

# Modeling of Rainfall Characteristics for Monitoring of the Extreme Rainfall Event in Makassar City

<sup>1</sup>Wahidah Sanusi, <sup>1</sup>Usman Mulbar, <sup>2</sup>Hendra Jaya, <sup>2</sup>Purnamawati and <sup>1</sup>Syafruddin Side

<sup>1</sup>Department of Mathematics, Faculty of Mathematics and Natural Sciences,

<sup>2</sup>Department of Electrical, Faculty of Engineering,

Universitas Negeri Makasar, 90224, Parangtambung Makassar, Sulawesi Selatan, Indonesia

## Article history

Received: 08-01-2017

Revised: 01-04-2017

Accepted: 08-04-2017

## Corresponding Author:

Wahidah Sanusi  
Department of Mathematics,  
Faculty of Mathematics and  
Natural Sciences, Universitas  
Negeri Makasar, 90224,  
Parangtambung Makassar,  
Sulawesi Selatan, Indonesia  
Email: wahidah.sanusi@unm.ac.id

**Abstract:** Flooding is a common problem that occurs in some regions of Indonesia, including Makassar city. In the planning of flood control, rainfall variables are very necessary as the frequency, intensity and duration of rainfall. The relationship of these variables can be expressed in a curve Intensity-Duration-Frequency (IDF). The objectives of this study are to identify the best fitting distribution of rainfall data of Makassar city and also to model the relationship between rainfall intensity, rainfall duration and rainfall frequency that is described through IDF curves. The annual maximum daily rainfall data from Ujung Pandang rainfall station of Makassar is used in this study for the period 1986-2015. Data collection was performed at the Department of Water Resources Management in South Sulawesi province. Five distributions which are considered are Gumbel, Generalized Extreme Value (GEV), Generalized Pareto (GPA), Generalized Logistic (GLO) and Pearson type III (P3) distributions. The study result found that the probability distribution of rainfall data in Makassar city has a generalized extreme value distribution. Meantime, based on IDF curves shown that the longer the rainfall duration, the rainfall intensity decreases for various return periods. The results of this study are expected to be valuable information for designers of water management.

**Keywords:** Probability Distribution, Rainfall Duration, Rainfall Frequency, Rainfall Intensity, IDF Curve

## Introduction

A flood event is a natural disaster that can cause losses for human life's and its environment. The other impacts of flood event are the emergence of tropical diseases as dengue fever disease (Syafuddin and Noorani, 2013). An effort to anticipate the impact of flood events is through wetness estimating techniques of an area. Du *et al.* (2013) investigated the spatiotemporal variation of dry/wet conditions with the Standardized Precipitation Index (SPI) in Hunan, China. Sanusi and Ibrahim (2012) predicted the wet class transition using a log linear model. Meanwhile, Cai (2010) estimated the wet and low water of precipitation with the weighted Markov chain methods.

Probability distribution model for hydrological data, such as rainfall data have a very important role in providing information about the patterns of behavior and characteristics of rainfall. Various models of probability distribution have been used to model the rainfall data in

an area. Some climate researchers have used the gamma distribution as a probability distribution for rainfall data (Aksoy, 2000; May, 2004). They choose this distribution, because of its ability to describe the characteristics of the rainfall. In addition to the gamma distribution, the other probability distributions are often used, such as the lognormal distribution (LN), Gumbel distribution, Pearson type III distribution (P3), Weibull distribution, Generalized Extreme Value distribution (GEV), Generalized Pareto distribution (GPA) and Generalized Logistic distribution (GLO) (Hosking and Wallis, 1997).

Kysely and Picek (2007) found that the fitted GEV distribution is suitable of the extreme rainfall amount in the northeast of the Czech Republic compared with other regions. The generalized Pareto distribution has been to study the extreme values (Singh and Guo, 1997). Similarly, Shabri and Ariff (2009) shown that the most appropriate distribution GLO used for annual maximum rainfall data in Selangor. LN distribution has also been used in agriculture and hydrology (Yang, 2000).

Meanwhile, the LP distribution also known as Gamma distribution with three parameters has also been used to describe annual one day maximum rainfall in Ludhiana, Punjab (Kumar and Bhardwaj, 2015).

The relationship of the rainfall Intensity-Duration-Frequency (IDF) is one of tools for planning and designing of flood event (Prodanovic and Simonovic, 2007). Indeed the IDF curves allow for the estimation of the return period of an observed rainfall event or conversely of the rainfall amount corresponding to a given return period for different aggregation times (Koutsoyiannis *et al.*, 1998). Norlida *et al.* (2011) estimated the IDF curve using Generalized Pareto Distribution in Klang region, Malaysia. Meanwhile, Soro *et al.* (2010) found that Gumbel and Lognormal distributions are suitable to estimate the IDF curve in Tropical area of West Africa. In this study, the Gumbel, GEV, GPA, GLO and P3 distributions are used for modeling the rainfall Intensity-Duration-Frequency relationship in Makassar city.

## Materials and Methods

### Data and Study Area

In this study, daily rainfall amount data (in mm) from Ujung Pandang rainfall station of Makassar for the period 1986-2015 were considered for analysis. The data are obtained from the Department of Water Resources Management in South Sulawesi province, Indonesia. The selected station was based not only on the completeness of data, but also on the longest period of data variability.

### Probability Distribution Models

In study, five types of distributions are selected in fitting for the daily rainfall amount, namely Gumbel, GEV, GPA, GLO and P3 distributions. Li *et al.* (2015) stated that advantages of those probability distributions are simple, superior and popular in frequency analysis of extreme events. Those distributions have a cumulative distribution function, respectively, as follows:

$$Gumbel : F(x) = \exp\left[-\exp\left(-\frac{x-\xi}{\alpha}\right)\right] \quad (1)$$

$$GEV : F(x) = \exp\left[-\exp\left(\kappa^{-1} \log\left(1 - \frac{\kappa}{\alpha}(x-\xi)\right)\right)\right] \quad (2)$$

$$GPA : F(x) = 1 - \exp\left(\kappa^{-1} \log\left(1 - \frac{\kappa}{\alpha}(x-\xi)\right)\right) \quad (3)$$

$$GLO : F(x) = \frac{1}{1 - \exp\left(-\kappa^{-1} \log\left(1 - \frac{\kappa}{\alpha}(x-\xi)\right)\right)} \quad (4)$$

$$P3 : F(x) = \frac{1}{\alpha\Gamma(\kappa)} \int_0^x \left(\frac{y-\xi}{\alpha}\right)^{\kappa-1} \exp\left(-\frac{y-\xi}{\alpha}\right) dy \quad (5)$$

Where:

- $x$  = Daily rainfall amount (mm)
- $\xi$  = Location parameter
- $\alpha$  = Scale parameter ( $\alpha > 0$ )
- $\kappa$  = Shape parameter
- $\Gamma(\cdot)$  = The Gamma function

### Parameter Estimation of Probability Distribution

Hosking and Wallis (1997) introduced the L-moments method for estimating the parameters of certain statistical distributions. The advantages of using method of the L-moments are that the parameter estimates are more reliable and more robust (Deka *et al.*, 2009; Eslamian and Feizi, 2007). The L-moments are the summary statistics for probability distributions and data samples and are analogous to ordinary moments (Hosking and Wallis, 1997). They provide measures of location, dispersion, skewness, kurtosis and other aspects of the shape of probability distributions or data samples.

L-moments are linear combinations of Probability Weighted Moments (PWM). Let  $x_{1:n} \leq \dots \leq x_{n:n}$  be the ordered sample and  $n$  is the sample size. Hosking and Wallis (1997) gave an estimator of PWB,  $\beta_r$ , as follows:

$$b_r = \hat{\beta}_r = \frac{1}{n} \sum_{j=r+1}^n \frac{(j-1)(j-2)\dots(j-r)}{(n-1)(n-2)\dots(n-r)} x_{j:n}, r = 1, 2, \dots \quad (6)$$

The first four L-moments are given by:

$$\begin{aligned} \lambda_1 &= b_0 \\ \lambda_2 &= 2b_1 - b_0 \\ \lambda_3 &= 6b_2 - 6b_1 + b_0 \\ \lambda_4 &= 20b_3 - 30b_2 + 12b_1 - b_0 \end{aligned}$$

where,  $\lambda_1$  is the measure of location (L-mean) and  $\lambda_2$  is the L-scale.

Hosking and Wallis (1997) defined the L-moment ratios in hydrological extreme analysis as follows:

$$\begin{aligned} \tau &= \frac{\lambda_2}{\lambda_1}, 0 \leq \tau < 1, \\ \tau_3 &= \frac{\lambda_3}{\lambda_2}, \\ \tau_4 &= \frac{\lambda_4}{\lambda_2}, \end{aligned}$$

where,  $\tau$  is the measure of coefficient of variation (L- $C_v$ ),  $\tau_3$  is the measure of skewness (L- $C_s$ ) and  $\tau_4$  is the measure of kurtosis (L- $C_k$ ). Details on the estimation of parameters

for the mentioned distribution using L-moments method can be found in Hosking and Wallis (1997).

### Selection of the Fitted Distribution Models

In this study, the best probability distribution for rainfall data is chosen based on three criteria that is Root Mean Square Error value (RMSE), Mean Absolute Error value (MAE) and Correlation Coefficient value (CC) (Zalina *et al.*, 2002; Zin *et al.*, 2009).

The Root Mean Square Error value (RMSE) is used to indicate the accuracy of the certain probability distribution in predicting the measured values. The minimum RMSE obtained will contribute to a more accurate probability distribution. Formula of RMSE is expressed as:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n \left( \frac{x_i - Q(F_i)}{x_i} \right)^2} \quad (7)$$

Where:

- $x_i$  = Annual maximum daily rainfall data (mm)
- $n$  = Number of observations
- $Q(F_i)$  = The  $i^{th}$  quantile estimation for the corresponding probability distribution,  $F_i = \frac{i - 0.44}{n + 0.12}$

The Mean Absolute Error value (MAE) is applied to calculate absolute value of the average different between the measured values and the predicted values. The best probability distribution model is chosen based on the MAE value that is closer to 0. The MAE is defined as:

$$MAE = \frac{1}{n} \sum_{i=1}^n \left( \left| \frac{x_i - Q(F_i)}{x_i} \right| \right) \quad (8)$$

Where:

- $x_i$  = Annual maximum daily rainfall data (mm)
- $n$  = Number of observations
- $Q(F_i)$  = The  $i^{th}$  quantile estimation for the corresponding probability distribution,  $F_i = \frac{i - 0.44}{n + 0.12}$

The Correlation Coefficient value (CC) is determined to describe how much of the measured value dispersion is explained by the predicted value. The CC value closer to one will be the better fit probability distribution model. Formula of CC is given as:

$$CC = \frac{\sum_{i=1}^n (x_i - \bar{x})(Q(F_i) - \bar{Q}(F_i))}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (Q(F_i) - \bar{Q}(F_i))^2}} \quad (9)$$

Where:

- $x_i$  = Annual maximum daily rainfall data (mm)

$\bar{x}$  = Average of  $x$

$n$  = Number of observations,  $F_i = \frac{i - 0.44}{n + 0.12}$

$Q(F_i)$  = The  $i^{th}$  quantile estimation for the corresponding probability distribution

$\bar{Q}(F_i)$  = Average of  $Q(F_i)$

Furthermore, the suitable probability distribution of rainfall data was used to derive rainfall intensity estimation for various return periods.

## Results and Discussion

The data used in this study is the daily rainfall amount data obtained from Ujung Pandang rainfall station which located in Makassar city for the period 1986-2015. Table 1 presents that during the period 1986-2015, the maximum daily rainfall amount is 376 mm that has occurred on February 2000. This month is the rainy season in Makassar city. In this period, the average daily rainfall is 38 mm. Meanwhile, on December 1996 has occurred the longest rainfall duration, i.e., 30 days. December is also the peak time of heavy rains in Makassar.

Figure 1 shown that the average daily of rainfall increases from November to January and decreasing on February. These results indicated that Makassar city experiences the rainy season from November to April and peak time from December to February. In contrast, Makassar city experiences the dry season from May to October and peak time on July to September.

Meanwhile, the best fitted probability distribution is determined due to closer to zero values of MAE, smaller values of RMSE and the CC value closer to one. In Table 2 presents that the Generalized Extreme Value distribution (GEV) is most suitable as compare with other distributions for daily rainfall at Ujung Pandang station.

The parameter values for the GEV probability distribution are given in Table 3. Based on the best distribution model and then determined the design rainfall estimation for various return periods, i.e., two years (T2), five years (T5), 10 years (T10), 30 years (T30), 50 years (T50) and 100 years (T100), respectively (Table 4). Further, these design rainfall values are used to estimate rainfall intensity for various rainfall durations and return periods as shown in Table 5.

Further, the rainfall intensity values in Table 5 are used to determine the rainfall Intensity-Duration-Frequency relationship. Based on the IDF curves in Fig. 2, the relationship between the rainfall Duration (D) and the rainfall Intensity (I) is exponential function. Figure 3 also shown that the longer the rainfall duration, the rainfall intensity is smaller.

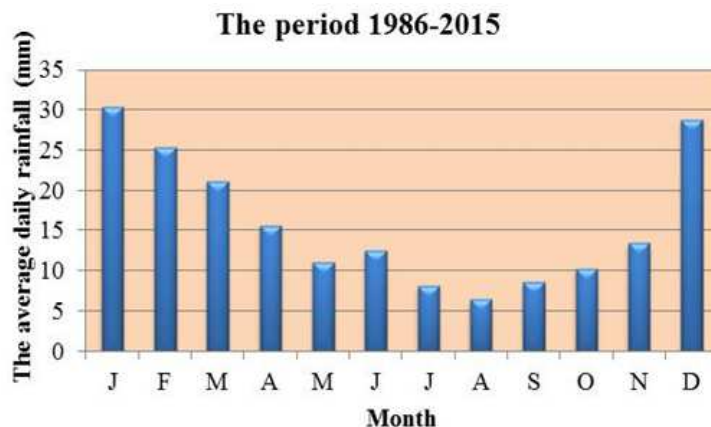


Fig. 1. Distribution of the average daily rainfall at Ujung Pandang station

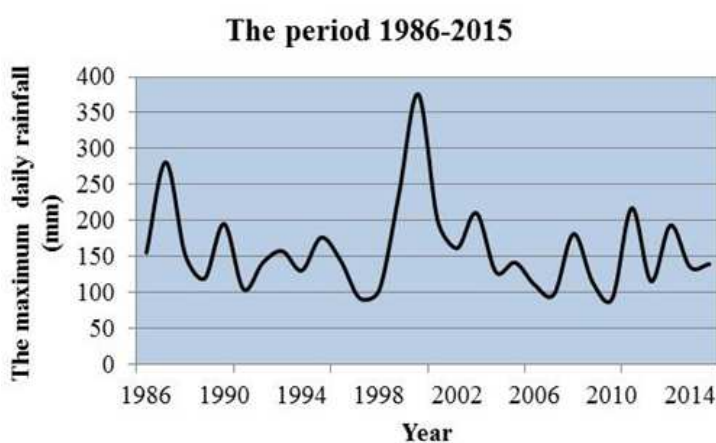


Fig. 2. Distribution of the annual maximum daily rainfall at Ujung Pandang station

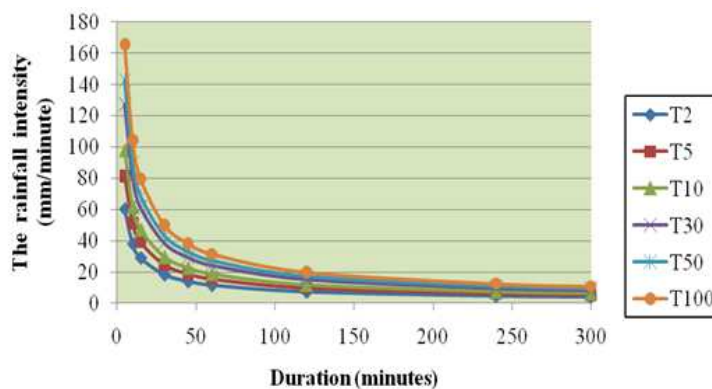


Fig. 3. The IDF curves of Ujung Pandang station

Statistics	Values
The maximum daily rainfall amount	376 mm
The minimum daily rainfall amount	0
The average daily rainfall	38 mm
The longest duration of daily rainfall	30 days

Statistics	Probability distribution				
	Gumbel	GEV	GPA	GLO	LP3
RMSE	0.047	0.030	0.042	0.035	0.382
MAE	0.037	0.023	0.032	0.025	0.303
CC	0.979	0.995	0.987	0.995	0.987

Table 3. Parameter estimation for the GEV distribution

Parameter		
Location ( $\xi$ )	Scale ( $\alpha$ )	Shape ( $\kappa$ )
29.69	8.82	-0.17

Table 4. Design rainfall estimation (mm) for GEV distribution

Return periods (years)					
2	5	10	30	50	100
33.03	44.76	53.87	70.04	78.52	91.22

Table 5. Rainfall intensity estimation (mm/minutes)

Duration (minutes)	Return periods (years)					
	2	5	10	30	50	100
5	60.01	81.33	97.89	127.27	142.69	165.75
10	37.80	51.24	61.66	80.18	89.89	104.42
15	28.85	39.10	47.06	61.19	68.60	79.69
30	18.17	24.63	29.65	38.54	43.21	50.20
45	13.87	18.80	22.62	29.41	32.98	38.31
60	11.45	15.52	18.68	24.28	27.22	31.62
120	7.21	9.78	11.76	15.30	17.15	19.92
240	4.54	6.16	7.41	9.64	10.80	12.55
300	3.92	5.31	6.39	8.30	9.31	10.81

## Conclusion

The purpose of this study is to determine the best fit distribution. Five distributions namely Gumbel, Generalized Extreme Value (GEV), Generalized Pareto (GPA), generalized logistic (GLO) and Pearson type III (P3) distributions are considered. Based on the result, the Generalized Extreme Value distribution (GEV) is the most suitable to describe the daily rainfall patterns in Makassar city.

The other aims this study is to construct IDF curves using the GEV distribution for describing the relationship between intensity, duration and return periods. Return periods are used as the measure of frequency of rainfall occurrence. Based on the IDF model, the results pointed out that the relationship has an exponential function form which is the long duration of rainfall is followed by decreasing intensity of rainfall for various return periods.

In this study, the best statistical distribution of rainfall and the IDF model have been identified. As at the moment, there are limited studies on probability distribution of extreme rainfall events in Makassar city, these results could provide useful information for estimating the design rainfall in frequency analysis, in particular for extreme rainfall events.

## Acknowledgement

The authors are grateful to Department of Water Resources Management in South Sulawesi province for supplying data. We would like also thank to Universitas Negeri Makassar for financial support with contract no: 2596/UN36/LT/2016.

## Author's Contributions

**Wahidah Sanusi:** Drafting of manuscript and data analysis.

**Usman Mulbar:** Designing of research methodology.

**Hendra Jaya:** Collecting and analyzing of data.

**Purnamawati:** Collecting and organizing of data.

**Syafuruddin Side:** Drafting of manuscript and data analysis.

## Conflict of Interest

All authors declare no conflict interest.

## References

- Aksoy, H., 2000. Use of gamma distribution in hydrological analysis. *Turkish J. Eng. Environ. Sci.*, 24: 419-428.
- Cai, M., 2010. Study on variation in wet and low water of precipitation prediction based on Markov with weights theory. *Proceedings of the 6th International Conference on Natural Computation*, Aug. 10-12, IEEE Xplore Press, pp: 4296-4300. DOI: 10.1109/ICNC.2010.5583626
- Deka, S., M. Borah and S.C. Kakaty, 2009. Distributions of annual maximum rainfall series of north-east India. *Eur. Water*, 27: 3-14.
- Du, J., J. Fang, W. Xu and P. Shi, 2013. Analysis of dry/wet conditions using the standardized precipitation index and its potential usefulness for drought/flood monitoring in Hunan Province, China. *Stochastic Environ. Res. Risk Assess.*, 27: 377-387.
- Eslamian, S.S. and H. Feizi, 2007. Maximum monthly rainfall analysis using L-moments for an arid region in Isfahan province, Iran. *J. Applied Meteorol. Climatol.*, 46: 496-503.
- Hosking, J.R.M. and J.R. Wallis, 1997. *Regional Frequency Analysis: An Approach Based on L-Moments*. 1st Edn., Cambridge University Press, London, ISBN-10: 0521430453, pp: 242.
- Koutsoyiannis, D., D. Kozonis and A. Manetas, 1998. A mathematical framework for studying rainfall intensity-duration-frequency relationships. *J. Hydrol.*, 206: 118-135. DOI: 10.1016/S0022-1694(98)00097-3
- Kumar, R. and A. Bhardwaj, 2015. Probability analysis of return period of daily maximum rainfall in annual data set of Ludhiana, Punjab. *Ind. J. Agric. Res.*, 49: 160-164. DOI: 10.5958/0976-058X.2015.00023.2
- Kysely, J. and J. Picek, 2007. Probability estimates of heavy precipitation events in a flood-prone central-European region with enhanced influence of Mediterranean cyclones. *Adv. Geosci.*, 12: 43-50.
- Li, Z., Z. Li, W. Zhao and Y. Wang, 2015. Probability modeling of precipitation extremes over two river basins in Northwest of China. *Adv. Meteorol.*, 1-12. DOI: 10.1155/2015/374127

- May, W., 2004. Variability and extremes of daily rainfall during the Indian summer monsoon in the period 1901-1989. *Global Planetary Change*, 44: 83-105.
- Norlida, M.D., I. Abustan, R. Abdullah, A.S. Yahaya and O. Sazali *et al.*, 2011. Intensity-duration-frequency estimation using generalized Pareto distribution for urban area in a tropical region. *Proceedings of 12th International Conference on Urban Drainage, (ICU' 11), Porto Alegre/Brazil*, pp: 11-16.
- Prodanovic, P. and S.P. Simonovic, 2007. Development of rainfall intensity duration frequency curves for the City of London under the changing climate. *Water Resour Res Report*, London.
- Sanusi, W. and K. Ibrahim, 2012. Application of loglinear models in estimating wet category in monthly rainfall. *Sains Malaysiana*, 41: 1345-1353.
- Shabri, A. and N.M. Ariff, 2009. Frequency analysis of maximum daily rainfalls via Lmoment approach. *Sains Malaysiana*, 38: 149-158.
- Singh, V.P. and H. Guo, 1997. Parameter estimation for 2-parameter generalized Pareto distribution by pome. *Stochastic Hydrol. Hydraul.*, 11: 211-227.
- Soro, G.E., T.A. Goula Bi, F.W. Kouassi and B. Srohourou, 2010. Update of intensity-duration-frequency curves for precipitation of short durations in tropical area of West Africa (Cote D'ivoire). *J. Applied Sci.*, 10: 704-715.
- Syafuruddin, S. and M.S.M. Noorani, 2013. Lyapunov function of SIR and SEIR model for transmission of dengue fever disease. *Int. J. Simulat. Process Model.*, 8: 177-184.
- Yang, Z., 2000. Predictive densities for the lognormal distribution and their applications. *Microelectr. Reliability*, 40: 1051-1059.
- Zalina, M.D., M.N.M. Desa, V.T.A. Nguyen and A.H.M. Kassim, 2002. Selecting a probability distribution for extreme rainfall series in Malaysia. *Water Sci. Technol.*, 45: 63-68.
- Zin, W.Z.W., A.A. Jemain and K. Ibrahim, 2009. The best fitting distribution of annual maximum rainfall in Peninsular Malaysia based on methods of L-moment and LQ-moment. *Theor. Applied Climatol.*, 96: 337-344.