

A Bivariate Gaussian–Gamma Mixture Model for Rainfall Distribution and Extreme Event Analysis over Karnataka, India

KUMUDHA HR

kumudhamayur@gmail.com

Bharathi College Post Graduate & Research Centre

KOKILA RAMESH

Jain University

ANGEL RICHARD

JAIN (Deemed-to-be University)

Research Article

Keywords: Rainfall, Monsoon, Mixture Model, Drought and flood prone regions, prediction

Posted Date: September 29th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10996172/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

A Bivariate Gaussian–Gamma Mixture Model for Rainfall Distribution and Extreme Event Analysis over Karnataka, India

KUMUDHA H. R.^{1*}, KOKILA RAMESH² and ANGEL RICHARD³

¹Professor, Bharathi College PG & RC, Mandya, India. Founder, StatMint Innovations LLP, Mandya (kumudhamayur@gmail.com)

²Associate Professor, Jain University, Bangalore, India (r.kokila@jainuniversity.ac.in)

³Research Scholar, JAIN (Deemed-to-be University), Bangalore, Karnataka, India

*Corresponding author's email: kumudhamayur@gmail.com

ABSTRACT

The study examines a statistical analysis of Southwest Monsoon rainfall over selected regions of Karnataka using long-term data from 1960 to 2025. Rainfall in monsoon-dominated regions often exhibits non-Gaussian characteristics such as skewness and extreme variability, which are not adequately captured by conventional Gaussian models. To address this limitation, a Gaussian–Gamma mixture model is applied to represent both normal rainfall behaviour and extreme events within a unified framework. Descriptive statistics and normalized rainfall analysis were carried out to examine regional variability. Model parameters were estimated using an Expectation–Maximization approach to ensure accurate representation of the observed data. The model performance was evaluated by comparing statistical moments and level-crossing characteristics with those of the observed rainfall series. The results indicate that the mixture model provides improved representation of skewness, kurtosis, and extreme rainfall when compared to the Gaussian model, particularly in high rainfall regions. A bivariate extension of the model was further used to examine temporal dependence in rainfall data, showing stronger persistence in coastal regions and weaker dependence in interior regions. In addition, the proposed framework was used to predict SWM rainfall for the year 2026 across the selected regions. The study demonstrates that the proposed approach effectively captures rainfall variability and extremes, making it useful for applications in flood risk assessment and drought monitoring, in monsoon regions.

Keywords: Rainfall, Monsoon, Mixture Model, Drought and flood prone regions, prediction

1. Introduction

Rainfall variability plays a crucial role in influencing agriculture, water resources, and disaster management in monsoon-dominated regions of India. In Karnataka, significant spatial contrasts exist between coastal and interior subdivisions, with coastal regions receiving high and persistent rainfall, while interior regions frequently experience irregular rainfall and drought conditions. The Southwest Monsoon contributes the major share of annual rainfall, and its variability directly affects cropping patterns, reservoir storage, and hydrological extremes. Understanding both rainfall variability and extreme behaviour is therefore essential for effective planning and risk management.

Traditional statistical approaches commonly assume that rainfall follows a Gaussian distribution. However, monsoon rainfall often exhibits non-Gaussian characteristics such as skewness, heavy tails, and the occurrence of extreme events, which limit the applicability of Gaussian models in hydrological analysis. These limitations are particularly evident when analysing high rainfall events and drought conditions. Several studies have highlighted these issues. Iyengar and Ramesh (2005) demonstrated that Indian monsoon rainfall deviates significantly from normality, while Katz and Parlange (1998) discussed

over-dispersion in precipitation processes. Wilks (2011) further emphasized the heavy-tailed nature of rainfall distributions. To address these limitations, alternative probabilistic approaches have been proposed. Glasbey and Nevison (1997) introduced latent Gaussian models for rainfall simulation, and Li et al. (2012) developed Bayesian approaches for rainfall estimation. More recent studies, including Ayar et al. (2020) and Hussein and Kadhém (2022), have applied advanced statistical techniques to capture spatial and temporal variability. In the Indian context, Kokila and Iyengar (2017) emphasized the importance of mixture-based models for representing monsoon rainfall behaviour more realistically.

In addition to statistical approaches, machine learning techniques have been widely applied for rainfall prediction. Artificial Neural Network (ANN) models have been used to forecast seasonal rainfall in Karnataka (Kumudha and Ramesh, 2023), and subsequent studies have extended these methods to drought and flood prediction using ANN and Kolmogorov–Arnold Networks with standardized drought indices (Kumudha et al., 2025a; Kumudha et al., 2025b). While these methods provide strong predictive performance, they primarily focus on forecasting accuracy and do not explicitly represent the underlying probability distribution of rainfall,

particularly higher-order moments and extreme behaviour. Studies published in *Mausam* have also highlighted the importance of probabilistic approaches for analysing monsoon rainfall over India, particularly in capturing regional variability and non-Gaussian characteristics. These studies emphasize the limitations of conventional Gaussian models and support the use of flexible statistical frameworks for improved representation of rainfall variability and extremes. Recent research has further focused on modelling rainfall dependence and variability using advanced probabilistic frameworks. Copula-based approaches have been used to capture dependence structures and joint behaviour in rainfall processes (Khedun, 2014; Serinaldi, 2009), while space–time stochastic models have demonstrated the importance of incorporating temporal and spatial dependence in rainfall modelling (Cantet et al., 2025). These developments highlight the need for models that can represent both variability and extreme rainfall behaviour.

In a study on flood-prone regions of Karnataka, Kumudha and Ramesh (2023) investigated the statistical properties and distributional behaviour of southwest monsoon rainfall data. The study employed descriptive statistics, stationarity analysis, maximum likelihood estimation, and goodness-of-fit testing to evaluate the suitability of different probability distributions. The findings revealed substantial skewness and variability in rainfall patterns, demonstrating that Gaussian assumptions were insufficient for modelling extreme rainfall events. The study therefore emphasized the need for more flexible distribution-based frameworks capable of capturing rainfall heterogeneity and extremes, providing a foundation for subsequent mixture-model approaches

2. Data and Methodology

2.1 Study Area and Data

Karnataka is one of the major monsoon-influenced states in southern India and exhibits considerable spatial variability in rainfall due to differences in topography, proximity to the Arabian Sea, and prevailing monsoon circulation patterns. Based on meteorological characteristics, the state is broadly divided into three subdivisions: Coastal Karnataka, North Interior Karnataka, and South Interior Karnataka.

These regions experience distinct rainfall regimes ranging from extremely high rainfall along the Western Ghats and coastal belt to comparatively low rainfall in the semi-arid interior districts. To investigate rainfall variability under contrasting climatic conditions, ten representative districts were selected from both flood-prone and drought-prone

to rainfall analysis. Recent studies have also demonstrated the effectiveness of Gaussian–Gamma mixture models in representing rainfall variability in both drought-prone and flood-prone regions of Karnataka (Kumudha and Ramesh, 2025). In these models, the Gaussian component captures moderate variations, while the Gamma component represents extreme rainfall behaviour, enabling improved characterization of skewness and tail properties. Despite these developments, most existing studies are limited to univariate modelling and do not explicitly capture temporal dependence in rainfall processes. Rainfall often exhibits persistence, particularly in coastal regions, where wet and dry conditions tend to continue over successive periods. This limitation motivates the need for models that jointly represent distributional characteristics and dependence structure.

In this study, a bivariate Gaussian–Gamma mixture model is proposed to analyse rainfall variability and extreme events across Karnataka. The model integrates both probabilistic distribution and temporal dependence, enabling a more comprehensive representation of rainfall behaviour. The proposed framework is applied to long-term Southwest Monsoon rainfall data, and its performance is evaluated using statistical moments and level-crossing characteristics. The specific contributions of this study are: (i) development of a bivariate mixture framework for rainfall modelling, (ii) improved representation of skewness and extreme rainfall behaviour, and (iii) assessment of temporal dependence across regions. The objective is to provide a robust statistical framework for analysing rainfall variability and extremes, with applications in drought assessment, flood risk analysis, and water resource management in monsoon regions.

regions of Karnataka. The flood-prone regions include Dakshina Kannada, Udupi, Uttara Kannada, Kodagu, Chikkamagaluru, and Shivamogga, which receive substantial Southwest Monsoon rainfall due to the orographic influence of the Western Ghats.

These regions frequently experience intense rainfall events, floods, and landslides during active monsoon periods. The drought-prone regions include Bagalkote, Chitradurga, Koppala, and Raichur, which are characterized by relatively low and highly variable rainfall and are often affected by moisture stress and recurring drought conditions. The geographical distribution of the selected study regions is shown in Fig. 1. The analysis is based on long-term Southwest Monsoon (SWM) rainfall data covering the period from 1960 to 2025.

The Southwest Monsoon season extends from June to September and contributes a major portion of the annual rainfall received in Karnataka. Monthly rainfall observations were obtained from reliable meteorological sources, including the India Meteorological Department (IMD), the Indian Institute of Tropical Meteorology (IITM), and the Karnataka State Natural Disaster Monitoring Centre (KSNDMC). For each district, monthly rainfall totals corresponding to the monsoon season were aggregated to obtain seasonal Southwest Monsoon rainfall values.

Prior to analysis, the dataset was subjected to basic quality-control procedures to ensure consistency and reliability. Missing observations were minimal and were appropriately handled using seasonal averages. The rainfall data were then examined using descriptive statistical measures to understand their distributional characteristics. Since rainfall data often exhibit skewness, heavy tails, and extreme events, a Gaussian–Gamma Mixture Model (GGM) framework was adopted to provide a more realistic representation

2.2 Data Transformation

Rainfall data are inherently non-negative and typically exhibit significant skewness and variability, particularly during monsoon seasons. These characteristics make direct application of standard statistical models difficult, especially those based on normality assumptions. In addition, the presence of occasional extreme rainfall values further emphasizes the need for appropriate data transformation.

To address these issues, the rainfall data (R_i) are normalized using the long-term mean rainfall (m_R) and transformed using a logarithmic transformation given by:

$$Z_i = \log\left(\frac{R_i}{m_R}\right) \quad \text{--- (1)}$$

of rainfall variability across both flood-prone and drought-prone regions of Karnataka.

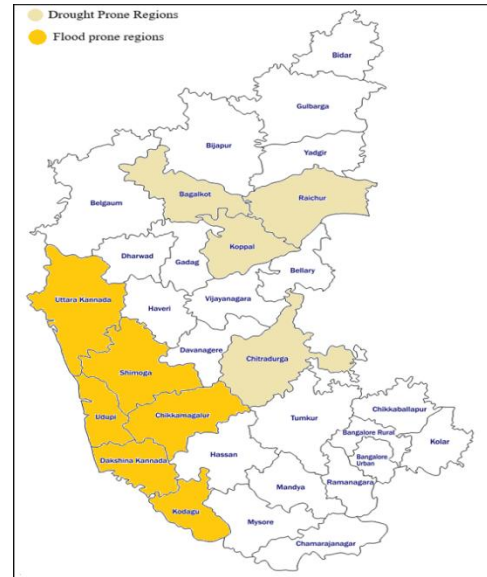


Fig. 1. Karnataka map with drought & Flood – prone regions

where (R_i) represents the observed rainfall and (m_R) denotes the long-term average rainfall.

This transformation stabilizes the variance, reduces skewness, and allows the data to take both positive and negative values. As a result, the transformed series becomes more suitable for probabilistic modelling and mixture distribution analysis.

The statistical characteristics of the original rainfall data are presented in Table 1, which indicate significant skewness and variability across regions, confirming the non-Gaussian nature of rainfall. After applying the logarithmic transformation, the normalized rainfall statistics are shown in Table 2. It is observed that the transformation reduces skewness and stabilizes variance, making the data more appropriate for statistical modelling.

TABLE 1

Basic statistics of South-West Monsoon (June–September) rainfall for the four drought and flood-prone districts of Karnataka (1960–2025). LTA: Long-Term Average rainfall LTD: Long-Term Deviation (standard deviation), Skewness and Kurtosis describe distributional shape.

Sub division	LTA (m_z) (in cm)	LTD (σ_z) (in cm)	Skewness (S_z)	Kurtosis (K_z)
Dakshina Kannada	177.39	61.19	1.65	2.37
Udupi	142.16	96.49	2.32	3.93
Uttara Kannada	272.21	38.26	0.50	0.23
Chikkamagalur	106.13	25.07	0.88	3.73
Kodagu	62.47	14.89	0.77	1.07
Shivamogga	97.30	24.93	0.69	0.52
Bagalkote	20.39	8.85	1.38	2.21
Chithradurga	21.33	14.69	4.56	2.70
Koppala	18.04	9.65	1.90	3.86
Raichur	25.25	9.94	1.51	2.81

TABLE 2

The table summarizes the first four statistical moments of the normalized SWM rainfall series, including Long-Term Average deviation (LTA), Long-Term Dispersion (LTD), skewness, and kurtosis. Normalization was performed using district-wise long-term mean rainfall to enable comparison of rainfall variability, asymmetry, and tail behavior among coastal and interior districts of Karnataka.

Sub division	LTA (m_{z_i})	LTD (σ_{z_i})	Skewness (S_{z_i})	Kurtosis (K_{z_i})
Dakshina Kannada	-0.05	0.30	0.85	0.64
Udupi	-0.14	0.46	1.85	2.37
Uttara Kannada	-0.01	0.14	0.09	-0.02
Chikkamagalur	-0.03	0.23	0.22	0.04
Kodagu	-0.03	0.23	0.01	0.06
Shivamogga	-0.03	0.25	0.03	-0.39
Bagalkote	-0.08	0.40	0.13	0.24
Chithradurga	-0.13	0.47	0.75	2.71
Koppala	-0.11	0.45	0.55	0.29
Raichur	-0.07	0.36	0.26	0.49

Descriptive measures such as mean, standard deviation, skewness, and kurtosis are computed for both the original and transformed datasets to evaluate the effectiveness of the transformation. The

transformed variable Z_i (for $i = 1, 2, \dots, n$) is subsequently used for model formulation and parameter estimation in the Gaussian–Gamma mixture framework.

2.3 Gaussian–Gamma Mixture Model

Rainfall data often exhibit a combination of moderate variations and extreme events, which cannot be adequately, represented using a single probability distribution. To address this, a Gaussian–Gamma

Mixture (GGM) model is employed to capture both normal rainfall behaviour and extreme rainfall characteristics within a unified probabilistic framework.

The probability density function of the mixture model is given by:

$$f(z) = wf_G(z; \mu, \sigma^2) + (1 - w)f_\Gamma(z; \alpha, \beta) \quad (2)$$

where $0 \leq w \leq 1$ is the mixing weight, $f_G(z; \mu, \sigma^2)$ denotes the Gaussian distribution with mean (μ) and variance (σ^2), and ($f_\Gamma(z; \alpha, \beta)$) denotes the Gamma distribution with shape parameter (α) and scale parameter (β). The conditional mean and variance of the process are given by:

distribution with shape parameter (α) and scale parameter (β).

$$m_{z|u} = wm_x + (1 - w)m_y \text{ and } \sigma_z^2 = w^2\sigma_x^2 + (1 - w)^2\sigma_y^2 \quad (3)$$

where m_x, m_y, σ_x^2 , and σ_y^2 represent the means and variances of the Gaussian and Gamma components

respectively. The Gaussian probability density function is given by:

$$f_G(z) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(z_i - \mu)^2}{2\sigma^2}\right)$$

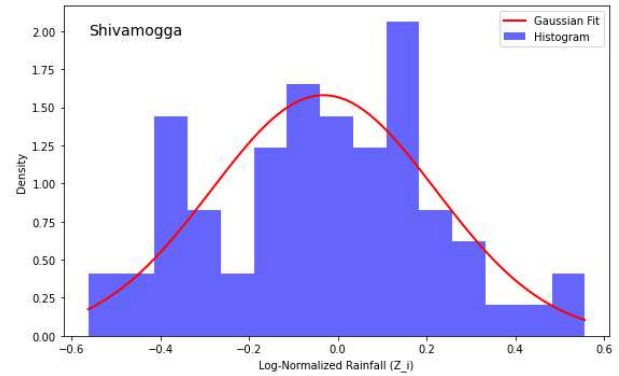
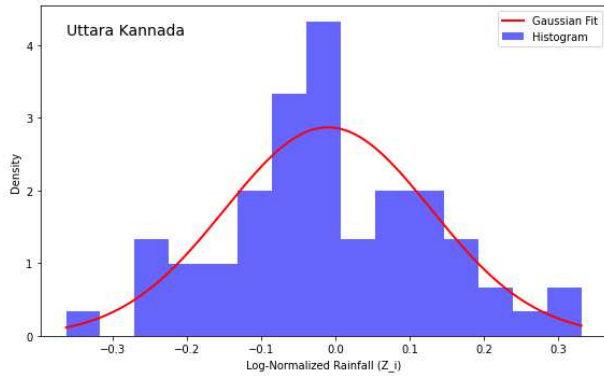
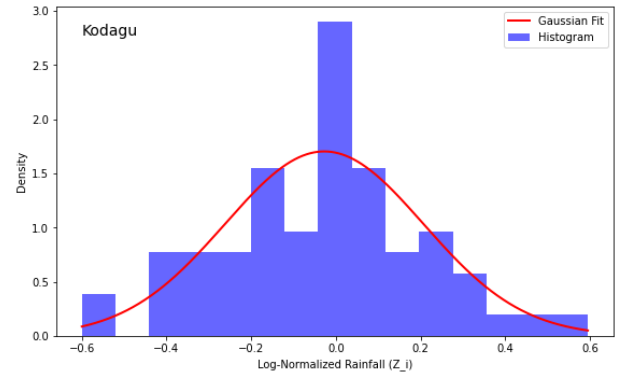
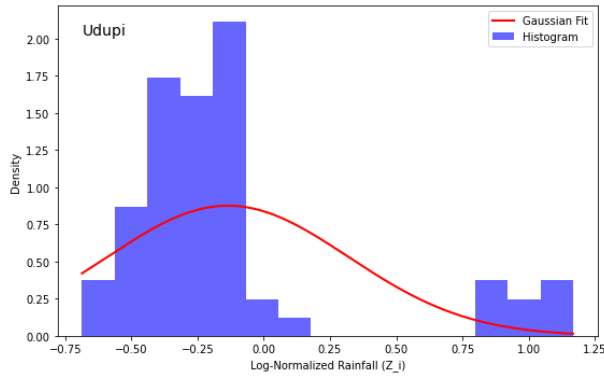
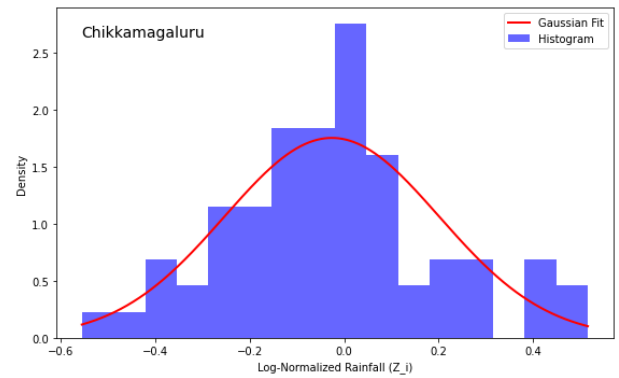
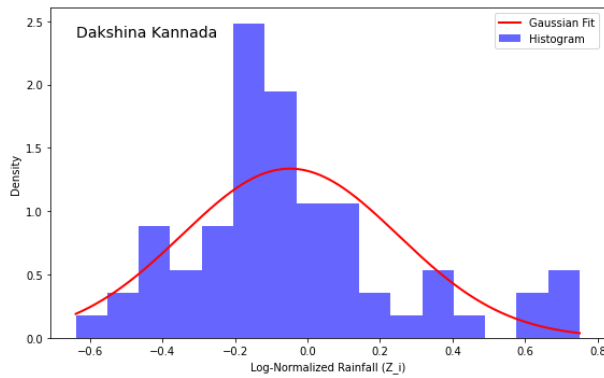
The Gamma probability density function is expressed as:

$$f_\Gamma(z) = \frac{1}{\Gamma(\alpha)\beta^\alpha} z_i^{\alpha-1} \exp\left(-\frac{z_i}{\beta}\right), z > 0$$

$$p(z) = \frac{w}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(z_i - \mu)^2}{2\sigma^2}\right) + \frac{(1 - w)}{\Gamma(\alpha)\beta^\alpha} z_i^{\alpha-1} \exp\left(-\frac{z_i}{\beta}\right), z > 0 \quad (4)$$

The Gaussian component represents moderate rainfall fluctuations around the mean, while the Gamma component captures the skewed and heavy-tailed behaviour associated with extreme rainfall events. The combination of these two components allows for a flexible representation of rainfall variability. In practice, the Gamma component is applied to the positive domain of the transformed rainfall data, while the Gaussian component accounts for the overall variability of the dataset. The parameters of the mixture model include the mixing weight (w), Gaussian parameters (μ, σ^2) and Gamma parameters (α, β) these parameters are estimated using an Expectation–Maximization (EM) algorithm, which

iteratively updates the parameter values to maximize the likelihood of the observed data. The Gaussian–Gamma mixture model provides a flexible and effective framework for modelling rainfall data, as it captures both central tendency and extreme behaviour. This makes it particularly suitable for analysing rainfall variability in monsoon-dominated regions. The frequency distribution of the transformed rainfall data is illustrated in Fig. 2. The histogram reveals that the data deviates from normality, exhibiting noticeable skewness and indications of bimodal behaviour. This departure from Gaussian assumptions justifies the use of a Gaussian–Gamma mixture model for effectively capturing the underlying rainfall distribution.



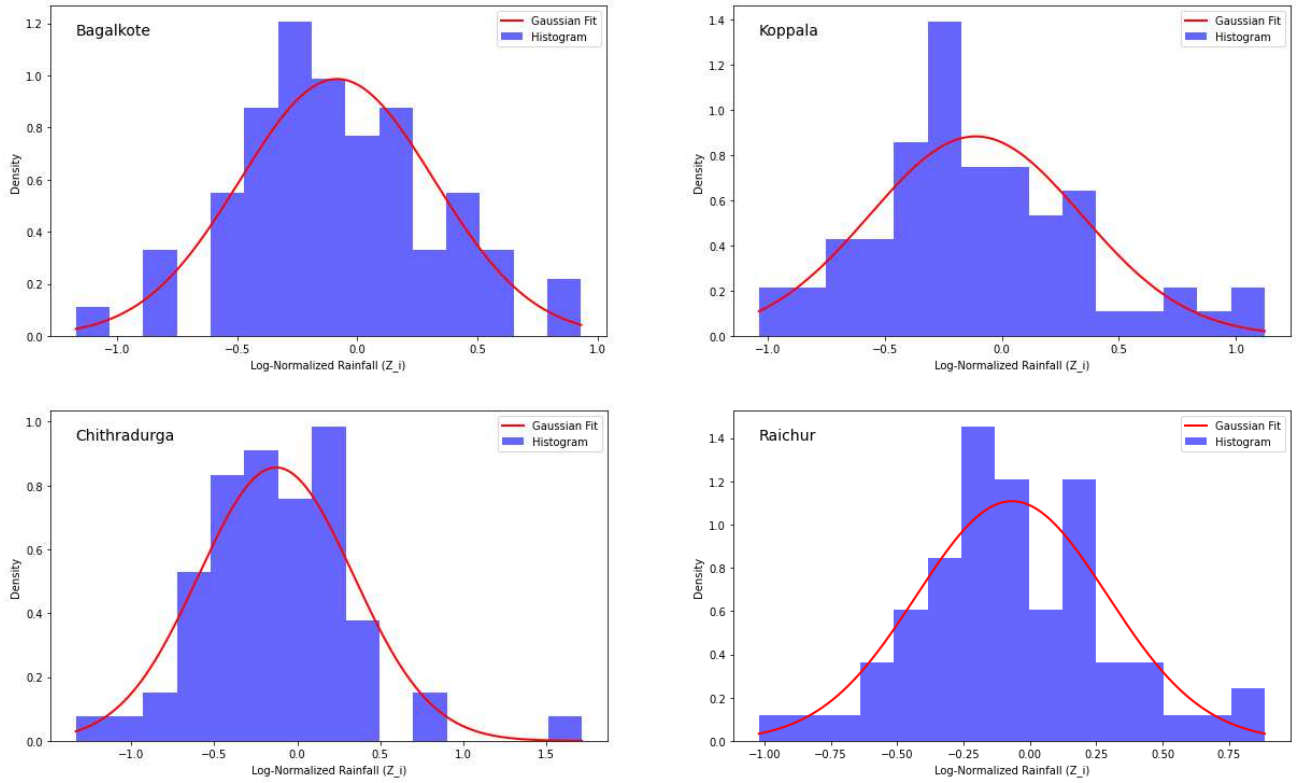


Fig. 2. Histograms of normalized Southwest Monsoon (SWM) rainfall data for Dakshina Kannada and Udupi (flood-prone/coastal) and Raichur and Chithradurga (drought-prone/interior) districts of Karnataka, overlaid with the fitted Gaussian–Gamma Mixture (GGM) distribution and a Gaussian density curve for comparison. The normalized data enable visualization of distributional characteristics such as skewness and tail behaviour, independent of district mean rainfall levels.

$$L(\mu, \sigma^2, \alpha, \beta) = \prod_{i=1}^n [w_i * f_{Gaussian}(x: \mu, \sigma^2) + (1 - w_i) * f_{Gamma}(y: \alpha, \beta)] \quad (5)$$

Maximizing the likelihood function directly can be computationally difficult due to the nonlinear nature of the mixture distribution. Therefore, it is convenient to

The log-likelihood function is expressed as:

$$\ln[L(\mu, \sigma^2, \alpha, \beta)] = \sum_{i=1}^n \ln \left[\frac{w}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(z_i - \mu)^2}{2\sigma^2}\right) + \frac{(1-w)}{\Gamma(\alpha)\beta^\alpha} z_i^{\alpha-1} \exp\left(-\frac{z_i}{\beta}\right) \right] \quad (6)$$

To estimate the optimal model parameters, the log-likelihood function is maximized using the Expectation–Maximization (EM) algorithm. The EM procedure iteratively updates the Gaussian and Gamma components until convergence is achieved. The estimated parameters include the mixing weight w , Gaussian mean μ , Gaussian variance σ^2 , and the Gamma shape and scale parameters α and β . The estimated parameter values obtained for the selected flood-prone and drought-prone regions of Karnataka are presented in Table 3. The parameter w represents the contribution of the Gaussian component in the mixture distribution. The parameters μ and σ^2 describe the mean and variance of the Gaussian component, while α and β characterize the shape and scale of the Gamma component associated with

consider the logarithm of the likelihood function, which simplifies the optimization process while preserving the maximum likelihood estimates.

extreme rainfall behaviour. The results indicate clear regional variability in rainfall characteristics across Karnataka. Coastal regions such as Udupi and Uttara Kannada exhibit higher Gaussian weights, indicating relatively persistent rainfall behaviour with moderate variability. In contrast, interior drought-prone regions including Bagalkote, Chithradurga, Koppala, and Raichur show comparatively larger variance values and stronger Gamma contributions, reflecting increased rainfall irregularity and extreme-event characteristics. The estimated parameters demonstrate that the Gaussian–Gamma mixture framework effectively captures both normal monsoon variability and non-Gaussian rainfall extremes across different climatic regions of Karnataka.

TABLE 3

Estimated Gaussian–Gamma Mixture Model Parameters for Selected Regions of Karnataka.

Region	w	μ	σ^2	α	β
Dakshina Kannada	0.74	-0.18	0.03	0.55	0.19
Udupi	0.85	-0.30	0.02	0.55	0.17
Uttara Kannada	0.98	-0.01	0.02	0.04	0.10
Chikkamagalur	0.65	-0.03	0.05	0.55	0.01
Kodagu	0.85	-0.05	0.06	0.90	0.14
Shivamogga	0.62	-0.05	0.07	0.90	0.04
Bagalkote	0.73	-0.18	0.14	0.90	0.23
Chithradurga	0.79	-0.29	0.12	0.74	0.49
Koppala	0.68	-0.33	0.08	0.90	0.26
Raichur	0.82	-0.12	0.12	0.90	0.21

Equation (7) satisfies the necessary conditions of a valid probability density function and is subsequently applied to the transformed Southwest Monsoon rainfall dataset. The resulting probability density

function $P(R)$ combines the Gaussian and Gamma components to represent both moderate rainfall variability and extreme rainfall behaviour within a unified framework.

$$P(R) = \frac{1}{R} \left[\frac{w}{\sigma\sqrt{2\pi}} \exp\left(-\frac{\left(\log\left(\frac{R_i}{m_z}\right) - \mu\right)^2}{2\sigma^2}\right) + (1-w) \frac{z_i^{\alpha-1} \exp\left(-\frac{z_i}{\beta}\right)}{\Gamma(\alpha)\beta^\alpha} \right] - \dots - (7)$$

Figure 4 illustrates the comparison between the observed rainfall histogram and the fitted Gaussian–Gamma Mixture (GGM) probability density function. The visual comparison indicates that the proposed GGM model closely follows the observed rainfall distribution across both coastal and interior regions of Karnataka. Table 4 presents the model selection statistics for the proposed Gaussian–Gamma Mixture Model (GGM) estimated using the Expectation–Maximization (EM) algorithm. The final log-likelihood measures how well the fitted model explains the observed rainfall data, with larger log-likelihood values indicating a better fit. The Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) assess model quality by balancing goodness-of-fit against model complexity. Lower AIC and BIC values indicate a more parsimonious model with improved predictive capability. Collectively, these statistics demonstrate that the proposed GGM provides an efficient representation of the rainfall distribution across the selected districts. In particular, the model effectively captures the skewed structure and tail behaviour associated with monsoon rainfall data, which are not adequately represented by a single Gaussian distribution. Table 5 summarizes the results of the Kolmogorov–Smirnov (KS) and Anderson–Darling (AD) goodness-of-fit tests for the proposed Gaussian–Gamma Mixture Model. The KS test evaluates the maximum difference between the empirical and fitted cumulative distribution functions, whereas the Anderson–Darling test places greater emphasis on the lower and upper tails of the distribution. For all study regions, the p-values exceed

the 5% significance level, indicating that the null hypothesis cannot be rejected. Therefore, there is no statistically significant difference between the observed rainfall data and the fitted GGM distribution, confirming that the proposed model adequately represents the rainfall characteristics across the selected districts. **Table 3** presents the Pearson Chi-square goodness-of-fit test comparing the observed rainfall frequencies with the expected frequencies obtained from the fitted Gaussian–Gamma Mixture Model. The Chi-square statistic quantifies the discrepancy between the observed and expected frequency distributions. For the regions considered, the obtained p-values exceed the 5% significance level, indicating that the null hypothesis cannot be rejected. Hence, the fitted GGM provides an acceptable representation of the observed rainfall distribution. These results further support the adequacy of the proposed model for describing long-term rainfall variability. To further evaluate the performance of the model, the first four statistical moments namely mean, standard deviation, skewness, and kurtosis are computed for both the observed rainfall data and the fitted GGM model. The comparison is presented in Table 7. The results show that the estimated model moments closely agree with the corresponding observed moments for most regions, indicating that the proposed model preserves the major statistical characteristics of the rainfall data. The coastal regions such as Dakshina Kannada and Udupi exhibit relatively larger skewness and kurtosis values, reflecting the occurrence of persistent heavy rainfall and extreme monsoon events. The GGM model

successfully reproduces these higher-order moments, demonstrating its capability to represent heavy-tailed rainfall distributions. Similarly, interior drought-prone regions such as Bagalkote, Koppala, and Raichur show moderate skewness with comparatively lower rainfall variability, and the model captures these characteristics with reasonable accuracy. The comparison further reveals that the Gaussian–Gamma framework performs substantially better than a conventional Gaussian model in reproducing non-

Gaussian rainfall behaviour. In particular, the skewness and kurtosis values obtained from the fitted model remain close to the observed data, confirming the suitability of the mixture distribution for modelling asymmetric and extreme rainfall characteristics. Overall, the agreement between the observed and model-generated moments demonstrates that the proposed GGM approach provides an effective statistical representation of Southwest Monsoon rainfall over Karnataka.

TABLE 4

Log-Likelihood, Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC) Values for the Proposed Gaussian–Gamma Mixture Model Across the Selected Rainfall Regions of Karnataka

Region	Log-Likelihood	AIC	BIC
Dakshina Kannada	-15.18	40.36	51.23
Udupi	-32.53	75.06	85.93
Uttara Kannada	36.85	-63.70	-52.83
Chikkamagalur	6.21	-2.42	8.46
Kodagu	2.39	5.23	16.11
Shivamogga	-2.81	15.62	26.49
Bagalkote	-33.40	76.80	87.67
Chithradurga	-38.40	86.80	97.67
Koppala	-43.75	97.49	108.36
Raichur	-25.84	61.68	72.56

TABLE 5

Kolmogorov–Smirnov (KS) and Anderson–Darling (AD) Goodness-of-Fit Test Results for the Proposed Gaussian–Gamma Mixture Model

Region	KS Statistic	p-value	AD Statistic
Dakshina Kannada	0.12	0.40	-0.11
Udupi	0.12	0.27	-0.52
Uttara Kannada	0.14	0.57	-1.06
Chikkamagalur	0.19	0.22	-0.15
Kodagu	0.12	0.71	-0.63
Shivamogga	0.12	0.71	-0.69
Bagalkote	0.08	0.99	-0.75
Chithradurga	0.17	0.31	0.60
Koppala	0.11	0.85	0.27
Raichur	0.11	0.85	-0.57

TABLE 6

Chi-square Goodness-of-Fit Statistics Comparing the Observed and Expected Rainfall Distributions

Region	Chi-square Statistic
Uttara Kannada	2.95
Shivamogga	5.21
Bagalkote	2.80
Raichur	5.84

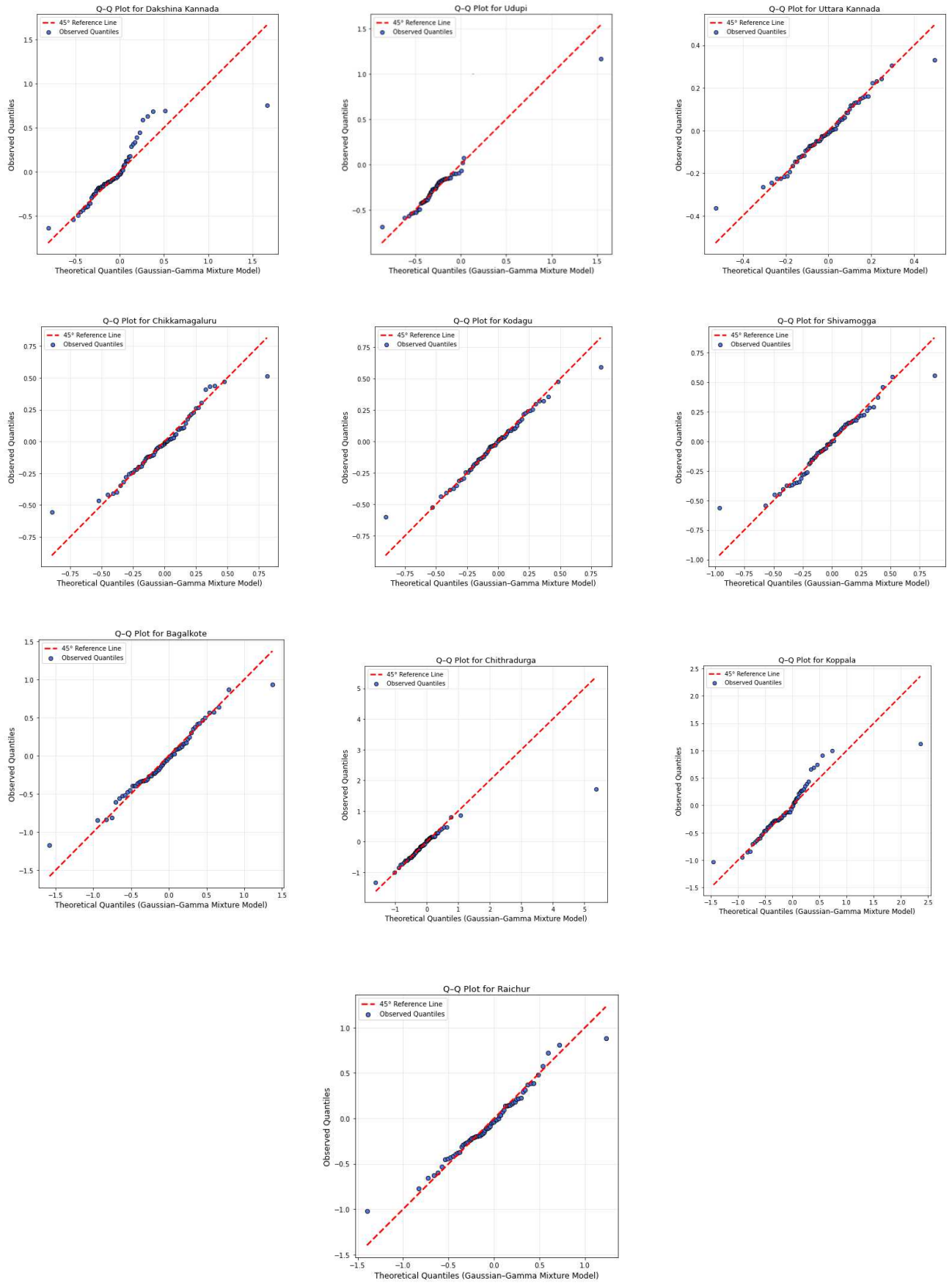


Fig 4. Quantile–Quantile (Q–Q) plot comparing the empirical rainfall quantiles with the theoretical quantiles obtained from the fitted Gaussian–Gamma Mixture Model.

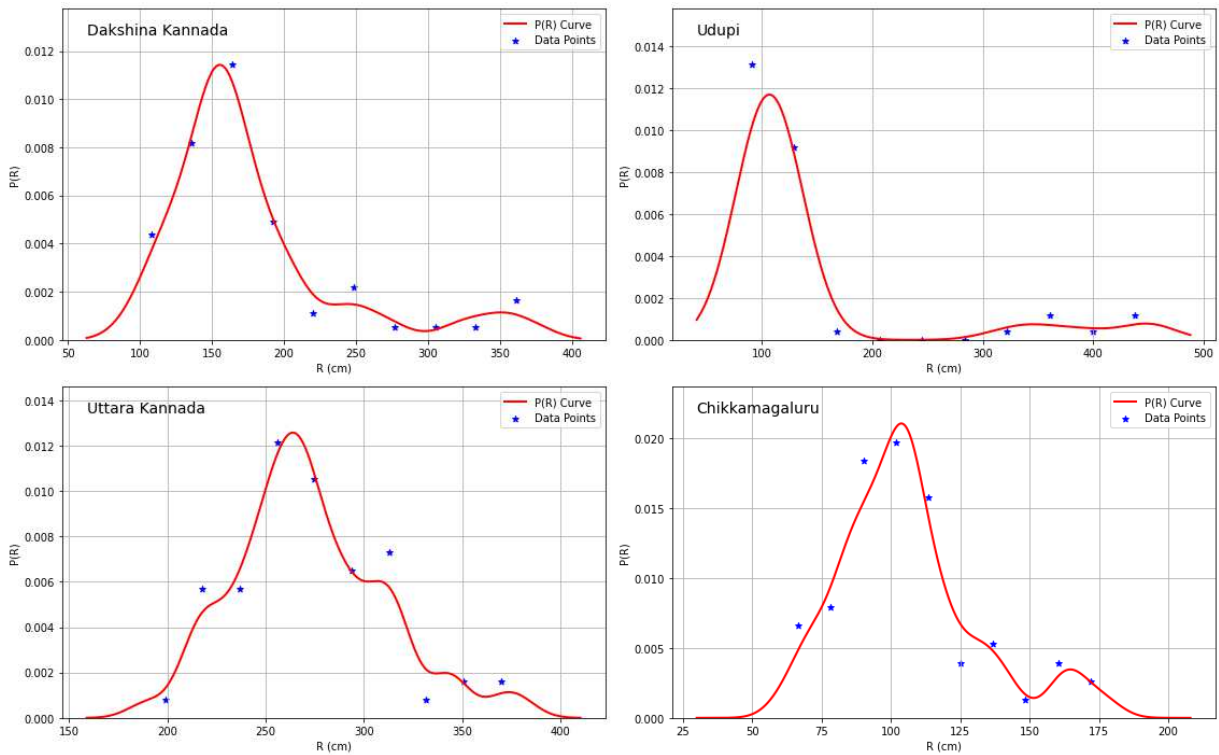
Fig 4 presents the Quantile–Quantile (Q–Q) plots comparing the empirical rainfall quantiles with the theoretical quantiles generated from the fitted Gaussian–Gamma Mixture Model. The majority of the observations lie close to the 45° reference line, indicating strong agreement between the observed rainfall distribution and the fitted model. Small deviations are observed at the extreme lower and

upper tails, reflecting the inherent variability associated with rare drought and extreme rainfall events. Overall, the close alignment of the empirical and theoretical quantiles confirms that the proposed GGM successfully captures the central tendency, dispersion, and overall distributional characteristics of long-term rainfall data.

TABLE 7

Comparison between the GGM model moments with the original data moments

Region/ Subdivision	Actual Moments (R)				Model Moments (R)			
	μ_R	σ_R	Sk_R	Ku_R	μ_{R_i}	σ_{R_i}	Sk_{R_i}	Ku_{R_i}
Dakshina Kannada	177.39	61.19	1.65	2.37	161.91	33.68	0.98	5.58
Udupi	142.16	96.49	2.32	3.93	135.93	25.43	1.39	3.85
Uttara Kannada	272.21	38.26	0.50	0.23	271.21	36.32	0.31	2.75
Chikkamagalur	106.13	25.07	0.88	3.73	105.74	22.29	0.52	3.13
Kodagu	62.47	14.89	0.77	1.07	62.48	13.89	0.52	3.28
Shivamogga	97.30	24.93	0.69	0.52	98.08	18.35	0.44	4.31
Bagalkote	20.39	8.85	1.38	2.21	20.09	7.21	0.77	4.06
Chithradurga	21.33	14.69	4.56	2.70	20.96	.33	2.96	2.89
Koppala	18.04	9.65	1.90	3.86	16.61	6.44	1.42	6.82
Raichur	25.25	9.94	1.51	2.81	24.99	8.30	0.81	3.04



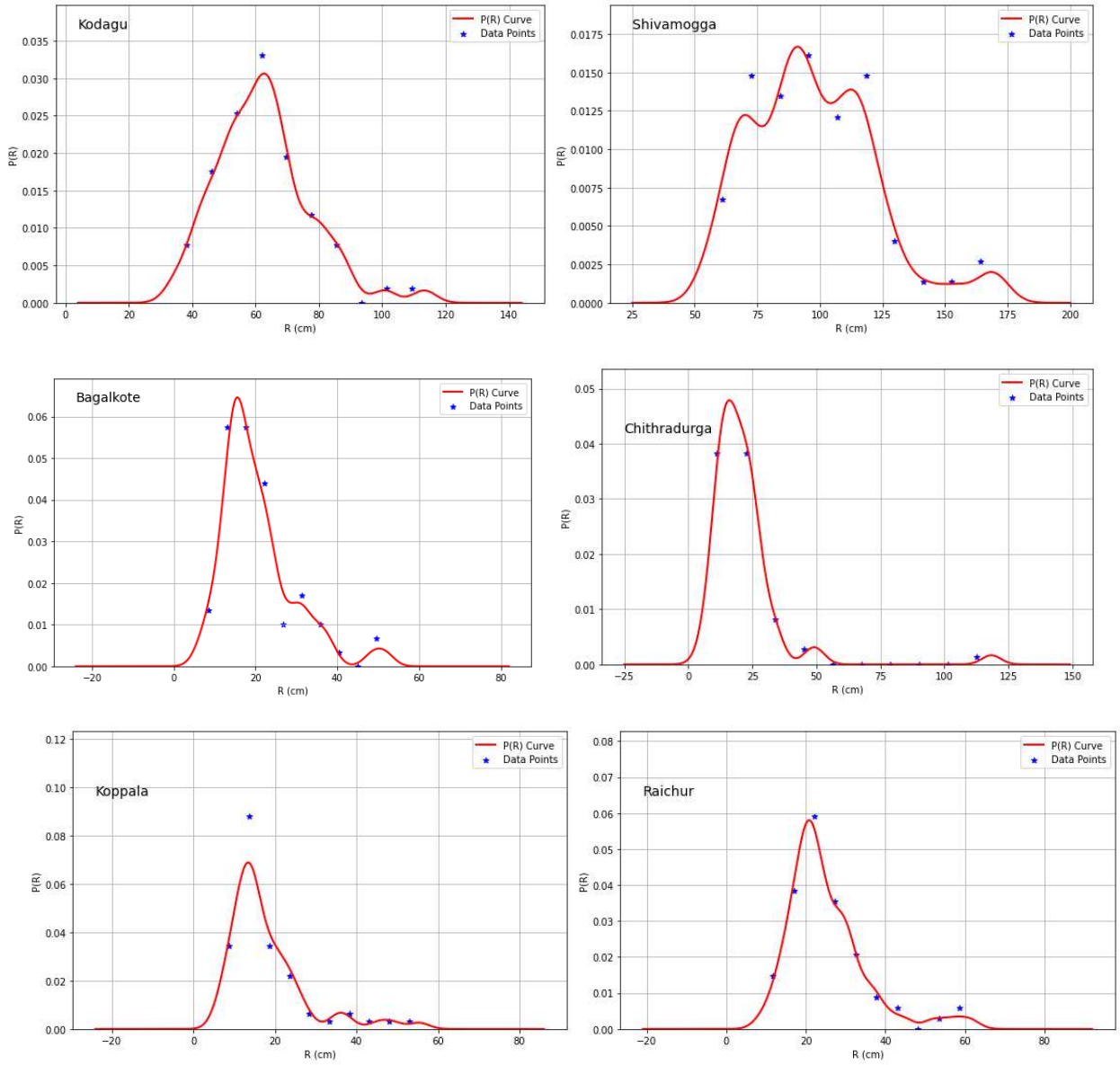


Fig. 3. Comparison between the observed Southwest Monsoon rainfall distributions (blue *) and the fitted Gaussian–Gamma Mixture (GGM) probability density functions (solid red curves) for selected flood-prone and drought-prone regions of Karnataka.

2.4 Level crossing statistics

In this study, the transformed rainfall data z has been analyzed using the Gaussian and Gamma Mixture Model (GGM) to examine the occurrence of extreme rainfall events through level crossing statistics. The effectiveness of the proposed model is evaluated by comparing its level crossing behaviour with that of the conventional Gaussian model. Level crossing analysis provides important information regarding the frequency with which rainfall exceeds or falls below specified threshold levels, thereby offering insight into flood and drought characteristics over different regions of Karnataka.

The rainfall process is assumed to be continuous, and the upward level crossing frequencies are computed for thresholds corresponding to +10% and +20% above the normal rainfall. These thresholds represent moderate and severe above-normal rainfall conditions respectively. Similarly, for drought-prone regions, thresholds corresponding to 10% and 20% below normal rainfall are considered to study deficit rainfall behaviour. The upward crossing frequency for the Gaussian and Gamma component are represented in equation (8) and equation (9) respectively.

$$N_u^{Gaussian} = \int_0^{\infty} \dot{z} \frac{w}{2\sigma\dot{\sigma}\sqrt{2\pi}} \exp\left[-\left(\frac{(z-\mu)^2}{2\sigma^2} + \frac{(\dot{z}-\dot{\mu})^2}{2\dot{\sigma}^2}\right)\right] d\dot{z} \text{ --- (8)}$$

$$N_u^{Gamma} = \int_0^{\infty} \dot{z} \frac{(1-w)}{\beta^\alpha \dot{\beta}^\alpha \Gamma(\alpha)\Gamma(\dot{\alpha})} z^{\alpha-1} \dot{z}^{\dot{\alpha}-1} e^{-\frac{z}{\beta}} e^{-\frac{\dot{z}}{\dot{\beta}}} d\dot{z} \text{ --- (9)}$$

Hence, the total upward level crossing rate is obtained as $N_u = N_u^{Gaussian} + N_u^{Gamma}$.

The upward and downward level crossings corresponding to a threshold a are defined by

$$N^+(a) = \int_0^{\infty} \dot{z} p(a, \dot{z}) d\dot{z} \text{ and } N^-(a) = \int_{-\infty}^0 (-\dot{z}) p(a, \dot{z}) d\dot{z} \text{ --- (10)}$$

To evaluate Equation (10), the joint probability density function $p(z, \dot{z})$ is required. Considering $z = wx + (1-w)y$ the derivative process is represented as $\dot{z} = w\dot{x} + (1-w)\dot{y}$. The conditional mean and variance of the derivative process are given by $\mu_{\dot{z}|w} = w\mu_{\dot{x}} + (1-w)\alpha_{\dot{y}}\beta_{\dot{y}}$ and $\sigma_{\dot{z}|w}^2 = w^2\sigma_{\dot{x}}^2 + (1-w)^2(\alpha_{\dot{y}}\beta_{\dot{y}}^2)$. The parameters $\mu_{\dot{z}}, \sigma_{\dot{z}}, \alpha_{\dot{z}}$ and $\beta_{\dot{z}}$ estimated using the Maximum Likelihood Estimation (MLE) method is presented in

Table 5. The results indicate noticeable regional variations in the derivative process, reflecting differences in rainfall variability and persistence among coastal, Malnad, and interior drought-prone regions. Coastal regions such as Udupi and Dakshina Kannada exhibit comparatively lower variability, whereas interior regions such as Chithradurga and Koppala show larger fluctuations in the derivative process.

TABLE 5

The parameter values of a derivative $\dot{z}$ of the process z

Region	$\mu_{\dot{z}}$	$\sigma_{\dot{z}}$	$\alpha_{\dot{z}}$	$\beta_{\dot{z}}$
Dakshina Kannada	0.01	0.28	0.12	-0.44
Udupi	0.02	0.25	1.83	1.41
Uttara Kannada	0.01	0.19	0.30	-0.54
Chikkamagalur	0.01	0.24	1.11	2.53
Kodagu	0.01	0.33	0.09	-0.33
Shivamogga	0.01	0.31	0.27	-0.46
Bagalkote	0.01	0.48	0.18	-0.86
Chithradurga	0.02	0.54	-0.31	-0.22
Koppala	0.02	0.53	0.17	-0.37
Raichur	0.01	0.45	-0.24	-0.47

The unconditional joint density function of $z, \dot{z}$ required for evaluating level crossing statistics, is expressed as

$$p(z, \dot{z}) = \frac{w}{2\sigma\dot{\sigma}\sqrt{2\pi}} e^{-\left(\frac{(z-\mu)^2}{2\sigma^2} + \frac{(\dot{z}-\dot{\mu})^2}{2\dot{\sigma}^2}\right)} + \frac{1-w}{\beta^\alpha \dot{\beta}^\alpha \Gamma(\alpha)\Gamma(\dot{\alpha})} z^{\alpha-1} \dot{z}^{\dot{\alpha}-1} e^{-\frac{z}{\beta}} e^{-\frac{\dot{z}}{\dot{\beta}}} \text{ --- (11)}$$

Using Equation (11), the upward and downward level crossing statistics are computed and summarized in Table 6. The comparison between observed crossings, Gaussian model crossings, and Gaussian–Gamma mixture model crossings clearly demonstrates the superiority of the proposed GGM approach. In most regions, the Gaussian model either overestimates or underestimates the number of crossings, particularly under extreme rainfall conditions. In contrast, the Gaussian–Gamma mixture model produces estimates that are much closer to the observed values. For coastal regions such as Dakshina Kannada and Udupi,

the GGM model accurately captures the occurrence of extreme positive rainfall crossings associated with flood conditions. Similarly, for drought-prone regions such as Bagalkote, Chithradurga, Koppala, and Raichur, the model successfully reproduces the observed downward crossings corresponding to deficient rainfall conditions. These results confirm that the proposed Gaussian–Gamma framework effectively captures both excess and deficit rainfall behaviour, making it suitable for analyzing monsoon variability and extreme rainfall events across Karnataka.

TABLE 6

Number of upward level crossings

Region	10% above			20% above		
	Observed	Guassian and Gamma	Gaussian	Observed	Guassian and Gamma	Gaussian
Dakshina Kannada	15	12	25	10	9	18
Udupi	8	7	29	8	9	25
Uttara Kannada	15	14	15	5	5	5
Chikkamagalur	17	18	22	11	11	13
Kodagu	18	19	22	11	11	13
Shivamogga	22	21	26	11	10	14
	10% below			20% below		
Bagalkote	20	21	27	14	15	21
Chithradurga	19	18	29	10	13	25
Koppala	19	16	28	15	10	23
Raichur	20	21	26	13	14	20

2.5 Joint Probability Density Analysis

To investigate the temporal dependence and persistence characteristics of Southwest Monsoon rainfall, the Gaussian and Gamma Mixture Model (GGM) is extended to a bivariate framework. The transformed rainfall variables corresponding to two consecutive time instances are represented by z_1 and z_2 , where the Gaussian component models the regular rainfall behaviour and the Gamma component captures the extreme rainfall variations. This

$$z_1 = ux_1 + (1 - u)y_1 \text{ and } z_2 = vx_2 + (1 - v)y_2 \text{ --- (12)}$$

The z_1 and z_2 represent the transformed rainfall data corresponding to two consecutive years. The variables x_t and y_t denote the Gaussian and Gamma components respectively. The conditional means of the transformed variables are given in equation (13)

$$m_{z_1|u} = um_{x_1} + (1 - u)m_{y_1} \text{ and } m_{z_2|v} = vm_{x_2} + (1 - v)m_{y_2} \text{ --- (13)}$$

$$\sigma_{z_1|u}^2 = u^2\sigma_{x_1}^2 + (1 - u)^2\sigma_{y_1}^2 \text{ and } \sigma_{z_2|v}^2 = v^2\sigma_{x_2}^2 + (1 - v)^2\sigma_{y_2}^2 \text{ --- (14)}$$

The conditional correlation structure between z_1 and z_2 is represented as a combination of Gaussian–Gaussian and Gamma–Gamma correlations. The conditional correlation coefficient is represented in equation (15) where $\rho_{xx}(\tau)$ and $\rho_{yy}(\tau)$ denote the lag- τ correlations corresponding to the Gaussian and

$$\rho_{z_1 z_2|(u,v)}(\tau) = \frac{uv \rho_{xx}(\tau) \sigma_{x_1} \sigma_{x_2} + (1 - u)(1 - v) \rho_{yy}(\tau) \sigma_{y_1} \sigma_{y_2}}{\sqrt{u^2 \sigma_{x_1}^2 + (1 - u)^2 \sigma_{y_1}^2} + \sqrt{v^2 \sigma_{x_2}^2 + (1 - v)^2 \sigma_{y_2}^2}} \text{ --- (15)}$$

$$P(u, v) = \alpha_1 \alpha_2 \delta(u - 1) \delta(v - 1) + \alpha_1 (1 - \alpha_2) \delta(u - 1) \delta(v) + (1 - \alpha_1) \alpha_2 \delta(u) \delta(v - 1) + (1 - \alpha_1)(1 - \alpha_2) \delta(u) \delta(v) \text{ --- (16)}$$

The conditional joint density $p(z_1, z_2|u, v)$ consists of four possible component combinations such as

formulation enables the study of both normal and extreme rainfall persistence within a unified probabilistic structure. Let x denote the Gaussian component and y denote the Gamma component. The latent indicator variables $u, v \in \{0, 1\}$ determine whether the rainfall process belongs to the Gaussian or Gamma regime. The transformed rainfall variables are expressed as below.

with $m_{x_t} = E[x_t]$, $m_{y_t} = E[y_t] = \frac{\alpha_t}{\beta_t}$, $t = 1, 2$. Similarly, the conditional variances are expressed in the equation (14) where $\sigma_{y_t}^2 = \frac{\alpha_t}{\beta_t^2}$, $t = 1, 2$.

Gamma regimes respectively. The latent indicator variables follow Bernoulli distributions with probabilities $P(u = 1) = \alpha_1$, $P(v = 1) = \alpha_2$. Hence, the joint probability distribution of (u, v) is expressed in equation (16).

Gaussian–Gaussian, Gaussian–Gamma, Gamma–Gaussian, and Gamma–Gamma components are

represented by $(u, v) = (1, 0), (1, 0), (0, 1)$ and $(0, 0)$ respectively. The Gaussian–Gaussian component is represented by the bivariate Gaussian density in equation (17). The remaining Gaussian–Gamma, Gamma–Gaussian, and Gamma–Gamma components

$$P_{GG}(z_1, z_2) = \frac{1}{2\pi\sigma_{x_1}\sigma_{x_2}\sqrt{1-\rho_{xx}^2}} \exp \left[-\frac{1}{2(1-\rho_{xx}^2)} \left(\frac{(z_1 - m_{x_1})^2}{\sigma_{x_1}^2} + \frac{(z_2 - m_{x_2})^2}{\sigma_{x_2}^2} - \frac{2\rho_{xx}(z_1 - m_{x_1})(z_2 - m_{x_2})}{\sigma_{x_1}\sigma_{x_2}} \right) \right] \quad (17)$$

$$p(z_1, z_2) = \sum_{u,v \in \{0,1\}} p(z_1, z_2 | u, v) p(u, v) \quad (18)$$

Thus, the final joint density becomes a weighted combination of all four component distributions, enabling simultaneous modelling of regular and extreme rainfall persistence. The parameters of the bivariate GGM estimated using the Maximum Likelihood Estimation (MLE) method and are presented in Table 7. The results indicate that coastal

are similarly formulated using the product of Gaussian and Gamma densities. The unconditional joint density is obtained using the law of total probability as represented in equation (18).

regions such as Dakshina Kannada and Udupi exhibit comparatively stronger Gaussian and cross-regime dependence, reflecting high rainfall persistence between consecutive years. The interior drought-prone regions such as Raichur and Uttara Kannada show weaker dependence structures, indicating comparatively lower rainfall persistence.

TABLE 7

The parameter values of equation (6.7) for monsoon months								
Region	w	μ_x	μ_y	σ_x	σ_y	ρ_{xx}	ρ_{yy}	ρ_{xy}
Dakshina Kannada	0.55	-0.18	0.10	0.17	0.14	0.31	0.70	0.73
Udupi	0.55	-0.30	0.095	0.15	0.13	0.15	0.91	0.92
Uttara Kannada	0.44	-0.01	0.01	0.14	0.02	0.12	-0.48	0.13
Chikkamagaluru	0.55	-0.03	0.01	0.23	0.01	0.23	0.47	0.19
Kodagu	0.90	-0.05	0.13	0.24	0.14	0.05	0.07	0.03
Shivamogga	0.90	-0.05	0.04	0.27	0.04	0.27	0.25	0.24
Bagalkote	0.72	-0.18	0.21	0.37	0.22	0.28	0.51	0.32
Chithradurga	0.73	-0.29	0.36	0.35	0.42	0.12	0.57	0.41
Koppala	0.90	-0.33	0.23	0.29	0.25	-0.07	0.62	0.40
Raichur	0.90	-0.12	0.19	0.34	0.20	0.19	0.24	0.22

The empirical, Gaussian-regime, Gamma-regime, and cross-regime correlations together with their confidence intervals and significance levels are summarized in Table 8. The results reveal that Udupi and Dakshina Kannada possess very strong empirical and cross-regime correlations, confirming persistent monsoon behaviour in coastal Karnataka. Regions

such as Kodagu and Uttara Kannada exhibit weaker or statistically insignificant correlations, indicating larger interannual variability. The Gamma-regime correlations are generally stronger in high rainfall regions, suggesting that extreme rainfall events exhibit higher persistence compared to normal rainfall behaviour.

TABLE 8

Region-wise Comparison of Empirical, Gaussian, Gamma, and Cross-Regime Correlations in Rainfall Time Series								
Region	Empirical r	p -value	ρ_{xx}	CI (ρ_{xx})	ρ_{yy}	CI (ρ_{yy})	ρ_{xy}	p-value (ρ_{xy})
Dakshina Kannada	0.67	0.01	0.25	[-0.10, 0.55]	0.6	[0.59, 0.77]	0.74	0.00
Udupi	0.84	0.01	0.18	[-0.06, 0.47]	0.9	[0.81, 0.94]	0.93	0.00
Uttara Kannada	0.12	0.34	0.12	[-0.11, 0.33]	-0.41	[-0.81, 0.08]	0.13	0.23
Chikkamagalur	0.23	0.07	0.24	[-0.00, 0.45]	0.32	[0.12, 0.55]	0.19	0.07
Kodagu	0.04	0.72	0.07	[-0.17, 0.29]	0.06	[-0.11, 0.19]	0.03	0.77
Shivamogga	0.25	0.05	0.28	[0.07, 0.47]	0.42	[0.23, 0.60]	0.24	0.05
Bagalkote	0.30	0.02	0.25	[0.01, 0.46]	0.53	[0.35, 0.68]	0.32	0.01
Chithradurga	0.33	0.01	0.11	[-0.09, 0.30]	0.61	[0.51, 0.69]	0.41	0.01
Koppala	0.3	0.02	-0.13	[-0.34, 0.10]	0.63	[0.51, 0.72]	0.40	0.01
Raichur	0.20	0.12	0.16	[-0.10, 0.40]	0.12	[-0.14, 0.32]	0.22	0.09

Using the conditional predictive density $p(z_2|z_1)$, rainfall predictions for the year 2026 are obtained for

all selected regions. The predictive density is represented in equation (19). The predicted Southwest

Monsoon rainfall values for 2026 are presented in Table 9. The results indicate comparatively higher predicted rainfall over the coastal regions such as Dakshina Kannada, Udupi, and Uttara Kannada, whereas drought-prone regions such as Bagalkote,

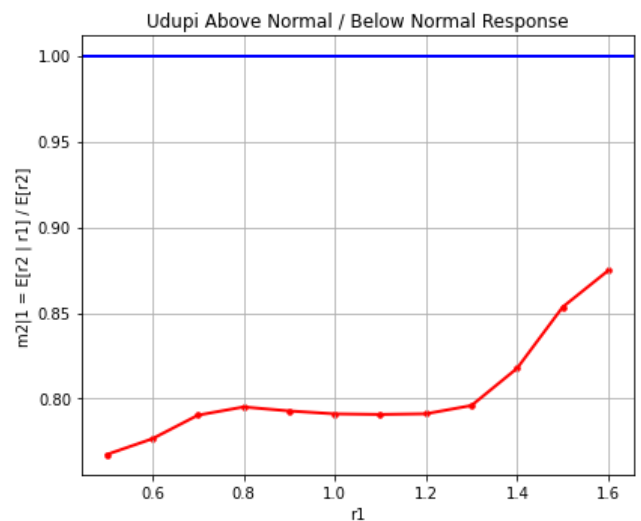
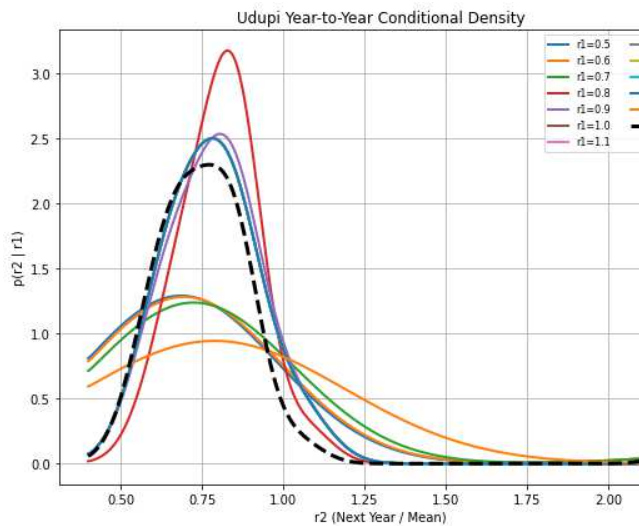
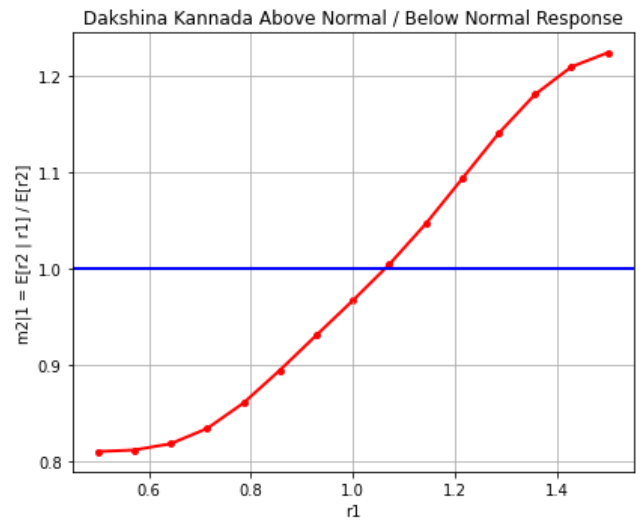
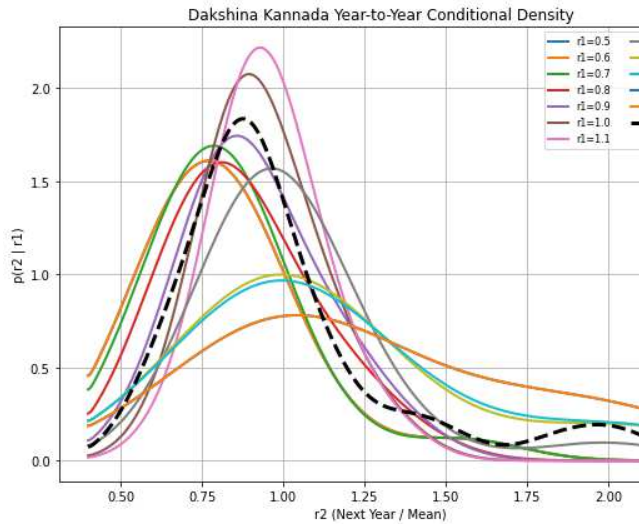
Koppala, and Raichur are expected to receive relatively lower monsoon rainfall. These predictions demonstrate the applicability of the proposed Gaussian–Gamma mixture framework for long-term rainfall forecasting and risk assessment.

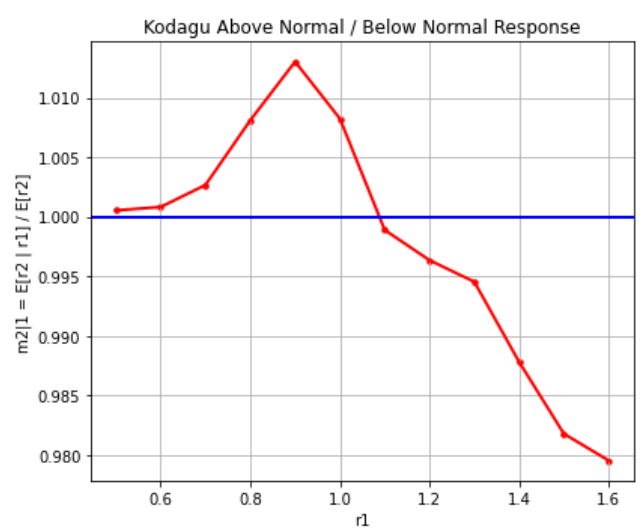
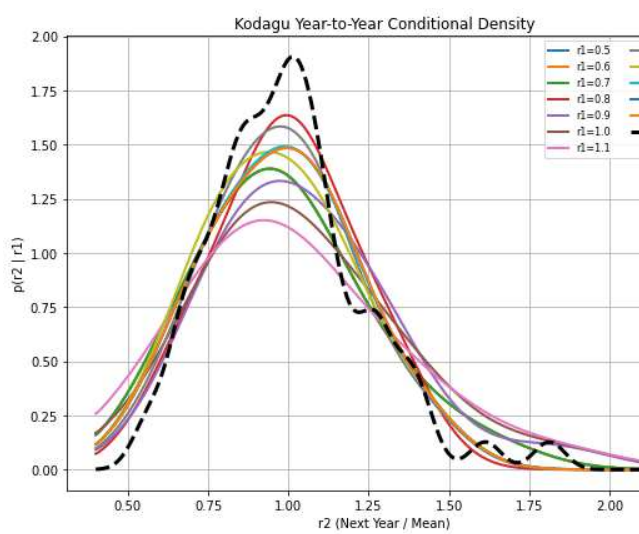
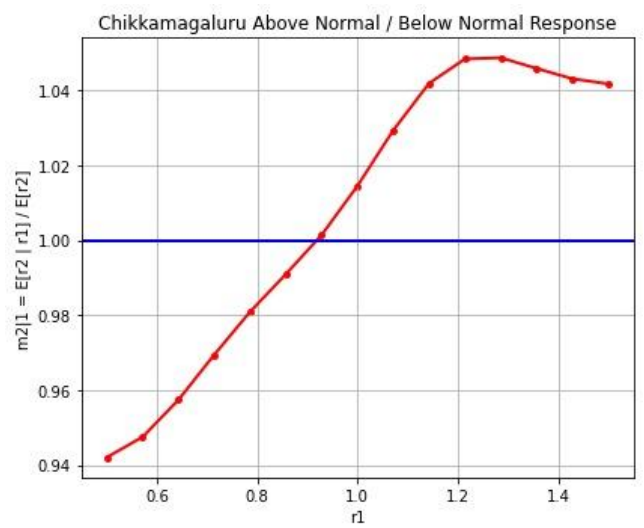
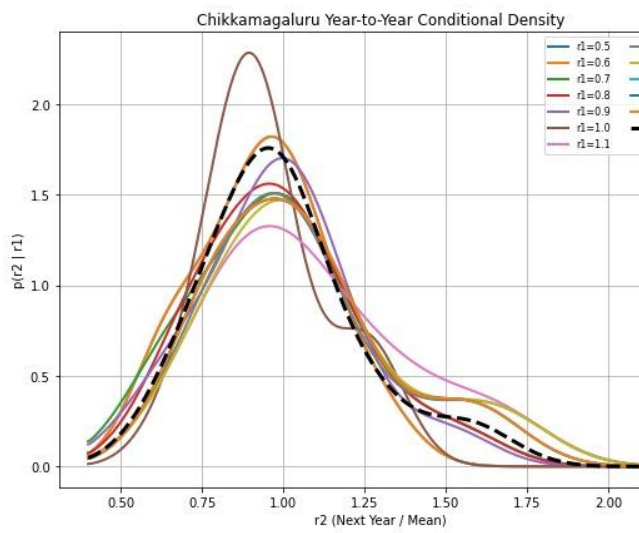
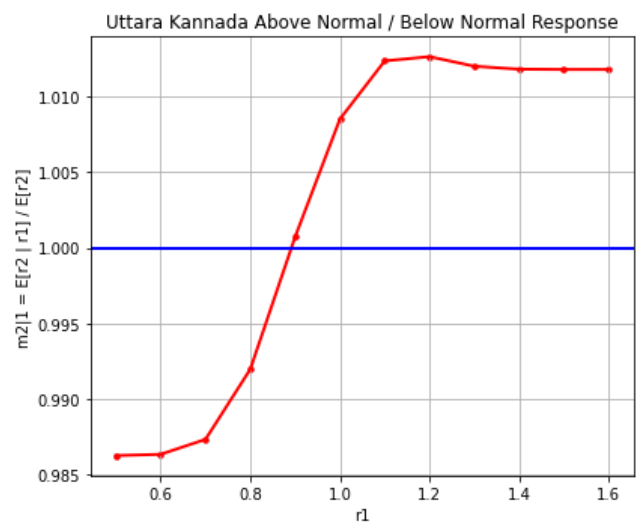
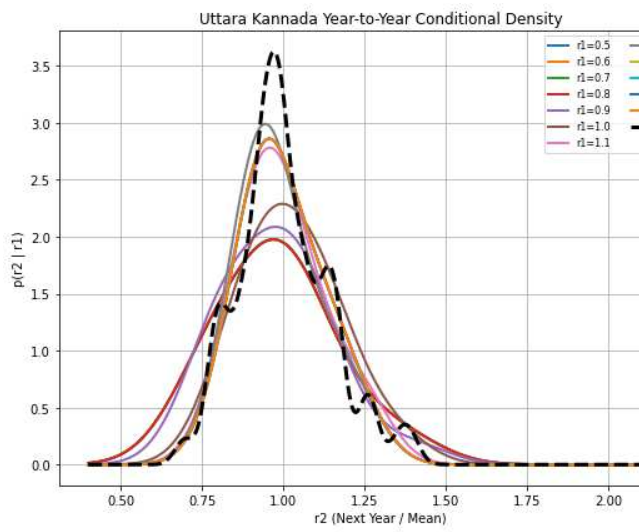
TABLE 9

Predicted Data in cm for 2026	
Region	Predicted Data
Dakshina Kannada	167.01
Udupi	116.65
Uttara Kannada	271.60
Chikkamagaluru	105.49
Kodagu	62.05
Shivamogga	97.63
Bagalkote	19.80
Chithradurga	20.77
Koppala	16.32
Raichur	24.61

Figures 4 and 5 present the year-to-year conditional density distributions and above-normal/below-normal response curves for flood-prone and drought-prone regions of Karnataka, respectively. The flood-prone regions exhibit broader conditional distributions and stronger persistence characteristics compared to the

drought-prone regions. In particular, Dakshina Kannada and Udupi display comparatively stronger temporal dependence, whereas interior regions such as Bagalkote and Raichur show weaker persistence and sharper conditional distributions.





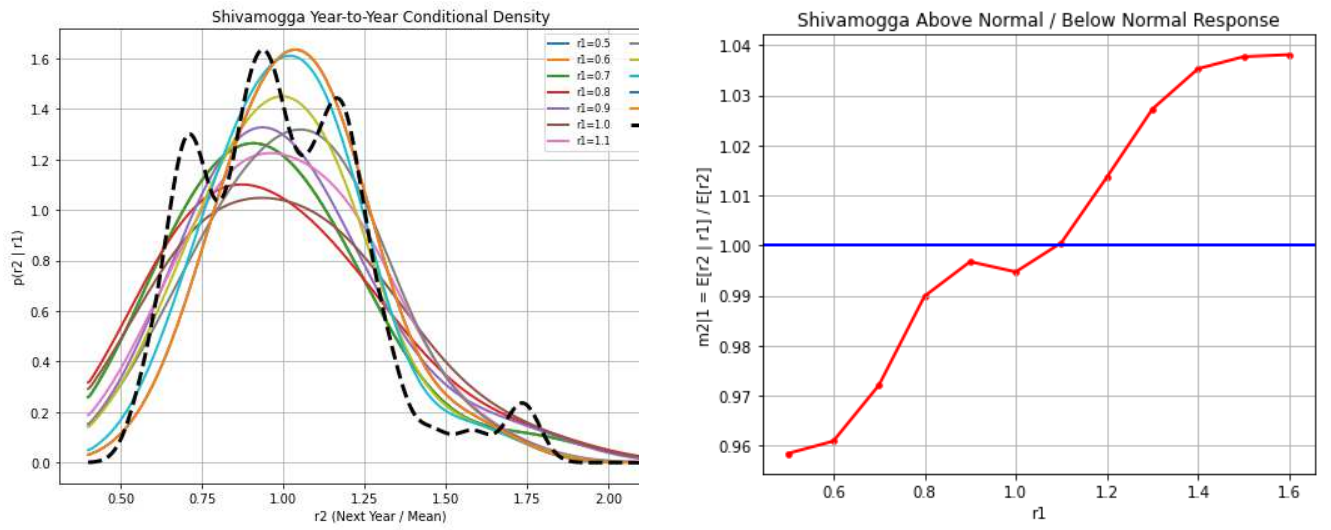
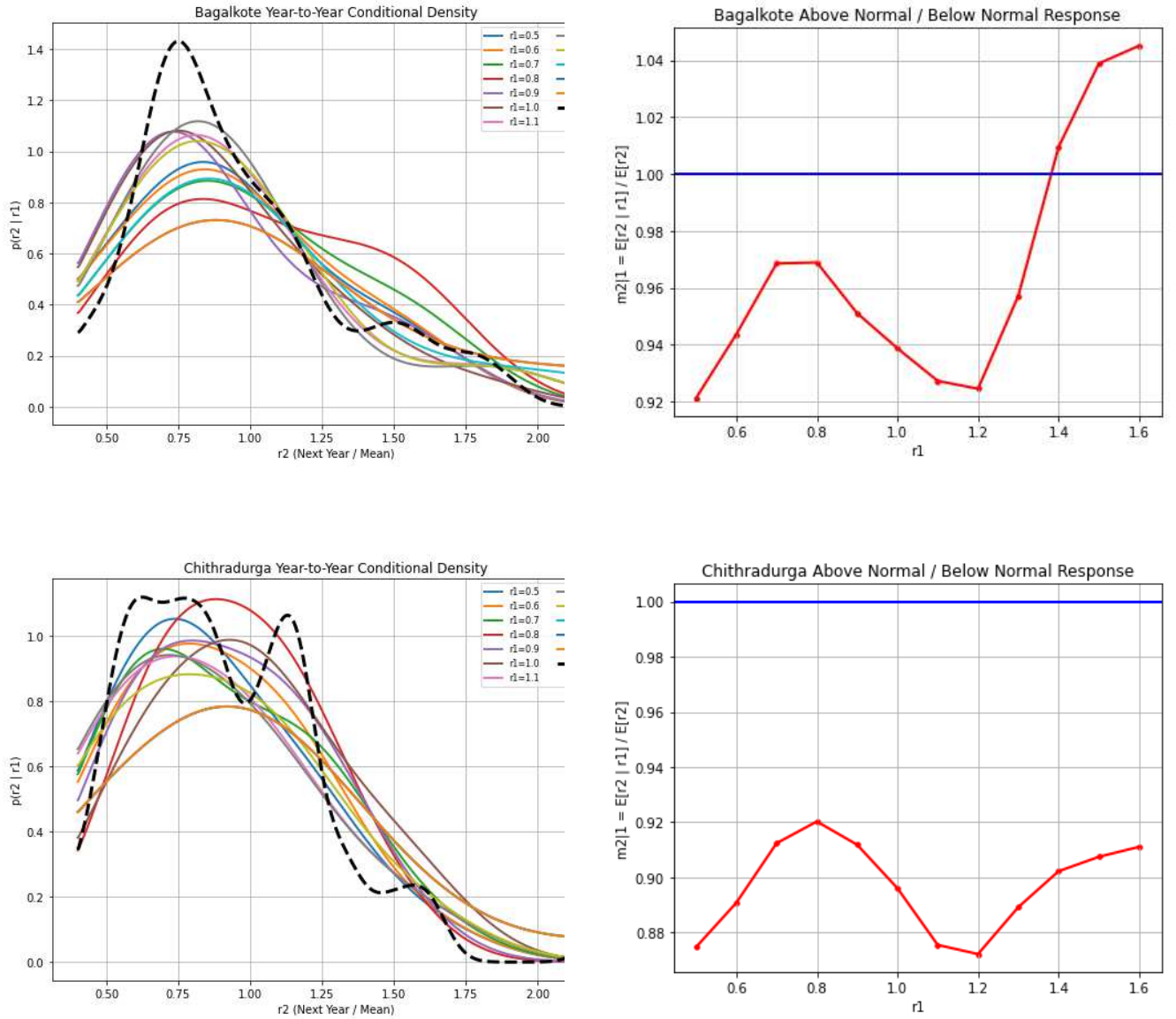


Fig. 4. Year-to-year conditional density and above-normal/below-normal response curves for flood-prone regions of Karnataka.



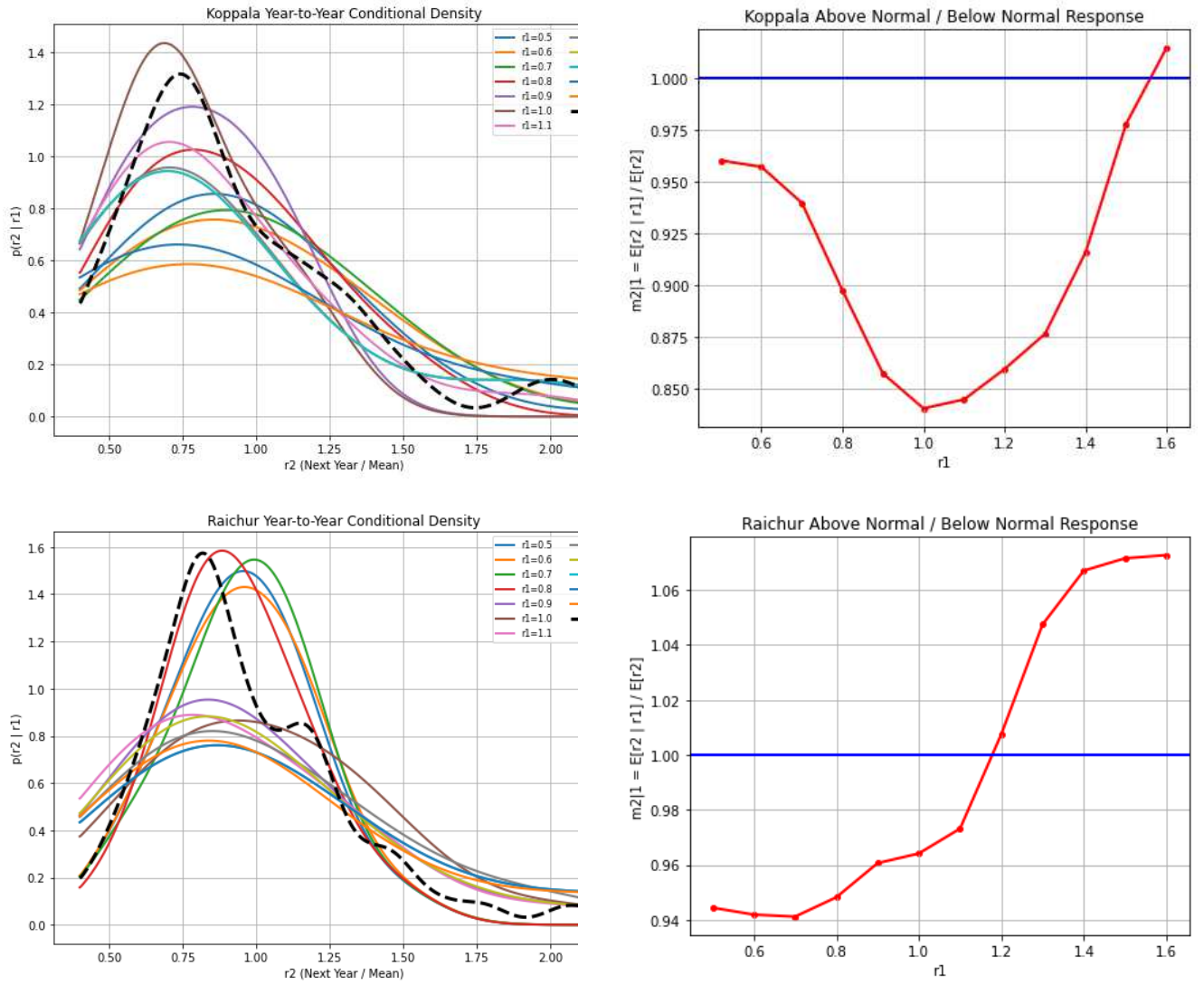


Fig. 5. Year-to-year conditional density and above-normal/below-normal response curves for drought-prone regions of Karnataka.

3. Discussion

The Gaussian–Gamma Mixture Model (GGM) developed in the present study successfully captured the non-Gaussian characteristics of Southwest Monsoon rainfall over Karnataka. The descriptive statistics revealed considerable regional variability in rainfall behaviour between the flood-prone coastal regions and the drought-prone interior regions. Coastal districts such as Dakshina Kannada and Udupi exhibited comparatively larger rainfall magnitudes together with higher skewness and kurtosis values, indicating the presence of persistent heavy rainfall and extreme monsoon events. In contrast, drought-prone regions such as Bagalkote, Koppala, and Raichur showed lower rainfall magnitudes with comparatively irregular rainfall occurrence. The estimated Gaussian–Gamma mixture parameters demonstrated that the Gaussian component primarily represented moderate rainfall fluctuations, whereas the Gamma component effectively captured extreme rainfall behaviour and

tail variability. The comparison between the observed rainfall moments and the model-generated moments showed good agreement for mean, standard deviation, skewness, and kurtosis across all selected regions. In particular, the proposed GGM framework reproduced higher-order moments more accurately than a conventional Gaussian model, confirming its suitability for modelling asymmetric and heavy-tailed rainfall distributions.

The level crossing analysis further demonstrated the effectiveness of the proposed model in representing above-normal and below-normal rainfall occurrences. The Gaussian model either overestimated or underestimated crossing frequencies in several regions, particularly under extreme rainfall conditions. However, the Gaussian–Gamma mixture model produced crossing frequencies that were much closer to the observed rainfall behaviour. Coastal regions

exhibited comparatively larger upward crossing frequencies, indicating stronger persistence of wet monsoon conditions and increased flood potential. Conversely, drought-prone interior regions displayed weaker upward crossings and greater variability in rainfall occurrence. The conditional density analysis revealed noticeable temporal persistence in Southwest Monsoon rainfall. The conditional density curves shifted systematically with increasing conditioning rainfall levels, particularly in coastal Karnataka. Dakshina Kannada and Udupi exhibited broader conditional density distributions and stronger persistence characteristics, indicating that above-normal rainfall years are more likely to be followed by subsequent wet years. In contrast, the drought-prone regions showed sharper conditional distributions and weaker persistence behaviour, reflecting greater interannual rainfall variability. The bivariate extension of the Gaussian–Gamma framework successfully captured the temporal dependence structure between consecutive rainfall years. Strong empirical and cross-regime correlations were observed in coastal regions, whereas weaker dependence structures were identified in drought-prone regions.

4. Conclusion

In the present study, a Gaussian–Gamma Mixture Model (GGM) was developed to analyse the statistical characteristics of Southwest Monsoon rainfall over selected flood-prone and drought-prone regions of Karnataka using long-term rainfall data from 1960–2025. The proposed framework successfully represented both moderate rainfall variability and extreme rainfall behaviour within a unified probabilistic structure. The descriptive statistical analysis revealed substantial regional differences in rainfall characteristics across Karnataka. Coastal regions exhibited higher rainfall variability together with larger skewness and kurtosis values, indicating the occurrence of persistent heavy rainfall and extreme monsoon events. The Gaussian–Gamma framework effectively captured these non-Gaussian rainfall characteristics and reproduced the observed rainfall moments with satisfactory accuracy. The level crossing analysis demonstrated that the proposed model provided improved estimates of above-normal and below-normal rainfall occurrences compared to the conventional Gaussian model. The conditional density and response curve analyses further revealed stronger rainfall persistence in coastal regions and weaker temporal dependence in drought-prone interior regions. The bivariate extension of the model successfully represented the temporal dependence structure between consecutive rainfall years through Gaussian–Gaussian, Gaussian–Gamma, Gamma–Gaussian, and Gamma–Gamma regimes. The correlation analysis indicated stronger persistence and cross-regime dependence in flood-prone regions,

The Gamma–Gamma regime correlations were comparatively stronger in high rainfall regions, suggesting that extreme rainfall events exhibit greater persistence than moderate rainfall fluctuations. The bootstrap confidence intervals and permutation tests further confirmed the statistical significance of rainfall persistence in several coastal regions. The predictive rainfall analysis for 2026 indicated comparatively higher Southwest Monsoon rainfall over Dakshina Kannada, Udupi, and Uttara Kannada, whereas lower rainfall values were predicted for Bagalkote, Koppala, and Raichur. These predictions are consistent with the historical rainfall persistence characteristics obtained from the conditional density and correlation analyses.

Overall, the results demonstrate that the proposed Gaussian–Gamma Mixture framework provides an effective probabilistic approach for analysing rainfall variability, persistence, and extreme rainfall behaviour over Karnataka. The integration of mixture modelling, level crossing statistics, conditional density analysis, and bivariate joint probability density offers a comprehensive statistical framework for studying monsoon rainfall in both flood-prone and drought-prone regions.

whereas weaker dependence structures were observed in drought-prone regions. The predictive analysis for the year 2026 indicated comparatively higher Southwest Monsoon rainfall over Dakshina Kannada, Udupi, and Uttara Kannada, while lower rainfall conditions were predicted for Bagalkote, Koppala, and Raichur. These predictions are consistent with the historical persistence characteristics obtained from the conditional density and joint probability analyses. Overall, the proposed Gaussian–Gamma Mixture framework provides an effective probabilistic methodology for modelling rainfall variability, persistence, and extreme rainfall behaviour over Karnataka. The integration of mixture modelling, level crossing statistics, conditional density analysis, and bivariate joint probability density offers a comprehensive statistical framework that may be useful for flood forecasting, drought monitoring, hydrological planning, and agricultural risk assessment in monsoon-dominated regions.

Acknowledgements

The authors gratefully acknowledge the India Meteorological Department (IMD), Indian Institute of Tropical Meteorology (IITM) and the Karnataka State Natural Disaster Monitoring Centre (KSNDMC) for providing the meteorological and drought data used in this study.

Funding

This research received no specific grant or funding from any funding agency in the public, commercial, or not-for-profit sectors.

Data Availability

The datasets used in this study are publicly available from the following sources:

Karnataka State Natural Disaster Monitoring Centre (KSNDMC) - <https://www.ksndmc.org>

Indian Institute of Tropical Meteorology (IITM) https://tropmet.res.in/static_pages.php?page_id=53

Declarations: Ethics, Consent to Participate, and Consent to Publish declarations is **not applicable**. This study is a statistical analysis of publicly available Southwest Monsoon rainfall data obtained from the Karnataka State Natural Disaster Monitoring Centre (KSNDMC) and the Indian Institute of Tropical Meteorology (IITM). The study does not involve human participants, animals, clinical trials, or personal data.

5. References

- Ayar, P. V., Blanchet, J., Paquet, E., & Penot, D., 2020, "Space-time simulation of precipitation based on weather pattern sub-sampling and meta-Gaussian model", *Journal of Hydrology*, **581**, 124518. <https://doi.org/10.1016/j.jhydrol.2019.124518>
- Angel R., Ramesh K., Chaturvedi A., & Kumudha H. R., 2025, "Application of Kolmogorov-Arnold Networks in Drought Prediction: A Comparison with Artificial Neural Networks using SPEI", *Journal of Harbin Engineering University*, **46**(12), 204–216. <https://harbinengineeringjournal.com/index.php/journal/article/view/4872/2661>
- Glasbey, C. A., & Nevison, I. M., 1997, "Rainfall modeling using a latent Gaussian variable. In V. N. Nair (Ed.)", *Modelling longitudinal and spatially correlated data*, pp. 233–242. Springer. https://doi.org/10.1007/978-1-4612-1882-1_18
- Hussein, A., & Kadhem, S. K., 2022, "Spatial modeling for analyzing a rainfall pattern: A case study in Ireland", *Open Engineering*, **12**, 1, 204–214. <https://doi.org/10.1515/eng-2022-0204>
- Iyengar, R. N., & Ramesh, K. (2005), "Forecasting the Indian summer monsoon rainfall: A nonparametric approach", *Earth and Planetary Science Letters*, **237**, 1–2, 59–68. <https://doi.org/10.1016/j.epsl.2005.06.025>
- Katz, R. W., & Parlange, M. B., 1998, "Over dispersion phenomenon in stochastic modeling of precipitation", *Journal of Climate*, **11**, 4, 591–601. [https://doi.org/10.1175/1520-0442\(1998\)011<0591:OPISMO>2.0.CO;2](https://doi.org/10.1175/1520-0442(1998)011<0591:OPISMO>2.0.CO;2)
- Khedun, C. P., Mishra, A. K., Singh, V. P., & Giardino, J. R., 2014, "A copula-based precipitation forecasting model: Investigating the interdecadal modulation of ENSO's impacts on monthly precipitation", *Water Resources Research*, **50**, 1, 580–600. <https://doi.org/10.1002/2013WR013763>
- Kumudha, H. R., & Ramesh, K., 2023, "Forecasting of Karnataka seasonal rainfall data using ANN approach", *Journal of Survey in Fisheries Sciences*, **10**, 3, 3431–3448. <https://doi.org/10.5281/zenodo.7709459>
- Kumudha, H. R., Ramesh, K., & et.al., 2025, "Gaussian and Gamma mixture model approach to rainfall analysis in flood-prone regions of Karnataka", *Communications on Applied Nonlinear Analysis*, **32**, 9s, 1–12.
- Kumudha H. R., & Ramesh K., 2025, "A gaussian and gamma mixture model approach to rainfall analysis in drought-prone regions of Karnataka", *MAUSAM*, **76**, 4, 1223–1234. DOI: <https://doi.org/10.54302/mausam.v76i4.7014>
- Kumudha H. R., Ramesh K., Gupta R., Chaturvedi A., & Richard A., 2025, "Forecasting and Prediction of Seasonal Rainfall in Drought and Flood Affected Zones of Karnataka Using Artificial Neural Network (ANN)", *Journal of Harbin Engineering University*, **46**, 12, 217–228. <https://harbinengineeringjournal.com/index.php/journal/article/view/4856>
- Kumudha H. R., & Ramesh K., 2023, "Statistical Analysis and Distribution Modeling of Rainfall Data in Flood-Prone Regions of Karnataka State", *Tuijin Jishu/Journal of Propulsion Technology*, **44**, 4. DOI: <https://doi.org/10.52783/tjpt.v44.i4.2478>
- Li, Z., Zhang, Y., & Giangrande, S. E., 2012, "Rainfall-rate estimation using Gaussian parameter estimator: Training and validation", *Journal of Atmospheric and Oceanic Technology*, **29**, 5, 861–877. <https://doi.org/10.1175/JTECH-D-11-00152.1>
- Ramesh K., & Iyengar R. N., 2017, "A non-Gaussian model for Indian monsoon rainfall", *International Journal of Research in Granthaalayah*, **5**, 4, 88–95. <https://doi.org/10.5281/zenodo.556423>
- Serinaldi, F., 2009, "A multisite daily rainfall generator driven by bivariate copula-based mixed distributions", *Journal of Geophysical Research: Atmospheres*, **114**, D10, D10103. <https://doi.org/10.1029/2008JD011258>
- Wilks, D. S., 2011, "Statistical methods in the atmospheric sciences", Academic Press, 3rd ed.