

Regional Frequency Analysis of Rainfall, using L-Moment Method, as A Design Rainfall Prediction

Devita Mayasari*, Pratiwi Setyaning Putri

Faculty of Infrastructure and Territorial Technology, Institut Teknologi PLN, Jakarta, INDONESIA

Jalan Lingkar Luar Barat, Cengkareng, Jakarta Barat

*Corresponding authors: devita@itpln.ac.id

SUBMITTED 9 October 2020 REVISED 27 November 2020 ACCEPTED 15 December 2020

ABSTRACT Frequency analysis is a method for predicting the probability of future hydrological events, based on historical data. Generally, frequency analysis of rainfall data and discharge data is performed using the moment method, but this method has a large bias, variant, and slope, thus there is a possibility of producing inaccurate hydrological design magnitudes. Meanwhile, the L-moment method is a linear combination of Probability Weighted Moment, with the ability to process data concisely and linearly. This study was therefore conducted to discover the L-moment method's capacity to obtain a regional probability distribution and design rainfall, used as a basis for calculating hydrological planning, in anticipation of disasters. The study location, Mount Merapi, was selected to enable a more accurate prediction of maximum rainfall with the capacity to cause cold lava in the area, and consequently, reduce the risk of loss for people living within close proximity. According to the results, the L-moment regional ratio results were $\tau_2^R = 0.203$, $\tau_3^R = 0.166$, and $\tau_4^R = 0.169$. The homogeneity and heterogeneity tests show all rainfall stations are uniform or homogeneous, and no data were released from the discordance test results. Also, the growth factor value increases in each return period design rainfall prediction. In this study, the suitable regional probability distribution for the research area is the Generalized Logistic distribution with formulated design rainfall equation. Regional design rainfall is able to predict possible rainfall within the area. The Test model showed the minimum $R_{Bias} = 0.45\%$, maximum $R_{Bias} = 41.583\%$, minimum $R_{RSME} = 0.45\%$, and maximum $R_{RSME} = 71.01\%$. Meanwhile, the L-moment method's stability was shown by the model test minimum error = 1.64% and maximum error = 16.60%. The higher error value in the higher return period shows L-moment is able to reduce bias data, however, this has limitations in the higher return period.

KEYWORDS Regional frequency analysis; probability weighted moment, L-moment method, and design rainfall.

© The Author(s) 2021. This article is distributed under a Creative Commons Attribution-ShareAlike 4.0 International license.

1 INTRODUCTION

Probability prediction of hydrological events based on historical data is possible, using frequency analysis. This analysis serves as the basis for calculating hydrological planning to anticipate any future possibilities and is used to predict extreme events, including future rainfall or design floods with the capacity to cause flooding. Thus, in flood protection structures, anticipation is required to minimize risks. Regional frequency analysis is performed in cases where data samples from the same variables in one region are used. Generally, for both rain data and discharge data, this analysis is performed using the moment method, a nonlinear transformation, where higher sample moments result in higher instability. In addition, the moment method has a considerable bias, variance, as well as skewness. Also, nonlinear transformations have the capacity to cause deviations and wrong parameter estimates due to large differences from the data majority.

However, the L-moment method performs data processing concisely as well as linearly, and nonlinear data transformations are avoided because there is no squared or cubic system calculation data. Thus, there is no bias, as is common with the moment method. The L-moment has the advantage of conventional moments, where there is a wide distribution range. Thus, more accurate data is produced in cases where estimation is carried out from a sample. The L-moment and probability-weighted moment (PWM) aim to conclude the probability distribution theory to be used in the regional frequency analysis. However, the application of these two methods for regional rainfall frequency analysis requires assessment to estimate design rainfall at specific return periods. This study, therefore, aims to utilize the L-moment study to provide more accurate results at Mount Merapi, the research area, selected to enable a more accurate prediction of

maximum rainfall with the capacity to cause debris flow in the area. Furthermore, analysis design rainfall is a possible basis for discharge analysis, planned as anticipation of flood protection waterworks, to reduce the risk of loss for people living near Mount Merapi.

2 METHODS

2.1 Study Area and Data

The study area is located at Mount Merapi, with rainfall stations having previously recorded daily rainfall data. In this study, 21 stations were selected, based on location distribution and rainfall data availability. Figure 1 shows the rainfall station distribution map in this study.

This study utilized daily rainfall records data from 1980 to 2018 at 21 rainfall stations, with a completed data length of 16 - 37 years. However, several years of data were incomplete and therefore not used in the analysis. The data selection process was conducted through the annual maximum series, and this selects data based on the maximum annual rainfall. In this study, Sorasan station was not involved in frequency analysis but was selected as a reference station in the test model, because this station possessed the most comprehensive as well as continuous recording data, and therefore possessed the most stable data.

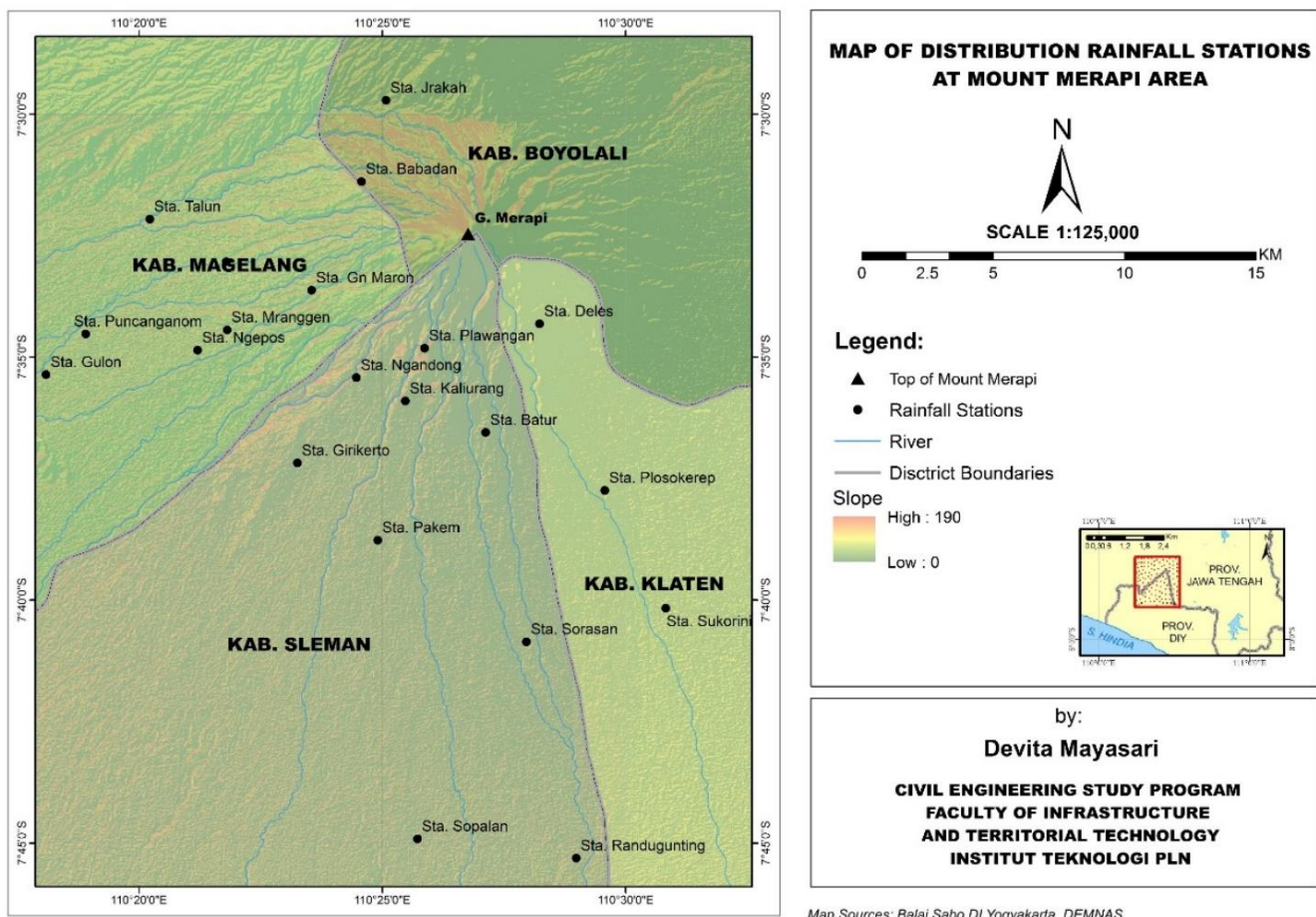


Figure 1. Distribution of rainfall station locations

2.2 Study Frameworks

Frequency analysis demands certain conditions data that must be uniform, independent, and representative (Haan, 2002). The moment method is a quantitative measure of geometric properties in a probability distribution. Furthermore, the moment is usually used to explain the sample's stability, and higher moments indicate instability and a need for further information (Takeleb, 2010). To estimate the amount of discharge or design rainfall in a return period, the analysis is performed using a statistical approach (Soewarno, 1995), including moment parameters, including mean value, standard deviation, variation coefficient, skewness coefficient, and kurtosis coefficient. Return period defined as average length of time for an event of given magnitude to be equalled or exceed in statistical sense (Maity, 2018). In hydrology, the most widely used probability distributions are the Normal, Log-Normal, Log Pearson III, and Gumbel distributions (Chow V.T., 2010). Meanwhile, in the moment method, the goodness-of-fit tests are Chi-Square and Smirnov-Kolmogorov tests (Harto, 2009). Hypothesis testing by goodness-of-fit test is needed to test the discrete variable (Kottegoda and Rosso, 2009). In addition, the homogeneity test is a method used to measure data uniformity in an area developed to analyze the flood but is also applicable to rainfall data (Darlymple, 1960).

The L-moment is the statistical result's linear length, as a conventional moment used to summarize the probability function's statistical parameters or the result of a data set. This method is an analogy to ordinary moment, estimated as a linear combination of Probability Weighted Moment (PWM). According to Hosking and Wallis (2009), L-moment using regional frequency analysis is performed through an environmental approach application because the data samples analyzed are observations of the same variables in one region. The method of regional frequency

analysis of hydrologic variables used either to provide quantile estimates at ungauged sites or to improve quantile estimates at poorly gauged sites (Naghetini, 2016)

The distribution selection using this method shows the relationship between each distribution's theoretical $L-C_s$ and $L-C_k$ parameters and the L-moment diagrams observational data. These parameters are identified for each type of theoretical probability distribution used in the L-moment diagram, to obtain a match with the data distribution. The L-moment parameter represents the inverse of the cumulative distribution function (cdf) for multiple distributions, and these are the uniform, Exponential, Logistic, Normal, Log-Normal, Generalized Pareto (GPA), Generalized Logistic (GLO), Generalized Extreme Value (GEV), Gumbel, and Pearson 3 Distributions. L-moment parameter is based on the PWM theory, Equations (1), (2), (3), (4), and (5) (Hosking and Wallis, 2009); (Maidment, 1993).

$$\beta_r = \frac{1}{n} \sum_{j=r+1}^n \frac{(j-1)(j-2)\dots(j-r)}{(n-1)(n-2)\dots(n-r)} X_j \quad (1)$$

$$\lambda_1 = \beta_0 \quad (2)$$

$$\lambda_2 = 2\beta_1 - \beta_0 \quad (3)$$

$$\lambda_3 = 6\beta_2 - 6\beta_1 + \beta_0 \quad (4)$$

$$\lambda_4 = 20\beta_3 - 30\beta_2 + 12\beta_1 - \beta_0 \quad (5)$$

With r ordo and X_j are the shorting data from largest to smallest, β data population parameter value and λ L-moment parameter, the L-moment ratio is based on Equations (6), (7), and (8) (Hosking and Wallis, 2009).

$$\tau_2 = \lambda_2 / \lambda_1 \quad (6)$$

$$\tau_3 = \lambda_3 / \lambda_2 \quad (7)$$

$$\tau_4 = \lambda_4 / \lambda_2 \quad (8)$$

Where, τ_2 signifies *L-variation* ($L-C_v$) coefficient, τ_3 denotes *L-skewness* ($L-C_s$) coefficient, and τ_4 represents *L-kurtosis* ($L-C_k$)

coefficient. Subsequently, data is filtered to determine the against of data and all data in an area. This filtering is based on the difference between the L-moment ratio of a data with the average L-moment ratio of all data regions, using Equations (9), (10), and (11) (Hosking and Wallis, 2009).

$$\bar{u} = N^{-1} \sum_{i=1}^N u_i, ; u_i = [\tau_{2i} \tau_{3i} \tau_{4i}]^T \quad (9)$$

$$A = \sum_{i=1}^N (u_i - \bar{u})(u_i - \bar{u})^T \quad (10)$$

$$D_i = \frac{1}{3} (u_i - \bar{u})^T A^{-1} (u_i - \bar{u}) \quad (11)$$

Where, u_i denotes the vector of sample L-moment ratio i location, N represents number of i locations i , \bar{u} signifies regional unweight mean of L-moment for each region, cross-product matrix A , D_i discordance test value i location. Also, i location is considered discordant in cases where the D_i value exceeds $D_{i-critic}$.

The heterogeneity test aims to assess whether the rainfall gauge station location is treatable as a uniform region, using Equations (12), (13), and (14) (Hosking and Wallis, 2009).

$$t^R = \sum_{i=1}^N n_i t^{(i)} / \sum_{i=1}^N n_i \quad (12)$$

$$V = \left\{ \sum_{i=1}^N n_i (t^{(i)} - t^R)^2 / \sum_{i=1}^N n_i \right\}^{1/2} \quad (13)$$

$$H = \frac{(V - \mu_v)}{\sigma_v} \quad (14)$$

Where t^R implies regional τ , n_i connotes length of data, $t^{(i)}$ τ represents the i^{th} order, while μ_v and σ_v are obtained from Kappa distribution methods (Karian Z.A, 2010; Hasby, 2014). The criteria were established by Hosking and Wallis (2009) for the assessment of heterogeneity in a region. These are, in cases where $H < 1$, the area is homogeneous, while in cases where $1 \leq H < 2$, the area is possibly homogeneous, and in cases where $H \geq 2$, the area is not heterogeneous. The selection of regional probability distribution of extreme rainfall using τ_3^R and τ_4^R statistical parameters through visual observation is to be

compared with L-moment diagram (Vogel and Fennessey, 1993). The regional frequency distribution is chosen by L-moment ratio diagram and the goodness-of-fit test Z^{DIST} . Subsequently, the selected distribution is subjected to a goodness-of-fit test, using Equation (15) (Hosking and Wallis, 2009; Malekinezhad, 2014).

$$Z^{DIST} = \frac{\tau_4^R - \tau_4^{DIST}}{\sigma_4} \quad (15)$$

Where τ_4^{DIST} denotes the fitted distribution and σ_4 represents the standard deviation of τ_4^R . Subsequently, the value of quantile at site is determined by matching all data on the selected area with the distribution, using Equation (16) (Hosking and Wallis, 2009).

$$X_i(F) = \lambda_i^{(i)} x(F) \quad (16)$$

Where, X_i indicates at-site quantile of the location, $\lambda_i^{(i)}$ denotes the average at i location, F represents the non-exceedance probability, and $x(F)$ signifies regional growth factor.

Meanwhile, the model test is using Equations (17) and (18) (Hosking and Wallis, 2009; Maidment, 1993).

$$B_i(F) = \frac{X_i(F)_{LM} - X_i(F)_M}{X_i(F)_{LM}} \times 100\% \quad (17)$$

$$R_i(F) = \left[\left(\frac{X_i(F)_{LM} - X_i(F)_M}{X_i(F)_{LM}} \right)^2 \right]^{1/2} \times 100\% \quad (18)$$

Where, $B_i(F)$ represents relative bias (R_{Bias}), $R_i(F)$ denotes relative root mean square error (R_{RSMSE}), $X_i(F)_{M,L}$ signifies at site quantile i location moment and L-moment methods, while $X_i(F)_L$ connotes at site quantile i location L-moment method.

3 RESULTS AND DISCUSSIONS

3.1 Homogeneity Test

This was conducted to determine the data uniformity in the study area. The homogeneity test's first step is to determine the return rainfall of 2.33 and 10 years, $R_{2.33}$ and R_{10} ,

followed by the ratio value of $R_{2.33}$ and R_{10} at each rain station, as well as the average ratio.

Subsequently, the average ratio is multiplied with R_{10} to discover the period year T in each rain station and plotted on a semi-log scale graph with the abscissa of the data length record and T years as ordinate, 95% confidence limit curve is drawn based on the confidence limit's value of Darlymple. Finally, visual observations were conducted made, and the data is regarded as heterogeneous in cases where some data are beyond the confidence limits.

Figure 2 provides the result of homogeneity test. According to the diagram, all rainfall stations have uniform or homogeneous data in the Mount Merapi area and are therefore suitable for analysis. However, in cases where the data is outside the confidence limit, the data is not included in the next analysis stage.

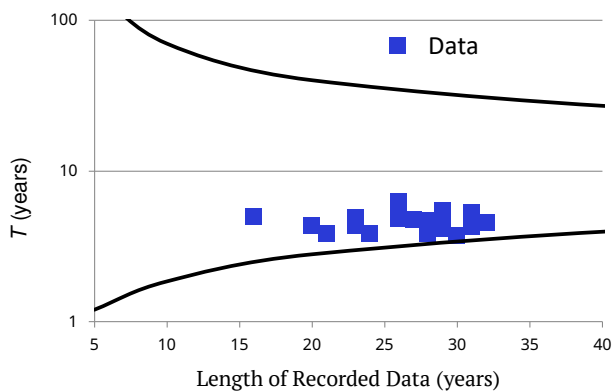


Figure 2. Homogeneity test result

3.2 Moment Method

Statistical parameters are used as the basis for determining the suitable probability distribution for the available data. In hydrological data analysis, the statistical measurement often used are the central tendency and dispersion measurements. The central tendency measure is the mean value considered as the central value and can be used to distribute measure. The dispersion measurement is the degree of variance distribution around the mean value and includes standard deviation, coefficient of variation, slope coefficient, and coefficient of

kurtosis. Standard deviation (S) is the second moment to the average value, showing data distribution. Thus, a greater standard deviation value implies the data is more scattered.

Coefficient of variation (C_v) is the comparison value between the standard deviation and average value. Meanwhile, the coefficient of skewness (C_s) is the mean value's third moment, showing the degree of distribution asymmetry form. Furthermore, the coefficient of kurtosis (C_k) is the mean value's fourth moment, measuring the distribution curve's ductility. The values of central tendency and disperse measurements are called moment parameters, as conventional statistical methods. Table 1 shows the moment parameter value at each rain station.

Table 1. Moment parameter

Rainfall Station	\bar{X}	S	C_v	C_s	C_k
Argomulyo	125.810	55.468	0.441	2.017	9.271
Babadan	103.259	27.182	0.263	0.098	3.365
Batur	97.214	30.436	0.313	0.412	2.900
Deles	99.929	27.290	0.273	0.775	3.224
Girikerto	122.570	38.128	0.311	0.529	3.427
Gn. Maron	154.350	70.335	0.456	0.943	3.678
Gulon	89.125	32.918	0.369	0.159	4.249
Jrakah	85.516	30.078	0.352	1.230	5.276
Kaliurang	105.658	31.851	0.301	0.390	2.685
Mranggen	123.594	35.129	0.284	0.273	2.881
Ngandong	145.550	31.951	0.220	0.111	2.900
Ngepos	131.232	52.936	0.403	1.661	7.088
Pakem	91.778	36.060	0.393	1.745	7.425
Plawangan	117.326	35.643	0.304	0.965	3.408
Plosokerep	87.757	42.944	0.489	1.112	3.840
Pucanganom	63.504	30.349	0.478	1.475	5.316
Randugunting	84.442	43.929	0.520	1.339	4.924
Sopalan	106.638	43.189	0.405	1.421	5.251
Sukorini	61.984	35.816	0.578	0.383	2.965
Talun	107.003	39.487	0.369	0.539	3.538

The moment parameter values obtained vary, with an average of 104.500 mm. Meanwhile, the average standard deviation, C_v , C_s , and C_k values were 39.502, 0.389, 0.885, and 4.736.

A study by Luknanto (2019) showed the moment method obtained using frequency analysis program. Based on the frequency analysis at each rainfall stations, 9 rainfall stations use Log

Pearson III Distribution, 4 use Log Normal Distribution, 4 use Gumbel Distribution, while 3 use Normal Distribution. The moment method design rainfall was obtained by return period of 2, 5, 10, 20, 50, and 100 years, based on the selected distribution at each rainfall stations. Figure 3 shows the moment method design rainfall plotted.

Figure 3. Moment method design rainfall

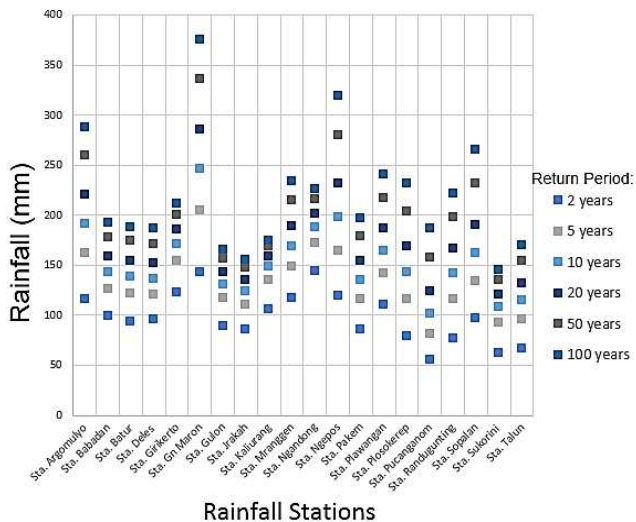


Figure 3 shows a higher return year implies higher design rainfall in moment method. The design rainfall at Argomulyo, Gn. Maron and Ngepos stations are higher, compared to other stations, because these stations have the highest standard deviation values in moment parameter. In addition, the standard deviation value of Argomulyo, Gn. Maron and Ngepos were 55.468, 70.335, and 52.936, respectively. These values show the data distribution, and a large standard deviation indicates the three station's maximum rainfall data is spread over in a range.

3.3 L-Moment Method

Regional frequency analysis using moment L-method's steps comprises parameters and ratio value determination, data filtering using discordance test, indicating data uniformity with heterogeneity test, selecting probability distribution using goodness-of-fit test, growth

factor value determination, and design rainfall prediction. The L-moment parameter determination values are based on the probability-weighted moment (PWM) theory and are be calculated using equations formulated by Hosking and Wallys (2009). Table 2 shows the L-moment parameter values obtained, then used to determine the L-moment ratio value, for use in the regional probability distribution selection.

According to Table 2, L-moment parameter produced average \bar{x}_1 , \bar{x}_2 , \bar{x}_3 , and \bar{x}_4 values of 105.685, 1.012, 3.539, 3.517. Also, the maximum parameter value at Gn. Maron is 154.350, and the minimum at Puncanganom station is 63.504. Meanwhile, the maximum value of \bar{x}_2 at Gn. Maron station is 39.329 350, and the minimum at Deles station is 15.297. Furthermore, the maximum value of \bar{x}_3 at Gn. Maron is 9.038 while the minimum \bar{x}_3 at Kaliurang station is -1.850.

Subsequently, the L-moment ratio was used to obtain the regional ratio value by calculating L-moment ratio value with each rainfall station formula's data amount, thus, gives regional values of $\tau_2^R = 0.203$, $\tau_3^R = 0.166$, and $\tau_4^R = 0.169$. In regional frequency analysis, the data is also checked to present the field's size, all data gives the same probability distribution results.

Meanwhile, the discordance test aims to filter out whether the data differs from the entire data in one region, based on the difference between the ratio of one data to the regional counterpart. Thus, a rainfall station location is considered discordant if a D_i value above the $D_{i-critic}$ value is produced. The calculations at each location and the value in criticism, based on the number of locations, with values previously determined by Hosking and Wallis (2009).

In the analysis, there are 20 different locations, thus, the critical value is 3. Table 2 shows all values are $D_i < 3$, therefore, no rainfall station is excluded in further analyses.

Table 2. Determination using L-Moment method.

Rainfall Station	Data Amount	PWM			L-Moment Parameter				L-Moment Ratio			D_i
		β_0	β_1	β_2	λ_1	λ_2	λ_3	λ_4	τ_2	τ_3	τ_4	
Argomulyo	29	125.810	76.889	57.138	125.810	27.967	7.308	8.172	0.222	0.261	0.292	1.172
Babadan	29	103.259	59.347	42.123	103.259	15.435	-0.082	2.676	0.149	-0.005	0.173	0.715
Batur	28	97.214	57.311	41.321	97.214	17.408	1.274	1.432	0.179	0.073	0.082	0.597
Deles	31	99.929	57.613	41.434	99.929	15.297	2.857	2.142	0.153	0.187	0.140	1.072
Girikerto	27	122.570	72.017	52.078	122.570	21.463	2.939	3.689	0.175	0.137	0.172	0.171
Gn. Maron	24	154.350	96.839	72.621	154.350	39.329	9.038	4.086	0.255	0.230	0.104	0.928
Gulon	20	89.125	53.672	38.627	89.125	18.219	-1.145	4.124	0.204	-0.063	0.226	2.338
Jrakah	32	85.516	50.822	37.093	85.516	16.129	3.138	3.930	0.189	0.195	0.244	0.499
Kaliurang	26	105.658	61.997	44.079	105.658	18.336	-1.850	1.855	0.174	-0.101	0.101	1.339
Mranggen	16	123.594	72.056	51.274	123.594	20.519	-1.101	1.907	0.166	-0.054	0.093	0.942
Ngandong	26	145.550	82.039	57.138	145.550	18.528	0.223	1.942	0.127	0.012	0.105	1.316
Ngepos	28	131.232	79.444	58.898	131.232	27.655	7.955	5.417	0.211	0.288	0.196	0.383
Pakem	23	91.778	55.093	40.618	91.778	18.409	4.927	6.275	0.201	0.268	0.341	2.229
Plawangan	23	117.326	68.507	49.840	117.326	19.688	5.322	2.501	0.168	0.270	0.127	1.571
Plosokerep	30	87.757	55.611	42.031	87.757	23.464	6.282	3.575	0.267	0.268	0.152	0.883
Pucanganom	28	63.504	39.756	30.091	63.504	16.008	5.516	2.492	0.252	0.345	0.156	0.866
Randugunting	33	84.442	54.025	41.080	84.442	23.607	6.772	3.008	0.280	0.287	0.127	1.381
Sopalan	29	106.638	64.739	48.081	106.638	22.839	6.694	4.766	0.214	0.293	0.209	0.409
Sukorini	21	71.400	44.593	33.061	71.400	17.786	2.206	1.951	0.249	0.124	0.110	0.967
Talun	31	107.003	64.579	47.164	107.003	22.154	2.511	4.392	0.207	0.113	0.198	0.224

The heterogeneity test aims to determine the non-uniformity, based on the difference between the ratios in each location and the area's counterpart. Meanwhile, the Kappa distribution method is used to determine the location parameters (ξ), scale parameters (α) and shape parameters (h, k). The coefficient was then determined using the Karian and Dudewicz (2010) formula, at study area regional ratio by interpolation, and values of $A = 0.815753$, $B = 0.428046$, $h = -0.55604$, and $k = 0.11634$, were obtained. In addition, the coefficient was calculated using the heterogeneity test formula by Hosking and Wallis (2009), as shown below.

$$t^R = \frac{\sum_{i=1}^N n_i t^{(i)}}{\sum_{i=1}^N n_i} = \frac{108.525}{534} = 0.203$$

$$V = \left\{ \sum_{i=1}^N n_i (t^{(i)} - t^R)^2 / \sum_{i=1}^N n_i \right\}^{1/2} = \left\{ \frac{0.937}{534} \right\}^{1/2} = 0.042$$

$$H = \frac{(V - \mu_v)}{\sigma_v} = \frac{0.042 - (-0.5247)}{1.2259} = 0.46$$

Based on the calculation, $H = 0.46$, meaning $H < 1$, thus, all rain station locations are homogeneous and therefore, no rainfall station is excluded.

Distribution selection using the L-moment method shows each distribution's theoretical L -

C_s and $L-C_k$ parameters, with the observed data. The L-moment diagram is used to determine the relationship of data L-moment ratio to the L-moment ratio of several distributions. In the diagram, the *cdf* function was depicted on ten probability distributions, Uniform, Exponential, Logistic, Normal, Log-Normal, Generalized Pareto (GPA), Generalized Logistic (GLO), Generalized Extreme Value (GEV), Gumbel, and Pearson 3. The probability distribution is based on the *cdf* value with parameters previously formulated by Hosking, Wallis (2009) and Maidment (1993). Furthermore, the L-moment diagram was identified using τ_3^R and τ_4^R , as a weighted average plotted on diagram. The L-moment diagram is plotted in a bid to determine the trend of average weighted data analysis towards probability distributions. Figure 4 shows the average weighted is close to the GEV and GLO distribution curves. The GEV and GLO distributions are to be included in the goodness-of-fit analysis.

Visual observation through the diagram must be tested with the goodness-of-fit test, an adaptation from L-moment statistics and Z probability (Z value is $|Z^{DIST}| \leq 1,6$). Data determination data with the selected probability distribution from Hosking and

Wallis formula obtained results of $|Z^{GEV}| = 0.295$ and $|Z^{GLO}| = 0.294$. The GLO distribution Z value is below the GEV distribution, therefore, the selected distribution for Mount Merapi area is Generalized Logistic (GLO). Furthermore, the regional equation is determined in L-moment parameter and cdf inverse for GLO probability distribution (Hosking, 2009).

$$x(F) = \xi + \frac{\alpha}{k} \left[1 - \left(\frac{1-F}{F} \right)^k \right]$$

$$k = -\tau_3^R$$

$$\alpha = \frac{\lambda_2 \sin k\pi}{k\pi}$$

$$\xi = \lambda_1 - \alpha \left(\frac{1}{k} - \frac{\pi}{\sin k\pi} \right)$$

Thus, the regional probability distribution equation below is obtained.

$$x(F) = 1.063 + \frac{0.194}{-0.166} \left[1 - \left(\frac{1-F}{F} \right)^{-0.166} \right]$$

Where, $x(F)$ represents the design rainfall's growth factor and F is probability of return period T . Based on the regional distribution equation, the growth factor value is determined at each return period. Table 3 shows the growth factor at the return periods of 2, 5, 10, 20, 50, and 100 years.

Figure 5 shows the design rainfall analysis at each rainfall station in Mount Merapi area is obtained by multiplying parameter L-moment λ_1 and each growth factor in every return period 2, 5, 20, 10, 50, and 100 years.

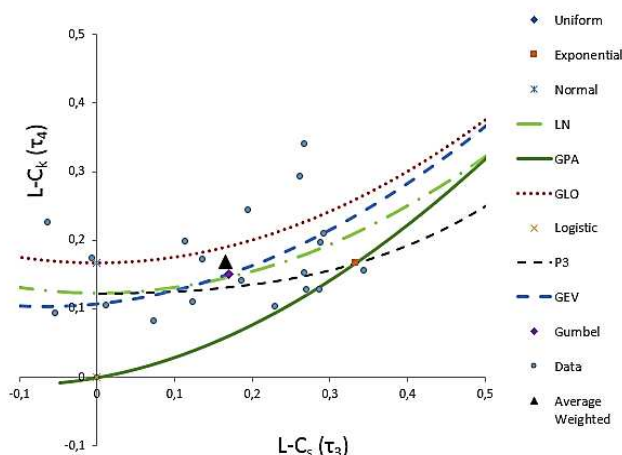


Figure 4. Plotting L-Moment diagram

Table 3. Growth factor values

Return Periods T (years)	1-F (1 - 1/T)	Growth Factor
2	0.50	1.063
5	0.80	1.365
10	0.90	1.577
20	0.95	1.800
50	0.98	2.124
100	0.99	2.400

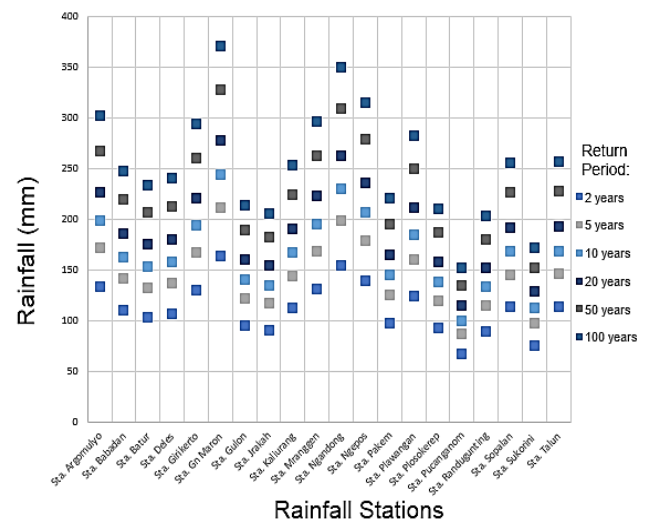


Figure 5. Design rainfall of L-Moment method

According to Figure 5, a higher return year leads to a higher design rainfall in L-moment method. As seen in the chart, several rainfall stations including as Argomulyo, Gn. Maron, Ngepos and Ngandong rainfall station have higher design rainfall higher, compared to other stations in Mount Merapi, because in these stations, parameter L-moment λ_1 values are higher, compared to others.

A comparison of regional and conventional method design rainfalls aims to determine trend of rainfall data. Figure 6 shows L-moment and moment method design rainfalls, as rainfall regional (R_{regional}) in ordinate, and rainfall at-site ($R_{\text{at-site}}$) in abscissa, respectively.

Based on the plotting data, trend data is very dense in the low return period but spread out in the high return period. This is due to data deviation in several rainfall stations with standard deviation value or quite high λ_1 parameter. The rainfall data distribution occurs

due to variation in historical data on several rainfall stations.

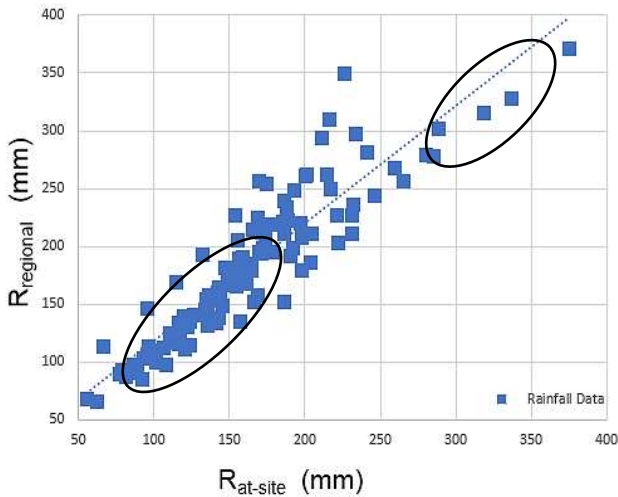


Figure 6. Relationship of R_{regional} and $R_{\text{at-site}}$

3.4 Model Test

Model test is performed by comparing the results of design rainfall at all rainfall stations for each return period with relative bias (R_{Bias}) and relative root mean square error (R_{RMSE}), using L-moment as reference. For instance, for Argomulyo station at a 2-year return period.

$$R_{\text{Bias}} = B_i(F) = \frac{X_i(F)_{\text{LM}} - X_i(F)_M}{X_i(F)_{\text{LM}}} \times 100\%$$

$$RBias = \frac{134 \text{ mm} - 116 \text{ mm}}{132 \text{ mm}} \times 100\% = 12,87\%$$

$$R_{\text{RSME}} = R_i(F) = \left[\left(\frac{X_i(F)_{\text{LM}} - X_i(F)_M}{X_i(F)_{\text{LM}}} \right)^2 \right]^{1/2} \times 100\%$$

$$R_{\text{RSME}} = \left[\left(\frac{134 \text{ mm} - 116 \text{ mm}}{134 \text{ mm}} \right)^2 \right]^{1/2} \times 100\% = 14,47\%$$

Figures 7 and 8 show the results of R_{BIAS} and R_{RSME} calculation. The test models in Figures 7 and 8 showed the minimum $R_{\text{Bias}} = 0.45\%$, maximum $R_{\text{Bias}} = 41.583\%$, minimum $R_{\text{RSME}} = 0.45\%$, and maximum $R_{\text{RSME}} = 71.01\%$.

Subsequently, a model test was performed at the Sorasan Station as reference rainfall station, as this had the most comprehensive data and continuous recording rainfall, and therefore, the

most stable data. The Sorasan rainfall station statistic parameters are outlined below.

$$\begin{aligned} \bar{X} &= 85,689 \text{ mm} \\ S &= 46,218 \text{ mm} \\ C_v &= 0,539 \\ C_s &= 2,227 \\ C_k &= 10,33 \end{aligned}$$

The Sorasan rainfall station has a Gumbel distribution, thus, the design rainfall for return periods of 2, 5, 10, 20, 50, and 100 years are 78, 119, 146, 172, 205, and 231 mm, respectively.

Figure 9 shows the design rainfall reference station, compared to regional design rainfall, calculated by multiplying design rainfall reference station with growth factor. Subsequently, the error was calculated with Sorasan station as design rainfall reference. For instance, in a 2-year return period, as shown below.

$$R_{\text{Sorasan}} = 78 \text{ mm}$$

$$\text{Growth factor return period 2 year} = 1,063$$

$$R_{\text{Regional L-moment}} = 78 \text{ mm} \times 1,063 = 91 \text{ mm}$$

$$|\text{Error}| = \left(\frac{91 \text{ mm} - 78 \text{ mm}}{78 \text{ mm}} \right) \times 100\% = 16,60\%$$

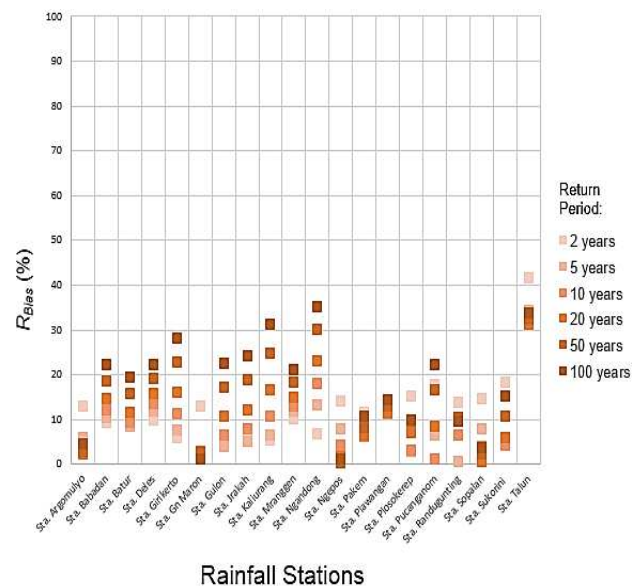


Figure 7. R_{Bias} moment relative to L-Moment

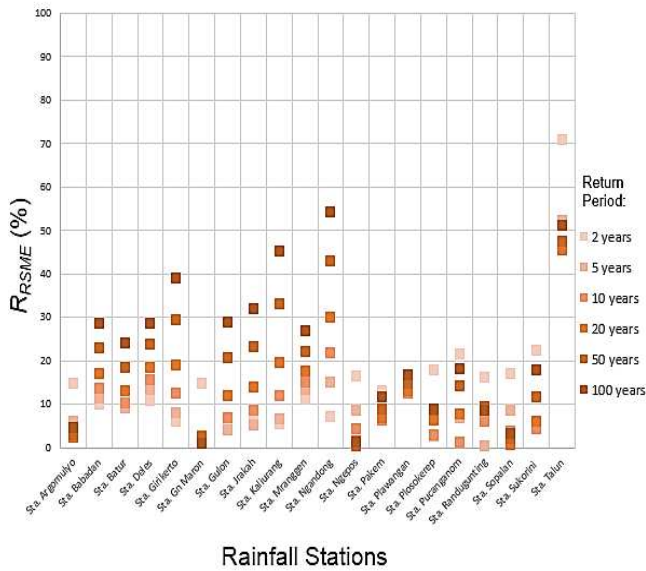


Figure 8. R_{RSME} moment relative to L-Moment

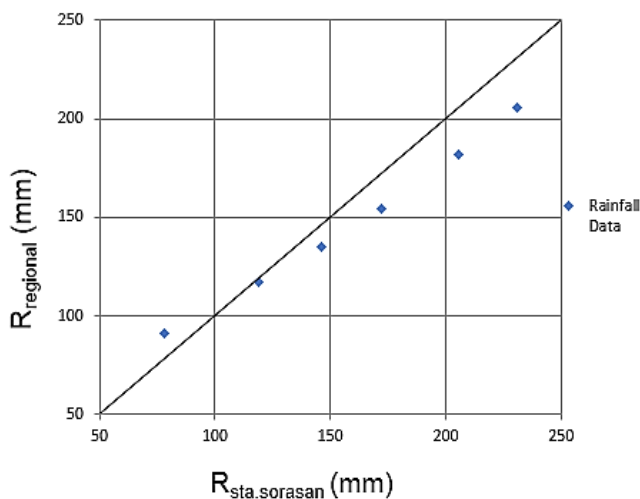


Figure 9. Model test of regional to rat-site

According to Figure 9, the calculation results for model test minimum and maximum errors are 1.64% and 16.60%, respectively. Based on the chart, a higher error value occurs at low return period and at an extremely high return period. In this case, the regional design rainfall L-moment method was able to reduce data bias. However, the method requires development to enable analysis in high return period.

4 CONCLUSIONS

The homogeneity and heterogeneity test shows all rainfall stations in the study area are uniform or homogeneous. Meanwhile, the discordance

test shows each location in the Mount Merapi area is below the critical value; thus, no data was excluded in further analysis. Parameters in moment method and L-moment method need to be able to determine appropriated probability distribution and design rainfall at return periods. The L-moment regional ratio in Mount Merapi area is $\tau_2^R = 0.203$, $\tau_3^R = 0.166$, and $\tau_4^R = 0.169$, and these values were used to determine the average weighted on L-moment diagram.

In addition, the regional probability distribution using a suitable L-moment method for the study area is Generalized Logistic (GLO) distribution, and the equation is

$$x(F) = 1.063 + \frac{0.194}{-0.166} \left[1 - \left(\frac{1-F}{F} \right)^{-0.166} \right].$$

Where, $x(F)$ and F represent design rainfall and return period probability, respectively. The higher standard deviation value in rainfall station data led to higher design rainfall because distribution data is quite scattered in historical rainfall data records. The test model showed the minimum $R_{Bias} = 0.45\%$, maximum $R_{Bias} = 41.583\%$, minimum $R_{RSME} = 0.45\%$, and maximum $R_{RSME} = 71.01\%$. Meanwhile, the L-moment method's stability showed the model test minimum and maximum errors were 1.64% and 16.60%, respectively. The L-moment method can reduce data bias, however, the higher error value in higher return periods indicates the method is currently inadequate for higher return period.

DISCLAIMER

The authors declare no conflict of interest.

AVAILABILITY OF DATA AND MATERIALS

All data are available from the author.

ACKNOWLEDGEMENTS

This study was supported by research fund from The Indonesian Ministry of Research and Technology, the scheme Beginner Lecturer Research Grant No. B/87/E3/RA.00/2020. The authors are grateful to Institut Teknologi PLN

and to the Balai Sabo Yogyakarta, as well as TechnoGIS, for supporting this study.

REFERENCES

- Chow, V.T., 2010. *Applied hydrology*, Tata McGraw Hill Education, New York.
- Darlymple, T., 1960. *Flood-frequency analysis*, Geological Survey Water-Supply, pp.1543-A.
- Haan, C.T., 2002. *Statistical method of hydrology*, Wiley, New Jersey.
- Harto, S., 2009. *Hidrologi: teori, masalah, penyelesaian*, Nafiri, Yogyakarta.
- Hasby, F., 2014. Analisis frekuensi regional hujan maksimum di DAS Ciujung dan DAS Cidurian menggunakan metode L-moment, *tesis*, Program Pascasarjana Fakultas Teknik UGM, Yogyakarta.
- Hosking, J.R.M. and Wallis, J.R., 2009. *Regional frequency analysis an approach based on L-moments*, Cambridge University Press, New York.
- Karian, Z.A., and Dudewicz, E.J., 2010. *Handbook of fitting statistical distribution with R*, CRC Press, New York.
- Kottegoda, N.T., and Rosso, R., 2009. *Applied statistics for civil and environmental engineers*, Wiley.
- Luknanto, D., *Analisis frekuensi untuk sumber daya air*, <http://luk.staff.ugm.ac.id/gapai/sw/anafrek/index.html>, access 12th April 2019.
- Maidment, D.R., 1993. *Handbook of hydrology*, McGraw-Hill Inc., New York.
- Maity, R., 2018. *Statistical methods in hydrology and hydroclimatology*, Springer Transactions in Civil and Environmental Engineering.
- Malekinezhad, H., and Garizi A.Z., 2014. *Regional frequency analysis of daily rainfall extremes using L-Moments approach*, *Atmosfera* 27(4), pp411-427.
- Naghattini, M., 2016. *Fundamental of statistical hydrology*, Springer.
- Soewarno, 1995. *Hidrologi – aplikasi metode statistik untuk analisa data*, Nova, Bandung.
- Takeleb, A.M., 2010. *Analisis hujan harian maksimum regional menggunakan metode L-Moment*, *Tesis*, FT UGM, Yogyakarta.
- Vogel, R.M., and Fennesey, N.M., 1993. *L-Moment diagrams should replace product moment diagrams*, *Water Resources Research*, 33(12), pp.2841-48.

[This page is intentionally left blank]