1989	0.6364
1990	0.5455
1991	0.4545
1992	0.4848
1993	0.1818
1994	0.4545
1995	0.3206
1996	0.5758
1997	0.303
1998	0.5152
1999	0.3333
2000	0.6644
2001	0.3636
2002	0.4848
2003	0.2424

ii) Find values of theta, θ for the student t copula

The Kendall's coefficient for the student t copula distribution is shown below,

$$\tau = \frac{2}{\pi}\arcsin(\theta) \qquad (1.35)$$

To find the theta, θ values, make the θ as the subject

$$\theta = \sin\left(\frac{\pi}{2}\tau\right)$$

Then substitute the value of Kendall's tau in order to find value of θ for every year. The range of θ values are shown to be from -1 to1.

Table 4.3: Values of θ in 33 Years Periods for Student t Copula

Year	Θ
2004	0.303
2005	0.5649
2006	0.7273
2007	0.1985

1975	-0.09505
1976	0.723787
1977	0.499955
1978	0.54876
1979	0.499955
1980	0.189233
1981	0.786094
1982	0.540593
1983	0.690024
1984	0.723787
1985	0.723787
1986	0.618107
1987	0.44008
1988	0.723787
1989	0.841284
1990	0.755796
1991	0.654807
1992	0.690024
1993	0.281705
1994	0.654807
1995	0.482579
1996	0.786094
1997	0.458184
1998	0.723787
1999	0.499955
2000	0.86424
2001	0.540593
2002	0.690024
2003	0.371627
2004	0.458184

2005	0.775397
2006	0.90965
2007	0.306775

iii) Find Degree of Freedom, *v*

In dependent t-test for paired samples, in which two samples are matched or paired, the degree of freedom used is n-1. Since there are 12 samples for every year, thus, the degree of freedom is 11.

iv) Find Real Values of *C(u,v)*

By substituting the real values of *u*, *v*, θ and degree of freedom for every year into Eq.1.34, the values of copula will be obtained.

$$\int_{-\infty}^{t_v^{-1}(u)} \int_{-\infty}^{t_v^{-1}(v)} \frac{1}{2\pi\sqrt{(1-\theta^2)}} \times \exp\left(1+\frac{(x^2-2\theta xy+y^2}{v(1-\theta^2)}\right)^{-1-\frac{v}{2}} dy$$

The values of copula must be in range between 0 and 1.

Table 4.4: The Real Values of Student t Copula for 33 Years Period

i	**Year**	***C_i(u,v)***
1	1975	0.7004
2	1976	0.4643
3	1977	0.0182
4	1978	0.3969
5	1979	0.4772
6	1980	0.077
7	1981	0.2471
8	1982	0.616
9	1983	0.9955
10	1984	0.8352
11	1985	0.0211
12	1986	0.6862
13	1987	0.7622
14	1988	0.1247
15	1989	0.6683
16	1990	0.2646
17	1991	0.4275
18	1992	0.1051

19	1993	0.823
20	1994	0.3194
21	1995	0.706
22	1996	0.0981
23	1997	0.8425
24	1998	0.3885
25	1999	0.3352
26	2000	0.5204
27	2001	0.6864
28	2002	0.1888
29	2003	0.5138
30	2004	0.0017
31	2005	0.4667
32	2006	0.0099
33	2007	0.0054

Simulation experiment

The procedure is divided into two parts: Part 1 – finding the theoretical copula values. Part 2 – comparing the analysis of empirical copula and theoretical copula.

i) Generate 100 simulated data for the multivariate-t distributed.

100 random numbers are generated. The degree of freedom is 11 using the formula of n-1. The correlation coefficients is calculated using the Kendall's tau of original data. The Kendall's tau between distribution in Malacca and Tangkak is calculated using eq. (1.32) which gives the value 0.405. The PDF for multivariate-t distribution is given by

$$f_{td}(v,\mu,P)(x) = \frac{\Gamma\left(\frac{v+d}{2}\right)}{\Gamma\left(\frac{v}{2}\right)\sqrt{(\pi v)^d \,|P|}}\left(1+\frac{(x-\mu)' P^{-1}(x-\mu)}{v}\right)^{\frac{-v+d}{2}}$$

where v is degree of freedom, d is dimensional random vector and Γ is the gamma function.

ii) Transform generated multivariate-*t* data to uniform data using marginals of univariate *t* distribution

The PDF for univariate student's *t* distribution is given by

$$f(x)=\frac{\Gamma\left(\frac{v+1}{2}\right)}{\sqrt{v\pi}\,\Gamma(v/2)}\left(1+\frac{x^2}{v}\right)^{-(v+1)/2}$$

where Γ defines a gamma function and v is degrees of freedom.

The Kendall's tau between distribution in Malacca and Tangkak is calculated using eq. (1.32) gives the value of 0.405. Then, substitute the value of Kendall's tau into eq. (1.33) to get the value of θ such as below:

$$\theta=\sin\left(\frac{\pi}{2}\tau\right)$$

$$\theta=\sin\left(\frac{\pi}{2}0.405\right)$$

$$\theta=0.594121$$

iii) Eq (1.34) belowis used to calculate t copula . The range of these copula values must be in between 0 and 1.

$$\int_{-\infty}^{t_v^{-1}(u)}\int_{-\infty}^{t_v^{-1}(v)}\frac{1}{2\pi\sqrt{(1-\theta^2)}}\times\exp\left(1+\frac{(x^2-2\theta xy+y^2)}{v(1-\theta^2)}\right)^{-1-\frac{v}{2}}dydx$$

Comparing The Analysis Of Empirical Copula And Theoretical Copula.

In this case, use Kolmogorov-Smirnov goodness of fit test to compare the distributions of the empirical and theoretical copula. The P-value is found to be greater than 0.05 significance level for this test, therefore the conclusion that the two distributions are not significantly different at 5% significance level can be made. For addition, 1000 and 5000 random numbers are generated to test the P-value. See table below:

Table 4.6: Goodness of fit test of theoretical and empirical copulas for the uniformised observed and generated data

	Kolmogorov-Smirnov		
Number of simulated data	100	1000	5000
P-value	0.9278	0.8524	0.8214

Summary

Two sets of data (two stations) were tested using Kolmogorov-Smirnov fit test to find the best fit marginal distributions. Lognormal provided the best fitted for station in Malacca and weibull was the best fitted for station in Tangkak. The parameters u, v, θ and degree of freedom were calculated in modelling the bivariate distribution. All calculated parameter values are within the acceptable range. Then, the bivariate joint distribution of rainfall data for student-t copula is computed. The simulation process of student-t copula has been carried out by using 100, 1000 and 5000 numbers of simulated data. All these simulations process give good results which their P-values are greater than 0.05 significance level.Therefore, every simulation distribution is not significantly different with the distribution of observed data.

1.6 CONCLUSION

Based on the Kolmogorov-Smirnov goodness-of-fit test, lognormal provided the best fitted distribution for station in Malacca and weibull was the best for station in Tangkak

The distributions of observed data and simulated data were not significantly different at 5% significance level Kolmogorov-Smirnov goodness of fit test

Results showed that all calculated parameter values were within the acceptable range and could be applied to compute the bivariate joint distribution of rainfall data for the student-t copula.

ACKNOWLEDGEMENT

The authors wish to thank Malaysia Meteorological Services for allowing us to do analysis on the climate data.

REFERENCES

Defu, L., Shuqin, W. and Liping, W. (2002). Poisson-Gumbel Mixed Compound Distribution and its application. *Chinese Science Bulletin.* 47(22): 1901-1906.

Kelly, K.S. and Krzysztofowicz, R. (1997). A bivariate meta-Gaussian density for use in hydrology. *Stochastic Hydrology and Hydrailics.* 11(1):17-31.

Nikoloulopoulos, A.K., H, Joe, H. and Haijun Li (2010). Vine copulas with asymmetric tail dependence and applications to financial return data. *Computational Statistics & Data Analysis.* 56(11), 3659-3673.

Quesada-Molina, J.J., Rodr´ıguez-Lallena, J.A., and Ubeda-Flores, M. (2003). What are copulas? *Monograf´ıas del Semin. Matem. Garc´ıa de Galdeano.* 27(3), 499-506.

Salarpour, M., Yusop, Z., Yusof, F., Shahid, S. and Jajarmizadeh, M. (2013). Flood frequency analysis based on t-copula for Johor River, Malaysia. *Journal of Applied Sciences.* 13(7), 1021-1028.

Salvadori, G. and De Michele, C. (2007). On the Use of Copulas in Hydrology: Theory and Practice. *Journal of Hydrologic Engineering.* 12(4), 369-380.

Schmidt, T. (2006). *Copulas-From Theory to Applications in Finance*. (1st ed.), London: Risk Books.

Sklar, A. (1959). *Functions de repartition an dimensionset leursmarges.* (8th ed.), Paris: PublInst Stat. Univ.

Yue, S., Ourda, T.B.M.J., Bobee, B., Legendre, P., and Bruneau, P. (1999). The Gumbel mixed model for flood frequency analysis. *Journal of Hydrology.* 226 (1-2): 88-100.

Yue, S. (2000). The Gumbel mixed model applied to storm analysis. *Water Resources Management.* 14(5): 377-389.

Zakaria, R., Metcalfe, A.V., Howlett, P., Piantadosi, J. and Boland, J. (2010). Using the skew-t copula to model bivariate rainfall distribution. *ANZIAM (Australia and New Zealand Industrial and Applied Mathematics) Journal*. 51, (c231-c246).