Temporal Probability Analysis of Flood Occurrence Using Peak Over Threshold Method for Extreme Rainfall Events

Pratnya Paramitha Oktaviana[1,2], *Arie Dipareza* Syafei[1], *Heri* Kuswanto[3] and *Joni* Hermana[1*]

[1]Department of Environmental Engineering, Faculty of Civil Planning and Geo Engineering, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia
[2]Department of Actuarial Science, Faculty of Science and Data Analytics, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia
[3]Department of Statistics, Faculty of Science and Data Analytics, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia

Abstract. Extreme rainfall often leads to flooding in various regions. This study conducts the analysis of temporal probability of flood risk due to extreme rainfall in Jakarta by using data from meteorological stations in Kemayoran (for Central Jakarta) and Tanjung Priok (for North Jakarta) over 30 years (1994-2023). Peak Over Threshold (POT) approach was applied to identify extreme rainfall events. Mean Residual Life Plot (MRLP) was used to determine thresholds for extreme rainfall, resulting in thresholds of 122 mm/day for Central Jakarta and 126 mm/day for North Jakarta. Using Generalized Pareto Distribution (GPD) modeling, the probability of flooding due to extreme rainfall exceeding these thresholds was calculated for various time frames. In Central Jakarta, the probability of flooding due to extreme rainfall exceeding 122 mm/day is 32.10% over 5 years, increasing to 97.90% over 50 years. In North Jakarta, the probability of flooding due to extreme rainfall exceeding 126 mm/day is 21.80% over 5 years and rises to 91.50% over 50 years. Extreme rainfall tends to result in higher probabilities of flooding over the long term. The results suggest that while extreme rainfall events are rare, it can cause the probability of flood occurrence to increase significantly over time.

1 Introduction

Indonesia is geographically prone to natural disasters. Climate change that frequently occurs makes the risk of natural disasters in Indonesia even higher. Shafqat et al. [1] shows a map of climate change vulnerability in the world and Indonesia is one of the countries in the Asian continent that is vulnerable to climate change and vulnerable to changes in environmental conditions due to the lack of preparedness and technology to deal with climate change. Indonesia is one of the countries participating in the Sustainable Development Goals (SDGs), a program that has 17 goals set by the United Nations (UN) in September 2015 [2]. One of

* Corresponding author: hermana@its.ac.id

© The Authors, published by EDP Sciences. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

the goals of the 17 SDGs goals, namely goal number 13, is Climate Action. Handling climate change and disaster management is very important to increase resilience and reduce the vulnerability of communities and several sectors due to the negative impacts of climate change.

The La Nina climate change indicator, namely the decrease/cooling of sea surface temperatures, can cause increased rainfall in Indonesia. One of the natural disasters that often occurs due to high rainfall is flooding. The phenomenon of flooding is one of the natural disasters that often occurs in Indonesia. Data on the frequency of flooding in the Indonesian Disaster Data Book-National Disaster Management Agency (BNPB) in the last 5 years (2019-2023) are 784, 1,518, 1,794, 1,531, and 1,255 incidents, respectively [3]. This shows that the frequency of flooding in Indonesia is very high. BNPB collected that flooding has the highest frequency of occurrence of other disasters over the past 20 years. Every year various regions in Indonesia experience damage and losses due to flooding, one of which is Jakarta. In March and April 2024, several areas in Jakarta were flooded and the Jakarta Regional Disaster Management Agency (BPBD) stated that this flooding event was caused by the intensity of rainfall, which was included in the extreme category, the highest rainfall recorded was 137.4 mm [4,5]. The high potential for flooding in DKI Jakarta results in the importance of disaster management and disaster mitigation efforts or mitigation efforts before a disaster occurs [6]. In 2020, the Jakarta area experienced a significant increase in flooded areas, reaching three times above normal [7]. In addition, Jakarta's rainfall was recorded at 377 mm/day, the highest in 24 years. This was the cause of the significant increase in flooding that year.

Previous studies on extreme rainfall in Indonesia, especially Jakarta, include: Hanugraheni and Iriawan [8], conducted hierarchical Bayesian modeling of extreme rainfall data in Jakarta. This study found that the characteristics of rainfall in Jakarta have a heavy tail pattern indicating extreme rainfall. Kiki and Wirahma [9] conducted an analysis of heavy rain in Jakarta on September 27, 2017, using the atmospheric dynamics analysis method and weather satellite imagery analysis. The results of the analysis showed that based on the recapitulation data of extreme weather events from the BMKG Jakarta Public Meteorology Center, the extreme rain event in the Jakarta area on September 27, 2017, was the second event throughout the month. Jannah et al. [10] conducted a study to understand the characteristics of rainfall in Jakarta at present and its changes in the future using the trend analysis method and Statistical Downscaling Model. The results of the study showed that rainfall in Jakarta has characteristics that are increasing in the future period for both annual rainfall parameters, maximum rainfall and number of rainy days. Several previous studies have shown that rainfall in Jakarta is classified as extreme rainfall.

Previous studies on the relationship between extreme rainfall and flooding in Jakarta include: Stephanie, Jimawan and Jayadi [11], studied the effect of rainfall on flooding in Jakarta using the time series method and the Gumbel method. This study produced a relationship between maximum daily rainfall and the recurrence period so that high rainfall will cause flooding that is almost evenly distributed in Jakarta. Then Damanik et al. [12] conducted a study on the evaluation of the rainfall index in flood conditions in Jakarta using Rclimdex. The study found that one of the factors causing flooding in Jakarta is extreme rainfall. The number of flood events in downstream Jakarta correlates more highly with the rainfall index than in upstream and central areas. Pratama and Boer [13] conducted an analysis of extreme rainfall events and their impact on the community in Cilandak and Jatinegara Villages, Jakarta, using the regression method. The results showed that the frequency of flood events in both areas was significantly related to extreme climate events. This shows that the risk of flooding in both areas tends to increase. The impact of flooding on the people of Jakarta is very large due to the damage to facilities owned, the emergence of health problems and disruption of work hours and causing quite large losses. The results of several previous studies indicate that flooding in Jakarta has a close relationship with extreme rainfall.

Based on the discussion, this study was conducted to analyze the temporal probability of flood risk due to extreme rainfall by using data from meteorological stations in Jakarta. Extreme Value Theory (EVT) method, especially Peak Over Threshold (POT) approach, was used to analyze extreme rainfall data. EVT is a statistical method that studies the behavior of the tail of a distribution of data containing extreme values and POT approach used a threshold to define those extreme values. The map of geographical context of this research is provided in Figure 1.

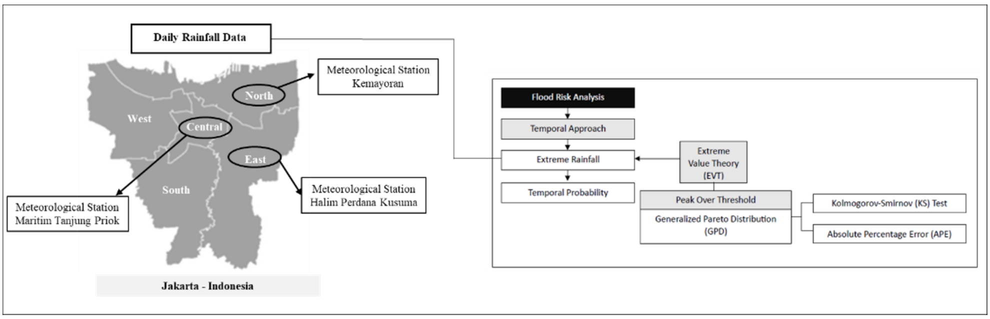

Fig. 1. The Map of Geographical Context of The Research

The results of analysis show that the probability of flooding due to extreme rainfall is high over long periods. It indicates that rainfall with an intensity of more than 122 mm/day has a more significant impact of flooding events in Central Jakarta, while rainfall with an intensity higher than 126 mm/day has considerable impact on flooding events in North Jakarta.

2 Literature Review

2.1 Extreme Value Theory (EVT)

Extreme Value Theory (EVT) is a statistical approach that is designed to analyze the behavior of the distribution's tail based on data characterized extreme values and it is commonly employed to assess the extreme behavior of environmental phenomena [14]. EVT can be employed to estimate rare event's probability. For example, it facilitates the estimation of return levels for return periods exceeding the duration of available observations [15]. Peaks Over Threshold (POT) is an EVT approach used by considering a predefined threshold or reference value by using Generalized Pareto Distribution (GPD). The GPD is well-suited to characterize the tail behavior of distribution particularly for extreme values that fall above the threshold. The values that exceed this threshold are the extreme values [16;17]. The GPD has the Cumulative Distribution Function that can be write as follows [18]:

$$F(x) = 1 - \left[1 + \xi\left(\frac{x-u}{\sigma}\right)\right]^{-1/\xi} \quad (1)$$

The Maximum Likelihood Estimation (MLE) is used to estimate the scale (σ) and shape (ξ) parameters of GPD [19]. These parameters define how extreme events evolve with an increasing return period [18]. The values exceeding the threshold u are effectively represented by GPD. The method for determining the threshold using graphs can be done in various ways, one of which is the Mean Residual Life Plot (MRLP) [20;21]. In general, if values exceeding the threshold are found and form or approximate a straight line, it can be chosen as the optimal MRLP. In MRLP, determining whether a line appears straight or not is highly subjective, leading to some degree of arbitrariness in threshold selection. To identify a more representative threshold, some criteria can be utilized to refine the threshold determination process [22]. For large values of x above a threshold u, the distribution can be approximated by the GPD, and its properties can be obtained as follows:

$$E(X-u|X>u)=\frac{\sigma-\xi u}{1-\xi}+\frac{\xi}{1-\xi}u \quad (2)$$

that is the expected value of the data exceeding threshold u. When the shape (ξ) parameter of GPD reaches a stable value, the resulting plot form a straight line. Specifically, when the threshold u is defined on the horizontal axis and the average value on the vertical axis, the line's slope and intercept are determined by the properties of the distribution in Equation 2, reflecting the relationship between threshold and extreme values. The core principle of MRLP is to achieve stability in parameter estimation by fitting the GPD within a specific threshold range [23].

2.2 Kolmogorov-Smirnov (KS) Test

The goodness-of-fit test can be conducted using the Kolmogorov-Smirnov test. This test aims to examine the extent to which a distribution fits a specific distribution. The hypotheses used in the Kolmogorov-Smirnov test can be formulated as follows:
H_0: The data distribution is the same as a specific distribution,
H_1: The data distribution is not the same as a specific distribution

In this context, the specific distribution is usually a theoretical distribution expected to fit the observed data. This test provides information about how well the observed data fits the expected distribution. For a set of N ordered data points $X_1, X_2, \dots, X_N$, the empirical distribution function is defined as:

$$E_N=n(i)/N \quad (3)$$

$n(i)$ represents the number of data points smaller than X_i, where the X_i values are arranged in descending order. Then in the KS test statistics, the calculation is performed using the equation as follows [24]:

$$D=\max_{1\le i\le N}\left(F(X_i)-\frac{i-1}{N},\frac{i}{N}-F(X_i)\right) \quad (4)$$

F represents the theoretical cumulative distribution of the dataset. The null hypothesis H_0 can be rejected if $D>D_\alpha$.

2.3 Mean Absolute Percentage Error (MAPE)

MAPE is a measure of relative accuracy used to determine the percentage of deviation from forecast results. MAPE is defined as the average of Absolute Percentage Error (APE). In other words, APE is a part of MAPE. The variable used in this case is rainfall. MAPE calculation uses the following equation [25]:

$$MAPE=\frac{1}{n}\sum_{i=1}^{n}APE_i \quad (5)$$

$$APE_i=\left|\left(\frac{x_i-\hat{x}_i}{x_i}\right)\times 100\%\right| \quad (6)$$

where :
x_i : actual value of highest rainfall
$\hat{x}_i$: estimated value (return period) of highest rainfall
n : number of meteorological stations

Sumari et al. [26] stated that MAPE is the average of the difference between actual data and predicted results shown in percentage, where the smaller the MAPE value, the closer the predicted results are to the original, with the classification of MAPE value interpretations listed in Table 1:

Table 1. The MAPE Value Interpretation

MAPE Value	Interpretation
< 10%	The model accuracy is very accurate
10% - 20%	The model accuracy is good
20% - 50%	The model accuracy is moderate
> 50%	The model accuracy is poor

2.4 Temporal Probability of Flooding

Temporal probability can be computed by the probability of rainfall exceeding a threshold. The probability (p) of exceeding a specified rainfall threshold in a particular year is defined as follows [27]:

$$p = 1 - F(x) \tag{7}$$

$F(x)$ is the cumulative distribution function of the rainfall, that is assumed to GPD with scale (σ) and shape (ξ) parameters. The temporal probability can be calculated using the following equation:

$$\Pr\,(x_{th}, N) = 1 - (F_{GPD}(x_{th}; \sigma, \xi))^N \tag{8}$$

where:
x_{th} : the rainfall threshold that triggers flooding
N : N year periods
F_{GPD} : CDF of Generalized Pareto Distribution (GPD)

3 Data and Methods

The data used in this study is daily rainfall data (mm/day) obtained per month from the Meteorology, Climatology, and Geophysics Agency (BMKG) of Indonesia from January 1, 1994, to December 31, 2023 (30 years: 10,957 days) for meteorological station Kemayoran - Central Jakarta, Tanjung Priok - North Jakarta, and Halim Perdana Kusuma - East Jakarta. This data collection is in accordance with the guidelines of the World Meteorological Organization (WMO) which indicate that climatology analysis is ideally carried out over a period of 30 years to obtain an accurate and stable climate representation [28], as well as reports from the Intergovernmental Panel on Climate Change (IPCC) which supports the use of a minimum period of 30 years for climate analysis to obtain long-term climate change trends [29]. The rainfall data structure is presented in Table 2.

The data analysis method used in this study is EVT analysis, including POT approach using GPD modeling and determination the threshold using MRLP. The steps of analysis are explained as follows:

1. Data preprocessing and exploration
2. Calculating the threshold value from rainfall data by using MRLP method

3. Selecting the extreme data (above the threshold) and fitting the extreme data using GPD
4. Next, the Kolmogorov Smirnov test will be used to test the suitability of the distribution. If it is suitable, the previous parameters can be used

Then the obtained parameters are used to calculate the temporal probability of flooding due to extreme rainfall in certain time periods.

Table 2. The Rainfall Data Structure

Observation	Date	Meteorological Station		
		Kemayoran	Tanjung Priok	Halim Perdana Kusuma
		X_1	X_2	X_3
1	January 1, 1994	$X_{1,1}$	$X_{1,2}$	$X_{1,3}$
2	January 2, 1994	$X_{2,1}$	$X_{2,2}$	$X_{2,3}$
	…	...	...	...
	…	...	...	...
n-1	December 30, 2023	$X_{n-1,1}$	$X_{n-1,2}$	$X_{n-1,3}$
n	December 31, 2023	$X_{n,1}$	$X_{n,2}$	$X_{n,3}$

X = rainfall (mm/day)

4 Results and Discussion

Preprocessing of rainfall data is carried out first so that the data is ready to be used for further analysis. Checking the missing and trace values (not measured) in the rainfall data is then carried out for each meteorological station. Rainfall data from BMKG is marked with '9999' for missing values (not measured) and '8888' for trace or not measured. Trace is an indication of rain with a value of less than 0.1 millimeters [30]. The Quality Control (QC) procedure applied by BMKG to validate rainfall data containing missing values, namely if the data contains more than 20% missing values, the data cannot be used [31]. To overcome missing and trace values, value imputation is carried out, with reference to trace values imputed with a value of 0 (zero) and missing values imputed with a chained equation using the 'mice' package in R software which is based on Monte Carlo Markov Chain (MCMC) simulation, where the imputation of missing values is carried out with an iteration process until the appropriate value is obtained [32]. A summary of the missing values and traces at each meteorological station based on daily rainfall data that has been collected from January 1, 2013 to December 31, 2023 is shown in Table 3:

Table 3. The Identification of Missing Values and Traces at Each Meteorological Station

Meteorological Station	City	%Trace	%Missing Value	Description
Kemayoran	Central Jakarta	6.1	3.9	Data can be used for analysis
Tanjung Priok	North Jakarta	3.6	5.9	Data can be used for analysis
Halim Perdana Kusuma	East Jakarta	2.4	66.3	Data cannot be used for analysis

Table 3 shows that of the three meteorological stations, rainfall data from the Halim Perdana Kusuma meteorological station cannot be used, because the percentage of missing values is 66.3%, so according to the BMKG QC for rainfall, the data is included in the criteria for data that cannot be used in the analysis because the % missing value is more than 20%. Therefore, the Halim Perdana Kusuma meteorological station and the city of East Jakarta were removed from the analysis. Furthermore, rainfall data at the Kemayoran and Tanjung Priok meteorological stations containing traces were imputed with a value of 0 (zero) and missing values were imputed with the average rainfall in the same month as the missing value. After data imputation, a re-check was carried out and it was ensured that there was no more missing data or traces in the rainfall data.

4.1 Descriptive Statistics Analysis of Rainfall Data

In accordance with the results of the missing value and trace checks, the rainfall data that can be continued for analysis is the daily rainfall data from the Kemayoran (for Central Jakarta) and Tanjung Priok (for North Jakarta) meteorological stations. Descriptive statistical analysis of daily rainfall data (mm/day) produces the following results:

Table 4. Descriptive Statistics of Rainfall Data (mm/day)

Deskriptive Statistics	Meteorological Stations	
	Kemayoran-Central Jakarta	Tanjung Priok-North Jakarta
Median	0.00	0.00
Mode	0.00	0.00
Minimum	0.00	0.00
Mean	5.46	5.09
Maximum	277.50	284.00
Variance	231.87	220.70

Table 4 shows that the median and minimum of the rainfall data are 0.00 mm/day. In addition, the mode (the most frequently occurring value) at both meteorological stations is also 0.00 mm/day, meaning that both Central and North Jakarta often have non-rainy (cloudy) days. However, the maximum rainfall value at both stations is very high, namely 277.50 mm/day for the Kemayoran meteorological station and 284.00 mm/day for the Tanjung Priok meteorological station, where according to BMKG, if the rainfall is > 150 mm/day then the rainfall is classified as extreme. Based on this information, it can be concluded that in the cities of Central and North Jakarta it rarely rains, but if it does rain it can be very high and even be classified as extreme in intensity. The average (mean) rainfall value is around 5 mm/day for both meteorological stations, meaning that on average there is light rain in the cities of Central and North Jakarta. The high variance value indicates that the rainfall data varies at both stations.

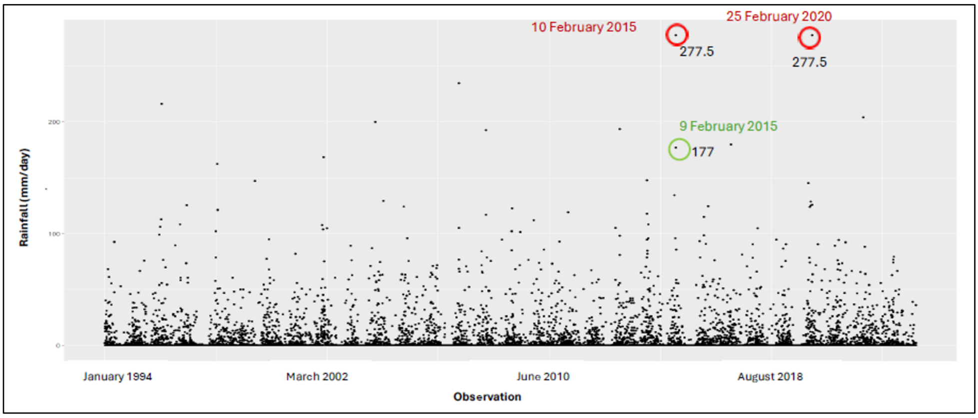

Fig. 2. Plot of Rainfall Data (mm/day) in Kemayoran Meteorological Station (Central Jakarta)

https://doi.org/10.1051/bioconf/202515711002

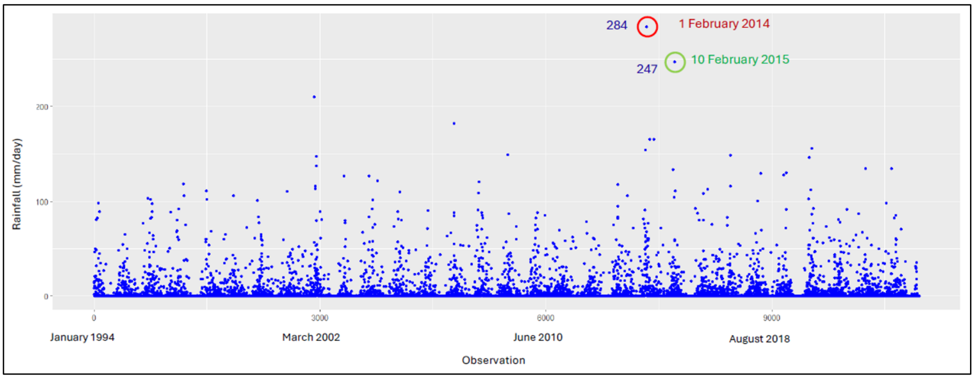

Fig. 3. Plot of Rainfall Data (mm/day) in Tanjung Priok Meteorological Station (North Jakarta)

Figure 2 and Figure 3 show the visualization of rainfall (mm/day) in Kemayoran and Tanjung Priok meteorological stations. From the figures, there were many days where there was no rain, which is indicated by many rainfall plots that are around 0.00 mm/day, this is in accordance with the results of previous descriptive statistics. Figure 2 shows that at the Kemayoran meteorological station (Central Jakarta), the highest rainfall occurred on February 10, 2015 and February 25, 2020, which was 277.50 mm/day. Meanwhile, Figure 3 shows the highest rainfall at the Tanjung Priok meteorological station (North Jakarta), which was 284.00 mm/day, occurred on February 1, 2014. Figure 3 also shows that on February 10, 2015, rainfall in North Jakarta was also monitored as high, which was 247 mm/day. This is in accordance with information from Indonesia's National Disaster Management Agency (BNPB) in the Kompas news [32] which noted that flooding occurred in Jakarta, especially the Central Jakarta area starting on February 9 2015, where this is in accordance with rainfall data from the Kemayoran Meteorological Station (Central Jakarta), namely 177 mm/day on that date, and immediately increased the following day.

4.2 Extreme Rainfall Analysis using Peak Over Threshold (POT)

In this analysis, data division is used, training data for model formation according to POT results and testing data to validate the results of the model that has been formed. Daily rainfall data for 1994-2021 is used as training data, while daily rainfall data for 2022-2023 is used as testing data. Model validation is carried out using the Mean Absolute Percentage Error (MAPE) which is calculated based on the return period results of the model and compared with the actual values of the testing data.

Extreme rainfall analysis using POT is carried out first by determining the threshold using MRLP. The threshold determination using MRLP can be seen in Figure 4. Plots that tend to form horizontal linear patterns (marked in yellow) will be considered as the selected threshold. The best threshold value is selected based on the smallest AIC value and the appropriate GPD distribution fit test results.

A summary of the threshold selection with MRLP from Figure 4 is presented in Table 5. Based on the threshold determination results in Figure 4 and Table 5, it can be concluded that the extreme rainfall threshold selected at the Kemayoran Meteorological Station - Central Jakarta and the Tanjung Priok Meteorological Station - North Jakarta are 122 mm/day and 126 mm/day, respectively.

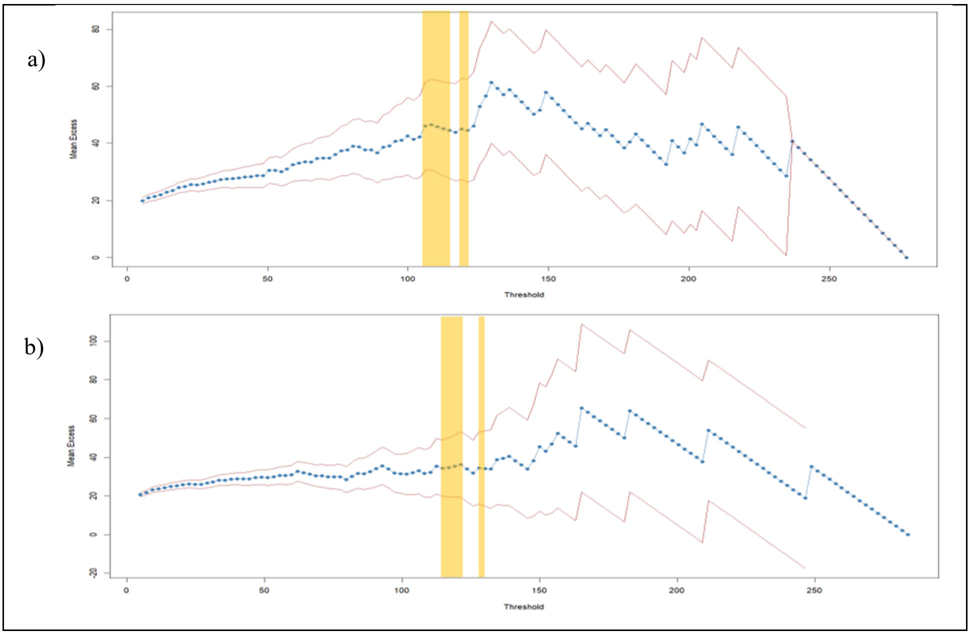

Fig. 4. Mean Residual Life Plot for Determining Threshold: a) Kemayoran Meteorological Station-Central Jakarta, b) Tanjung Priok Meteorological Station-North Jakarta

Table 5. Extreme Rainfall Threshold Selection using Mean Residual Life Plot (MRLP)

Rainfall	Threshold	The amount of data above the threshold	AIC	KS test of GPD
Kemayoran-Central Jakarta	110	31	301.77	Failed to reject H_0 (suitable)
	112	30	291.24	Failed to reject H_0 (suitable)
	114	29	280.92	Failed to reject H_0 (suitable)
	116	28	270.64	Failed to reject H_0 (suitable)
	120	25	243.24	Failed to reject H_0 (suitable)
	122	24	233.35	Failed to reject H_0 (suitable)
Tanjung Priok-North Jakarta	112	29	275.12	Failed to reject H_0 (suitable)
	114	26	250.93	Failed to reject H_0 (suitable)
	116	23	240.58	Failed to reject H_0 (suitable)
	118	23	223.63	Failed to reject H_0 (suitable)
	120	22	213.78	Failed to reject H_0 (suitable)
	126	20	191.11	Failed to reject H_0 (suitable)

Fitting the distribution of extreme rainfall was carried out using the Generalized Pareto Distribution (GPD). The results of the Generalized Pareto Distribution (GPD) distribution fitting for extreme rainfall data according to the threshold at both meteorological stations are shown in Table 6. The results show that the GPD distribution for extreme rainfall at both stations have parameters $\xi > 0$. This indicates that the probability of extreme rainfall events is still high. Meanwhile, the scale parameter (σ) shows that extreme rainfall data is quite diverse at both meteorological stations.

Table 6. GPD Parameter Estimation Results for Extreme Rainfall According to Threshold at Each Meteorological Station

Extreme Rainfall	Threshold	Parameter	Estimated Value
Kemayoran Meteorological Station-Central Jakarta	122 mm/day	ξ = *shape* parameter	0.204
		σ = *scale* parameter	35.599
Tanjung Priok Meteorological Station -North Jakarta	126 mm/day	ξ = *shape* parameter	0.310
		σ = *scale* parameter	24.199

Based on the results of the parameter estimation, the CDF for each extreme rainfall data according to the threshold at both meteorological stations can be written as follows:

a. The CDF of GPD distribution for extreme rainfall (> 122 mm/day) at Kemayoran Meteorological Station - Central Jakarta:

$$F_1(x) = 1 - \left(1 + \frac{0.204}{35.599}x\right)^{-\frac{1}{(0.204)}} \tag{9}$$

b. The CDF of GPD distribution for extreme rainfall (> 126 mm/day) at Tanjung Priok Meteorological Station - North Jakarta:

$$F_2(x) = 1 - \left(1 + \frac{0.310}{24.199}x\right)^{-\frac{1}{(0.310)}} \tag{10}$$

The calculation of return period – 2 years (the highest rainfall value expected to be exceeded in the next 2 years) compared to the actual value of the testing data (the highest rainfall value for 2 years). The results of the analysis are shown in Table 7:

Table 7. Extreme Rainfall Model Validation using APE Value

Meteorological Station	Return Period (2 years)	Actual Value (2 years)	APE	Description
Kemayoran-Central Jakarta	142.28 mm/day	204.00 mm/day	30.25%	The model accuracy is moderate
Tanjung Priok-North Jakarta	135.13 mm/day	134.70 mm/day	0.32%	The model accuracy is very accurate

The results of the model validation show that the model produced in Eq. (9) is not good but is still suitable for use based on the APE value (the APE criteria follows MAPE 20%-50% indicating that the model has moderate accuracy, meaning that the model is acceptable but needs improvement), while the model in Eq. (10) is very accurate because it has an APE value <10% and is even very close to the actual value.

The results of the calculation of the temporal probability of flooding due to extreme rainfall (exceeding the threshold) are presented in Table 8. The analysis results in Table 8 show that the probability of at least one flood event in Central Jakarta due to extreme rainfall above the threshold of 122 mm/day is predicted to increase over time. For example, the probability of a flood occurring in 1 year is 0.074, and continues to increase to 0.979 over 50 years. With a threshold of 122 mm/day, the probability of a flood occurring is quite high in a longer period, indicating that extreme rainfall with an intensity of 122 mm/day and above has a more significant impact on flood events. In North Jakarta, the probability of a flood due to extreme rainfall above 126 mm/day also increases from 0.048 in 1 year to 0.915 over 50 years. With a threshold of 126 mm/day, the probability of a flood occurring is high in the long term, indicating that rainfall with this intensity has quite an impact on flood events in

North Jakarta. The thresholds of 122 mm/day and 126 mm/day tend to provide a higher probability of flood events, especially in the long term in the Central Jakarta and North Jakarta areas.

Table 8. The Temporal Probability of Flooding Due to Extreme Rainfall

Meteorological Station	$F(x_{th})$	Threshold (x_{th})	N	Pr (x_{th}, N)
Kemayoran - Central Jakarta	$F_1(x) = 1 - \left(1 + \frac{0.204}{35.599}x\right)^{-\frac{1}{(0.204)}}$	122 mm/day	1	0.074
			2	0.143
			5	0.321
			10	0.538
			15	0.686
			20	0.787
			25	0.855
			50	0.979
Tanjung Priok - North Jakarta	$F_2(x) = 1 - \left(1 + \frac{0.310}{24.199}x\right)^{-\frac{1}{(0.310)}}$	126 mm/day	1	0.048
			2	0.094
			5	0.218
			10	0.389
			15	0.522
			20	0.626
			25	0.708
			50	0.915

5 Conclusion

The results of the temporal probability analysis of flooding due to extreme rainfall show that the temporal probability of flooding increases over time, with the following provisions: In Central Jakarta (with extreme rainfall > 122 mm/day), the probability of flooding increases from 0.074 over 1 year to 0.979 over 50 years; In North Jakarta (with extreme rainfall > 126 mm/day), the probability of flooding increases from 0.048 over 1 year to 0.915 over 50 years. The probability approaching 1 over a 50-year period indicates that flooding due to extreme rainfall is almost certain to occur in the long term in both regions. In can be conclude that EVT method is essential to risk analysis about climate change and natural disaster. This method produces good results and can be applied in other environmental fields.

The calculation of the probability of flooding is not yet complete and will be continued to the calculation of the spatial probability of flooding according to the spatial vulnerability of each area in Central and North Jakarta in the next research.

References

1. W.Shafqat, S.A. Naqvi, R. Maqbool, M.S. Haider, M.J. Jaskani and I.A. Khan, Climate Change and Citrus (IntechOpen, Rijeka, 2021)
2. Bappenas, SDGs Knowledge Hub: Agenda 2030 untuk Pembangunan Berkelanjutan. https://sdgs.bappenas.go.id/ (2023)
3. BNPB, inaRISK: Buku Risiko Bencana Indonesia. https://inarisk.bnpb.go.id/BUKU-RBI-2022/mobile/index.html#p=44 (2022)
4. T.A. Azzahra, Curah Hujan Tinggi, 66 RT di Jakarta dilanda Banjir selama 3-4 hari. https://news.detik.com/berita/d-7279677/curah-hujan-tinggi-66-rt-di-jakarta-dilanda-banjir-selama-3-4-april (2024)
5. S. Saragih, BPBD DKI: Banjir di Jakarta Dampak Hujan Ekstrem.

https://www.metrotvnews.com/read/kpLCWaJ6-bpbd-dki-banjir-di-jakarta-dampak-hujan-ekstrem (2024)
6. S. Dahlia and Fadiarman, Analisis Risiko Banjir Terhadap Fasilitas Pendidikan di DKI Jakarta. *J. Geo. Gea*, **20**(**2**), 185-196 (2020) https://doi.org/10.17509/gea.v20i2.24113
7. A. Ahdiat, Jumlah Wilayah Rukun Warga (RW) yang Terdampak Banjir di DKI Jakarta (2018-2022).https://databoks.katadata.co.id/datapublish/2023/02/27/riwayat-banjir-jakarta-5-tahun-terakhir-membaik-atau-memburuk (2023)
8. J.S.I. Hanugraheni and N.Iriawan, Pemodelan Bayesian Hirarki Data Curah Hujan Ekstrem di Jakarta. *J. Sains dan Seni ITS* **5**(**1**), 2337-3520. (2016) https://doi.org/10.12962/j23373520.v5i1.1437
9. Kiki and S. Wirahma, Analisis Hujan Lebat Tanggal 27 September 2017 di DKI Jakarta. *J. Sains & Teknologi Mod. Cuaca,* **18**(**2**), 51 – 59. (2017)
10. M. Jannah, J. Sujono and A.P. Rahardjo, Kajian Perubahan Iklim di DKI Jakarta berdasarkan data curah hujan. *Teknisia,* **28**(1), 043-054. (2023) https://doi.org/10.20885/teknisia.vol28.iss1.art5
11. Stephanie, O.N. Jimawan and D. Jayadi, Analisis Statistika Pengaruh Curah Hujan Terhadap Banjir di Jakarta Melalui Pemodelan Matematika. *J. Met. Klimatologi dan Geofisika*, **5**(**2**), 22-18. (2018)
12. A.A. Damanik, I.P. Santikayasa and M. Taufik, Evaluasi Indeks Curah Hujan pada Kondisi Banjir di Wilayah Jakarta-Indonesia. (IPB University Scientific Repository, Bogor, 2020)
13. M.F. Pratama and R. Boer, Kejadian Curah Hujan Ekstrem dan Dampaknya terhadap Masyarakat (Studi Kasus: Kelurahan Cilandak Timur dan Kelurahan Jatinegara, DKI Jakarta). (IPB University Scientific Repository, Bogor, 2024)
14. A. Lang and B. Poschlod, Updating catastrophe models to today's climate – An application of a large ensemble approach to extreme rainfall. *Climate Risk Management* (**44**) 100594. (2024) https://doi.org/10.1016/j.crm.2024.100594
15. P. Rivoire, P. Le Gall, A.C. Favre, P. Naveau and O. Martius, High return level estimates of daily ERA-5 precipitation in Europe estimated using regionalized extreme value distributions. *Weather and Climate Extremes,* **38**, 100500. (2022) https://doi.org/10.1016/j.wace.2022.100500
16. E. Afuecheta and M. H. Omar, Characterization of variability and trends in daily precipitation and temperature extremes in the Horn of Africa. *Climate Risk Management,* **32**, 100295. (2021) https://doi.org/10.1016/j.crm.2021.100295
17. B. Moccia, C. Mineo, E. Ridolfi, F. Russo and F. Napolitano, Probability distributions of daily rainfall extremes in Lazio and Sicily, Italy, and design rainfall inferences. *J. of Hydro.: Reg. Studies,* **33**, 100771. (2021) https://doi.org/10.1016/j.ejrh.2020.100771
18. M.G. Sanderson, M. Teixeira, N. Fontes, S. Silva and A. Graça, The probability of unprecedented high rainfall in wine regions of northern Portugal. *Climate Services,* **30**, 100363. (2023) https://doi.org/10.1016/j.cliser.2023.100363
19. Y. Shin and J.S. Park, Modeling climate extremes using the four-parameter kappa distribution for R-largest order statistics. *Weather and Climate Extremes, 39*, 100533. (2023) https://doi.org/10.1016/j.wace.2022.100533
20. S. Coles, An introduction to statistical modeling of extreme values. *Technometrics*, **44**(**4**), 397. (Springer, 2001)
21. Y. Karimah, A.H. Wigena and A.M. Soleh, Penentuan Nilai Ambang Batas Sebaran Pareto Terampat Dengan Measure of Surprise. *Indonesian Journal of Statistics and Its Applications*, **3**(**2**), 161–168. (2019) https://doi.org/10.29244/ijsa.v3i2.284
22. F. Mazas and L. Hamm, A multi-distribution approach to POT methods for determining extreme wave heights. *Coastal Engineering*, **58**(**5**), 385–394. (2011) https://doi.org/10.1016/j.coastaleng.2010.12.003

23. G. Liu, Z. Gao, B. Chen, H. Fu, S. Jiang, L. Wang and Y. Kou, Study on Threshold Selection Methods in Calculation of Ocean Environmental Design Parameters. *IEEE Access*, **7**, 39515-39527. (2019) https://doi.org/10.1109/ACCESS.2019.2906935
24. P.P. Oktaviana and Irhamah, Kolmogorov-Smirnov Goodness-of-Fit test for identifying distribution of the number of earthquakes and the losses due to earthquakes in Indonesia. *J. Phys.: Conf. Ser.*, **1821**, 012045. (2020) https://doi.org/10.1088/1742-6596/1821/1/012045
25. S. Kim and H. Kim, A new metric of absolute percentage error for intermittent demand forecasts. *Int. Journal of Forecasting*, 3**2**(**3**), 669-679. (2016) https://doi.org/10.1016/j.ijforecast.2015.12.003
26. A.D.W. Sumari, M.B. Musthafa, Ngatmari and D.R.H. Putra, Comparative Performance of Prediction Methods for Digital Wallet Transactions in the Pandemic. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, **4**(**4**), 642–647. (2020) https://doi.org/10.29207/resti.v4i4.2024
27. H. Kim, J.H. Lee, H.J. Park and J.H. Heo, Assessment of temporal probability for rainfall-induced landslides based on nonstationary extreme value analysis. *Engineering Geology*, **294**, 106372. (2021) https://doi.org/10.1016/j.enggeo.2021.106372
28. WMO, WMO Guidelines on the Calculation of Climate Normals. (WMO, Geneva, 2017)
29. IPCC, Synthesis Report of the Sixth Assessment Report: A Report of the Intergovernmental Panel on Climate Change. (2023) https://www.ipcc.ch/ar6-syr/
30. Karyati, M. Syafrudin, R. Mulyadi and S. Sarminah. Panduan Praktikum Agroklimatologi. (Fakultas Kehutanan, Universitas Mulawarman Samarinda, 2022)
31. P.J. Wahab, Sutikno and A. Sopaheluwakan, Homogenisasi Data Unsur Iklim Sebagai Tahap Pra-Pemrosesan Kajian Perubahan Iklim dengan Metode Acmant dan Climatol. (Jurusan Statistika, Institut Teknologi Sepuluh Nopember, Surabaya, 2016).
32. S. Van Buuren & K. Groothuis-Oudshoorn, mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software*, **45**(**3**), 1–67. (2011) https://doi.org/10.18637/jss.v045.i03