Global Journal of Pure and Applied Mathematics.
ISSN 0973-1768 Volume 12, Number 6 (2016), pp. 4749-4758

http://www.ripublication.com/gjpam.htm

Relation Model of Storm Wet Duration and Storm Intensity for Various Rainfall Aggregation Levels using Copula Method

Rado Yendra[1], Hartono[2], Ari Pani Desvina[1], Nizam Muhaijir[1] and Dedi Irawan[3]

[1]Department of Mathematics, Faculty of Science and Technology, UIN Suska Riau, Indonesia.
[2]Department of Mathematics Education, Faculty of Education, UIN Suska Riau, Indonesia.
[3]Department of Industrial Engineering, Faculty of Science and Technology, UIN Suska Riau, Indonesia.

Abstract

Studies on various hydrological fields such as floods and agriculture are greatly dependent on rainfall data recorded over long periods of time (at least 30 years). Hourly rainfall data recorded over a period of time would result in storm studies which are useful especially in identifying variable relationship models of wet durations and intensity and are able to provide effective information for the aforementioned hydrological studies. In this study, the relation of both variables at various rain aggregation levels are tabled namely in 1 hour, 4 hours, 6 hours, 8 hours, 12 hours and 24 hours using various Archimedean Copula methods. Survey results achieved show that Archimedean Frank Copula method gives the best model for relation of both storm variables mentioned above at all levels of aggregation rainfall, while benefits of the model will continue to decrease with longer aggregation rain levels. This is due to the number of storms that continue to decline as the variable data decreases.

Keyword : Aggregation, Archimedean Copula, Frank Copula, storm rainfall

INTRODUCTION

Studies in various hydrological fields such as agriculture, floods and highways depend heavily on information obtained from rainfall. Rainfall data that is recorded at a long period of time would be able to provide effective information in the above studies. Hourly rainfall data will be able to produce effective study on storm rainfall

relationship variables. This study is mainly focused on variables in every storm, namely depth, wet storm durations, dry storm durations, dry periods between storms, storm durations and intensity. A number of researchers have defined a storm; Huff (1976) defines a storm as the interval between rain and no rain, which contains no rain time or non-continuous rain time of up to 6 hours or more. Similarly, Salvadori and Michelle (2006) states that the interval for each storm can be defined if there is no rain or non-continuous rain for at least or up to 7 hours.
The case will be clarified through figure 1. The Copula Method is used in this study to obtain the relation between two out of the three variables which usually are the focus in storm rainfall studies (depth, duration, intensity). The Copula Method has been applied by other researchers of storm rainfall studies, namely [Hutchinson and Lai(1990); Chakak and Koehle (1995); Joe (1997); Nelsen (1999); De Michele and Salvadori (2003); Fevre and partner (2004a); Salvadori and Michele (2006); Salvadori and partner (2007)].

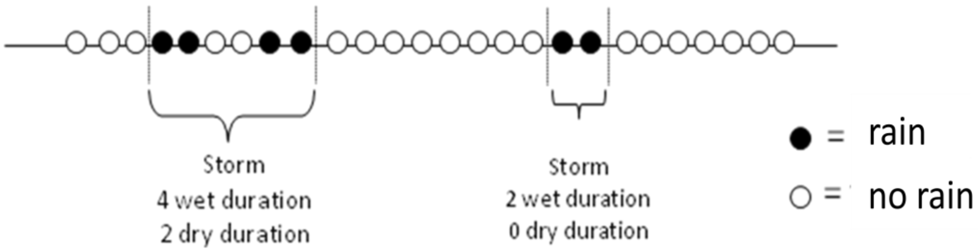

Figure 1: Theory of storm

The study is structured as such : Part 2 is a summary discussion the theories from the Copula and Archimedean Copula Methods, Part 3 is a discussion on the rainfall data at every aggregation levels and early stage statistics, including the suitable distribution models for wet storm duration and intensity variables, while part 4 discusses the Kendal correlation values and parameter estimation on Archimedean Copula for the various aggregation levels. Next, Part 5 is a discussion on methods in identifying the accurate Archimedean Copula model, including the application of the Akaike Information Criterion value (AIC), and finally in Section 6 is a discussion on the conclusion, and methods to generate the Archimedean Copula variables, **Estimator** and copula parameter as in the attachment.

THE THEORY OF ARCHIMEDEAN COPULA

Copula is the function which connects from multivariate distribution function to unvaried marginal distribution function (Nelsen ,2006). Say that X and Y are variable couple in haste with $Fx(x)$ function's distribution and $Fy(y)$, and say that function distribution with $H_{X,Y}(x,y)$, given two uniform random variables U and V, said $U = F_X(x)$ and $V = F_Y(y)$, in every pair point (x, y) in four sides unit [0,1]x[0,1] and every pair of this value connects at $H_{X,Y}(x,y)$ in [0,1]. This relationship is cited as a Copula Function. The Copula function is further illustrated in Figure 2.

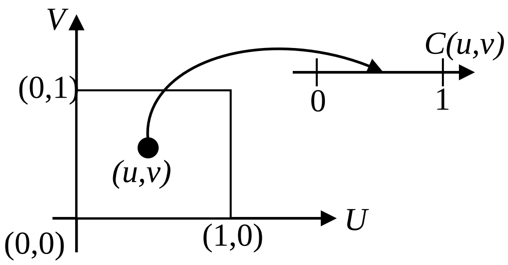

Figure 2. Theory of Copula

Application of Copula in probability and statistics are based on Teorema Sklar (Sklar ,1959), which states that there is copula *C (u,v)* such for each x, y in $\overline{R} \in (-\infty, \infty)$, $H_{X,Y}(x,y) = C(F_X(x), F_Y(y)) = C(u,v)$

Archimedean copula is a part of the Copula section, that definable as follows $C(u,v) = \varphi_\theta^{[-1]}(\varphi_\theta(u) + \varphi_\theta(u))$ (Genest and MacKay, 1986a). $\varphi_\theta(\bullet)$ is cited as Copula generator and θ is the suitable parameter of Archimedean Copula, meanwhile $\varphi_\theta^{[-1]}(\bullet)$ is cited as inverse pseudo that is defined as

$$\varphi_\theta^{[-1]}(t) = \begin{cases} \varphi_\theta^{-1}(t), & 0 \le t \le \varphi(0) \\ 0, & \varphi(0) \le t \le \infty \end{cases}$$

The basic nature of the Archimedean Copula makes it often used by many researchers is relation between Kendall (τ) and Copula generator as follows:

Table 1. Basic Information for Archimedean Copula Theory

Archimedean Copula	**Statistics**
Frank	$C_\theta(u,v) = -\frac{1}{\theta}\ln\left(1 + \frac{(e^{-\theta u}-1)(e^{-\theta v}-1)}{e^{-\theta}-1}\right), \quad \theta \neq 0$
	$\varphi_\theta(t) = -\ln\left(\frac{e^{-\theta t}-1}{e^{-\theta}-1}\right)$
	$\tau = 1 - \frac{4}{\theta}[D_1(-\theta)-1]$, $D_1(\theta) = \frac{1}{x}\int_0^\theta \frac{t}{\exp(t)-1}dt \quad \theta > 0$
Gumbel-Hougard	$C_\theta(u,v) = \left\{-\left[(-\ln u)^\theta + (-\ln v)^\theta\right]^{\frac{1}{\theta}}\right\}, \quad \theta \in [1, \infty)$
	$\varphi_\theta(t) = (-\ln t)^\theta$
	$\tau = 1 - \frac{1}{\theta}$

Clayton	$C_\theta(u,v)=\left[u^{-\theta}+v^{-\theta}-1\right]^{\frac{1}{\theta}},\quad \theta\geq 0$ $\varphi_\theta(t)=t^{-\theta}-1$ $\tau=\dfrac{\theta}{\theta+2}$
Ali-Mikhail-Haq	$C_\theta(u,v)=\dfrac{uv}{1-\theta\,(1-u)\,(1-v)},\quad \theta\in[-1,1)$ $\varphi_\theta(t)=\ln\dfrac{1-\theta(1-t)}{t}$ $\tau=\left(\dfrac{3\theta-2}{\theta}\right)-\dfrac{2}{3}\left(1-\dfrac{1}{\theta}\right)^2\ln(1-\theta)$

$$\tau=1+4\int_0^1\frac{\varphi(u)}{\varphi'(u)}du$$

In this study, four Archimedean Copula functions will be discussed to identify therelational model between wet duration wet and stomr rainfall intensity variables. The Archimedean functions are namely Frank (Frank, 1979), Clayton (Clayton, 1978), Ali-Mikhail-Haq (AMH) and Gumbel- Hougard (Gumbel, 1960 and Hougard ,1968). Important information from the 4 Archimedean Copula functions are stated in Table 1.

Method to generate pair value (u, v) in two variables U and V for Archimedean Copula was generated by Genest and Rivest, (1993) namely:

- Generate two random variables

  s = u(0,1) and q = u(0,1)

- Construct

  $$t-\frac{\phi(t)}{\phi'(t)}-q=0$$

  Obtain

  $$u=\phi^{-1}\left(s\phi(t)\right)$$
  $$v=\phi^{-1}\left[(1-s)\phi(t)\right]$$

The methods above are very useful in generating software programs as an example, in this study, big size data is used where 30 years of rainfall data in every hour for each rainfall is recorded. In order to study such sizable data using these methods in a study such as this one, which takes into account historical hourly rainfall data over a period of 30 years, computer programming is required.

VARIOUS BASIC STATISTICS OF STORM WET DURATION AND STORM INTENSITY RAINFALL

Hourly rainfall data is recorded for 33 years at Kampar Station is aggregately processed in order to obtain rainfall records for a duration of 4 hours, 6 hours, 8 hours, 12 hours and 24 hours. Such processing must be done using computer software; S PLUS software program is generated to resolve the above matter.In the initial part of this study, statistical values such as averages, standard deviations, kurtosis, skewness and number of storms for each aggregation will be issued. These results are listed in Table 2. Various wet duration and rainfall intensity rainfall statistics at the various levels of aggregation that have been listed show that the number of storms are decreasing while rainfall aggregation continues to increase.

This study can also be used as a solid foundation for several hydrology fields in order to applying the precise aggregation in any study. Table 2 has presented that average intensity value of storm rainfalls has declined identically with the increase of rainfall aggregation levels, while average duration increases identically with the increase of rainfall aggregation levels. Research that requires average value of two storm rainfall variables that is not obviously different is recommended to utilize hourly rainfall data, whilst research that requires average variable values that are mutually opposite should use rainfall data recorded every 24 hours.

From the standard deviation, it can be discussed that storm rainfall intensity variables decline identically with the increase in rainfall aggregation levels, whereas in wet storm duration variables, the standard deviation continues to increases parallel to the increase in rainfall aggregation levels. As for hydrological studies which require rain depth that is homogeneous, it is recommended to use rainfall observation data that is recorded hourly. Table 3 presents various distributions which can be applied to model second storm rainfall variables, namely exponential density, wibull, gamma and lognormal distribution functions. Accuracy tests of the function density distributions is presented using AIC and BIC (Bayesian Information Criterion) methods, where both methods are produced by using the formula:

$$AIC = l - 2p$$

$$BIC = l - \frac{p}{2}\log n$$

l is likelihood log value while p is the parameter number for a distribution, both of these formulas state that the biggest value can be used as the best model to explain such density function distribution of available observation data. These AIC and BIC value will be used to acquire the best distribution in explaining the distribution data of wet duration and storm intensity variables that are scheduled in Table 3. While the other formula for AIC value is as stated in (Zhang and Shingh, 2006)

Table 2: Various basic statistics of storm wet duration and storm intensity rainfall

	1 hour	**4 hour**	**6 hour**	**8 jam**	**12 hour**	**24 hour**
No. of Storm	5983	5562	5728	5265	4373	2308
Mean						
Intensity	3.946038	1.721616	1.283862	1.08041	0.813802	0.468568
Duration wet	4.137724	8.494786	10.60056	13.70256	21.3684	62.93241
Standard Deviation						
Intensity	5.324987	2.236514	1.668237	1.39169	1.001073	0.4724265
Duration wet	9.344824	10.58163	10.87393	12.41912	18.40683	59.98292
Skewness						
Intensity	3.686453	2.935033	2.683071	2.741318	2.78784	2.570746
Duration wet	11.95496	9.64049	9.074871	6.985847	5.552318	3.659315
Kurtosis						
Intensity	23.31966	13.68445	9.894022	11.24486	12.16276	11.57783
Duration wet	211.065	150.239	138.2417	87.32901	65.5991	23.3069

This formula states that, the lowest value is the best distribution to explain distribution of observed data. This formula is then used to get the best Archimedean copula model in order to explain the relationship of wet duration and storm rainfall intensity tabled in Part 5. In table 3, at various rainfall aggregation levels, almost all rainfall aggregation levels at the intensity levels state that lognormal distribution is the best in explaining the distribution of observed rainfall data except rainfall that is recorded in 24 hours showed that exponential distribution is the best to explain distribution of observation data. Next, storm rainfall duration variables is strongly shows that lognormal distribution is the most precise way in explaining distribution of observation data for all levels of aggregation rainfall in this study.

Table 3: Various function density distribution of intensity and duration wet storm in various rainfall aggregation level

	1 hour	4 hours	6 hours	8 hours	12 hours	24 hours
INTENSITY						
exponent						
λ	0.254	0.581	0.778	0.925	1.228	2.134
AIC	-14197.94	-8585.632	-7163.276	-5676.197	-3475.996	-562.3651
BIC	-14200.28	-8587.944	-7167.929	-5680.766	-3480.379	-566.1092
weibull						
λ	0.298	0.694	0.948	1.109	1.410	2.166
γ	1.146	1.062	1.034	1.025	1.039501	1.145475
AIC	-14824.7	-8880.907	-7398.348	-5840.921	-3589.568	-584.7713
BIC	-14829.39	-8885.531	-7403.001	-5845.49	-3593.951	-588.5154
gamma						
α(ldd)	0.549	0.592	0.592	0.602	0.660	0.983
β(muu)	0.139	0.344	0.461	0.557	0.812	2.099
AIC	-14556.56	-8688.61	-7192.858	-5710.893	-3521.619	-564.8462
BIC	-14561.26	-8693.234	-7197.511	-5715.462	-3526.002	-568.5903
log-normal						
μ	0.706	-0.179	-0.504	-0.666	-0.899	-1.277
σ^2	1.361889	1.575	1.661	1.669	1.600	1.258
AIC	-13712.54	-8231.58	-6791.842	-5389.845	-3367.015	-641.9252
BIC	-13717.24	-8235.58	-6796.495	-5394.414	-3371.399	-645.6693
DURATION						
exponent						
λ	0.241	0.117	0.094	0.072	0.046	0.015
AIC	-14481.73	-17463.74	-19255.31	-19051.05	-17767.39	-11875.64
BIC	-14484.08	-17466.05	-19259.96	-19055.62	-17771.77	-11879.38
weibull						
λ	0.296	0.122	0.096	0.073	0.047	0.016
γ	1.828	2.770	3.378	3.472	3.246	2.232
AIC	-38923.16	-91164.49	-231430.2	-142349.5	-59502.42	-14293.73
BIC	-38927.85	-91169.11	-231434.9	-142354.1	-59506.81	-14297.47
gamma						
α	0.196	0.644	0.950	1.217	1.347	1.100
β	0.0473	0.075	0.089	0.088	0.063	0.017
AIC	-18334.27	-18420.77	-19371.29	-18663.73	-17319.24	-11810.81
BIC	-18338.96	-18425.39	-19375.95	-18668.3	-17323.62	-11814.56
log-normal						
μ	0.901	1.887	2.171	2.439	2.869	3.862
σ^2	0.619	0.288	0.201	0.191	0.219	0.416
AIC	-12884.06	-15720.28	-17095.45	-17067	-16322.07	-11378.81
BIC	-12888.76	-15724.91	-17100.11	-17071.57	-16326.45	-11382.55

ESTIMATION ON KENDALL CORRELATION AND ARCHIMEDEAN COPULA PARAMETER

In early stages of Copula research, the value for Kendall (τ) correlation must be indentified first, as this value is important in determining the copula parameter(θ). Kendall correlation can be produced using the following formula :

$$\tau_n = \binom{n}{2}^{-1} \sum_{i<j} sign\left[\left(X_{1i} - X_{1j}\right)\left(X_{2i} - X_{2j}\right)\right]$$

This proximity value on correlation value (τ) as tabled in Table 1 will facilitate in estimating the copula parameter copula (θ), using the method that is recommended by Genest and Rivest (1993) in order to obtain θ as follows:

1. Get the value of Kendall correlation (τ_n)
   1. Approximate this value with value (τ) as scheduled in Table 1. $\tau \approx \tau_n$, until the parameter value θ as for various Archimedean copula in this study can be obtained. The value of Kendall correlation and parameter copula estimation (θ) at various aggregation rainfall levels used in this study is presented in table 4.

Table 4: Value of Kendal Correlation and Archimedean Copula parameter estimation at various level of rainfall aggregation

	Aggregation Scale	**τ**	**Gumbel (θ)**	**Clayton (θ)**	**Frank (θ)**	**AMH (θ)**
Rainfall Intensity-Duration	1 hour	0.1177044	1.133407	0.266814	4.4752	0.6476384
	4 hour	0.09516131	1.105169	0.2103387	4.35822	0.6412965
	6 hour	0.09694722	1.107355	0.21471	4.367282	0.641795
	8 hour	0.07358818	1.079434	0.1588671	4.25142	0.6353285
	12 hour	0.08148858	1.088718	0.1774362	4.28997	0.6375028
	24 hour	0.183808	1.225202	0.4504038	4.854076	0.6668772

GOODNESS OF FIT TEST

Goodness of fit test for distributions using the AIC and BIC methods. The AIC method is given different formulas for section 5, the AIC formula (*) is used to determine the best Archimedean copula to explain the best distribution function model for the variable relation of wet duration and storm rainfall intensity variables. AIC values for every level of rainfall aggregation for the Archimedean copula is presented in Table 5.

Table 5: AIC value of various Archimedean Copula at various rainfall aggregation levels.

	Aggregation Scale	**AIC Gumbel**	**AIC Clayton**	**AIC Frank**	**AIC AMH**
Rainfall Intensity-Duration	1 hour	-321.8485	-346.0164	-5378.547	-428.6809
	4 hours	-207.9047	-223.1324	-5210.831	-489.4852
	6 hours	-202.2675	-243.9673	-5255.447	-467.4159
	8 hours	-158.7329	-131.3929	-4746.534	-450.8656
	12 hours	-87.37104	-176.8019	-3943.175	-280.914
	24 hours	-311.7053	-228.2283	-1999.150	-183.0173

In Table 5, all levels of rainfall aggregation have the least AIC value as indicated by the Archimedean Copula Frank for all levels of rainfall aggregation and it is strongly stated that Archimedean copula frank is the best model for the relation of storm wet duration and storm intensity variables. This table also shows that the model is less significant as the time of rainfall recorded increases. It also records that rainfall at aggregation levels of 1 hour, 4 hours and 6 hours have naturally identical, as the AIC value produced is clearly different. This case needs to be presented as the rainfall research through stochastic field process especially in using Neyman-Scoot method, where parameter estimation for this method is only used in testing for various rainfall aggregation levels.

CONCLUSION

It can be a solid foundation for this study, that storm rainfall variables especially wet duration and intensity both have similar nature at rainfall aggregation levels of 1 hour, 4 hours and 6 hours. This matter can be used for Neyman-Scoot method in using rainfall level as mentioned above in estimating model parameter. This study strongly suggests that this study be continued as it is known that wet duration and intensity at every rainfall cell is 2 out of 5 variables parameter that will be estimated in this study. Study to obtain variable relational of storm rainfall model at various aggregation levels provide important information that can be used for hydrological field studies. This study which started at basic statistical level to an extended level has generated meaningful findings, with the findings in Tables 1 to 5 showing beneficial data for the study. In this study, plotted data is purposely limited in order to increase focus on storm rainfall events at the various rainfall aggregration levels, and to encourage further study on this topic.

REFERENCE

[1] Cowpertwait, P.S.P., 1991a. Further developments of the Neyman-Scott clustered point process for modeling rainfall. Water Resour. Res., 27(7): 1431-1438
[2] Cordova, J.R., Rodriguez-Iturbe, I., 1985. On Probabilistic structure of storm surface runoff. Water Resources Research 21 (5), 755-763
[3] Favre, A. C., El Adlouni, S., Perreault, L., Theimonge, N., Bobee, B., 2004. Multivariate hydrological frequency analysis using copulas. Water Resources Research 40, W01101
[4] Genest, C., Mackay, l., 1986. The joy of copulas: Bivariate distribution with uniform marginal. The American Statistician 40 (4), 280-283
[5] Genest, C., Rivest, L., 1993. Statistical Inference procedures for bivariate archemedian copulas. Journal of the American Statistical Association 88 (424), 1034-1043
[6] Nelsen, R.B., 1999. An Introduction to Copulas. Springer.
[7] Rodriguez-Iturbe, I., Cox, D.R. and Isham, V., 1987a. Some Models for Rainfalls based on stochastic point processes. Proc. R. Soc. London, Ser. A, 410: 269-288
[8] Salvadori, G., De Michele, C., 2004. Frequency analysis via copulas : Theoritical aspects and application to hydrological events. Water Resources Research 40 (12), W12511
[9] Yue, S., 2000a. Joint probability distribution of annual maximum storm peaks and amounts as represented by daily rainfalls. Hydroscience Journal 45 (2), 315-326
[10] Yue, S., 2000c. The Gumbel Mixed model applied to storm frequency analysis. Water Resources Management 14 (5), 377-389
[11] Zhang, L., Singh, V.P. 2006. Bivariate rainfall frequency distributions using Archimedean Copulas