International Journal of Mathematics Trends and Technology
ISSN: 2231-5373/ https://doi.org/10.14445/22315373/IJMTT-V69I10P501

Volume 69 Issue 10, 1-9, October 2023
© 2023 Seventh Sense Research Group®

Original Article

Hourly Storm Analysis: Understanding Precipitation and Duration of Short-Term Rainfall in Alor Setar

Voni Apriana Dewi[1*], Arisman Adnan[1], Rado Yendra[2]

[1]*Department of Mathematics, University of Riau, Pekanbaru, Indonesia.*
[2]*Department of Mathematics, State Islamic University of Sultan Syarif Kasim, Pekanbaru, Indonesia.*

[*]*Corresponding Author : voniad05@gmail.com*

Received: 07 August 2023 Revised: 14 September 2023 Accepted: 01 October 2023 Published: 24 October 2023

Abstract *- Global climate change has brought about a surge in extreme weather events, with storm rainfall being one of the most impactful occurrences. Understanding and modeling storm rainfall is crucial for predicting and preparing for future extreme events. In this study, we analyzed hourly rainfall data in Alor Setar, Malaysia, from 1970 to 2008, defining storm events as consecutive periods of rainfall and no rainfall with a minimum duration and precipitation of 3 hours and 3 mm, respectively, and a time interval between storm events of at least 3 hours. We utilized the Generalized Extreme Value (GEV) and Generalized Pareto (GP) distributions with Maximum Likelihood Estimation (MLE) to model the storm rainfall data. Four models were created, and the comparison revealed that GEVt and GEVt models were the best fit for the data. Additionally, we calculated the return period values to estimate the recurrence interval of extreme storm rainfall events, providing valuable insights for better disaster preparedness*

Keywords *- Generalized Extreme Value, Generalized Pareto, Return period, Short-term rainfall, Storm analysis.*

1. Introduction

The world is currently grappling with climate change, a phenomenon that poses a significant ecological threat and jeopardizes human safety. One of the detrimental consequences of climate change is the alteration and reduction in rainfall patterns. Insufficient rainfall can lead to devastating droughts, while an excessive increase in rainfall can result in flooding disasters. The development of a short-scale rain model is imperative to comprehend rainfall patterns comprehensively and mitigate the risk of flood disasters.

Alor Setar, the capital of the state of Kedah in Malaysia, is particularly susceptible to flooding due to its geographical location. Situated at a confluence of major rivers and merely ten miles from the sea, the city rests at a mere two meters above sea level. To address these concerns, an extreme short-scale rain model is employed to predict the occurrence of extreme storm rainfall in Alor Setar, with the ultimate goal of minimizing potential damages. Various studies on extreme rain models have been conducted in Malaysia. In 2002, Zalina [15] conducted a study on extreme rainfall, specifically the maximum annual rainfall, at 13 rain stations on the Malaysian Peninsula using a 20-year period of rainfall data (1975–1995). They employed several extreme models, such as the Generalized Extreme Value (GEV) and the Log-Normal 3 parameters (LN3). The study concluded that the GEV model was the most accurate in describing the frequency patterns. In 2009, Wan et al. [13] utilized an extreme probability distribution model, which included the Generalized Extreme Value (GEV) and the generalized Pareto (GP) distribution, to analyze rainfall occurring on the Malaysian Peninsula. Their study concluded that the GP model is the most suitable. Also in 2009, Shabri and Arrif [11] demonstrated that the Generalized Logistic model was the most appropriate choice to describe the frequency of extreme rainfall occurrences in Selangor.

In this research, a modeling of daily rainfall data spanning 38 years in Alor Setar, Malaysia, is conducted. The study employs storm analysis techniques, which are crucial for categorizing extreme rainfall data, with the aim of uncovering valuable insights into the nature and behavior of rainfall patterns in the region. To achieve this objective, the research utilizes the Generalized Extreme Value (GEV) and generalized Pareto (GP) distributions, both renowned for their effectiveness in modeling extreme weather events. Through the application of these methodologies, the researchers aim to unravel the intricacies of short-term rainfall in Alor Setar and make a significant contribution to the broader understanding of its hydrological dynamics.

Fig. 1 Alor setar

To comprehensively understand the modeling of hourly rainfall data, a selection of research papers is referenced. In 2015, Dyrrdal et al. [5] conducted a study developing a spatial model for extreme precipitation in Norway, particularly essential for managing the risks associated with flooding and infrastructure planning in regions with hydrological power facilities. The study employs Bayesian hierarchical modeling, generalized extreme value distributions, and Gaussian fields to account for spatial variability and uncertainty in extreme precipitation estimates. Additionally, in 2015 Shaffie et al [14] investigated the spatial distribution of extreme rainfall in Peninsular Malaysia at various hourly intervals, highlighting the impact of short-duration extreme rainfall events compared to daily data.

It identifies the most appropriate probability distribution and emphasizes the need for planning measures to mitigate disasters associated with shorter-duration extreme rainfall events. Furthermore, the research Yeo et al. [12] in 2020 compare three fitting methods for characterizing extreme rainfall and constructing confidence intervals for IDF (Intensity-Duration-Frequency) curves using the Scaling-GEV distribution model, demonstrating its feasibility and accuracy in analyzing sub-daily rainfall data in Dorval, Quebec (Canada), and Seoul (South Korea). These selected references collectively provide the research with a solid foundation to build upon existing knowledge and contribute to a more comprehensive understanding of short-term rainfall modeling in Alor Setar, Malaysia. However, it's worth noting that none of the studies mentioned above utilized storm event analysis, setting this research apart in its methodology.

In contrast to the previous study, this research adopts a different approach by employing storm event analysis with distinct criteria. Specifically, this study focuses on identifying storm events based on the minimum requirements of a 3-hour duration and 3 millimeters of precipitation. Moreover, an additional criterion is set to ensure that there is a minimum time interval of at least 3 hours between consecutive storm events. This methodology divergence implies a more stringent selection process for storm event identification, as it necessitates a more prolonged duration and a higher threshold for precipitation intensity. As a result, the storm events selected for analysis in this study are expected to be of a more pronounced and potentially severe nature, offering a different perspective on extreme weather events and their characteristics compared to the previous research. This modified approach is crucial for gaining insights into the dynamics of more intense storm events and their implications in various research applications.

2. Storm Event Analysis

Storm Event Analysis is a measurement of the external rainfall of a storm precipitation event adjusted to its probability density function . In storm analysis, the definition of storm incidents will be created based on specific features. There is not a specific description of storm events. to study and define the actual storm event, separate events between storm events must be identified first using the time model of the interval between events. The interval time is the time between the event of no rain and rain. There are several interval time models between events that have been used in some storm rain studies, such as Inter-Event Time Duratin (IETD) and Minimum Inter-event Time (MIET). IED model defined a storm event is a consecutive hourly rainfall preceded and followed by a minimum period of time greater than or equal to 6 hours. The six-hour time is a benchmark based on frequent storm events that have a concentration time of less than six hours. While on the MIET model, the interval time used is the minimum interval time between rainfall per hour of 0-6 hours. Palynchuk [11] defines storm occurrences as rain circumstances with a time restriction between 6 hours of rainfall with a minimum storm width of 2.54 mm.

In this study, hourly rainfall data was categorised as storm rainfall events based on particular criteria. Storms were determined when the precipitation of the rainfall exceeded a minimum threshold of 3 mm per hour. A continuous storm also required a minimum of three hours between subsequent wet spells and dry intervals. Furthermore, each storm has to endure at least three hours. The rainfall data was divided into discrete storm occurrences using these severe criteria, allowing for a thorough analysis of storm characteristics and their impact on numerous hydrological and meteorological phenomena. Our research focused on two crucial variables: the maximum annual precipitation (*p*) and the maximum annual duration (*t*)of each storm occurrence. The use of these stringent criteria and the examination. The utilization of these rigorous criteria and the examination of these key variables provide a solid foundation for enhanced forecasting capabilities and informed decision-making in managing extreme weather events.

3. Data Set

The data used in this study is a short-scale hourly rainfall data of Alor Setar, taken from July 1, 1970 until July 31, 2008. The data was recorded by the JPS Stor Station Alor Setar with station number 6103047. The location of the station is at the latitude 6006'202" and longitude 100023'30". We get 329.469 hourly rainfall data for 38 years.

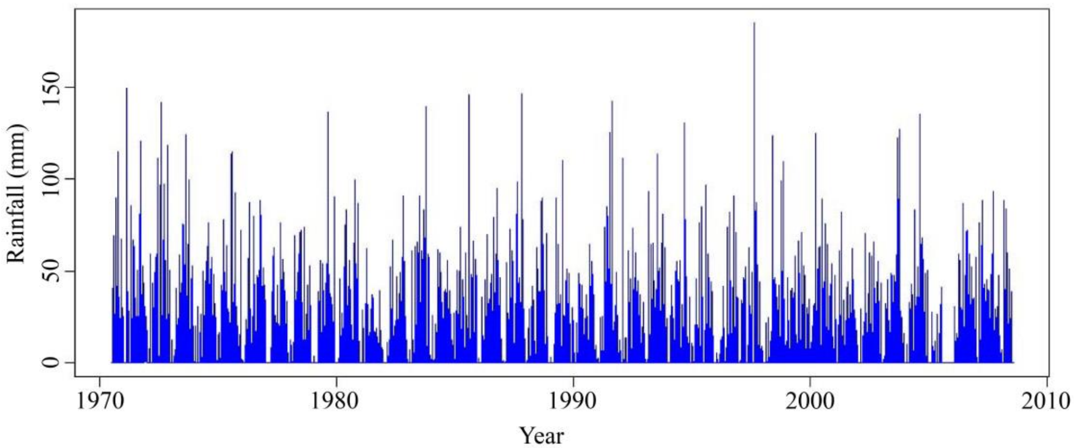

Fig. 2 Hydrological time series rainfall plot

In Figure. 2, annual rainfall patterns in Alor Setar from 1970 to 2008 vary, with some rainless days. Rainfall hit 150 mm in 1970, decreased for a few years, and returned to 150 mm in 1980. The highest rainfall was in 1997. In the final decade (2000-2008), there's a slight decline in rainfall intensity. Monthly data in Figure 3 displays differences in rainfall through color variations.

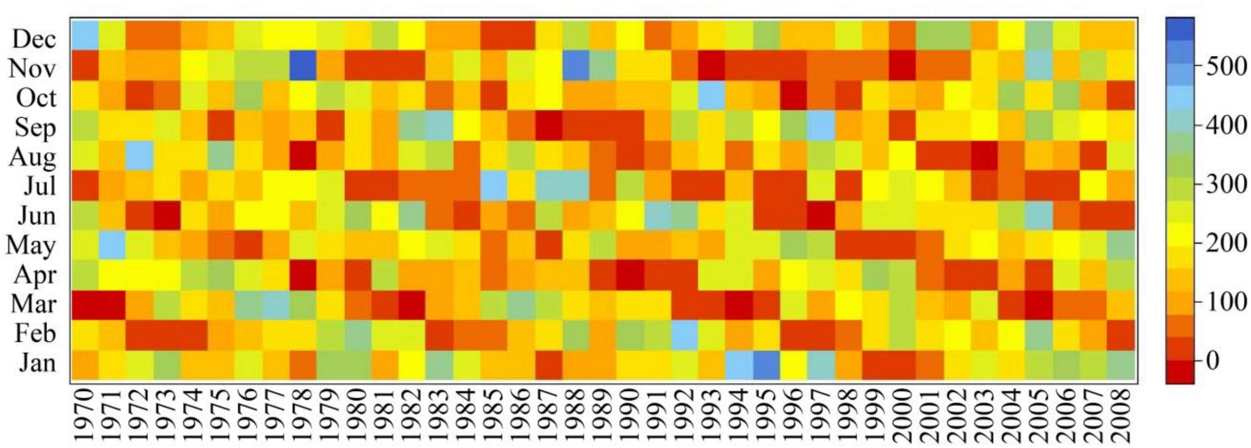

Fig. 3 Monthly preparation distribution

Extreme rainfall, indicated in dark blue, exceeding 500 mm, occurred in November 1978, and similar events in November 1988 and January 1995, albeit with slightly lower totals. Most monthly rainfall falls within the 100-300 mm range. The data suggests that rain is more frequent than not in Alor Setar between 1970 and 2008, necessitating short-scale rainfall modeling.

4. Probability Distribution

Probability distribution is useful for determining behavioral patterns and properties of hydrological data, such as rainfall and storm data. The rainfall probability distribution model is useful during the modeling phase of hydrological data simulation, particularly for short-scale and restricted rainfall data. In order to effectively predict storm rainfall data, we must use the General Extreme Value (GEV) and General Pareto distributions in our research. These distributions have shown to be useful tools for capturing the extreme occurrences and heavy-tailed behaviour seen in storm rainfall records. The GEV distribution is especially well-suited for modeling data with an upper bound, which is frequently present in storm-related phenomena. At the same time, the General Pareto distribution gives a valid representation of the tail end of the data, allowing us to effectively assess and predict extreme rainfall events. The probability density functions of each distribution are presented in Table 1.

Table 2 List of Distributions Used in This Study

Distributions	The Formula
GEV	$f(x;\xi,\mu,\sigma)=\frac{1}{\sigma}\left(1+\xi\left(\frac{x-\mu}{\sigma}\right)\right)^{-\frac{1}{\xi}} exp\left(-\left(1+\xi\left(\frac{x-\mu}{\sigma}\right)\right)^{-\frac{1}{\xi}}\right),\xi \neq 0.$
GP	$f(x;\xi,\sigma)=\frac{1}{\sigma} exp\left(1+\frac{\xi}{\sigma}x\right)^{\left(-\frac{1}{\xi-1}\right)},\xi \neq 0.$

5. Methods

Maximum Likelihood Estimation (MLE) is widely employed for parameter estimation in statistical models. MLE plays a pivotal role due to its ability to provide precise parameter estimates and facilitate model selection, especially in the analysis of rainfall data. It ensures that estimated parameters closely align with the characteristics of the rainfall distribution by maximizing the likelihood function. This method's significance is particularly pronounced when standard assumptions, such as normality, do not adequately capture the unique characteristics of rainfall data, which frequently exhibits skewness and heavy-tailed distributions.

In the realm of statistical modeling, the AIC (Akaike Information Criterion) and BIC (Bayesian Information Criterion) serve as invaluable tools for assessing goodness-of-fit and selecting the most appropriate model for a given dataset. Utilizing AIC and BIC, researchers can systematically evaluate competing models and strike a balance between model fit and complexity. This approach is essential to prevent the adoption of unnecessarily intricate models while ensuring that the chosen model not only fits the data well but also possesses the capacity to make robust predictions and generalize effectively across diverse research disciplines and applications.

6. Results and Discussion

6.1. Results

By applying storm identification criteria in the analysis of storm events, a total of 2000 storm events were successfully extracted from the dataset. After the collection of storm data, the subsequent research phase involved the determination of annual maximum values for each variable as shown in Table 2. These yearly maxima represent the highest values observed for each variable within the span of a calendar year. Calculating annual maximum values is essential for conducting an in-depth analysis of the magnitude and precipitation associated with storm events over time.

Examining the precipitation data, one can observe significant variations in the annual amounts of rainfall. The values range from as low as 41 ml in 2004 to as high as 303 ml in 1996. This substantial variability in precipitation levels suggests that the region experiences both dry and exceptionally wet years. The highest recorded precipitation in 1996 may indicate an extreme weather event, while the low of 41 ml in 2004 might signify a relatively dry year. Analyzing the duration of these storm events, we can see that they vary as well, with some events lasting as long as 244 hours in 1980 and others as short as 8 hours in 2007.

Table 2. The maximum annual precipitation and duration

No	Year	Precipitation (*p*)	Duration (*t*)	No	Year	Precipitation (*p*)	Duration (*t*)
1	1970	114	120	20	1990	91.6	71
2	1971	266	92	21	1991	206	34
3	1972	142	53	22	1991	112	13
4	1973	139	57	23	1992	114	18
5	1974	65.4	38	24	1993	128	28
6	1975	115	29	25	1994	97	22
7	1976	93.4	43	26	1995	90.5	23
8	1977	98	200	27	1996	303	40
9	1978	73	155	28	1997	132	18
10	1979	136	13	29	1998	81.5	26
11	1980	137	244	30	1999	125	34
12	1981	72	31	31	2000	81.8	15
13	1982	137	96	32	2001	81	15
14	1983	140	15	33	2002	128	30
15	1984	61.5	24	34	2003	148	18
16	1985	150	24	35	2004	41	9
17	1986	94.9	18	36	2005	87	21
18	1987	147	120	37	2006	99.5	25
19	1988	104	38	38	2007	88.5	8
20	1989	131	48				

In the context of the provided in Table 2, an analysis is being undertaken to construct probability distribution models using the Generalized Extreme Value (GEV) and Generalized Pareto (GP) distributions. The application of probability distribution models, like GEV and GP, plays a pivotal role in the field of hydrology and extreme value analysis This analysis aims to generate four distinct models, two for each variable: precipitation (p) and duration (t), namely GEV*p*, GEV*t*, GP*p*, and GP*t*. These models, once established, will serve as valuable tools for risk assessment, disaster management, and infrastructure planning.

The parameters for these four models (a, b, c, and d) were determined through the utilization of the Maximum Likelihood Estimation (MLE) technique.. The MLE approach, which maximizes the likelihood function, is widely recognized for its accuracy in parameter estimation and is particularly valuable in fitting statistical models to data. To facilitate this estimation process, the statistical software program R was utilized, and the resulting parameter estimates are presented in Table 3. These estimates play a crucial role in characterizing the behavior of the models, allowing for a more comprehensive understanding of the underlying distributions and their applicability in various research contexts.

Table 3. Parameters of different distribution

Paramater	GEVp	GEVt	GPp	GPt
μ	97.58656	21.463	—	—
ξ	0.06018	0.4229	-0.5334	-0.2255
σ	33.81608	11.9757	170.14751	43.3833

6.2. Discussion

The objective of this study is to evaluate the performance of these models based on graphical representations and assessing the goodness fit test using AIC and BIC values in Table . The analysis involved visualizing the model fits through graphical representations, with each graph depicted in Figure 4,5,6 and 7.

These criteria play a crucial role in determining the most suitable model for storm data modeling, considering both model complexity and accuracy. By comparing the graphical results and AIC/BIC values for each model, we aim to identify the most appropriate model that can accurately represent storm characteristics and enhance flood forecasting and water resource management strategies.

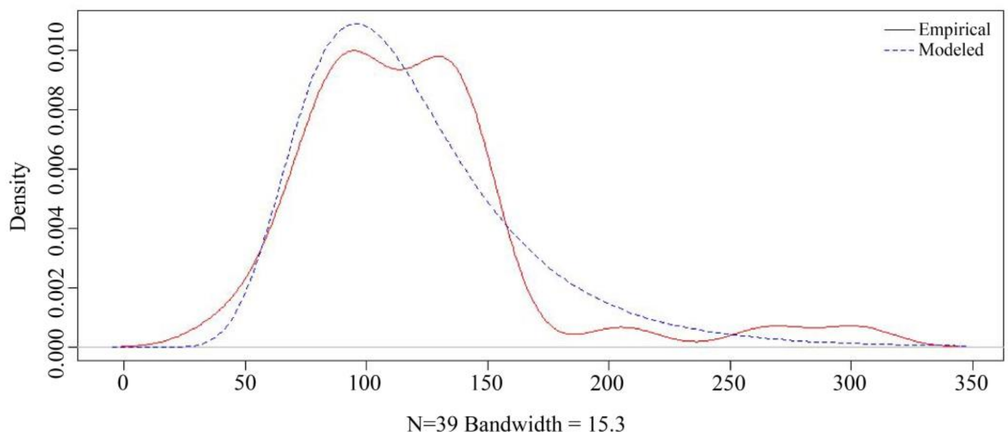

Fig. 4 GEV*p* Model

In Figure 4. while the peak of the GEV*p* model is slightly higher and exhibits a single peak, in contrast to the data's density function graph, which displays two peaks. The GEV*p* model aligns closely with the data graph when the range of maximum intensities falls between 50 and 100.

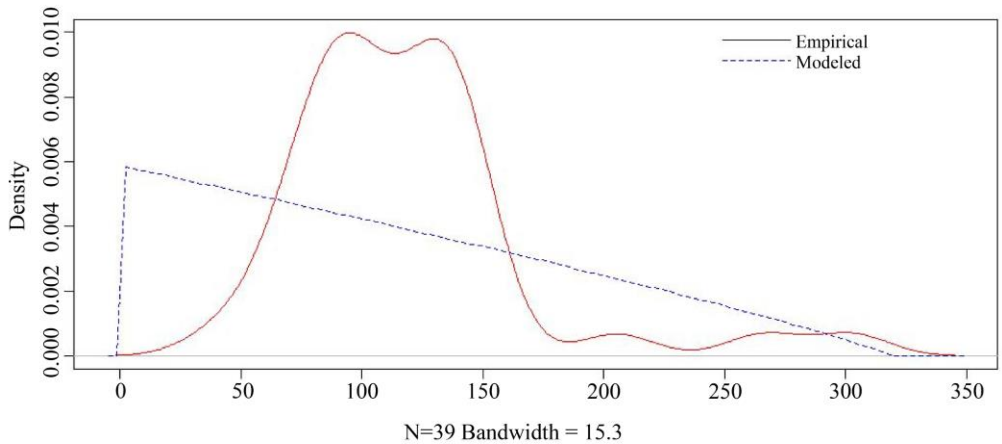

Fig. 5 GEV*t* Model

Figure 5 clearly illustrates that the GPp model falls short in approximating the research data. The peak of the GP*p* model is notably lower than the data density curve. The GP*p* model exhibits a triangular shape, in contrast to the data's density graph, which forms a curve.

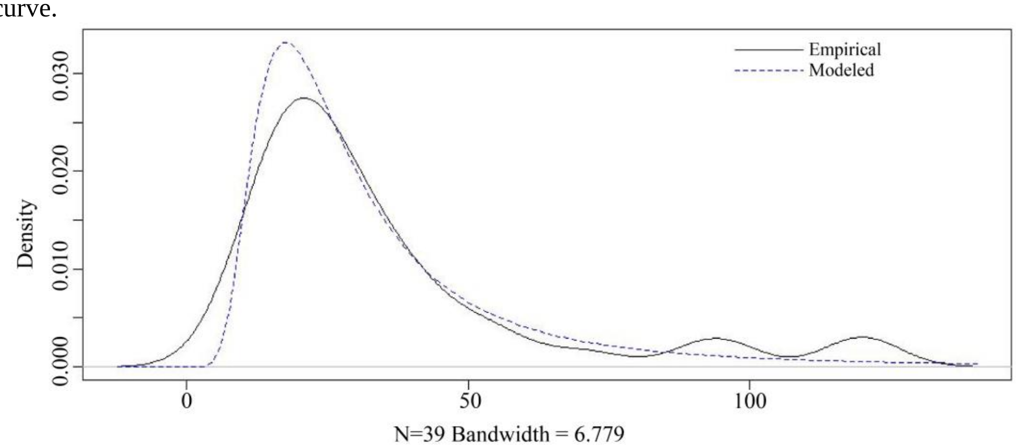

Fig. 6 GP*p* Model

Based on Figure 6, it is evident that the GEV*t* graph effectively approximates the research data. Although the GEV*t* graph is sharper in shape, there are several points where both graphs closely coincide.

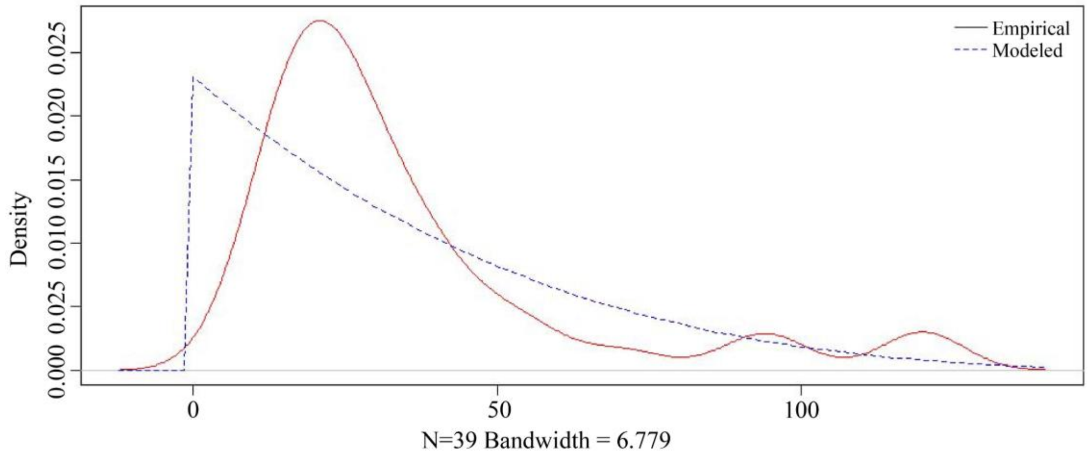

Fig. 7 GP*t* Model

It is evident from Figure 7 that the GP*t* model falls short in approximating the research data. This discrepancy arises because the data's density graph takes on a curved shape, while the density function graph of GP*t* appears more triangular in nature. Subsequently, AIC and BIC tests were conducted using the R program, resulting in the respective AIC and BIC values for each annual maximum storm duration model, as displayed in Table 4.

Table 4. The goodness fit test

Model	AIC	BIC
GEV*p*	405.9629	410.9536
GP*p*	441.3032	444.6303
GEV*t*	341.758	346.7487
GP*t*	358.5482	361.8753

The models utilizing the GEV probability distribution exhibit lower AIC and BIC values compared to the GP models. Consequently, based on the comparison of graphs and the results of AIC and BIC tests, it can be concluded that both the GEV*p* and GEV*t* models are the most effective in approximating the annual maximum precipitation and duration of storms in Alor Setar from 1970 to 2008.

Extreme weather events such as storm rainfall in Alor Setar are recurring incidents with future implications that can be quantified in terms of return periods. The return period provides insight into the frequency with which similar storm events are expected to recur in Alor Setar over the upcoming years. The study predicts return periods at intervals of 2, 5, 10, 20, 50, 100, 200, 500, and 1000 years. The return periods for each best-fitting model are illustrated in Figures 8 and 9.

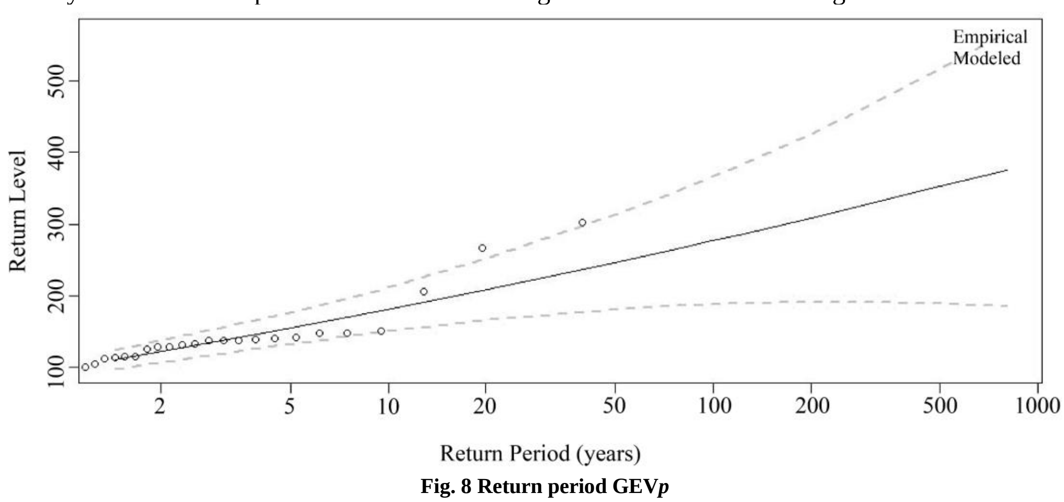

Fig. 8 Return period GEV*p*

Analyzing Figure 8 alongside the data from Table 2 for GEV*p*, several notable findings emerge regarding the return periods of storm events in Alor Setar. The most extreme storm, exceeding 303 mm in 1997, is anticipated to reoccur approximately every 40 years, around 2037. In contrast, the intense 266 mm storm of 1971 is projected to return roughly every 20 years, with a high level of confidence, around 1991. The 206 mm storm in 1991 is expected to recur approximately every 13 years. For storm events with intensities within the 80-200 mm range, the return periods tend to be shorter, typically spanning 0-10 years. These estimates offer valuable insights into the recurrence patterns of extreme storm events in Alor Setar, serving as essential information for disaster preparedness and risk assessment.

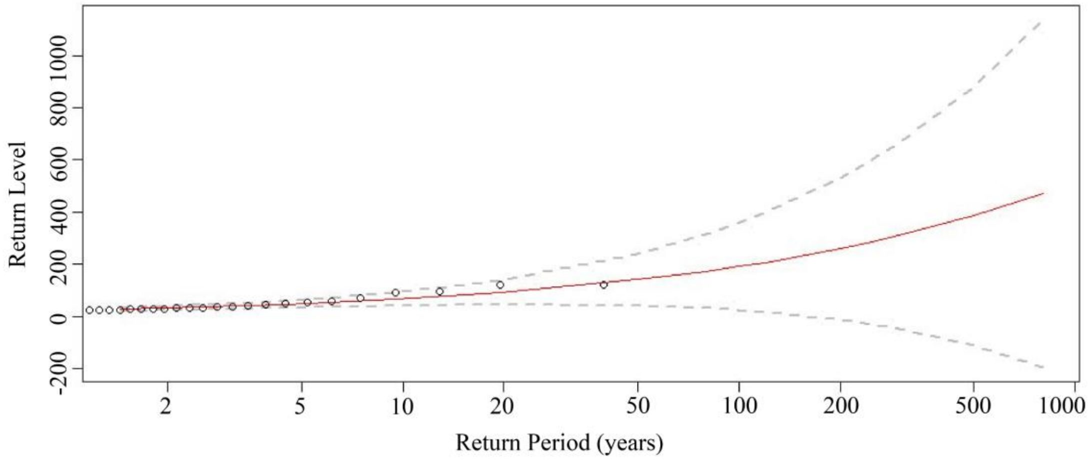

Fig. 9 Return period GEV*t*

Figure 9 and Table 3 reveals valuable insights. Firstly, there are two storm events with the highest maximum durations of 120 hours, occurring in 1970 and 1987. One of these events has an approximate return period of 40 years, while the other has a return period of about 20 years. For instance, a storm event in 1970 with a 20-year return period suggests a similar event around 1990. However, in the data, it's observed that a storm event with the same duration occurred in 1987, and in 1990, a storm with a 71-hour duration was recorded. Although not exact, this prediction is still reasonable with a 95% confidence level, given the minor three-year difference. Considering a 40-year return period for 1987 implies a likelihood of a similar event occurring in 2027. Additionally, for durations less than 100 hours, return periods are relatively short and range from 0 to 14 hours. In summary, it can be inferred that shorter maximum durations correspond to shorter return periods, and with a 95% confidence level allowing for a 5% margin of error, the return periods from the GEVt model are reliable for predicting storm events with varying maximum durations in Alor Setar.

7. Conclusion

In conclusion, our study into storm data modeling using the General Extreme Value (GEV) and General Pareto (GP) distributions has yielded valuable insights. Through the comparison of four distinct models, GEV*p*, GEV*t*, GP*p*, and GP*t*, we identified the GEV*p* and GEV*t* models as the most suitable for representing the maximum annual precipitation and duration variables, respectively. These models demonstrate superior performance in capturing the extreme behavior of storm events, as evidenced by their closer fit to the observed data compared to other models. The lower BIC and AIC values associated with the GEV*p* and GEV*t* models further validate their suitability in accurately representing storm characteristics. These findings suggest that the General Extreme Value approach are robust choices for modeling storm events based on hourly rainfall data in Alor Setar, Malaysia. The utilization of these models can significantly enhance our understanding of extreme weather events, thereby improving flood forecasting and water resource management strategies. Such insights are valuable for meteorologists, hydrologists, and policymakers in effectively addressing the challenges posed by extreme weather phenomena.

References

[1] Noratiqah Mohd Ariff, and Abdul Aziz Jemain, "Comparisons between the Window-Based and Storm-Event Analysis," *Sains Malaysiana*, vol. 41, no. 11, pp. 1377-1387, 2012. [Google Scholar]

[2] Abdolreza Bahremand, and Florimond De Smedt, "Distributed Hydrological Modeling and Sensitivity Analysis in Torysa Watershed, Slovakia," *Water Resource Management*, vol. 22, no. 3, pp. 393-408, 2008. [CrossRef] [Google Scholar] [Publisher Link]

[3] Kenneth P. Burnham, and David R. Anderson, "Multimodel Inference: Understanding AIC and BIC in Model Selection," *Sociological Methods & Research*, vol. 33, no. 2, pp. 261-304, 2004. [CrossRef] [Google Scholar] [Publisher Link]

[4] Stuart Coles, “Basics of Statistical Modeling,” *An Introduction to Statistical Modeling of Extreme Values*, pp. 18-44, 2001. [CrossRef] [Google Scholar] [Publisher Link]

[5] Anita Verpe Dyrrdal et al., “Bayesian Hierarchical Modeling of Extreme Hourly Precipitation in Norway,” *Environmental and Ecological Statistics*, vol. 26, no. 2, pp. 89-106, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[6] Gilli Manfred, and Evis Kellez, “An Application of Extreme Value Theory for Measuring Financial Risk,” *Computational Economics*, vol. 27, 207-228, 2006. [CrossRef] [Google Scholar] [Publisher Link]

[7] Kotz Samuel, and Saralees Nadarajah, *Extreme Value Distributions: Theory and Applications*, Imperial College Press, pp. 1-185, 2000. [Google Scholar] [Publisher Link]

[8] Richard H. McCuen, *Hydrologic Analysis and Design*, “Pearson Prentice Hall,” 2005. [Google Scholar] [Publisher Link]

[9] Douglas C. Montgomery, Elizabeth A. Peck, and G. Geoffrey Vining, *Introduction to Linear Regression Analysis*, 6th ed, “John Wiley & Sons,” 2021. [Google Scholar] [Publisher Link]

[10] Palynchuk Barry, and Yiping Guo, “Threshold Analysis of Rainstorm Depth and Duration Statistics at Toronto, Canada,” *Journal of Hydrology*, vol. 348, no. 3-4, pp. 535-545, 2008. [CrossRef] [Google Scholar] [Publisher Link]

[11] Shabri Ani, and Nor Atiqah Mohd Ariff, “Frequency Analysis of Maximum Daily Rainfalls via L-moment Approach,” *Sains Malaysiana*, vol. 38, no. 2, pp. 149-158, 2009. [Google Scholar]

[12] Mardhiyyah Shaffie et al., “The Spatial Return Level of Aggregated Hourly Extreme Rainfall in Peninsular Malaysia,” *Theoretical and Applied Climatology*, vol. 121, pp. 61-69, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[13] Wan Zawiah Wan Zin, Abdul Aziz Jemain, and Kamarulzaman Ibrahim, “The Best Fitting Distribution of Annual Maximum Rainfall in Peninsular Malaysia Based on Methods of L-moment and LQ-moment,” *Theoretical and Applied Climatology*, vol. 96, pp. 337-344, 2009. [CrossRef] [Google Scholar] [Publisher Link]

[14] Myeong-Ho Yeo, Van-Thanh-Van Nguyen, and Theodore A. Kpodonu, “Characterizing Extreme Rainfalls and Constructing Confidence Intervals for IDF Curves using Scaling-GEV Distribution Model,” *International Journal of Climatology*, vol. 41, no.1, pp. 456-468, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[15] M.D. Zalina et al., “Selecting a Probability Distribution for Extreme Rainfall Series in Malaysia,” *Water Science and Technology*, vol. 45, no. 2, pp. 63-68, 2002. [CrossRef] [Google Scholar] [Publisher Link]

[16] N. Vivekanandan, R.S. Jagtap, and C. Srishailam, “Comparison of Probability Distributions for Extreme Value Analysis of Wind Speed Data,” *SSRG International Journal of Geoinformatics and Geological Science*, vol. 5, no. 3, pp. 16-23, 2018. [CrossRef] [Publisher Link]