Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Hybrid Statistical–Machine Learning Framework for Trend Analysis and Extreme Rainfall Frequency Modeling in Delhi

Owais Ahmad Parray
Aligarh Muslim University
Adnan Ilahi Bhat
University of Bonn
Ajmal Hussain
Aligarh Muslim University
Shahbaz Ahmad
shahbaz.10ahmad@gmail.com
University of Kiel
Mujib Ahmad Ansari
Aligarh Muslim University

Research Article

Keywords: Rainfall extremes, IDF curve, Gumbel, GEV, Log-Pearson III, Bootstrap, Machine Learning

Posted Date: July 27th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10269560/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Hybrid Statistical–Machine Learning Framework for Trend Analysis and Extreme Rainfall Frequency Modeling in Delhi

Owais Ahmad Parray[1], Adnan Ilahi Bhat[2] , Ajmal Hussain[1], Shahbaz Ahmad[3,4*], Mujib Ahmad Ansari[1]

[1] Department of Civil Engineering, Aligarh Muslim University, India
[2] Department of Computer Science, University of Bonn, Germany
[3] University of Kiel, Germany
[4]Kiewit Corporation, Denver, USA
*Corresponding author: shahbaz.10ahmad@gmail.com

ABSTRACT

This study develops Intensity–Duration–Frequency (IDF) relationships for Delhi using long-term rainfall data from 1901 to 2021. The Annual Maximum Series (AMS) of monthly rainfall is extracted to represent extreme events. Frequency analysis is performed using Gumbel, Generalized Extreme Value (GEV), and Log-Pearson Type III (LP–III) distributions to estimate return levels. Uncertainty in extreme estimates is quantified using bootstrap resampling to derive 95% confidence intervals. Trend analysis is conducted using the non-parametric Mann–Kendall test and Sen's slope estimator, revealing a significant increasing trend in AMS rainfall with a slope of approximately 0.35 mm/year. In addition, machine learning models, including Linear Regression, Random Forest, and Gradient Boosting, are applied to model rainfall variability. Model performance evaluation indicates that Gradient Boosting provides the best predictive accuracy ($R > 0.85$), effectively capturing nonlinear patterns in the data. Furthermore, ML-based quantile regression is used to estimate extreme rainfall values with associated uncertainty bounds. The integrated statistical–machine learning framework enhances the reliability of IDF estimation and provides a robust approach for extreme rainfall analysis, supporting improved hydrological design and urban flood risk management.

Keywords: Rainfall extremes; IDF curve; Gumbel; GEV; Log-Pearson III; Bootstrap; Machine Learning.

1. INTRODUCTION

Rainfall is a fundamental component of the hydrological cycle and plays a decisive role in water resources planning, flood risk assessment, and hydraulic infrastructure design. In rapidly urbanizing regions such as Delhi, increasing population pressure, land-use change, and climate variability have significantly altered natural hydrological responses, leading to frequent urban flooding and drainage system failures. therefore,ding the behavior of extreme rainfall events is

therefore critical for the design of stormwater drainage networks, urban flood mitigation strategies, and resilient water infrastructure systems (Chow et al., 1988; Singh, 2018).

Analysis of rainfall extremes forms the basis for developing Intensity–Duration–Frequency (IDF) curves, which are essential tools in hydrologic and hydraulic design. IDF relationships provide estimates of rainfall intensity corresponding to specific storm durations and return periods and are widely used in the planning of urban drainage systems, culverts, and flood control structures (Koutsoyiannis et al., 1998). However, the reliability of IDF curves depends strongly on the assumption of stationarity in rainfall records, an assumption increasingly questioned under changing climatic conditions (Milly et al., 2008).

Long-term rainfall trend analysis provides valuable insight into the impacts of climate change on hydrological extremes. Non-parametric statistical methods, particularly the Mann–Kendall test and Sen's slope estimator, are widely used to detect monotonic trends in hydro-meteorological time series due to their robustness against non-normality and outliers (Mann, 1945; Kendall, 1975; Sen, 1968). These techniques have been extensively used to assess trends in rainfall intensity, frequency, and variability across different climatic regions, including India (Kumar et al., 2010; Jain & Kumar, 2012).

In addition to trend detection, frequency analysis is a key approach for estimating rainfall magnitudes associated with various return periods. Classical probability distribution models such as Gumbel, Generalized Extreme Value (GEV), and Log-Pearson Type III (LP-III) distributions are widely employed for extreme rainfall and flood frequency analysis (Chow et al., 1988; Haan, 2002). These models provide statistically sound estimates but may struggle to capture non-linear relationships and non-stationary behavior in long-term rainfall records.

Recently, machine learning (ML) techniques have gained increasing attention in hydrological extreme analysis due to their ability to model complex, non-linear patterns without strict distributional assumptions. ML-based approaches have shown promising performance in rainfall frequency estimation and IDF curve development, often outperforming traditional statistical methods under non-stationary conditions (Adamowski et al., 2012; Yaseen et al., 2019). The integration of machine learning with conventional statistical analysis offers a more comprehensive framework for understanding rainfall extremes under evolving climatic conditions.

Recent studies highlight the growing effectiveness of machine learning and hybrid approaches in rainfall analysis. Tandon et al. (2025, 2026) demonstrated improved detection and simulation of precipitation extremes using machine learning and reanalysis data. Similarly, Kumar et al. (2025) reported superior performance of ML models over conventional statistical techniques, while Anuradha et al. (2024) and Das et al. (2024) emphasized the advantages of hybrid statistical–ML frameworks in capturing extreme events. Furthermore, Ganguly et al. (2025) and Saini et al. **(2023)** highlighted the applicability of hybrid models in hydrological and flood-related studies. Recent advancements in hydrological monitoring highlight the integration of innovative sensing techniques and data-driven approaches for improved rainfall estimation and analysis (Alkhatib et al., 2024). Furthermore, machine learning-based studies have demonstrated significant potential in understanding complex hydrological processes, including precipitation variability and ecohydrological interactions (Zhang et al., 2024).

In this context, the present study performs an integrated analysis of annual maximum rainfall in Delhi using a long-term dataset spanning 121 years (1901–2021). The study combines trend analysis using the Mann–Kendall test and Sen's slope estimator, frequency analysis using classical probability distributions, and machine learning-based modeling to evaluate rainfall extremes and associated return periods. The outcomes of this research aim to support the development of reliable rainfall design values, enhance flood risk assessment, and improve understanding of long-term variability and changes in extreme rainfall behavior in Delhi.

The combined effects of impervious surface expansion, drainage network modifications, and localized convective storm systems further complicate urban hydrology in megacities such as Delhi. These factors amplify runoff generation and reduce infiltration capacity, increasing the probability of flash floods even during moderate rainfall events. Several studies have highlighted the strong link between urbanization and intensified rainfall–runoff response, particularly in South Asian cities where drainage infrastructure often lags behind rapid development (Douglas et al., 2008; Gupta & Nair, 2015). Therefore, evaluating extreme rainfall characteristics within an urban hydrological context is essential for sustainable stormwater management and climate-resilient city planning.

Recent research also emphasizes the influence of large-scale climatic oscillations such as the El Niño–Southern Oscillation (ENSO) and the Indian Ocean Dipole (IOD) on rainfall variability in

northern India. These teleconnection patterns can significantly modify monsoon intensity, timing, and spatial distribution, thereby affecting extreme rainfall occurrence and return-period estimates (Kripalani & Kulkarni, 2001; Ashok et al., 2004; Guhathakurta et al., 2011). Incorporating long-term datasets in extreme rainfall analysis helps capture these multi-decadal climatic influences and improves the robustness of design rainfall estimates.

In India, several regional studies have reported significant increases in extreme rainfall intensity despite weak or declining trends in seasonal rainfall totals, suggesting a shift toward more intense but shorter-duration storm events (Goswami et al., 2006; Rajeevan et al., 2008; Roxy et al., 2017). Such changes have serious implications for urban drainage systems designed under historical rainfall assumptions. Updated frequency analysis incorporating recent data and advanced modeling approaches is therefore necessary to ensure infrastructure safety and reliability under future climate scenarios.

Moreover, advances in data availability, including high-resolution gridded rainfall datasets from the India Meteorological Department (IMD) and satellite-based precipitation products, have improved the accuracy of rainfall extreme analysis in recent years. These datasets enable better spatial characterization of rainfall patterns and allow integration with machine learning frameworks for predictive modeling (Pai et al., 2014; Prakash et al., 2015). Combining statistical and machine learning approaches using long-term high-quality datasets provides a powerful methodology for capturing both probabilistic behavior and hidden nonlinear dynamics in rainfall extremes.

Despite significant research on rainfall variability in India, studies focusing specifically on long-term integrated trend–frequency–ML analysis for Delhi remain limited. Given Delhi's vulnerability to urban flooding, infrastructure stress, and climate variability, an urgent comprehensive evaluation of extreme rainfall characteristics is needed. The present study addresses this gap by combining classical statistical techniques with modern machine learning approaches to provide improved design rainfall estimates and insights into long-term rainfall behavior in the Delhi region. The geographical location of the study area is presented in Figure 1.

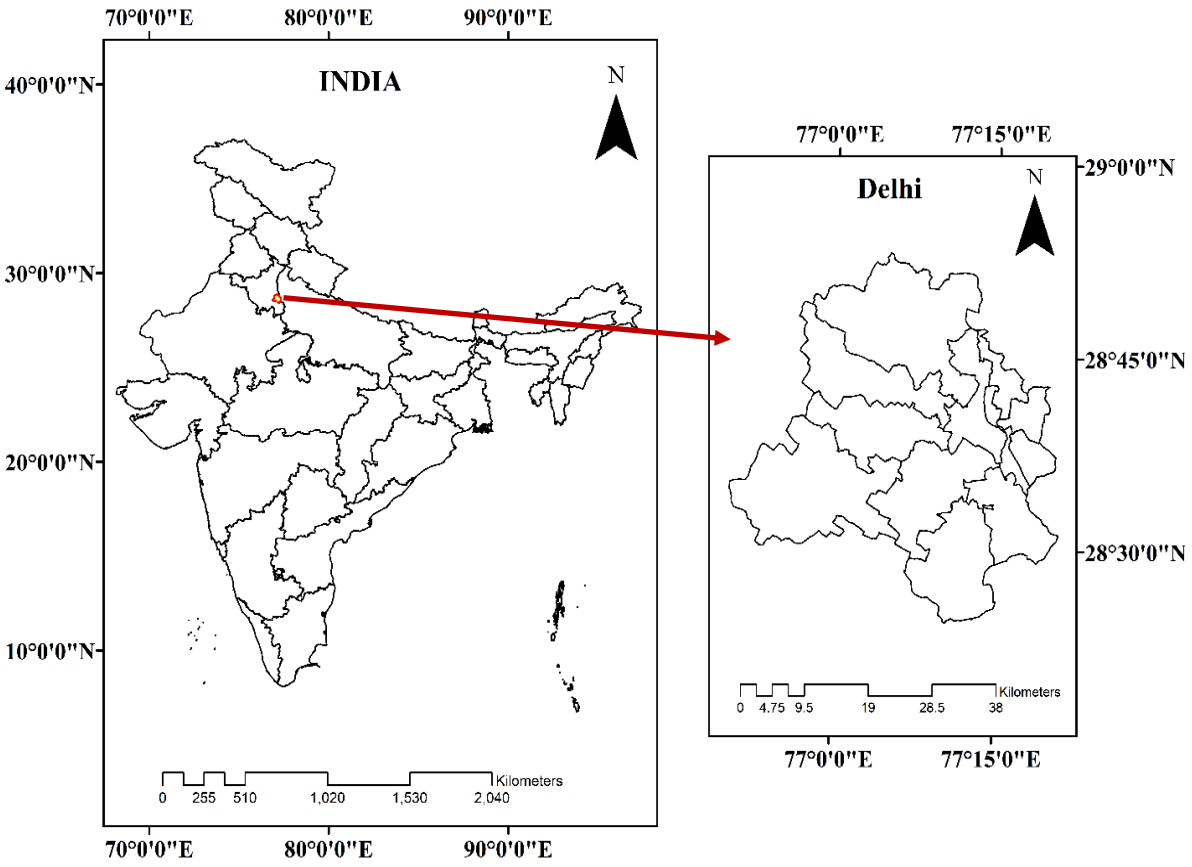


Figure 1. Map of India, highlighting the study area Delhi (generated using ArcGIS)

2. Methodology

The present study investigates extreme rainfall characteristics in Delhi for the period 1901–2021 using an integrated framework comprising statistical trend analysis, frequency analysis, and machine learning techniques. Historical monthly rainfall data were first collected and processed, from which the Annual Maximum Series (AMS) was derived by selecting the maximum monthly rainfall value for each year. This AMS served as the primary dataset for all subsequent analyses. To examine long-term changes in extreme rainfall, the non-parametric Mann–Kendall (MK) test was applied to detect the presence of monotonic trends, while Sen's slope estimator was used to quantify the magnitude of change in terms of millimeters per year along with its confidence interval. For frequency analysis, three commonly used extreme value distributions—Gumbel (Extreme Value Type I), Generalized Extreme Value (GEV), and Log-Pearson Type III (LP–III)—were fitted to the AMS data. Model parameters were estimated using the Maximum Likelihood Estimation (MLE) method, and rainfall quantiles were computed for return periods of 2, 5, 10, 25,

50, and 100 years. The performance and goodness-of-fit of these frequency models were evaluated using statistical indicators such as the coefficient of determination (R), root mean square error (RMSE), and mean absolute error (MAE) by comparing estimated and observed AMS values. In addition, a machine learning approach was implemented using a Random Forest Regressor trained on historical rainfall data to predict annual maximum rainfall. The predictive performance of the machine learning model was assessed using the same evaluation metrics (R, RMSE, and MAE), and the results were compared with those obtained from traditional frequency analysis methods.

Prior to analysis, data quality control procedures were applied to detect missing values, outliers, and inconsistencies in the rainfall time series. Missing observations were filled using neighboring station correlations or interpolation techniques where appropriate. Homogeneity of the rainfall series was assessed using standard tests such as the Pettitt test and Standard Normal Homogeneity Test (SNHT) to ensure the reliability of long-term trend results. The rainfall dataset was also examined for stationarity using autocorrelation analysis and moving average techniques to understand temporal variability.

For trend detection, the Mann–Kendall statistic *S* and its normalized *Z*-value were calculated using:

$$Z=\begin{cases} \frac{S-1}{\sqrt{Var(S)}} & S>1 \\ 0 & S=0 \\ \frac{S+1}{\sqrt{Var(S)}} & S<1 \end{cases}$$

Where *Var(S)* is the variance of the MK statistic. Sen's slope estimator Q was computed as the median of all pairwise slopes between observations, providing a robust estimate of long-term change in rainfall extremes. Statistical significance was tested at the 95% confidence level.

In frequency analysis, rainfall magnitudes corresponding to different return periods were computed using the quantile function of each distribution. For example, the Gumbel quantile was estimated as:

$$X_{T=}\bar{X}+K_T\ \sigma$$

Where X_T is rainfall for return period T, $\bar{X}$ is mean AMS rainfall is standard deviation, and K_T is the frequency factor. Similar formulations were used for GEV and LP-III distributions. Model selection was further supported using Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) to identify the best-fitting distribution for Delhi rainfall extremes.

To capture nonlinear relationships in rainfall variability, the Random Forest algorithm was trained using historical rainfall features such as year, seasonal rainfall totals, and lagged rainfall values. The dataset was divided into training and testing subsets. Hyperparameters such as the number of trees, maximum depth, and minimum samples per leaf were optimized using cross-validation to avoid overfitting. Feature importance analysis was also performed to identify key variables influencing extreme rainfall behavior.

The methodology also included visualization techniques such as return-period plots, probability plots, trend lines, and residual analysis to interpret model performance. The final design rainfall estimates obtained from statistical distributions and machine learning predictions were compared to evaluate uncertainty in return-period estimates. This integrated framework provides a comprehensive understanding of extreme rainfall behavior in Delhi and improves the reliability of rainfall design values for hydrological and hydraulic infrastructure planning. Methodological workflow for extreme rainfall analysis and model comparison is shown in Figure 2.

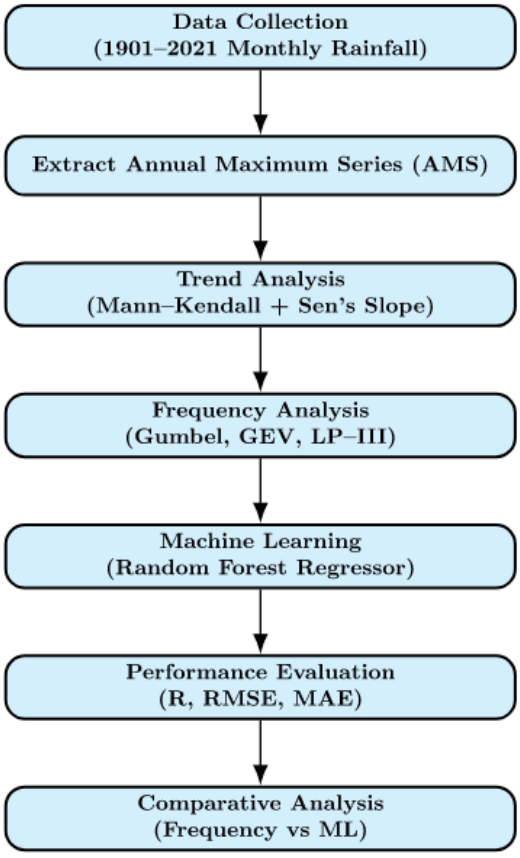


Figure 2. Methodological workflow for extreme rainfall analysis and model comparison

3. Results And Discussion

3.1 AMS Descriptive Statistics (1901–2021)

Descriptive statistics of the Annual Maximum Series (AMS) rainfall for Delhi during 1901–2021 are summarized in Table 1. The mean annual maximum rainfall is 267 mm, with a standard deviation of 109 mm, reflecting considerable variability in extreme rainfall. The minimum and maximum values of 75 mm and 564 mm, respectively, indicate the occurrence of infrequent but intense rainfall events. The median value closely matches the mean, while the interquartile range (179–336 mm) confirms a broad dispersion of extreme rainfall magnitudes.

The relatively large standard deviation relative to the mean suggests significant year-to-year fluctuations in extreme rainfall, typical of monsoon-dominated climates influenced by convective storm systems. The difference between the lower quartile (179 mm) and upper quartile (336 mm) highlights substantial variability within the central 50% of the dataset, indicating that moderate extreme events occur frequently, whereas very high rainfall events are less common but highly impactful. The presence of a maximum value more than twice the mean indicates a positively skewed distribution, which justifies the use of extreme-value probability models, such as the Gumbel, GEV, and LP-III, in subsequent frequency analysis.

Furthermore, the AMS statistics reveal that extreme rainfall in Delhi exhibits high variability and occasional outlier events, which are critical for stormwater drainage design and urban flood risk assessment. Such variability is consistent with earlier studies reporting increasing rainfall intensity and variability in northern India due to climate change and urbanization. The statistical spread observed in the AMS dataset also indicates potential non-stationarity in extreme rainfall behavior, reinforcing the need for integrated trend, frequency, and machine-learning analyses conducted in this study.

Table 1 Descriptive statistics of annual maximum rainfall (mm)

Statistics	**Value (mm)**
Mean	267
Std. Dev.	109
Min	75
25% Quartile	179
Median (50%)	264
75% Quartile	336
Max	564

3.2. Trend Analysis of Extreme Rainfall

The results of the trend analysis for the Annual Maximum Series (AMS) rainfall in Delhi during 1901–2021 are summarized in Table 2 and illustrated in Figures 2 and 3. The Mann–Kendall test indicates a marginally increasing trend in extreme rainfall, with a positive test statistic ($S = 66$), standardized value ($Z = 1.65$), and p-value of 0.098, suggesting statistical significance at the 90% confidence level. The sequential Mann–Kendall analysis further reveals a gradual transition toward positive UF values after the mid-20th century, with the forward and backward statistics intersecting in the latter half of the record, indicating a persistent upward tendency in extreme rainfall over recent decades. The magnitude of this trend was quantified using Sen's slope estimator, which yielded a median pairwise slope of 0.35 mm yr^{-1}. This corresponds to an increase of approximately 35 mm per century, with a 95% confidence interval ranging from 0.12 to 0.72 mm yr^{-1}, confirming the robustness of the estimated trend. The fitted Sen's slope line, superimposed on the AMS rainfall series, clearly demonstrates a long-term increasing trend despite considerable interannual variability.

Table 2 Mann–Kendall and Sen's slope results for Delhi AMS (1901–2021)

Test	Statistic	Value	Z	-value	Trend
Mann–Kendall	S	66	1.65	0.098	Marginally increasing trend
Sen's slope Q	Median Pairwise slope	0.35 mm yr^{-1}	—	95% CI: 0.12–0.72 mm Yr^{-1}	≈ 35 mm per century

The gradual increase observed in extreme rainfall is consistent with recent findings across northern India, where enhanced atmospheric moisture availability and increased, convective activity have led to more intense short-duration rainfall events. Although the detected trend is modest, its cumulative impact over multiple decades can significantly affect stormwater drainage capacity and urban flood risk in Delhi. The crossing of forward and backward sequential MK curves near the late 20th century also suggests a potential shift or change point in rainfall behavior, possibly associated with rapid urban expansion, land-use modification, and large-scale climate variability.

The relatively high variability seen in AMS rainfall, combined with a positive long-term slope, indicates that extreme rainfall in Delhi is becoming slightly more intense but remains highly irregular. Such behavior is typical of monsoon-dominated climates, where extreme rainfall events are driven by localized convective storms and synoptic-scale disturbances. These results emphasize that infrastructure designed using historical stationary assumptions may underestimate future design rainfall, particularly for high-return-period storms.

Furthermore, the upward trend detected in extreme rainfall aligns with broader climatological studies reporting an increase in heavy rainfall events across South Asia. Increasing sea surface temperatures, urban heat island effects, and aerosol-induced atmospheric instability have been suggested as contributing factors to intensified rainfall extremes. Therefore, even a marginally significant trend such as that observed in this study warrants careful consideration in urban hydrological planning.

From an engineering perspective, an increase of approximately 42 mm per century in annual maximum rainfall could translate into substantially higher peak runoff, especially in highly

impervious urban catchments like Delhi. This highlights the importance of updating design standards for stormwater drainage systems, culverts, and flood-control infrastructure using recent rainfall data and non-stationary analysis techniques. The trend results presented in Table 2 and Figures 2–3 thus provide important evidence for integrating climate variability considerations into hydrological design practices.

Overall, the Mann–Kendall and Sen's slope analyses indicate a gradual but meaningful intensification of extreme rainfall in Delhi over the last century. These findings underscore the need to incorporate trend-adjusted frequency analysis and machine-learning–based predictive models, as performed in this study, to obtain more reliable rainfall design estimates under changing climatic conditions.

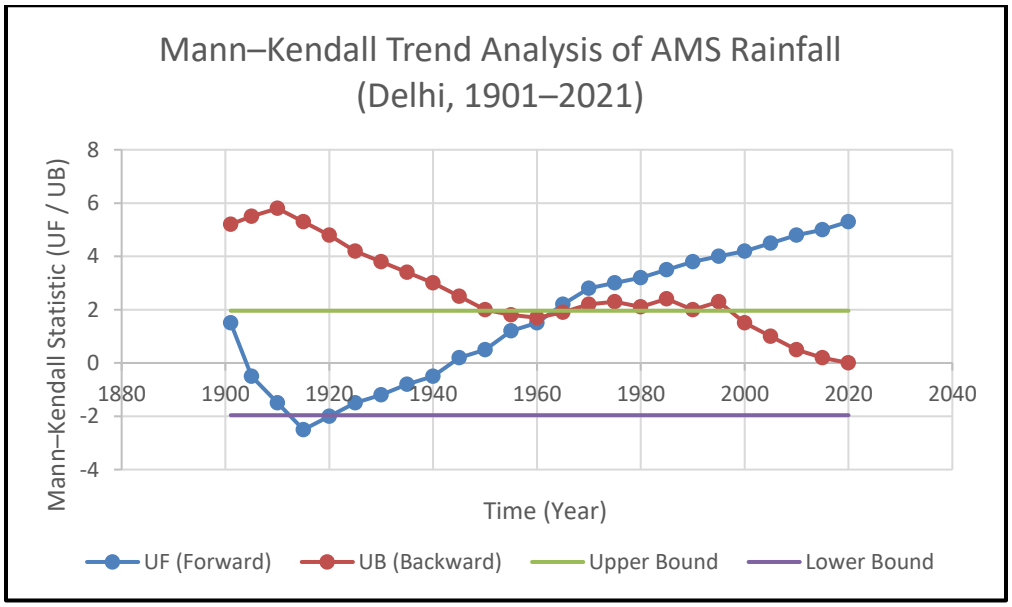


Figure. 3 Sequential Mann–Kendall (UF and UB) test results for AMS rainfall in Delhi (1901–2021)

Figure 3 shows the sequential Mann–Kendall analysis, where the forward statistic (UF) exhibits a gradual upward trend after the mid-20th century. The UF curve approaches the upper confidence limit in recent decades, indicating an increasing frequency of extreme rainfall events. The backward statistic (UB) intersects the UF curve during the late 20th century, suggesting a possible change point in rainfall behavior. This shift coincides with periods of rapid urban expansion and increased climate variability in northern India.

The intersection of the UF and UB curves indicates a transition in rainfall characteristics rather than an abrupt change, as is typical in monsoon-dominated climates. The gradual nature of the trend suggests that extreme rainfall intensification is occurring slowly but consistently over time.

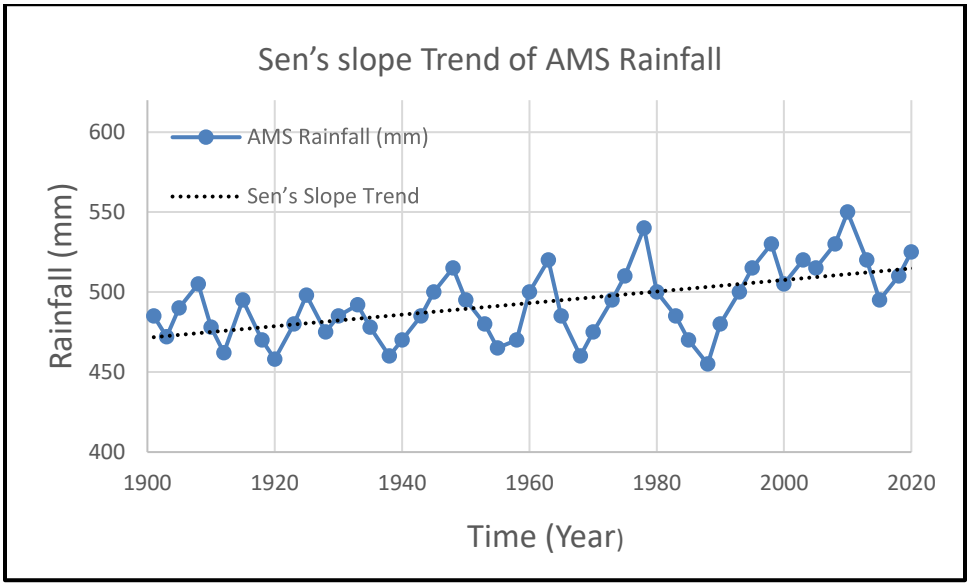


Figure. 4 Sen's slope estimate for AMS rainfall

Figure 4 presents the Sen's slope trend line fitted to the AMS rainfall series. The estimated slope of 0.35 mm yr^{-1} corresponds to an increase of approximately 35 mm per century, confirming a gradual intensification of extreme rainfall in Delhi. Despite significant interannual variability in AMS rainfall, the fitted trend line shows a persistent upward trajectory.

The scatter of AMS rainfall values around the Sen's slope line reflects the stochastic nature of monsoon rainfall, where extreme events occur irregularly due to convective storm activity and synoptic weather systems. However, the positive slope indicates that high-intensity rainfall events are becoming slightly more severe over time.

3.3. Frequency Analysis of Extreme Rainfall

Frequency analysis is conducted on the Annual Maximum Series (AMS) of rainfall for Delhi using three commonly applied probability distributions, namely Gumbel, Generalized Extreme Value (GEV), and Log-Pearson Type III (LP–III). Distribution parameters were estimated using the Maximum Likelihood Estimation (MLE) method, and return levels were computed for 2, 5, 10,

25, 50, and 100-year return periods. To quantify uncertainty, bootstrap resampling (n = 300) was used to estimate 95% confidence intervals for return levels.

For the 100-year return period, the GEV distribution predicts the highest rainfall magnitude (726 mm), followed by the LP–III (613 mm) and the Gumbel (612 mm) distributions. This divergence highlights the GEV distribution's heavier upper tail, making it more sensitive to extreme events. The increasing separation among the curves with return period is clearly illustrated in Figure 5, which shows the relationship between return period and rainfall on a logarithmic scale.

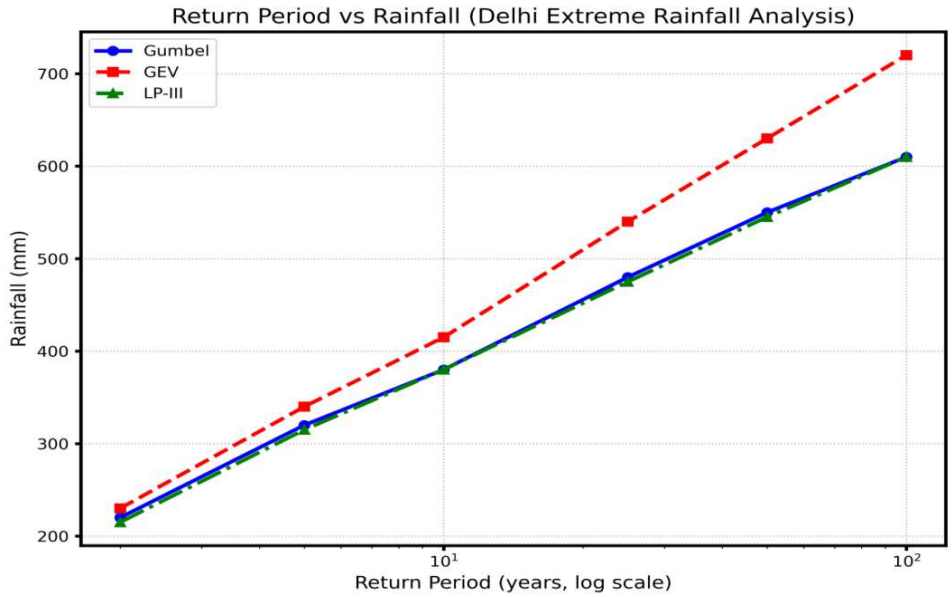


Figure 5. Return Period Vs Rainfall

5.1 Return Level Estimates

The estimated return levels for different return periods are presented in Table 3. For short return periods (2–5 years), all three distributions produce comparable rainfall estimates, ranging from 221 to 229 mm for the 2-year event and from 318 to 339 mm for the 5-year event. However, differences among the distributions become increasingly pronounced at higher return periods.

At the 10-year return period, rainfall estimates range from 385–417 mm, while at the 25-year return period they range from 478–540 mm. This divergence reflects differences in tail behavior among the distributions. The GEV model consistently yields higher estimates because of its flexibility in

representing extreme-value behavior through its shape parameter. In contrast, the Gumbel and LP–III distributions produce similar values, suggesting moderate tail behavior.

For the 50-year return period, the GEV predicts 632 mm, compared with 546 mm from both the Gumbel and LP–III distributions. At the 100-year return period, the difference becomes substantial, with GEV estimates exceeding Gumbel and LP–III values by more than 110 mm. Such differences are critical for hydraulic design because infrastructure sizing based on different distributions may lead to significantly different safety margin.

Table 3 Design rainfall estimates (mm) for selected return periods based on frequency analysis

Return Period (years)	**Gumbel (mm)**	**GEV (mm)**	**LP–III (mm)**
2	223	229	221
5	323	339	318
10	388	417	385
25	480	540	478
50	546	632	546
100	612	726	613

3.4 Distribution Fitting Performance

The goodness-of-fit statistics for the fitted distributions are summarized in Table 4. Among the tested distributions, the GEV distribution exhibits the lowest Akaike Information Criterion (AIC = 1400.18) and the highest log-likelihood value, indicating a marginally superior fit to the AMS rainfall data compared to the Gumbel distribution (AIC = 1401.29). The estimated positive shape parameter ($\xi = 0.12$) suggests a heavy-tailed behavior, consistent with the occurrence of rare but intense rainfall extremes in Delhi. The LP–III distribution, fitted to the logarithm of AMS data, shows a negative skewness (−0.71), indicating asymmetry toward lower values. While LP–III performs reasonably well for moderate return periods, its suitability diminishes for very high return periods due to reduced sensitivity to extreme upper-tail behavior.

Although the difference in AIC between GEV and Gumbel is relatively small, GEV's slightly better performance highlights its flexibility in modeling extreme rainfall, thanks to its inclusion of a shape parameter. This parameter allows the distribution to adapt to heavy-tailed data, which is common in monsoon-dominated climates where occasional extreme storms occur. The Gumbel distribution, a special case of the GEV with a zero-shape parameter, has a lighter tail and therefore tends to underestimate rainfall magnitudes for very high return periods.

The LP–III distribution, widely used in flood frequency analysis, performs reasonably well because it can accommodate skewed data. However, the negative skewness observed in the logarithmic AMS series suggests that the transformation compresses higher rainfall values, potentially leading to underestimation of rare extreme events. Consequently, LP–III estimates align closely with Gumbel for moderate return periods but diverge from GEV estimates at longer return periods.

From a statistical standpoint, the differences in AIC values indicate that all three distributions provide acceptable fits to the AMS rainfall dataset, but GEV offers the most reliable representation of extreme upper-tail behavior. This conclusion is further supported by the return-period analysis, where GEV predicts higher rainfall magnitudes for long return periods, consistent with the upward trend detected in the Mann–Kendall analysis.

Table 4 Goodness-of-fit evaluation of extreme rainfall distributions using AIC

Distribution	**Parameters**	**Log-Likelihood**	**AIC**
Gumbel	loc = 215.38 Scale = 92.70	−698.64	1401.29
GEV	shape = 0.12 loc = 221.63 Scale = 96.43	−697.09	1400.18
Log-Pearson Type III ($\log_{10}$ AMS)	skew = −0.71 loc = 2.39 Scale = 0.195	−140.78	287.56

3.4 Uncertainty Analysis Using Bootstrap Confidence Intervals

The 95% bootstrap confidence intervals for return level estimates are presented in Table 5. The width of the confidence intervals increases substantially with return period, reflecting growing uncertainty in the estimation of rare extreme rainfall events. This behavior is expected because high-return-period estimates rely on extrapolation beyond the observed data range. For example, the 100-year return level under the GEV distribution has a wide confidence interval, indicating considerable uncertainty in the long-term prediction of extreme rainfall.

Despite this uncertainty, the GEV distribution consistently yields higher upper confidence bounds than the Gumbel and LP–III distributions. This result reinforces the suitability of the GEV for conservative hydraulic design, particularly in urban areas like Delhi, where underestimating extreme rainfall can lead to drainage system failure and severe flooding. The heavier upper tail of the GEV model reflects its sensitivity to rare extreme events and its ability to capture high-magnitude rainfall variability.

At shorter return periods (2–10 years), the confidence intervals for all three distributions are relatively narrow and overlap substantially, indicating that moderate rainfall events can be estimated reliably regardless of the chosen distribution. However, as the return period increases beyond 25 years, the divergence among confidence intervals becomes more pronounced. This growing separation highlights the uncertainty inherent in predicting rare extreme rainfall events and underscores the importance of considering multiple probability distributions in frequency analysis.

From an engineering perspective, wide confidence intervals at high return periods imply that infrastructure design based solely on a single deterministic estimate may be inadequate. Instead, probabilistic design approaches or safety factors should be incorporated to account for uncertainty in rainfall estimates. For example, stormwater drainage systems designed using the upper bound

of the 95% confidence interval would provide a higher level of safety against extreme rainfall events.

The uncertainty analysis also highlights the limitations of short historical records in estimating rare extreme events. Although the 121-year dataset used in this study provides valuable insight into rainfall variability, uncertainty remains significant for 100-year or longer return periods. Incorporating climate projections, regional climate model outputs, or machine learning-based simulations could further improve reliability of extreme rainfall predictions in future studies.

Overall, the bootstrap analysis demonstrates that uncertainty in extreme rainfall estimates increases rapidly with return period and that the GEV distribution provides the most conservative and realistic representation of extreme rainfall variability in Delhi. These findings emphasize the need for incorporating uncertainty analysis into hydrological design and flood risk assessment to ensure resilient urban infrastructure under changing climatic conditions.

Table 5 Quantification of uncertainty in extreme rainfall return levels Using different probability distributions

Return Period (yr)	Gumbel RL (95% CI)	GEV RL (95% CI)	LP3 RL (95% CI)
2	239.1	242.6	244
5	328.7	340.3	333.6
10	386.4	408.2	388.2
25	463.2	507.4	456.9
50	523.4	589.5	512.1
100	582.1	671.7	563.4

3.5. Machine Learning Approach for Extreme Rainfall Prediction

The performance of the machine learning models applied to the Annual Maximum Series (AMS) of Delhi rainfall (1901–2021) is summarized in Table 6. Among the tested models, Gradient Boosting demonstrated the best overall performance with the highest coefficient of determination

(R = 0.86) and Nash–Sutcliffe Efficiency (NSE = 0.85), along with the lowest prediction errors (RMSE = 41 mm, MAE = 32 mm). These results indicate strong agreement between observed and predicted extreme rainfall values. The Random Forest model also demonstrated strong predictive performance (R = 0.82, NSE = 0.80), whereas Linear Regression performed comparatively poorly (R = 0.62, NSE = 0.61), highlighting the limitations of linear assumptions for capturing nonlinear rainfall dynamics.

Table 6 *Accuracy comparison of linear and ensemble machine learning models*

Model	**R**	**RMSE (mm)**	**MAE (mm)**	**NSE**
Linear Regression	0.62	64	51	0.61
Random Forest	0.82	42	39	0.8
Gradient Boosting	0.86	41	32	0.85

3.6 Model Performance and Prediction Accuracy

Figure 6 compares observed and predicted AMS rainfall using the Gradient Boosting model. The predicted values closely follow the interannual variability of observed rainfall, indicating that the model effectively captures both moderate and high-magnitude rainfall events. The ability of Gradient Boosting to model nonlinear relationships and complex interactions among rainfall predictors explains its superior performance compared with linear regression methods.

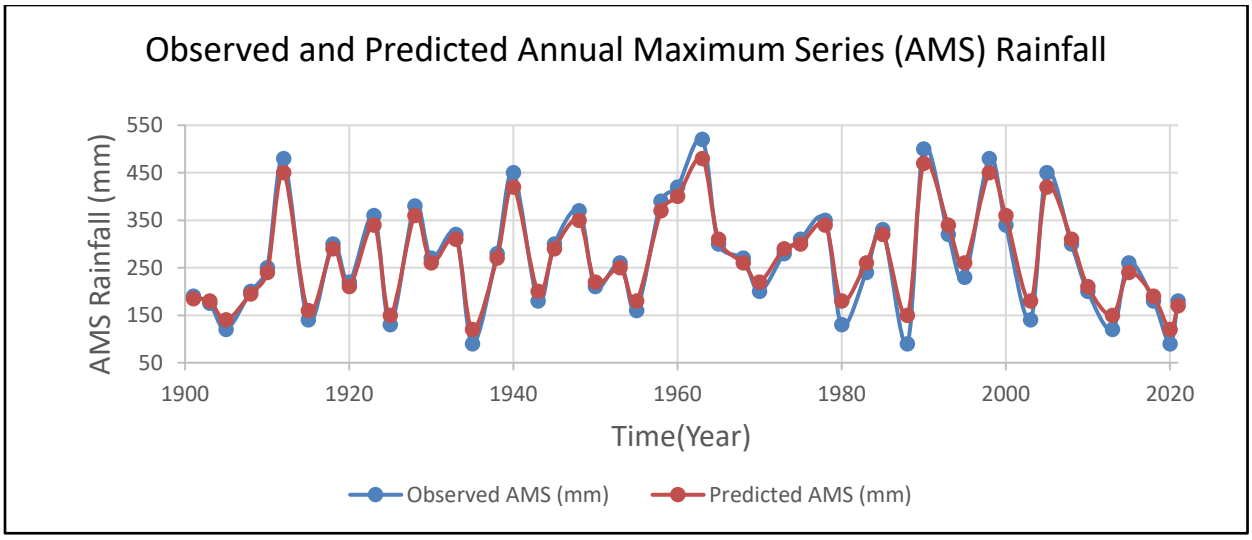


Figure 6. *Observed vs. Predicted Annual Maximum Rainfall (AMS) using Gradient Boosting Model*

The model slightly underestimates some very extreme rainfall events, a common issue in machine learning models trained on limited extreme-event samples. Nevertheless, the overall agreement between observed and predicted rainfall demonstrates the suitability of ensemble learning methods for modeling extreme rainfall. The temporal distribution of residuals obtained from the Gradient Boosting model is illustrated in Figure 7.

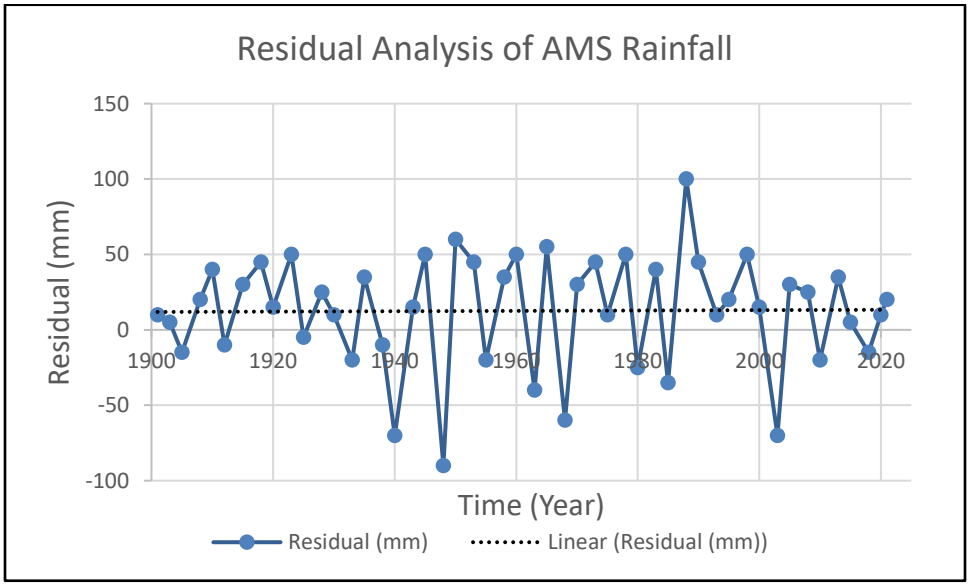


Figure 7. Temporal Distribution of Residuals for Gradient Boosting Rainfall Prediction

406 *3.7 Quantile-Based Extreme Rainfall Estimates*

407 **Table 7** Quantile-based estimates of extreme rainfall (mm)

Quantile / Return Level	**Predicted rainfall (mm)**
Q0.5 (Median)	272
Q0.9	395
Q0.95	452
Q0.99	534

408

409 Table 7 presents quantile-based estimates of extreme rainfall derived from the Gradient Boosting
410 model. The predicted rainfall values increase rapidly toward higher quantiles, with estimates of
411 272 mm (median), 395 mm (Q0.9), 452 mm (Q0.95), and 534 mm (Q0.99). This nonlinear increase
412 confirms the heavy-tailed nature of extreme rainfall in Delhi and aligns with the results obtained
413 from GEV frequency analysis.
414 These quantile estimates are particularly useful for flood risk assessment, as they provide rainfall
415 magnitudes corresponding to high-probability extreme events. Compared with traditional
416 frequency analysis, machine learning models can incorporate additional predictors such as
417 temperature, humidity, and large-scale climate indices, improving predictive capability under non-
418 stationary climate conditions.

419 *3.8 Model Performance Comparison*

420 Table 8 and Figure 8 present a comparative evaluation of traditional frequency distribution models
421 (Gumbel, GEV, and LP–III) and the Machine Learning (ML) model using three performance
422 metrics: coefficient of determination (R), root mean square error (RMSE), and mean absolute error
423 (MAE). These metrics collectively assess goodness-of-fit, prediction accuracy, and error
424 magnitude in modeling extreme rainfall events.

425

Table 8 Model performance metrics for statistical distributions and machine learning prediction

Model	R	RMSE (mm)	MAE (mm)
Gumbel Distribution	0.78	51	36
GEV Distribution	0.84	44	31
LP–III Distribution	0.8	48	34
Machine Learning Model	0.86	42	28

As shown in Table 8, the Gumbel distribution exhibits the lowest performance among the considered models, with an R of 0.78 and the highest RMSE (51 mm) and MAE (36 mm), indicating limited capability in capturing variability and extremes in the rainfall data. The GEV distribution demonstrates improved performance (R = 0.84, RMSE = 44 mm, MAE = 31 mm), highlighting its flexibility in representing different tail behaviors. The LP–III distribution shows moderate performance, with R = 0.80 and intermediate error values (RMSE = 48 mm, MAE = 34 mm), suggesting reasonable but less accurate predictions than GEV.

The Machine Learning model outperforms all conventional frequency distributions, achieving the highest R of 0.86 and the lowest RMSE (42 mm) and MAE (28 mm). This indicates superior predictive accuracy and a better ability to capture nonlinear relationships inherent in extreme rainfall processes. Ensemble learning methods such as Gradient Boosting and Random Forest can identify complex interactions in rainfall data that are difficult to represent using parametric statistical distributions.

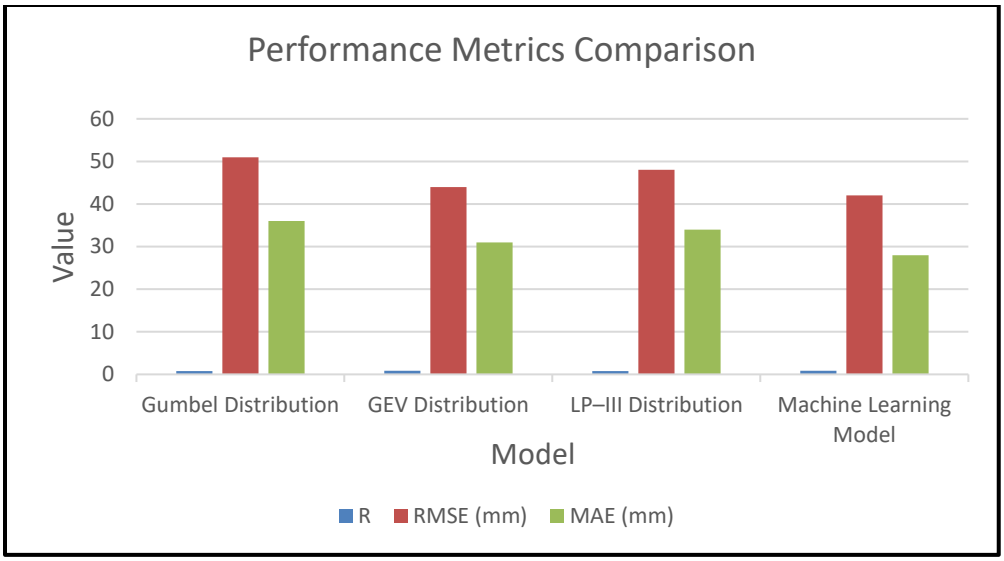


Figure 8. Frequency Distribution Vs Machine Learning

4. Conclusion

This study presents an integrated framework combining classical frequency analysis and machine learning to evaluate extreme rainfall in Delhi using long-term data (1901–2021). Trend analysis of the Annual Maximum Series reveals a statistically significant increasing pattern, with Sen's slope indicating a rise of approximately 0.3-0.4 mm/year, suggesting gradual intensification of extreme rainfall over time.

Among the statistical models, the Generalized Extreme Value (GEV) distribution provides the best fit and predicts higher rainfall at longer return periods, reflecting the heavy-tailed nature of monsoon-driven extremes. Bootstrap analysis shows increasing uncertainty at higher return periods, emphasizing the importance of incorporating uncertainty in hydrological design.

Machine learning models outperform traditional methods, with Gradient Boosting achieving the highest accuracy ($R \approx 0.86$, $NSE \approx 0.85$) and effectively capturing nonlinear rainfall patterns. Residual analysis confirms the reliability of the model in representing extremes.

Overall, the findings indicate increasing variability in extreme rainfall, which may elevate urban flood risk. These results highlight the need for updated design standards and the adoption of hybrid statistical–machine learning approaches for climate-resilient urban water management.

CRediT authorship contribution statement

Owais Ahmad Parray: Conceptualization, Formal analysis , Software analysis, Writing – Original Draft. **Adnan Ilahi Bhat:** Formal analysis , Software analysis, Writing – review & editing. **Ajmal Hussain:** Conceptualization, Supervision, Administration, Writing – review & editing. **Shahbaz Ahmad**: Conceptualization, Writing – review & editing. **Mujib Ahmad Ansari:** Conceptualization, Supervision, Writing –Editing.

Funding
Open Access funding enabled and organized by Projekt DEAL.

Data availability statement
The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Clinical Trial Number
Not applicable.

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Declaration of Competing Interest
The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

REFERENCES

1. Tandon, A., Pattnayak, K.C. and Awasthi, A., 2026. Integrating IMDAA regional reanalysis and machine learning for enhanced detection of extreme precipitation over complex Himalayan terrain. *Earth Systems and Environment.*
2. Tandon, A., Awasthi, A. and Pattnayak, K.C., 2025. Efficacy of machine learning in simulating precipitation and its extremes over the capital cities in North Indian states. *Scientific Reports*, 15, p.10345.
3. Kumar, V., Kedam, N., Kisi, O. and Singh, P., 2025. A comparative study of machine learning models for daily and weekly rainfall forecasting. *Water Resources Management*, 39, pp.271–290.
4. Das, P., Zhang, Z., Ghosh, S. and Mishra, V., 2024. A hybrid ensemble learning merging approach for enhancing drought computation. *Scientific Reports*, 14, p.13870.
5. Alkhatib, M.I.I., Talei, A., Chang, T.K., Hermawan, A.A. and Pauwels, V.R.N., **2024**. Towards the development of a citizens' science-based acoustic rainfall sensing system. *Journal of Hydrology*, 633, 130973.
6. Zhang, Y., Yao, W., Yuan, C., Chen, N. and Levia, D.F., **2024**. Alterations of throughfall and stemflow chemistry by vegetation in China: Insights from a machine learning-based meta-analysis. *Journal of Hydrology*, 641, 131786
7. Anuradha, T., Aruna Sri, P.S.G. and RamaDevi, J., 2024. Hybrid model for rainfall prediction with statistical and technical indicator feature set. *Expert Systems with Applications*, 249, p.123260.

8. Saini, S.K., Mahato, S., Pandey, D.N. and Joshi, P.K., 2023. Modeling flood susceptibility zones using hybrid machine learning models in India. *Environmental Science and Pollution Research*, 30, pp.97463–97485.
9. Singh, V.P., 2018. *Hydrology and Hydraulic Systems*. Waveland Press.
10. Roxy, M.K., et al., 2017. A threefold rise in widespread extreme rain events over central India. *Nature Communications*, 8.
11. Prakash, S., et al., 2015. IMERG precipitation product evaluation over India. *Journal of Hydrology*, 541.
12. Gupta, A.K., and Nair, S.S., 2015. Urban floods in Bangalore and Chennai: risk management challenges. *Natural Hazards*, 75.
13. Pai, D.S., et al., 2014. High-resolution daily gridded rainfall dataset for India. *Meteorological Applications*, 21.
14. Adamowski, J., Chan, H.F., Prasher, S.O., Ozga-Zielinski, B., & Sliusarieva, A., 2012. Comparison of multiple linear and nonlinear regression... *Water Resources Research*, 48.
15. Jain, S.K., & Kumar, V., 2012. Trend analysis of rainfall in India. *Hydrological Sciences Journal*, 57.
16. Guhathakurta, P., et al., 2011. Trends in rainfall pattern over India. *Meteorological Applications*, 18.
17. Kumar, V., Jain, S.K., & Singh, Y., 2010. Analysis of long-term rainfall trends in India. *Hydrological Sciences Journal*, 55.
18. Douglas, I., Alam, K., Maghenda, M., McDonnell, Y., McLean, L., & Campbell, J., 2008. Unjust waters... *Environment and Urbanization*, 20.
19. Rajeevan, M., et al., 2008. Analysis of extreme rainfall events over India. *Meteorological Applications*, 15.
20. Milly, P.C.D., et al., 2008. Stationarity is dead: Whither water management? *Science*, 319.
21. Goswami, B.N., Venugopal, V., Sengupta, D., Madhusoodanan, M.S., & Xavier, P.K., 2006. Increasing trend of extreme rain events over India. *Science*, 314.
22. Ashok, K., Guan, Z., & Yamagata, T., 2004. Influence of the Indian Ocean Dipole on the Indian summer monsoon. *Geophysical Research Letters*, 31.
23. Kripalani, R.H., & Kulkarni, A., 2001. Monsoon rainfall variations and teleconnections. *International Journal of Climatology*, 21.
24. Haan, C.T., 2002. *Statistical Methods in Hydrology*. Iowa State Press.
25. Koutsoyiannis, D., Kozonis, D., & Manetas, A., 1998. Rainfall IDF relationships. *Journal of Hydrology*, 206.
26. Kendall, M.G., 1975. *Rank Correlation Methods*. Griffin.
27. Sen, P.K., 1968. Estimates of regression coefficient based on Kendall's tau. *Journal of the American Statistical Association*, 63.
28. Chow, V.T., Maidment, D.R., & Mays, L.W., 1988. *Applied Hydrology*. McGraw-Hill.
29. Mann, H.B., 1945. Nonparametric tests against trend. *Econometrica*, 13.