International Journal of Academic and Applied Research (IJAAR)
ISSN: 2643-9603
Vol. 6 Issue 7, July - 2022, Pages: 68-72

Extreme Rainfall Model in Palangkaraya City Using Probability Weight Moments Estimation

Hasnita

Faculty of Islamic economics and business, State Islamic institute Palangkaraya, Indonesia
Hasnita@iain-palangkaraya.ac.id

Abstract*: Rainfall is one of the climate elements that is interesting to study in order to reduce the impact caused by extreme rainfall. The method used in analyzing extreme values is Extreme value Theory with a maximum block approach of Generalized Extreme Value. And using the Probability Weighted Moments method in estimating the parameters. The results of this study indicate that December has the longest tail compared to other months, because it has the smallest shape value. The estimation result in December shows the shape value -0.025, location 309.355 and scale 98.959. And the return level results show that December has the greatest probability of occurrence of extreme values because it produces the largest maximum estimated value compared to the others. The maximum estimated values in December for the next 2,3,4,5 and 6 years period are 346,4808 mm, 394,0712 mm, 424,2779 mm, 446,5129 mm, and 464,1209 mm.*

Keywords: Extreme Rainfall, Extreme value Theory, Generalized Extreme Value, the Probability Weighted Moments

1. INTRODUCTION

Rainfall is the amount of water that falls on a flat surface during a certain period. Rainfall is measured in millimeters (mm) height above the horizontal surface. Rainfall can also be interpreted as the height of rainwater that collects in a flat place, does not evaporate, does not seep and does not flow [9]. Rainfall is one of the climate elements that is interesting to study in order to reduce the impact caused by extreme rainfall. This extreme weather with high rainfall can potentially cause hydrometeorological disasters, in the form of floods, flash floods, landslides, strong winds, especially for people who live and live in areas prone to hydrometeorological disasters.

Extreme value theory (EVT) is one method that can be used in analyzing extreme values [11]. One approach that can be used in EVT is the maximal block of Generalized Extreme Value (GEV) which identifies extreme values based on the maximum value of grouped data in a certain period. Return level is a very important part of EVT. The purpose of this return level is to estimate the maximum value in a certain period. However, before determining the return level value, you must first know the estimated parameters used. One of the parameter estimates that can be used in the EVT method of the most concentrated distribution or GEV is Probability Weighted Moments (PWM).

Based on the description above, the researchers used the method Extreme value theory in analyzing rainfall in the city of Palangkaraya 2011-2019. Researcheruse method Probability Weighted Moments on the extreme distribution involved in estimating the parameters.

2. LITERATUR REVIEW

2.1 EXTREME VALUE THEORY

Extreme Value Theory (EVT) is a theory to model and measure an event, where the probability of the occurrence of the event is very small [1]. There are generally two ways to identify extreme values. The first method, BM, is to take the maximum values in a period, for example monthly or yearly periods. Observations of these values are considered as extreme values. The second method, POT is to take values that exceed a threshold value. All values that exceed the threshold are considered as extreme values [10]. In Wahyudi's research (2011) in the agricultural production center area in Ngawi Regency which identified extreme rainfall. In his research, the behavior of the distribution tail shows that in some cases the climate (rainfall, temperature, wind speed).

2.2 Generalized Extreme Value Distribution

Modeling for sampling using block maxima, which is denoted by M_n is :

$$M_n = maks(x_1, x_2, \dots x_n) \quad (1)$$

Where is $x_1, x_2, \dots x_n$ a sequence of independent random variables x_i with $i = 1, \dots n$ states the values of the observations in a regular time period and has a distribution function F. Then M_n is the maximum value in n units time. The distribution function of M_n is given as follows [2]:

$$P_r\{M_n \le x\} = P_r\{X_1 \le x, \dots, X_n \le x\}$$
$$= \{F(x)\}^n \quad (2)$$

Distribution function F^n is the degenerate distribution function which at in practice it is not useful because we cannot know the boundary function distribution. So the distribution family of the set of extreme values cannot be known. Thus, linear normalization is performed for find out the distribution family of F^n that is with :

$$M_n{}^* = \left(\frac{M_n - b_n}{a_n}\right) \quad (3)$$

in constant sequence $\{a_n > 0\}$ and $\{b_n\}$. So that we get the distribution approach for $M_n{}^*$ given by Theorem 1.

Theorem 1 [1]:

If there is a constant sequence $\{a_n > 0\}$ and $\{b_n\}$ such that

$$P_r = \left\{\frac{M_n - b_n}{a_n} \leq x\right\} \to F(x), n \to \infty \quad (4)$$

Where F is a non-degenerate distribution function, then F is included in one of the following distribution families:

$$\text{I} : F(x) = \exp\left\{-\exp\left[-\frac{x-b}{a}\right]\right\}, -\infty < x < \infty \quad (5)$$

$$\text{II} : F(x) = \begin{cases} 0, x \leq b \\ \exp\left[-\frac{x-b}{a}\right]^{-a}, x > b \end{cases} \quad (6)$$

$$\text{III} : F(x) = \begin{cases} \exp\left[\frac{x-b}{a}\right]^{a}, x < b \\ 1, x \geq b \end{cases} \quad (7)$$

For parameters $a > 0$, b and in distributions I and II $a > 0$

The theorem above states that the maximum sample measurement of $M_n{}^*$ is convergent in the distribution to a variable that has a distribution among the distribution families I, II or III. So based on definition, $M_n{}^*$ has a distribution limit of one of the distribution families I , II, III.

Good analysis can be done with the formulation reformulation of the model in Theorem 1. The reformulation is intended to examine the family Gumbel, Frechet and Weibull because they can be combined into one family model which has the following cumulative distribution function :

$$G(z) = \begin{cases} \exp\left\{-\left[1 + \xi\left(\frac{x-b}{\sigma}\right)\right]^{\frac{-1}{\xi}}\right\} \xi \neq 0 \\ \exp\left\{-exp\left[-\left(\frac{x-b}{\sigma}\right)\right]\right\}, \xi = 0 \end{cases} \quad (8)$$

With $\left\{z: 1 + \xi\left(\frac{x-\mu}{\sigma}\right) > 0\right\}$; $-\infty < \mu < \infty, \sigma > 0$ dan $-\infty < \xi < \infty$

2.3 Probability Weighted Moments

According to Greenwood et al. (1979) in Rao (2000) Probability Weighted Moments is defined as follows [8] :

$$M_{p,r,s} = E\left[X^p F^r (1-F)^s\right] = \int_0^1 [x(F)]^p F^r (1-F)^s \, dF \quad (9)$$

There are two moments to consider $M_{1,0,s}$ and $M_{1,r,0}$ as follows :

$$M_{1,0,s} = a_s = \int_0^1 x(F)(1-F)^s \, dF \quad (10)$$

$$M_{1,r,0} = B_r = \int_0^1 x(F) F^r \, dF \quad (11)$$

With p, r and s is rill number.

Based on a random sample with $x_1 \leq x_2 \leq \cdots \leq x_n$, $n > r, n > s$ the PWM unbias estimate is obtained as follows :

$$a_s = \hat{a}_s = \widehat{M}_{1,0,s} = \frac{1}{n}\sum_{i=1}^{n} \frac{\binom{n-i}{s} x_i}{\binom{N-1}{s}} \quad (12)$$

$$b_r = \hat{\beta}_r = \widehat{M}_{1,r,0} = \frac{1}{n}\sum_{i=1}^{n} \frac{\binom{i-r}{r} x_i}{\binom{n-1}{r}} \quad (13)$$

According to Hosking et al. (1985) Probability Weighted Moments from the GEV distribution for $\xi \neq 0$ as follows [5]:

$$\beta_r = \frac{1}{r+1}\left\{\mu + \frac{\sigma}{\xi}\left(1 - (r+1)^{-\xi}\Gamma(1+\xi)\right)\right\}, \xi > -1 \quad (14)$$

Parameter estimation PWM $\hat{\mu}, \hat{\sigma}, \hat{\xi}$ as follows :

$$\hat{\xi} = 7{,}85890\, c + 2{,}9554\, c^2 \quad (15)$$

With :

$$c = \frac{2b_1 - b_0}{3b_2 - b_0} - \frac{\log 2}{\log 3} \quad (16)$$

$$\hat{\sigma} = \frac{(2b_1 - b_0)\hat{\xi}}{\{\Gamma(1+\hat{\xi})(1-2^{-\hat{\xi}})\}} \quad (17)$$

$$\hat{\mu} = b_0 + \frac{\hat{\sigma}}{\hat{\xi}}\{\Gamma(1+\hat{\xi}) - 1\} \quad (18)$$

2.4 Distribution Conformity Test

Distribution checks with quantile plots are generally easy to do because only see the pattern of distribution of extreme values that follow a linear line. If the quantile plot follows a linear line then the distribution is appropriate [7].

In addition to the quantile plot, the Kolmogorov-Smirnov test can be used. The steps of the Kolmogorov-Smirnov test according to Daniel (1989) are [3]

1. Hypothesis

$H_0 : S(x) = F_0(x)$ The data have followed the theorical distribution

$H_1 : S(x) \neq F_0(x)$ The data does not follow the theorical distribution

2. Test statistics

$$D = sup_x |S(x) - F_0(x)| \quad (19)$$

3. Test criteria

Reject H_0 if $D > D_{1-\alpha}$

Where $D_{1-\alpha}$ is the critical value obtained from the Kolmogorov-Smirnov table

at the level of significance α.

2.5 Return Level

According to Gilli and Kellezi (2006), the maximum expected value will be exceeded once in a period of k with a period of n will follow the equation [4]:

$$\hat{R}_n^k = \begin{cases} \hat{\mu} - \frac{\hat{\sigma}}{\hat{\xi}}\left[1 - \left\{-\log(1 - \frac{1}{k})\right\}^{-\hat{\xi}}\right], \hat{\xi} \neq 0 \\ \hat{\mu} - \hat{\sigma}\log\left\{-\log\left(1 - \frac{1}{k}\right)\right\}, \hat{\xi} = 0 \end{cases} \quad (20)$$

With k is return level period and n is Length of time in the period of return level Number equations consecutively.

3. Result and Discussion

3.1 Descriptive data

The general description of the rainfall in Palangkaraya City is carried out as initial information to determine the characteristics and patterns of rainfall used for subsequent analysis. The general description of the rainfall in the Palangkaraya method is presented in Table 1.

Table 1. Average Value, Standard Deviation, Minimum and Maximum Value of Rainfall in Palangkaraya City.

N	120
Minimum	7.00
Maximum	604.7

International Journal of Academic and Applied Research (IJAAR)
ISSN: 2643-9603
Vol. 6 Issue 7, July - 2022, Pages: 68-72

Average	253.31
Standard Deviation	142,20

Table 1 shows that the minimum value of rainfall in Palangkaraya City is 7.00. For a maximum value of 604.7 mm. In addition to the maximum and minimum values, it can also be seen that the standard deviation value used to express the diversity of rainfall indicates that the standard deviation of rainfall in Palangkaraya City is of 142.20. The standard deviation value shows a value smaller than the average value, which means that the rainfall in Palangkaraya City has a homogeneous data distribution.

Table 2. Average Value, Standard Deviation, Minimum and Maximum Value of Monthly Rainfall in Palangkaraya City

Month	Minimum	Average	Maximum
January	138.3	347.1	485.5
February	149.4	328.6	522.4
March	248.5	355.2	511.1
April	235.7	346.9	575.9
May	70.0	268.0	475.7
June	35.0	173.7	436.6
July	7.0	125.9	244.3
August	23.0	88.8	188.5
September	28.8	123.8	280.4
October	60.0	206.7	414.9
November	133.0	310.8	430.8
December	217.7	364.0	604.7

Table 2 shows that the maximum value of rainfall is in December, which is 604.7. And the highest average monthly rainfall is also in December, which is 364.0. This indicates that the peak of the rainy season in the city of Palangkaraya occurs in December.

3.2 Rainfall pattern of Palangkaraya City

The pattern of rainfall in Palangkaraya City is presented in the following graphic form:

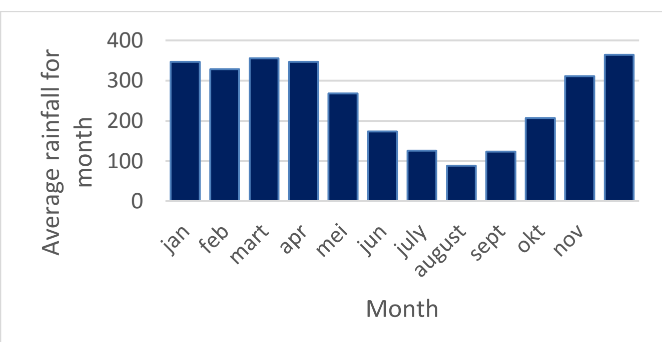


fig 1. Average Rainfall For month

The figure above shows that the average monthly rainfall in Palangkaraya City in 2011-2019 has a monsoon pattern.The monsoon pattern is a rainfall pattern that forms the letter U or in other words has one peak of the rainy season (unimodal)..The peak of the rainy season in Palangkaraya City occurs in December because it has the highest average monthly rainfall frequency compared to other months.

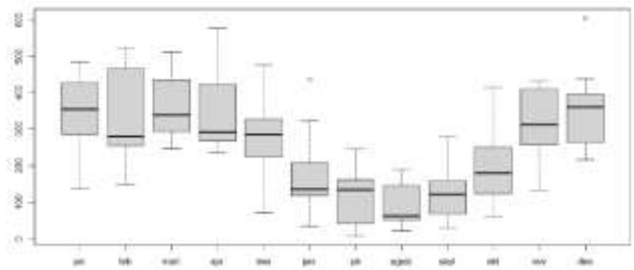


Fig 2. Boxplot

The picture above shows that there are extreme values in June and December. This shows that the extreme values are not only in the rainy season but also in the dry season. Which means that even in the dry season months, it still rains.

3.3 Rainfall Criteria

The classification of monthly rainfall aims to determine the criteria for rain that occurs in a certain month within a period of 9 years so that it is known that there are many criteria for monthly rainfall which are included in the low, medium, high and very high categories. The classification of monthly rainfall, October, November, December, January, February, March, April and May is presented in the following table 3.

Table.3. Monthly Rainfall Criteria

Month	low (0-100)	medium (100-300)	high (300-600)	very high (>600)
November	0	4	5	0
December	0	3	5	1
January	0	3	6	0
February	0	5	4	0
March	0	3	6	0
April	0	5	4	0
May	1	5	3	0

The table above shows that October and December have the criteria for the largest Monthly Rainfall compared to other months. This shows that October and December have extreme values.

3.4 Long-tail data identification

Identification of tailed data on rainfall in Palangkaraya City can be seen by using a normality probability plot.

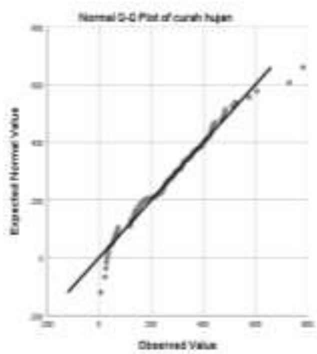


Fig 3. Histogram Image Of Rainfall In Palangkaraya City.

Based on the picture above, it can be seen that the condition of the distribution of dots (gray) does not follow a linear line (black), which means that the data is not normally distributed. And it can also be interpreted that the pattern of rainfall data in Palangkaraya City contains tailed data (extreme data).

3.5 Taking extreme values using max blocks

Taking extreme values is only done onperiodThe rainy seasons are November, December, January, February, March and April. Based on rainfall data in Palangkaraya City in 2011-2019, within a period of 9 years for October, November, December, January, February, March, April and May there are 9 blocks with each block containing 4 observations.

3.6 Parameter Estimation Using PMW method

After obtaining the extreme value in each month for a period of 10 years, the value is processed using the R 4.1.3 software with the fExtremes and TLMoments software packages. The aim is to find out the type of distribution in each month using the most extreme value distribution with weighted probability moment parameter estimation. The results of parameter estimation using weighted probability moments are presented in the following table:

Table.4.Parameter Estimation

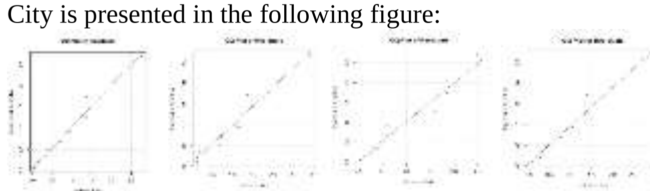
Month	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\varepsilon}$
October	151,246	91.468	0.029
November	280.961	109.281	-0.414
December	309,355	98,959	-0.025
January	325,307	121,038	-0.595
February	270,932	137.511	-0.186
March	306.083	78,100	0.049
April	284,328	70.138	0.244
May	223,325	128,929	-0.292

The table above shows that December has the longest tail compared to other months because it has the smallest estimated shape parameter $\hat{\varepsilon}$of -0.025. The cumulative distribution function for December is as follows:

$$G(z) = \exp\left\{-\left[1+(-0{,}025)\left(\frac{z-309{,}355}{98{,}959}\right)\right]^{\frac{-1}{-0{,}025}}\right\} \quad (21)$$

3.7 Model fit test

In this study, to test the suitability of the distribution of whether the data has followed the distribution of the most extreme values, quantile plots and the Kolmogorov-Smirnov test are used. The plot of rainfall quantiles in Palangkaraya City is presented in the following figure:



(October) (November) (December) (January)

(February) (March) (April) (May)

Fig 4. The plot of rainfall quantiles

The picture above shows that in the eight months the distribution of points follows a linear line, which means that the data follows the distribution of the highest extreme values.

Kolmogorov-Smirnov test.

Significance level:$\alpha = 0{,}05$

Table 5. values for the Kolmogorov-Smirnov test D_{count}

Month	D_{count}
October	0,121
November	0,256
December	0,221
January	0,125
February	0,233
March	0,166
April	0,153
May	0,127

Conclusion :
Reject if value with H_0 $D_{count} > D_{table}$. $D_{table} = 0{,}318$. Based on the test statistics, it can be seen that in the six months the value of $D_{count} < D_{table}$. Which means there is enough evidence to say that the eight months were accepted. so that

the data H_0 has followed the distribution of the most extreme values.

3.8 return level

The analysis of the estimated maximum value in period k with period n on rainfall data in Palangkaraya City in 2011-2019 is intended to find out how much maximum value will occur in period k with period n in certain months. The estimated maximum rainfall values for the next 2, 3, 4 and 5 years can be seen in the following table:

Table 6. Return Level

Month	Estimated maximum value in k years				
	2	3	4	5	6
October	188,7072	233.2920	261,6900	282.6434	299.2665
November	341,1938	377.8335	392.9914	401.3549	406,6730
December	346,4808	394.0712	424,2779	446.5129	464,1209
January	368,6887	408,2264	426.8188	437,9964	445.5826
February	328,7616	367,9272	395,6612	417,6541	436,0900
March	300.5676	341.8415	376,3168	404.9258	434.9029
April	268,6012	319,7328	348,1979	367,3702	381.5582
May	268,6012	319,7328	348,1979	367,3702	381.5582

Based on the table above, it shows that the estimated maximum rainfall value in December is the largest compared to other months, so that December has the greatest chance of extreme values.

4. CONCLUSION

Based on the analysis and discussion that has been described previously, the following conclusions can be drawn:

1. Rainfall in Palangkaraya City has a monsoon pattern because it follows the letter U. With extreme values found in December.
2. Based on the calculation of parameter estimates, it shows that December has the longest tail compared to other months. Because in December has a value of $\hat{\varepsilon}$ the smallest is -0,025 compared to other months.
3. Based on the calculation of the estimated maximum value of rainfall for the next 2, 3, 4 and 5 years, it shows that Decemberr has the greatest chance of extreme values occurring compared to the other six months.

5. References

[1] Alves, Isabel Fraga & Neves, Claudia. Extreme Value Distribution. Portugal: University of London, 2010

[2] Coles, S. An Introduction to Statistical Modelling of Extreme Values. London: Verag, 2001.

[3] Daniel, W.W., Statistika Nonparametrik Terapan, PT Gramedia, Jakarta, 1989.

[4] Gilli, M. and Kellezi, E. An Application of Extreme Value Theory for Measuring Financial Risk. Switzerland: University Of Geneva And FAME. 2006

[5] Hastaryta, R dan Effendie, A.R., Estimasi Value-At-Risk dengan Pendekatan Extreme Value Theory- Generalized Pareto Distribution (Studi Kasus IHSG 1997- 2004), Jurnal Fakultas Matematika Ilmu Pengetahuan Alam, Vol 16. No. 2. 2006.

[6] Hosking, J.R.M., Wallis, J.R. dan Wood, E.F., Estimation of The Generalized Extreme-Value Distribution by the Method of Probability-Weighted Moments, Techometrics, Vol 27. No. 3. August 1985.

[7] Mallor, Nualart dan Omey, An Introduction to Statistical Modelling Of Extreme Value Application to Calculate Extreme Wind Speeds, Hogeschool Universitei Briscel, 2009

[8] Rao, A.R. dan Khaled H.H. Flood Frequency Analysis. New York : CRC Press. 2000.

[9] Suroso. 2006. Analisis Curah Hujan untuk Membuat Kurva Intensity Duration Frequency (IDF) di Kawasan Rawan Banjir Kabupaten Banyumas. Jurnal Teknik Sipil Vol. 3.

[10] Wahyudi & Sutikno. Identifikasi Curah Hujan Ektstrem di Kabupaten Ngawi Menggunakan Generalized Extreme Value dan Generalized Pareto Distribution. Surabaya: Institut Tekhnologi Surabaya, diunduh pada 10 Februari 2013 pukul 11:10.

[11] Wahyudi. Identifikasi Curah Hujan Ekstrem di Kabupaten Ngawi Menggunakan Generalized Extreme Value dan Generalized Pareto Distribution. Jurnal Sains dan Seni ITS. 2011.