IOCMA 2026 Conference

The 2nd International Online Conference on Mathematics and Applications

10–12 June 2026 | Online

MDPI

Evaluation of Probability Distributions for Flood Frequency Analysis Using PDS

M. L. P. Anuruddhika[1], K. K. K. R. Perera[1], L. P. N. D Premarathna[1], T. H. W. Pathiranage[1] and V. P. A. Weerasinghe[2],

[1]Department of Mathematics, Faculty of Science, University of Kelaniya, Kelaniya, Sri Lanka, [2]Department of Zoology and Environmental Management, Faculty of Science, University of Kelaniya, Kelaniya, Sri Lanka

INTRODUCTION & AIM

Flood Frequency Analysis (FFA) is a statistical approach used to estimate the magnitude and occurrence probability of floods for different return periods. While the traditional Annual Maximum Series (AMS) considers only the largest flood in each year, the Partial Duration Series (PDS) method includes all flood events above a selected threshold, providing a more comprehensive representation of flood behavior, especially in regions experiencing multiple floods annually.

MATERIAL AND METHODS

1. Daily water level data (2010–2024) were collected from the Dunamale gauging station of Attanagalu Oya.
2. Data were cleaned, formatted, and arranged chronologically.
3. PDS method was used for flood event extraction.
4. Thresholds from 3.5 m to 5.0 m were tested at 0.1 m intervals.
5. The optimal threshold was selected using the mean excess plot.
6. A 10-day declustering method ensured independent flood events.
7. Five distributions (Five probability distributions: Gumbel, Generalized Extreme Value (GEV), Lognormal, Weibull, and Generalized Pareto Distribution (GPD)) were fitted to the data.
8. Parameters were estimated using Moments, Maximum Likelihood Estimation, and L-moments.
9. Nash–Sutcliffe Efficiency (NSE), Percent Bias (PBIAS), Root Mean Square Error (RMSE), and the RSR index

RESULTS & DISCUSSION

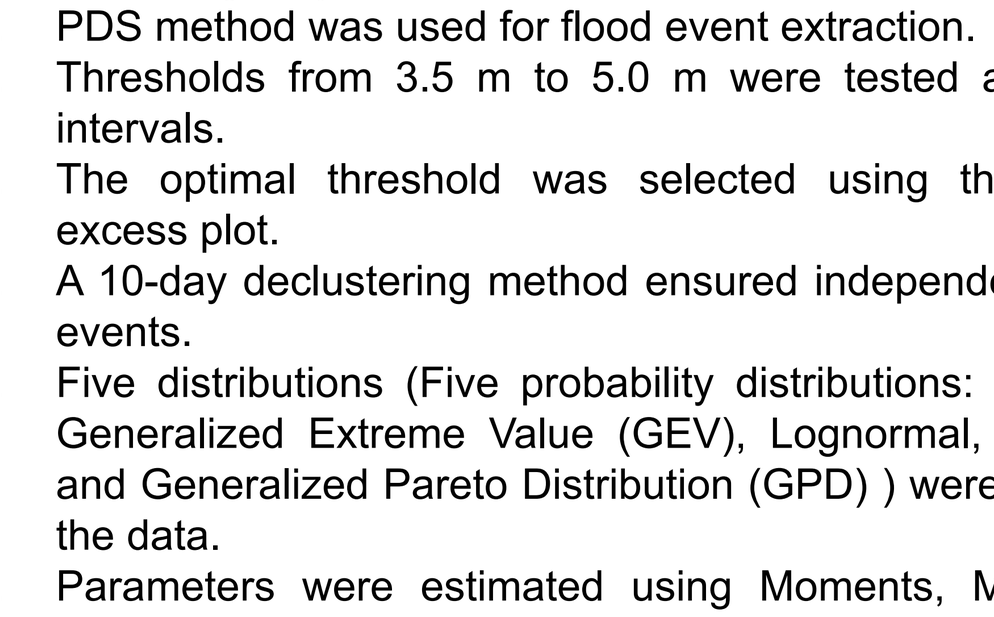

Figure 1: Mean excess plot for the time series

Figure 1 indicates that 3.5 m was selected as the optimal threshold from the mean excess plot, resulting in 68 declustered flood peaks above the threshold.

Table 1 Fitted Parameters for each distribution

Distribution	Parameters		
	Location	Scale	Shape
Gumbel	4.3025	0.5828	
Weibull	0	4.9105	8.0222
Log Normal	0	4.5793	0.1424
GEV	4.3804	0.6321	0.2304
GPD	$Threshold = 3.5$	1.4240	−0.5000

GEV and GPD distributions showed better performance in modeling extreme floods due to their flexible shape parameters. GEV indicated heavy-tailed behavior with higher probabilities of extreme events, while GPD suggested bounded extremes above the threshold. Gumbel, Lognormal, and Weibull provided reasonable overall fits but were less effective in representing extreme flood behavior.

The K–S test results showed that all distributions except GPD fitted the observed flood data adequately, with p-values above 0.05. Weibull showed the best fit with the lowest K–S statistic and highest p-value, followed closely by GEV. In contrast, GPD showed a very poor fit, indicating that threshold selection may need further investigation.

Validation metrics showed that the GEV distribution provided the best overall performance with the highest NSE (0.9714)and lowest RMSE (0.0252) and RSR (0.1691)values, indicating high accuracy and reliability. Lognormal and Weibull also performed well, while Gumbel showed acceptable results with nearly unbiased predictions.

CONCLUSION

The GEV distribution showed the best performance for flood frequency analysis, while Gumbel and Weibull provided acceptable but less reliable results for extreme events. Overall, the findings emphasize the importance of selecting suitable distributions and parameter estimation methods to improve the reliability of flood frequency analysis and support effective hydrological planning and flood risk management.

FUTURE WORK / REFERENCES

Future work should improve GPD threshold selection and apply multivariate approaches for more comprehensive flood risk analysis.