International Journal of Hydrology Research
2025 Vol. 10, No. 1, pp. 16-38
ISSN: 2771-5590
DOI: 10.18488/ijhr.v10i1.4679

Modelling daily precipitation occurrence and amount using a first-order Markov chain with distribution fitting: A case of black bush polder and Ebini, Guyana

Bunnel Bernard[1+]
Linda Francois[2]
Dwayne Shorlon Renville[3]

[1,2,3]*Department of Mathematics, Physics and Statistics, University of Guyana, Turkeyen Campus, Georgetown, Guyana.*
[1]*Email: bunnel.bernard@uog.edu.gy*
[2]*Email: linda.francois@uog.edu.gy*
[3]*Email: dwayne.renville@uog.edu.gy*

(+ Corresponding author)

Article History
Received: 14 November 2025
Revised: 22 December 2025
Accepted: 24 December 2025
Published: 31 December 2025

Keywords
Guyana
Markov chain application
Rainfall distribution fitting
Rainfall variability
Stochastic precipitation modeling.
Black Bush Polder
Ebini

ABSTRACT

Precipitation simulation models are crucial for understanding, decision-making, and responding to phenomena related to hydrological, agricultural, and water resource management. This is particularly true for these climate-sensitive sectors in countries with high average annual rainfall, as well as those that depend on rainfall for food security and economic resilience. The application of precipitation simulation models in Guyana remains largely unexplored, despite the country's high average annual precipitation and its population residing mainly along the low-elevation coastal zone. This study aimed to develop and evaluate a stochastic precipitation model capable of simulating daily rainfall patterns for two climatically distinct regions of Guyana. Daily rainfall occurrence was modeled using a first-order Markov chain, while wet-day rainfall amounts were fitted to Gamma, Weibull, and Lognormal probability distributions. The analysis used daily rainfall records from 1981 to 2022, with monthly stratification applied to capture Guyana's bimodal rainfall regime. The model accurately reproduced key precipitation characteristics, showing high agreement between observed and simulated data. Projections for 2023–2030 closely align with established seasonal patterns, replicating the primary wet season (May–August) and the secondary wet season (November–January). The Gamma and Weibull distributions provided superior fits for most months, reflecting the skewed nature of daily rainfall. This study provides the first empirical framework for stochastic rainfall modeling in Guyana, offering a foundation for hydrological forecasting, climate risk assessment, and agricultural planning. The modeling framework also holds transferability for other Caribbean and South American regions facing comparable climatic variability and limited observational data.

Contribution/Originality: This study presents the first empirical application of stochastic precipitation modeling for Guyana. It integrates Markov chain rainfall occurrence and probabilistic distribution fitting. It establishes a transferable methodological framework for data-scarce tropical regions, enhancing rainfall simulation, hydrological forecasting, and climate risk assessment across the Caribbean and South America.

1. INTRODUCTION

The significance of precipitation in the context of global climate dynamics and environmental stability is profound. As a fundamental component of the Earth's hydrological cycle, precipitation directly affects hydrological, agricultural, and ecological systems, particularly in regions where rainfall strongly determines food security and water availability. However, paradoxically, precipitation also serves as a principal driver of disasters: floods, landslides, droughts, waterborne disease outbreaks, and widespread economic losses, which, combined, can lead to

severe socioeconomic and environmental impacts [1, 2]. Furthermore, with current evidence and forecasts of extreme weather patterns associated with climate change [3], this duality of precipitation becomes sharply focused for both scholars and practitioners. In particular, the central role of precipitation across many societal domains and the rising risks associated with its variability have led to heightened interest in understanding and predicting it. Modelling has become indispensable for anticipating hazards, planning infrastructure, managing water resources, and enhancing resilience.

Historically, modeling precipitation has attracted considerable research attention due to its complex and variable nature, influenced by geographic, atmospheric, climatic, and anthropogenic factors [4, 5]. Moreover, as various modeling techniques evolve, the necessity for accurate precipitation predictions continues to grow. Traditional deterministic models often struggle to account for the inherent uncertainty in precipitation events. This has led to a shift towards more sophisticated probabilistic or stochastic models to capture the uncertainty, dependence structure, temporal dynamics, and randomness inherent in precipitation processes, thereby enhancing the reliability and applicability of rainfall forecasts [6]. The incorporation of stochastic elements into precipitation modeling has emerged as a robust approach, enriching traditional methodologies with improved adaptability and predictive power, which are essential for effective resource management and risk assessment [7, 8].

Recent decades have seen significant advancements in precipitation modeling technology and methodology. A particular focus has been on probabilistic and stochastic models, which have proven effective for temporal and spatial rainfall modeling. One of the most widely applied approaches is the two-step stochastic framework, in which a Markov chain models rainfall occurrence and probability distributions are fitted to wet-day rainfall amounts. This framework has been successfully applied across a range of climatic zones, including Asia, Europe, and Africa [9-11]. The first-order Markov chain can model precipitation occurrences based on previous states, enabling the construction of rainfall data for hydrological simulations [12]. Such models capture the dependency of rainfall events not only on their immediate predecessors but also on broader patterns, providing a nuanced understanding of precipitation dynamics. Moreover, increased computational resources facilitate the application of more complex statistical methods, including Bayesian models that support intricate probabilistic inference and parameter estimation [13]. Because of these properties, first-order Markov chain models remain widely used in rainfall modeling and generation exercises across diverse climates [14].

First-order Markov chain models are not without issues. They tend to underestimate the duration of dry spells, and as a result, subsequent refinements have been made [15, 16]. To address these shortcomings, scholars introduced higher-order Markov chain models that extended the 'memory' of previous wet and dry days, thereby improving simulation fidelity. For example, optimization of model order and structure was employed to determine the most suitable Markov chain order for daily precipitation occurrence in the United States [17] and for pairing a Markov chain with a gamma distribution to enhance rainfall intensity modeling in Argentina [18]. Additionally, scholars have leveraged the flexibility of stochastic methods to model precipitation and other phenomena across diverse geographical and climatic conditions [11, 19, 20].

Aside from precipitation occurrence, stochastic models are also used to model wet-day rainfall amounts. These include the Gamma, Weibull, Lognormal, and Exponential distributions. Research has consistently highlighted the Gamma distribution as particularly effective for daily rainfall in tropical climates [8, 21] while alternative heavy-tailed distributions have been suggested for capturing extremes [22]. The integration of probability distributions with Markov chain–based occurrence models allows for the generation of synthetic rainfall series that accurately replicate both the frequency and intensity of observed rainfall. These models are highly valuable for hydrological planning, crop yield modeling, runoff estimation, and climate risk assessment. Despite their more advanced forecasting adaptations, first-order Markov chain models remain widely used and applicable, particularly for baseline studies.

Despite these substantial global advances, significant disparities exist in where precipitation modelling research is conducted. Much of the empirical literature is concentrated in North America, Europe, Asia, and parts of Africa. More proximate to Guyana, a growing body of research across the Caribbean continues to develop, highlighting regional differences and the need for localized studies to improve understanding and predictive capabilities in precipitation modelling [23] Central America [24, 25] and parts of South America [26, 27] demonstrates the value of stochastic rainfall simulation for hydrological planning, flood risk assessment, agricultural modelling, and climate-change impact studies. However, while countries within these regions share many climatic, geographic, socioeconomic, and environmental characteristics with their neighbors, precipitation variability means models are context-specific. Therefore, the absence of stochastic modelling in a particular context not only constitutes a gap in the global literature but also hampers local capacity to prepare for and manage the impacts of climatic variability [28].

A thorough literature search proved the scarcity of precipitation modeling within the Guyana context. Moreover, while there is evidence of descriptive climatology and deterministic modeling in Guyana [29], there are no peer-reviewed studies of stochastic precipitation modelling. This represents a meaningful gap in the literature. Without stochastic rainfall models calibrated to local data, decision-makers must rely on general hydrological assumptions, foreign models, or climate-model outputs that may not reflect local rainfall behaviour. This creates challenges for national planning, including water resource management, agricultural scheduling and irrigation planning, urban drainage design, flood risk assessment, drought preparedness, and climate adaptation policy.

To bridge this gap, the present study develops the first stochastic precipitation model tailored for Guyana. It focuses on two ecologically distinct regions: 'Black Bush Polder' in the low coastal plain and 'Ebini' in the hinterland. Rainfall occurrence is modeled using a first-order Markov chain, while wet-day rainfall amounts are characterized through probability distribution fitting. Monthly stratification is employed to capture Guyana's bimodal rainfall regime, and model performance is evaluated against observed data from 1981–2022, with simulations extended to 2030. By integrating occurrence and intensity modeling, this research establishes a foundational stochastic rainfall framework for Guyana, one that complements previous trend-based investigations and equips water managers, agricultural planners, and climate scientists with a robust tool for scenario generation, resource optimization, and risk-informed decision-making.

2. METHODOLOGY

2.1. Study Area and Data

The study areas for this research are Black Bush Polder, Guyana, specifically at coordinates 6°4'58" N and 57°15'57" W, and Ebini, Latitude 5°38′39.26″N and Longitude 57°46′17.51″W. A hybrid dataset, combining both observed and reanalysis data, was used due to limited historical data availability for the Black Bush Polder area and the specific requirements of the precipitation model. This dataset merged information from local stations and reanalysis data from the Climate Engine's CHIRPS dataset, spanning from 1981 to 2022. The CHIRPS dataset, offering the highest resolution among the available options (4.8 km, 4800 m, 1/20-deg), was used to extract data from 1981 to 2000. Data from 2001 to 2022 primarily stemmed from weather stations. Approximately 3.4% of the data for the years 2001 to 2022 were identified as missing for the Black Bush Polder region. A direct imputation method was employed to address the absence of this data. This method utilized corresponding data retrieved from the CHIRPS dataset.

Due to the lack of precipitation data for the Ebini area, reanalysis data from the Climate Engine for the period 1981-2022 were used in the precipitation simulations. All analyses were performed using a combination of R and Microsoft Excel. The stochastic rainfall model, including the Markov chain simulation and probability generation processes, was developed and executed in R, while the distribution fitting and graphical evaluations were conducted in Excel using an add-on package designed for probability distribution analysis.

2.2. Modeling Framework

2.2.1. Precipitation Model

Here, we present a precipitation model designed to simulate rainfall patterns from 1981 to 2030, drawing on historical precipitation data from Black Bush Polder and Ebini. The simulation uses a Markov chain to determine precipitation events, while a probability distribution is used to estimate precipitation amounts on wet days. To determine the precipitation amounts, we explore four distinct probability distributions: Lognormal, Weibull, and Gamma (2 and 3 parameters). These proposed distributions serve as tools for quantifying precipitation amounts.

Given the influence of seasonal shifts on rainfall patterns, we fit distributions for each month. This analysis helps identify the most appropriate distribution that accurately aligns with the precipitation characteristics for that specific month.

Our selection of suitable distribution models for each month is guided by a combination of criteria, including the Akaike Information Criterion (AIC), and the prevalence of these distributions within the existing literature. These methodologies collectively contribute to determining the optimal distribution for estimating precipitation amounts.

2.2.2. Precipitation Occurrence Model

In this two-step precipitation model, precipitation occurrence is modeled using a first-order Markov chain. This is achieved by categorizing the data into two states: "ones" representing wet days and "zeroes" representing dry days. The transitions between these states are then counted to determine transition probabilities and final states.

2.2.3. Markov Chain

In hydrology, modelling precipitation data at appropriate times for various purposes has been an important issue over the past 30 years [30]. Using a higher-order Markov chain can improve the model's performance in replicating precipitation occurrences. However, parameter estimation becomes more computationally intensive as the order increases, requiring more computations. In numerical analysis, the improvements in model performance achieved by higher-order methods are relatively minimal. As a result, it is more advantageous to utilize a first-order approach. A first-order Markov chain is a type of stochastic process characterized by the fact that the value of the process at time t, X_t, is solely determined by its value at the previous time step $t-1, X_{t-1}$. The value of X_t does not depend on the specific sequence of values that the process went through to reach X_{t-1} . This means that the likelihood of precipitation occurring tomorrow is contingent solely upon the current weather conditions, namely, whether it is currently raining or not. The explanation can be attributed to the Markov property, which can be described as follows:

$$P(X_{t+1} = s_{t+1} \mid X_t = s_t, X_{t-1} = s_{t-1}, \cdots, X_0 = s_0) = P(X_{t+1} = s_{t+1} \mid X_t = s_t) \quad (1)$$

Where the time variable, denoted as t, takes on values from the set {0, 1, 2, · · ·, T}. The state space, denoted as s, consists of elements from the set {1, 2, 3, · · ·, S}. The Markov property in a first-order Markov chain states that the previous state is irrelevant for the next state. The present value is the sole significant value for determining future values. In the context of a Markov chain, a transition matrix is essential for specifying the probability of each event. In this context, the composition of transition probabilities refers to the conditional probability of state j given state i. The transition matrix is given as:

$$\mathbf{P} = \begin{bmatrix} p_{11} & \cdots & p_{1j} \\ \vdots & \ddots & \vdots \\ p_{i1} & \cdots & p_{ij} \end{bmatrix} \text{ for } i, j \in S \quad (2)$$

Where $P_{ij} = \frac{n_{ij}}{\sum_{j=1}^{n} n_{ij}} = P(X_{t+1} = j \mid X_t = i)$. The transition matrix has the property that each row sums to 1.

$$\sum_{j=1}^{S} P_{ij} = \sum_{j=1}^{S} P(X_{t+1} = j \mid X_t = i) = \sum_{j=1}^{S} P_{\{X_t = i\}}(X_{t+1} = j) = 1 \quad (3)$$

Precipitation incidence may be classified into two states: dry and wet. Hence, the transition matrix is determined by two conditional probabilities, which are

$P_{dd} = P($ dry on day $t+1 \mid$ dry on day $t)$.

$P_{dw} = P($ wet on day $t+1 \mid$ dry on day $t)$.

$P_{wd} = P($ dry on day $t+1 \mid$ wet on day $t)$.

$P_{ww} = P($ wet on the day $t+1 \mid$ wet on the day $t)$.

Given that only two states exist, the transition probabilities within a given state are complementary. It is optional to estimate all four transition probabilities; instead, it suffices to estimate a single probability for each transition pair. Given this, we can define the transition probability matrix as follows:

$$\mathbf{P} = \begin{bmatrix} p_{dd} & p_{dw} \\ p_{wd} & p_{ww} \end{bmatrix} \quad 0 \le p_{ij} \le 1 \qquad i,j = \{d,w\} \quad (4)$$

Using the transition matrix allows computation of the stationary state vector. The stationary state vector π, has the unique property of $\pi = \pi\mathbf{P}$. It gives the long-run relative frequency of being in a wet and dry state. It has the property: $\sum_{i=1}^{S} \pi_i = 1$, where $\pi_i \ge 0$ for all i. property: π_i, d probability of being in state i.

$$\pi = (\pi_d \quad \pi_w)$$

$$(\pi_d \quad \pi_w) = (\pi_d \quad \pi_w) \begin{bmatrix} 1 - p_{dw} & p_{dw} \\ p_{wd} & 1 - p_{wd} \end{bmatrix} \quad (5)$$

For a two-state Markov chain, the stationary probabilities π_d *and* π_w are given as follows:

$$\pi_d = \frac{p_{wd}}{p_{wd} + p_{dw}} \quad (6)$$

$$\pi_w = \frac{p_{dw}}{p_{wd} + p_{dw}} \quad (7)$$

2.2.4. Determining Precipitation Amount

The second part of the stochastic precipitation model aims to determine the precipitation amount. This is achieved by identifying a suitable generalized probability distribution that fits the wet-day counts in the precipitation data stratified by month. According to existing literature, the two-parameter gamma distribution is the most widely used distribution for fitting precipitation data on wet days [17, 31]. Insights can be drawn from density plots and histograms presented in Figures 1, 2, 3, and 4. Distributions skewed to the right will be considered for modeling precipitation quantities. In this study, the following distributions will be tested: Weibull, two-parameter gamma, three-parameter gamma, and lognormal. Log-likelihood estimation was used to estimate the parameters of these distributions.

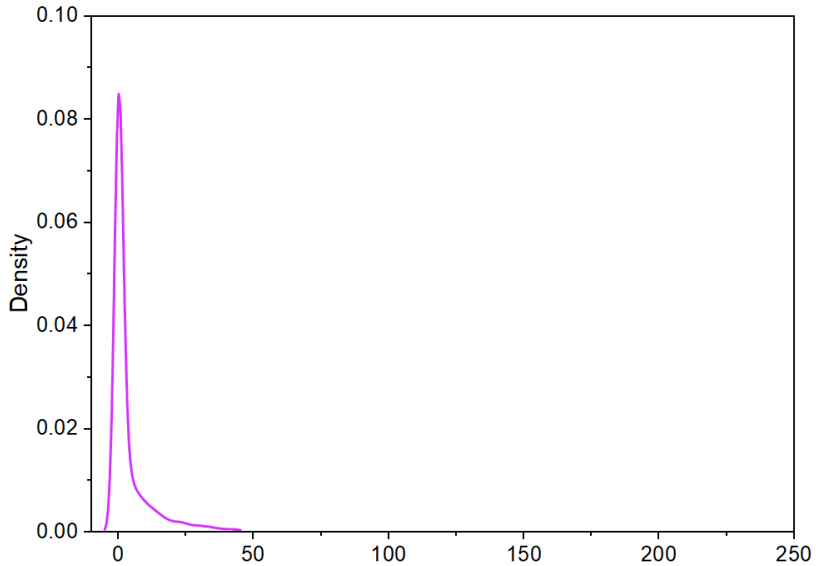


Figure 1. Overall Density plot of precipitation over Black Bush Polder (1981-2022).

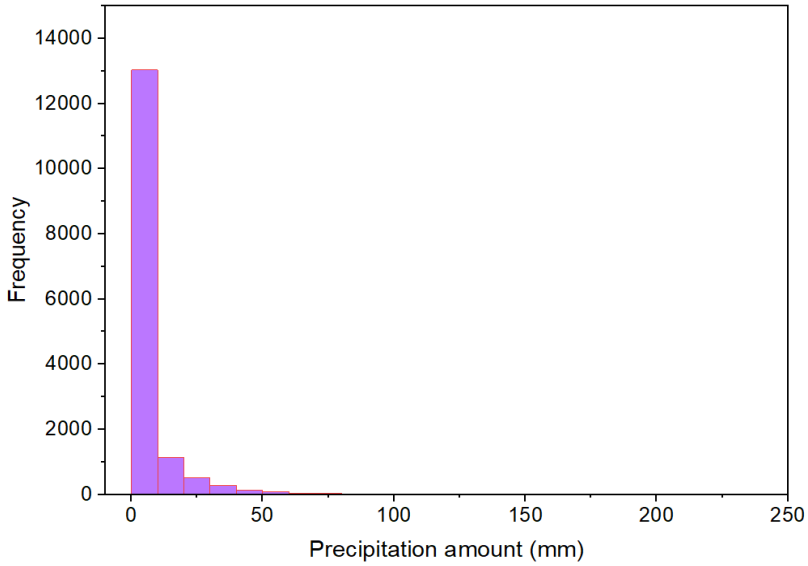


Figure 2. Precipitation histogram for Black Bush Polder (1981-2022).

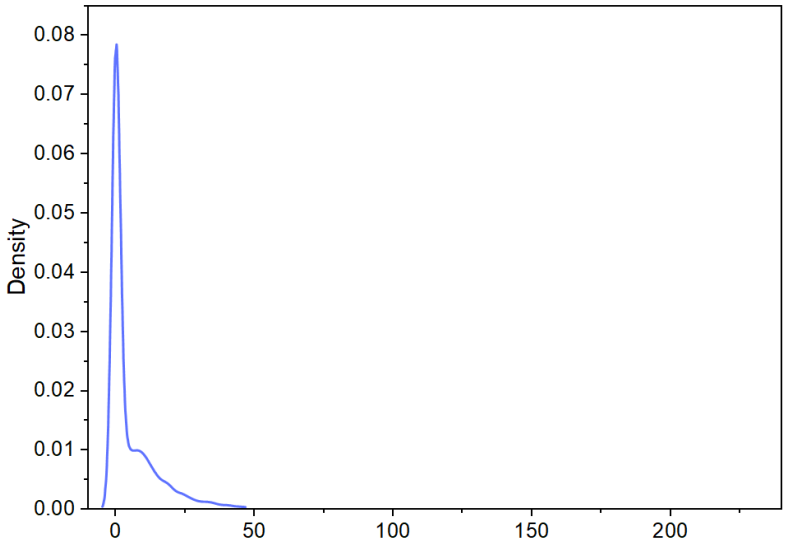


Figure 3. Overall Density plot of precipitation over Ebini (1981-2022).

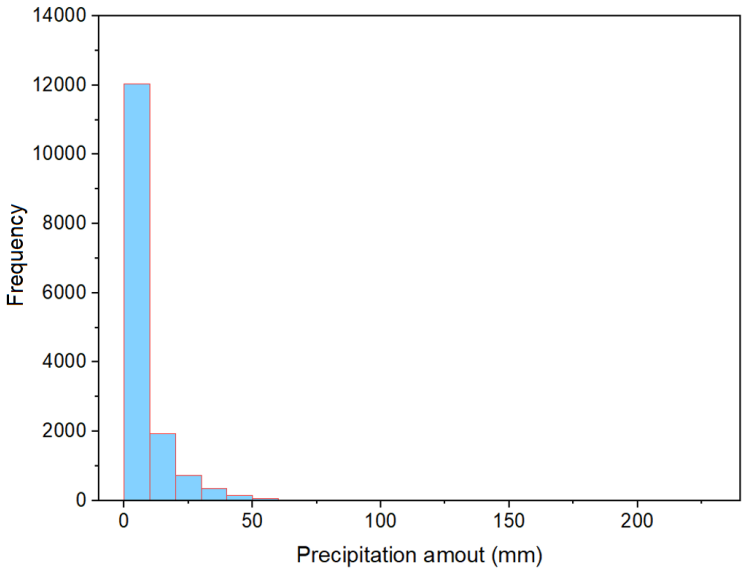


Figure 4. Precipitation histogram for Ebini (1981-2022).

2.2.5. Probability Distributions

Four probability distributions, namely the lognormal, Weibull, gamma, and three-parameter gamma distributions, will be evaluated to identify the one that best fits the monthly precipitation data. The selection of these distributions was motivated by the pronounced positive skewness evident in both the histograms and kernel density plots of the rainfall data, as illustrated in Figures 1, 2, 3, and 4 for the study locations. Such right-skewed behavior is characteristic of precipitation data and aligns with the underlying assumptions of these distributions, which have also been widely applied and shown to perform effectively in the literature.

2.2.6. Lognormal Distributions

The lognormal distribution is a continuous probability distribution that exhibits right skewness, with elongated tails to the right. It can be generated through a variable-method transformation from the normal distribution. As a result, if the random variable X follows a lognormal distribution, then Y = ln(X) follows a normal distribution. This distribution is widely used in modeling various socio-economic and natural processes, including but not limited to revenue distributions and precipitation levels.

The standard lognormal distribution has two parameters: location and scale. This can be expanded by introducing a third parameter corresponding to a threshold value. This expansion is termed the three-parameter lognormal distribution. The three-parameter lognormal distribution is a refinement of the two-parameter version, incorporating a shift in location. The probability density function (PDF) and parameter estimations are given below:

$$f(x;\mu,\sigma,\gamma) = \frac{1}{(x-\gamma)\sigma\sqrt{2\pi}} \exp\left\{-\frac{[\ln(x-\gamma)-\mu]^2}{2\sigma^2}\right\} \quad (8)$$

Where $0 \le \gamma < x,\ -\infty < \mu < \infty,\ \sigma > 0.\ \mu, \sigma$ and γ are the parameters of the distribution. When $\gamma = 0$, the distribution becomes the two-parameter lognormal distribution:

$$f(x;\mu,\sigma) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(\frac{-(\ln x-\mu)^2}{2\sigma^2}\right) \quad (9)$$

In these equations, the parameters μ, σ and γ are referred to as the location, scale, and threshold parameters, respectively. The mean of the lognormal distribution is given by $\exp\left(\mu+\frac{1}{2}\sigma^2\right)$, and its variance is given by $(\exp(\sigma^2)-1)\exp(2\mu+\sigma^2)$ The maximum likelihood estimate of the two parameters of the lognormal distribution is given below.

$$\hat{\mu} = \frac{1}{n}\sum_1^n \ln(X_i - \hat{\gamma}) \quad (10)$$

and

$$\hat{\sigma}^2 = \frac{1}{n}\sum_1^n \ln^2(X_i - \hat{\gamma}) - \left[\frac{1}{n}\sum_1^n (X_i - \hat{\gamma})\right]^2 \quad (11)$$

Where $\hat{\gamma}$ is assumed to be zero for the two parameters log normal. The value of $\hat{\gamma}$ is are often estimated using numerical means.

2.2.7. Generalized Gamma Distribution (Two and Three Parameters)

The generalized gamma distribution is a continuous probability distribution characterized by two shape and scale parameters. In the literature, the gamma distribution is one of the most popular distributions for modeling precipitation amounts. The probability density of the generalized gamma distribution $(\mathrm{GG}(\alpha,\tau,\lambda))$ is given by

$$f(y \mid \alpha,\tau,\lambda) = \frac{\tau}{\lambda\Gamma(\alpha)}\left(\frac{y}{\lambda}\right)^{\alpha\tau-1} e^{-\left(\frac{y}{\lambda}\right)^{\tau}} \quad y \ge 0,\ \tau,\alpha,\lambda > 0 \quad (12)$$

where $\Gamma(.)$ *is the gamma function*, α and τ are shape parameters, and λ is the scale parameter.

2.2.8. Weibull Distribution

The Weibull distribution is a widely used continuous probability distribution that has been applied to examine product dependability, analyze life statistics, and model failure times across several fields, such as biology, economics, and engineering [32].

The probability density function of a Weibull random variable is:

$$f(x;\lambda,k)=\begin{cases}\frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1}e^{-(x/\lambda)^k}, & x\geq 0,\\ 0, & x<0,\end{cases} \quad (13)$$

where $k>0$ is the shape parameter and $\lambda>0$ is the scale parameter of the distribution.

The mean and variance of the Weibull distribution are given as:

Mean= $\lambda\Gamma(1+1/k)$.

Variance = $\lambda^2\left[\Gamma\left(1+\frac{2}{k}\right)-\left(\Gamma\left(1+\frac{1}{k}\right)\right)^2\right]$.

2.3. Model Evaluation

2.3.1. Validation of Distribution Fit

When modeling precipitation amounts, choosing the distribution that best fits the data for days with precipitation is crucial. This process is commonly known in the literature as goodness-of-fit testing and is extensively employed to determine distribution fitting. This research uses the Akaike Information Criterion to assess the fit of the distribution.

2.3.2. Akaike Information Criterion (AIC-Hirotugu Akaike [33])

The Akaike Information Criterion (AIC) was among the first model selection criteria to gain widespread recognition in the statistical community as a method for evaluating the effectiveness of linear regression models. The fundamental concept of the Akaike Information Criterion (AIC) is to identify a model that minimizes the loss of information in the given data. The formulation for the AIC is given as: $AIC=-2\ln(L)+2K$.

Where L represents the log-likelihood function in this equation, commonly used in statistical modeling, K denotes the number of parameters in the model being analyzed. A penalty term, represented as K, is associated with the number of parameters. When deciding on distribution fit, smaller AIC values are preferred.

2.3.3. Simulation of Precipitation

The subsequent steps are used to simulate precipitation for the periods 1981-2022 and 2023-2030. The initial three steps encompass the processes required to generate precipitation occurrences. The final step involves the procedure needed for generating precipitation amounts. The combination of these steps is essential to the formulation of the precipitation model.

Step 1: Transform your data into a binary representation of wet and dry states (0 for dry, 1 for wet). Employ this binary state to construct a transition probability matrix, which outlines the transitions from a present state to a future state.

Step 2: Use the transition probability matrix results to construct the final state vector.

Step 3: For generating precipitation occurrences throughout the projected period, employ a random number generator sourced from a uniform distribution with a range of 0 to 1. If the generated random number is smaller than π_d, the state is considered dry; otherwise, it is deemed wet. Next, produce another random number within the range of 0 to 1 and compare it to the transition probability for transitioning from the current state to an alternative state. If the number is equal to or less than the probability of transitioning to the next state, the subsequent state will be that state. Otherwise, you will be in the other state.

Step 4: Repeat step 3 until you have generated all the days required.

Step 5: Once the precipitation occurrence has been determined, random values are generated from the probability distribution fitted to the precipitation amount on wet days. These values correspond to the projected amount of rain expected on days identified as wet, as obtained from the precipitation occurrence model.

3. RESULTS

3.1. Results of the Precipitation Occurrence Model

Tables 1 and 2 present the transition matrix results for each month from 1981 to 2022 across the Black Bush Polder and Ebini regions. As indicated by the transition matrix and consistent with traditional seasonal rainfall patterns, May, June, and July exhibited the lowest transition probabilities from a wet day to a dry day at both study locations, with values below 0.3. Conversely, these three months also showed the highest likelihood of transitioning from a wet day to another wet day, with probabilities exceeding 0.7. Throughout this time span, January, February, September, and October consistently demonstrated the least likelihood of transitioning from a wet day to a dry day, with probabilities ranging from 0.48 to 0.64.

Table 1. Transition probability matrix for each month over Black Bush Polder from 1981-2022.

State	Dry	Wet	Month	State	Dry	Wet	Month
Dry	0.760	0.240	**January**	Dry	0.563	0.437	**July**
Wet	0.360	0.640		Wet	0.283	0.717	
Dry	0.742	0.258	**February**	Dry	0.696	0.304	**August**
Wet	0.368	0.632		Wet	0.359	0.641	
Dry	0.776	0.224	**March**	Dry	0.784	0.216	**September**
Wet	0.372	0.628		Wet	0.475	0.525	
Dry	0.736	0.264	**April**	Dry	0.778	0.222	**October**
Wet	0.353	0.647		Wet	0.496	0.504	
Dry	0.670	0.330	**May**	Dry	0.750	0.250	**November**
Wet	0.259	0.741		Wet	0.350	0.650	
Dry	0.528	0.472	**June**	Dry	0.729	0.271	**December**
Wet	0.261	0.739		Wet	0.291	0.709	

Table 2. Transition probability matrix for each month over Ebini from 1981-2022.

State	Dry	Wet	Month	State	Dry	Wet	Month
Dry	0.754	0.246	**January**	Dry	0.442	0.558	**July**
Wet	0.509	0.491		Wet	0.287	0.713	
Dry	0.780	0.220	**February**	Dry	0.510	0.490	**August**
Wet	0.518	0.482		Wet	0.380	0.620	
Dry	0.811	0.189	**March**	Dry	0.666	0.334	**September**
Wet	0.503	0.497		Wet	0.445	0.555	
Dry	0.761	0.239	**April**	Dry	0.681	0.319	**October**
Wet	0.389	0.611		Wet	0.471	0.529	
Dry	0.570	0.430	**May**	Dry	0.708	0.292	**November**
Wet	0.293	0.707		Wet	0.424	0.576	
Dry	0.378	0.624	**June**	Dry	0.749	0.251	**December**
Wet	0.261	0.739		Wet	0.443	0.556	

From the transition probabilities reported in Tables 1 and 2, the likelihood of transitioning from a wet day to another wet day is consistently greater than the probability of transitioning from a dry day to a wet day. Figure 5 visually presents this relationship for Black Bush Polder, while Figure 6 visually exhibits the same behaviour for Ebini, with $P(W \rightarrow W)$ remaining larger than $P(D \rightarrow W)$ throughout the year. Hence, $p_{DW} < \pi_W < p_{WW}$.

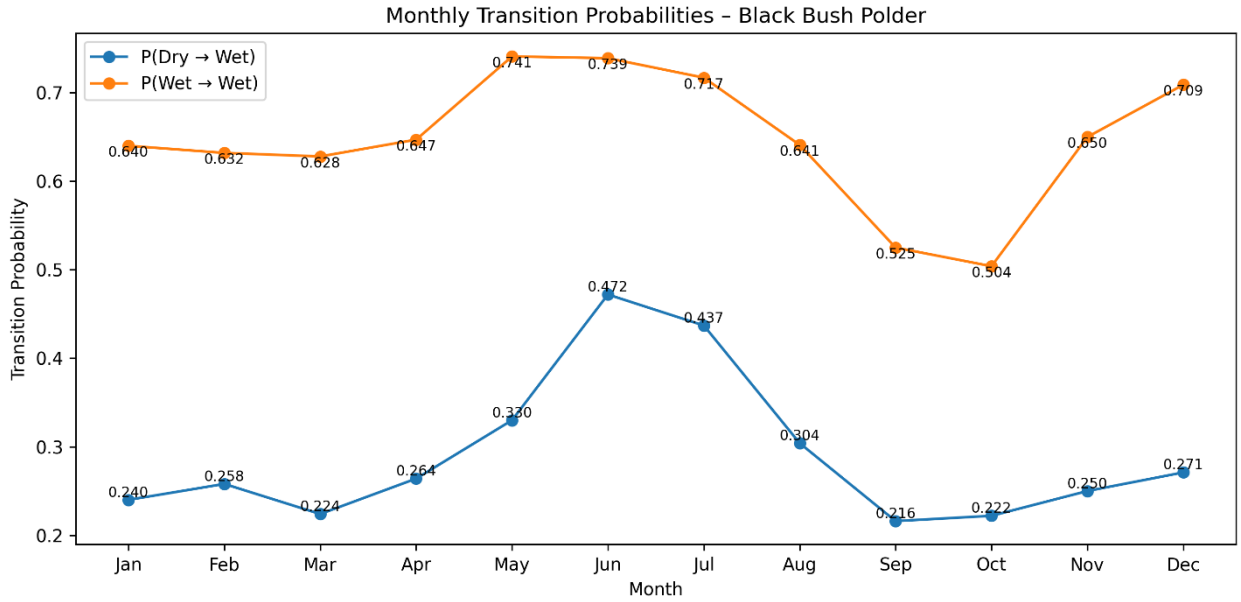


Figure 5. Line graph showing transition states (DW-dry to wet, WW- wet to wet) over Black Bush Polder.

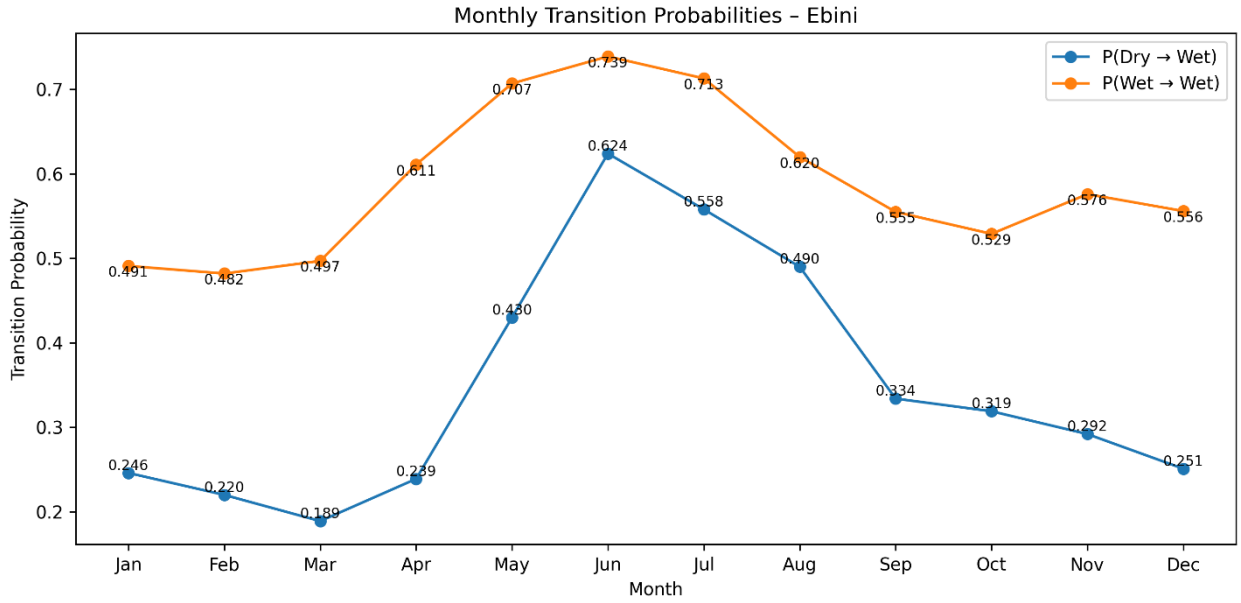


Figure 6. Line graph showing transition states (DW-dry to wet, WW- wet to wet) over Ebini.

3.1.1. Steady State Vector

The transition probability matrix provides essential data for computing the steady-state vector, which offers insights into the long-term average occurrence of precipitation events. Analysis of the steady-state vectors reveals that May, June, and July consistently showed probabilities exceeding 0.55 for encountering wet conditions across both study locations. These values suggest that these months experienced wet days at least 55% of the time over an extended period. This observation aligns with Guyana's established seasonal rainfall pattern, which occurs in May-June/July. Furthermore, noticeable indications of increased likelihood of wet conditions are evident in Black Bush Polder during August and December, and in Ebini, specifically during August.

Table 3. Final state probability vectors over Black Bush Polder from 1981-2022.

January	February	March	April
[0.600, 0.400]	[0.588, 0.413]	[0.624, 0.376]	[0.572, 0.428]
May	**June**	**July**	**August**
[0.439, 0.561]	[0.356, 0.644]	[0.393, 0.607]	[0.542, 0.458]
September	**October**	**November**	**December**
[0.687, 0.313]	[0.691, 0.310]	[0.583, 0.417]	[0.518, 0.482]

Table 4. Final state probability vectors over Ebini from 1981-2022.

January	February	March	April
[0.675, 0.325]	[0.702, 0.298]	[0.726, 0.274]	[0.620, 0.381]
May	**June**	**July**	**August**
[0.405, 0.595]	[0.295,0.706]	[0.340, 0.660]	[0.437, 0.563]
September	**October**	**November**	**December**
[0.5701, 0.429]	[0.596, 0.404]	[0.591, 0.409]	[0.639, 0.361]

Table 5. Descriptive statistics of observed precipitation over Black Bush Polder 1981-2022.

Months	Maximum	Mean	Std. Deviation	Variance	Rain days
January	244.590	10.480	20.120	404.900	520
February	160.870	9.670	16.650	277.360	489
March	137.900	9.090	17.460	304.890	490
April	90.000	11.180	15.490	239.800	539
May	154.600	13.520	15.450	238.620	732
June	131.100	13.200	15.370	236.070	827
July	107.100	13.240	15.880	252.180	789
August	200.000	11.920	18.100	327.410	596
September	72.000	7.070	10.540	111.100	396
October	144.800	6.690	12.100	146.340	402
November	72.000	7.350	8.900	79.270	525
December	221.480	10.490	18.720	350.250	631
Total					**6936**

Table 6. Descriptive statistics of observed precipitation over Ebini, 1981-2022.

Month	Maximum	Mean	Std. Deviation	Rain days
January	236.900	17.200	28.440	423
February	76.820	10.360	12.440	353
March	127.810	13.610	17.170	357
April	83.890	14.320	14.790	481
May	74.110	14.200	10.840	775
June	86.010	13.590	10.540	890
July	116.770	15.480	12.840	859
August	61.090	11.850	9.520	734
September	51.490	9.670	8.360	542
October	45.910	8.280	7.350	525
November	59.240	9.540	9.590	514
December	94.120	16.250	18.350	471
Total				**6924**

The final-state probability vectors in Table 3 show that, for Black Bush Polder, the steady-state probabilities of a dry and a wet day were 0.6867 and 0.3131, respectively, during the 1981–2022 period. These values indicate that, on average, approximately 69 percent of days were dry, while 31 percent were wet. Such a distribution is characteristic of coastal Guyana's rainfall regime, which alternates between extended dry periods and concentrated wet spells. The

relatively high persistence of dry days also suggests that rainfall events tend to occur in short bursts rather than evenly throughout time, emphasizing the importance of temporal clustering in stochastic modeling.

The monthly transition behavior for Ebini is summarized in Table 4 and provides further insight into the seasonal dynamics of rainfall occurrence. During the first four months of the year (January to April), the probability of remaining dry ranged from 0.62 to 0.73, with corresponding wet-day probabilities below 0.40.

This dry dominance reflects the secondary dry season typically observed along Guyana's coast. In contrast, from May to August, the likelihood of a wet day exceeded 0.50, peaking at 0.7055 in June. These months coincide with the primary rainy season, when the Intertropical Convergence Zone (ITCZ) reaches its northernmost position, bringing increased convection and moisture inflow.

A gradual decline in wet-day probability is seen from September onward, with December (0.3609) marking a return to drier conditions. The bimodal pattern in the data closely corresponds to the well-documented double wet season across much of the country.

Descriptive statistics of observed precipitation for Black Bush Polder, presented in Table 5, reinforce these findings. The mean monthly rainfall ranged from 6.69 mm in October to 13.52 mm in May, while the standard deviation ranged from 8.90 mm to 20.12 mm, indicating notable variability within each month. The highest monthly maximum (244.59 mm) occurred in January, reflecting occasional intense rainfall even during the relatively dry season.

Rain-day frequencies also showed strong seasonality, peaking in June (827) over the 42-year period, followed by July (789), and declining sharply in September (396). The coefficient of variation (standard deviation divided by the mean) exceeded 1.0 in most months, indicating high dispersion and supporting the suitability of skewed distributions, such as the gamma and lognormal, for modeling rainfall amounts.

In contrast, rainfall characteristics at Ebini (Table 6) show a slightly wetter but more irregular regime. Mean monthly rainfall ranged from 8.28 mm in October to 17.20 mm in January, and standard deviations were generally higher than those observed at Black Bush Polder. The maximum recorded rainfall of 236.90 mm in January indicates occasional heavy events even outside the peak rainy season.

The total number of rainy days (6,924) was comparable to that of Black Bush Polder (6,936), but with greater month-to-month variation. This pattern reflects the influence of localized convection and interior climatic dynamics, which the ITCZ influences less than coastal regions. The rainfall regime at Ebini is therefore characterized by more frequent but less predictable rainfall events, consistent with inland tropical climates.

These results underscore the seasonal rainfall distribution in Ebini over the past four decades and highlight the advantage of fitting distributions to model the frequency of rainy days within each documented month.

3.1.2. Parameter Estimates of Each Distribution

Maximum Likelihood Estimation (MLE) is employed to estimate the distribution's parameters. The parameter estimation for each month is shown in Tables 4 and 5.

The most suitable distribution for fitting the data on wet days in each month was determined using the AIC method. Lower AIC values are considered optimal. Tables 7 and 8 present the selected distributions for each month are based on the dual criteria of low AIC values.

Good fits were obtained for almost all months in Ebini; however, this was not the case in New Amsterdam, where some months had moderate fits (September, October, January, February). As a result, multiple simulations had to be run to achieve alignment with observed data.

Table 7. Best fitting distributions by months over Black Bush Polder from 1981-2022.

Month	Distribution	Parameters			
		Location	Shape	Scale	Threshold
January	log normal	1.293		1.649	
February	log normal	1.184		1.618	
March	log normal	1.006		1.634	
April	log normal	1.423		1.583	
May	Gamma (3P)		0.672	19.980	0.099
June	Gamma (3P)		0.693	18.920	0.099
July	Gamma (3P)		0.730	18.090	0.046
August	Weibull		0.735	9.720	
September	Weibull		0.694	5.456	
October	Log normal	0.860		1.596	
November	Weibull		0.805	6.491	
December	lognormal	1.264		1.593	

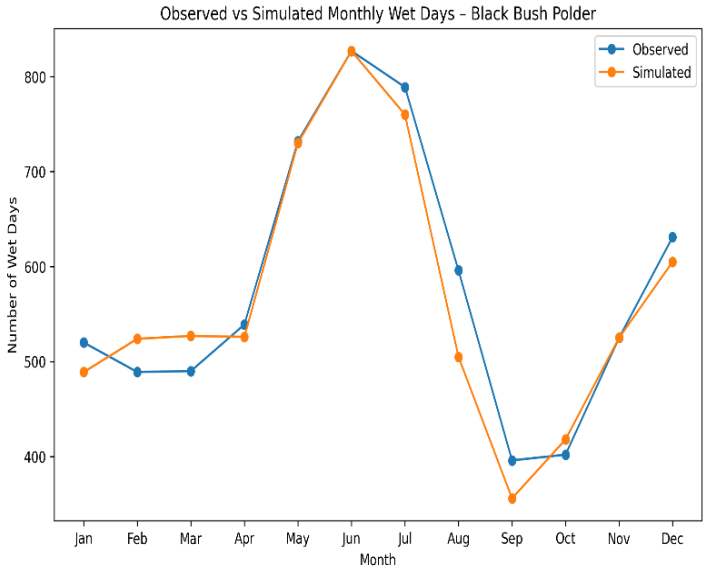
Table 8. Best fitting distributions by months over Ebini from 1981-2022.

Month	Distribution	Parameters			
		Location	Shape	Scale	Threshold
January	lognormal	1.846		1.541	
February	Gamma (3P)		0.721	14.170	0.141
March	Weibull		0.787	11.870	
April	Gamma (3P)		0.848	16.840	0.035
May	Weibull		1.263	15.220	
June	Weibull		1.307	14.700	
July	Weibull		1.195	16.410	
August	Weibull		1.214	12.600	
September	Weibull		1.119	10.060	
October	Gamma		1.088	7.612	
November	Gamma (3P)		0.889	10.680	0.045
December	Gamma (3P)		0.736	22.070	0.008

3.2. Simulation Results

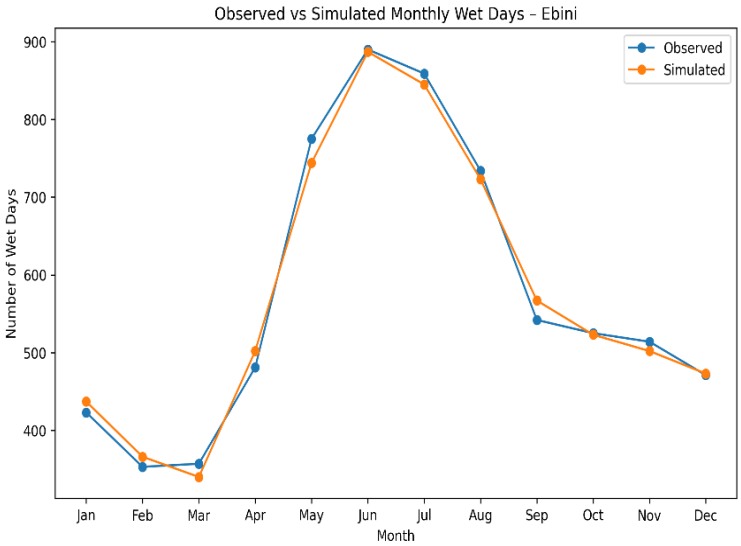
Simulation data spanning 50 years (1981-2030) was generated using the Stochastic Precipitation Model employing a Markov chain approach across both study locations. This data was divided into two periods: simulated data from 1981 to 2022 and from 2023 to 2030. Figures 7 and 8 compare simulated data with the observed number of rainy days from 1981 to 2022 across both study locations.

The figures indicate that the first-order Markov chain model effectively generates precipitation occurrences in Black Bush Polder and Ebini regions. Remarkably, the simulation of precipitation occurrences demonstrated high accuracy for most months, with May, June, July, and November showing the highest accuracy for Black Bush Polder, and June, July, August, and October showing high accuracy for Ebini.



Month	Observed	Simulated
Jan	520	489
Feb	489	524
Mar	490	527
Apr	539	526
May	732	730
Jun	827	827
Jul	789	760
Aug	596	505
Sep	396	356
Oct	402	418
Nov	525	525
Dec	631	605

Figure 7. Line graph showing the number of observed vs simulated wet days over Black Bush from 1981-2022.



Month	Observed	Simulated
Jan	423	437
Feb	353	366
Mar	357	340
Apr	481	502
May	775	744
Jun	890	887
Jul	859	845
Aug	734	723
Sep	542	567
Oct	525	523
Nov	514	502
Dec	471	473

Figure 8. Line graph showing the number of observed vs simulated wet days over Ebini from 1981-2022.

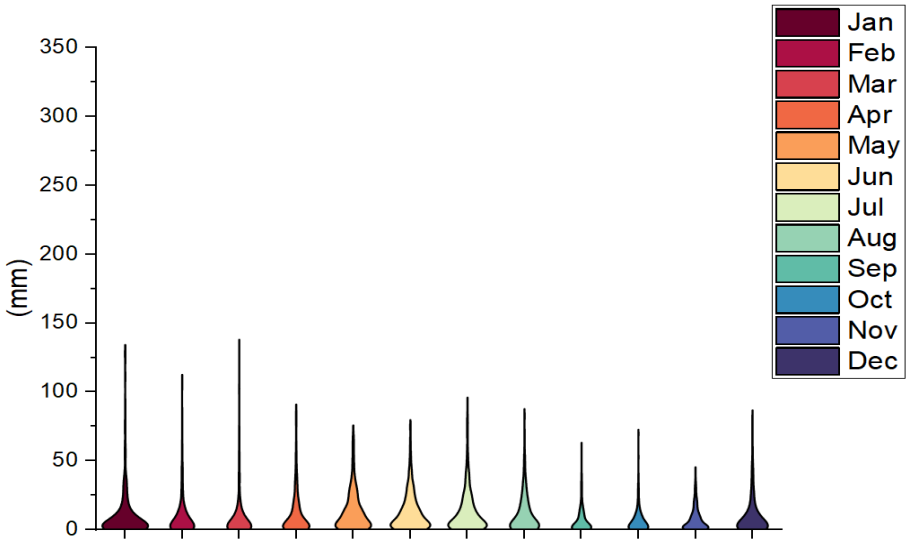


Figure 9. Violin plot showing observed rainfall distribution by months over Black Bush Polder 1981-2022.

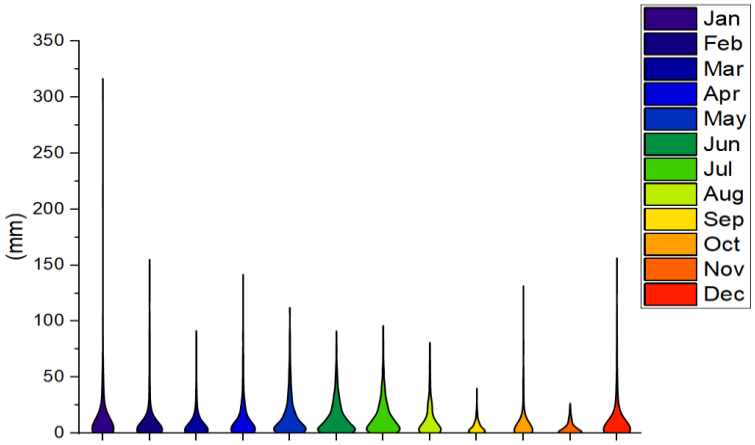


Figure 10. Violin plot showing simulated rainfall distribution by months over Black Bush Polder 1981-2022.

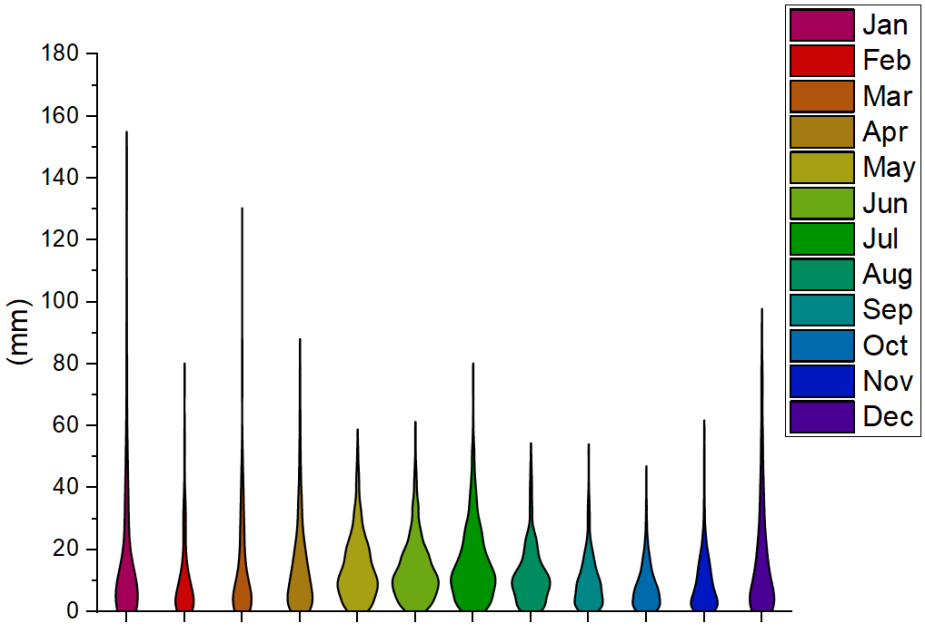


Figure 11. Violin plot showing observed rainfall distribution by months over Ebini 1981-2022.

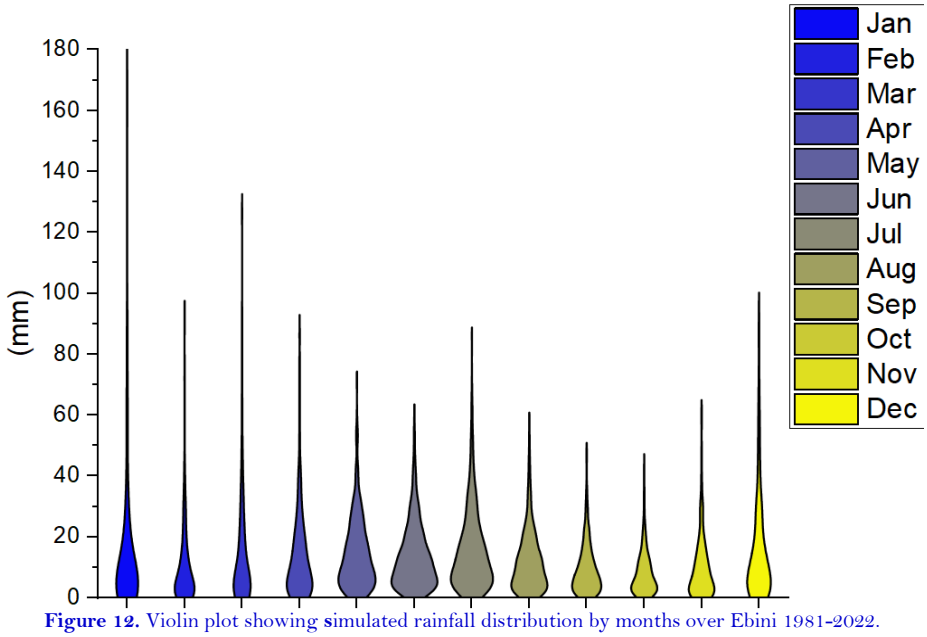


Figure 12. Violin plot showing simulated rainfall distribution by months over Ebini 1981-2022.

3.3. Comparison Between Observed and Simulated Precipitation Occurrence and Amount

Various descriptive measures have been computed to analyze the simulated and observed data. These measures include the mean, median, maximum, minimum, and the lower and upper quartiles. The outcomes of these descriptive measures for both simulated and observed data are presented in the analysis are presented in Tables 9 & 10.

Comparing observed and simulated precipitation data from 1981 to 2022 in the Black Bush Polder and the Ebini observed and simulated data, except for September and October in New Amsterdam. The overall alignment in average precipitation demonstrates the simulation model's ability to represent it accurately. This suggests that, on average, the simulated data captures the central tendency of precipitation for both regions reasonably well. Moreover, in the Ebini region, the simulated data displays closer proximity to observed mean and median values across months, highlighting a potentially more accurate representation compared to the Black Bush Polder.

Table 9. Descriptive statistics for Observed and simulated precipitation over Black Bush Polder 1981-2022.

Month	Mean	Median	Minimum	Maximum	Q1	Q3
January_sim	12.370	3.890	0.010	265.440	1.100	11.460
January_obs	11.810	4.260	0.100	244.590	1.000	11.500
February_sim	10.230	3.110	0.040	155.950	1.160	9.420
February_obs	9.670	3.840	0.070	160.870	1.000	10.950
March_sim	9.290	3.110	0.020	204.820	1.100	8.640
March_obs	9.090	2.880	0.100	137.900	0.700	9.130
April_sim	12.880	4.660	0.050	133.510	1.730	13.690
April_obs	11.170	4.720	0.100	90.000	1.200	15.300
May_sim	13.930	6.910	0.100	118.490	2.360	18.120
May_obs	13.520	8.660	0.100	154.600	2.460	19.500
June_sim	13.220	8.000	0.100	135.330	2.540	18.480
June_obs	13.200	7.930	0.100	131.100	2.530	18.140
July_sim	14.030	8.980	0.050	93.740	3.080	19.460
July_obs	13.240	7.240	0.050	107.100	2.050	18.810
August_sim	11.780	5.970	0.000	109.160	1.800	15.230
August_obs	11.920	5.400	0.100	200.00	1.440	16.150
September_sim	5.500	3.550	0.020	38.980	1.540	7.640
September_obs	7.070	3.00	0.010	72.000	0.700	9.710
October_sim	8.740	2.280	0.020	205.390	0.700	6.030
October_obs	6.690	3.000	0.030	144.800	0.700	8.220
November_sim	7.520	3.980	0.000	60.060	1.220	9.900
November_obs	7.340	3.800	0.100	72.000	1.000	10.550
December_sim	12.410	3.270	0.020	291.590	1.110	9.860
December obs	10.490	4.020	0.100	221.480	1.000	11.250

Table 10. Descriptive statistics for Observed and simulated precipitation over Ebini from 1981-2022.

Month	Mean	Median	Minimum	Maximum	Q1	Q2	Q3
January_sim	17.230	6.060	0.030	208.230	2.110	6.060	17.720
January_obs	17.200	7.350	0.050	236.900	2.160	7.350	17.710
February_sim	10.520	5.580	0.140	95.090	1.620	5.580	14.000
February_obs	10.360	6.560	0.140	76.820	1.660	6.560	13.000
March_sim	14.640	7.370	0.010	127.380	2.240	7.370	20.220
March_obs	13.610	6.930	0.060	127.810	2.420	6.930	19.290
April_sim	14.080	9.170	0.030	117.000	3.240	9.170	19.300
April_obs	14.320	9.910	0.030	83.890	3.500	9.910	19.520
May_sim	14.500	11.770	0.050	71.630	5.640	11.770	20.840
May_obs	14.200	11.710	0.140	74.100	6.320	11.710	19.830
June_sim	13.600	10.860	0.050	62.078	5.350	10.860	19.220
June_obs	13.570	11.140	0.140	86.010	6.230	11.140	18.250
July_sim	15.910	12.150	0.110	95.380	5.910	12.150	21.540

Month	Mean	Median	Minimum	Maximum	Q1	Q2	Q3
July_obs	15.480	12.450	0.130	116.770	6.540	12.450	21.410
August_sim	12.100	9.560	0.020	60.720	4.078	9.560	17.230
August_obs	11.850	10.080	0.110	61.090	5.320	10.080	15.960
September_sim	9.500	7.120	0.060	48.930	3.410	7.120	13.140
September_obs	9.670	8.090	0.070	51.490	3.470	8.090	13.760
October_sim	8.220	5.960	0.000	46.480	2.720	5.960	11.680
October_obs	8.280	6.610	0.090	45.910	2.730	6.610	11.370
November_sim	9.950	6.800	0.050	61.540	2.240	6.800	13.910
November_obs	9.540	6.830	0.050	59.240	2.550	6.830	13.720
December_sim	16.520	10.010	0.010	126.350	3.360	10.010	23.430
December obs	16.250	9.760	0.010	94.120	2.680	9.760	23.120

This alignment indicates that the simulation model better predicts average precipitation in this region.

Quartile analyses show consistent distributions in both the observed and simulated sets, underscoring the simulation model's ability to capture the central 50% of precipitation values. Despite occasional variations in extreme values, which may indicate outliers or model limitations, the overall consistency of quartile ranges suggests a degree of reliability in representing typical precipitation patterns.

Figures 9–12 present violin plots of observed and simulated monthly rainfall distributions at Black Bush Polder and Ebini. At Black Bush Polder, months such as May, June, and July exhibit broader distributions and higher upper tails, consistent with elevated means (≈15–16 mm) and large maxima reported in Table 9. In contrast, September and October show more concentrated distributions at lower rainfall values, reflecting reduced means and quartiles. At Ebini, the violin plots indicate consistently wetter conditions, particularly during January, May–July, and December, where higher means, medians, and extreme values are observed (Table 10). Across both locations, the elongated upper tails visible in several months confirm the presence of extreme rainfall events and the strongly right-skewed nature of the rainfall distributions.

These findings indicate that the simulation model performs well in approximating average precipitation trends in both regions. This supports the potential utility of these models for forecasting and environmental management but requires continued validation and refinement to improve accuracy during extreme precipitation events.

3.4. Precipitation Projections for 2023-2030

The two-step precipitation model was employed to simulate precipitation from 2023 to 2030 over both study locations. The descriptive measures are presented in Table 11 and Table 12. As observed in the previous simulation and consistent with historical data, the months of May, June, and July had the highest number of wet days across both locations, recording 138, 150, and 126 wet days for Black Bush Polder, and 139, 180, and 177 days for the Ebini region (See Figures 13 and 14). Over this period, September and October once again had the fewest wet days for Black Bush Polder, while February and March had the lowest number of wet days for Ebini.

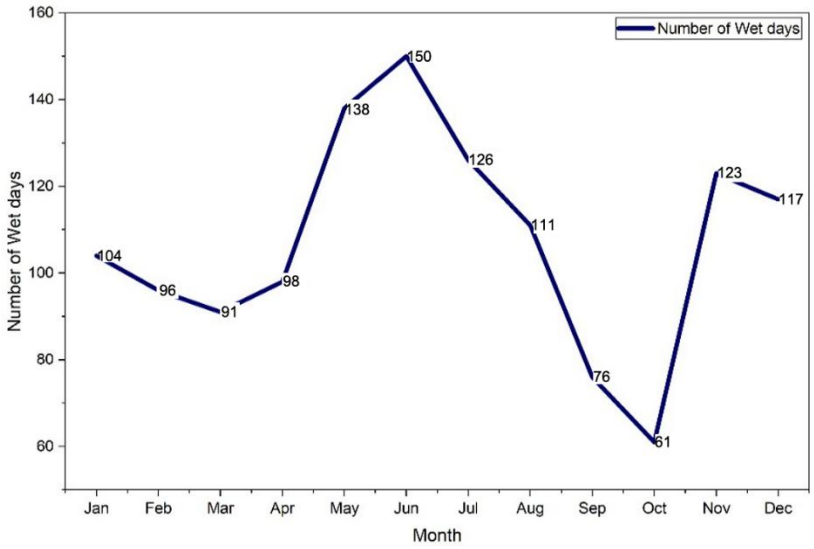


Figure 13. Line graph showing the simulated number of wet days from 2023-2030 over Black Bush Polder.

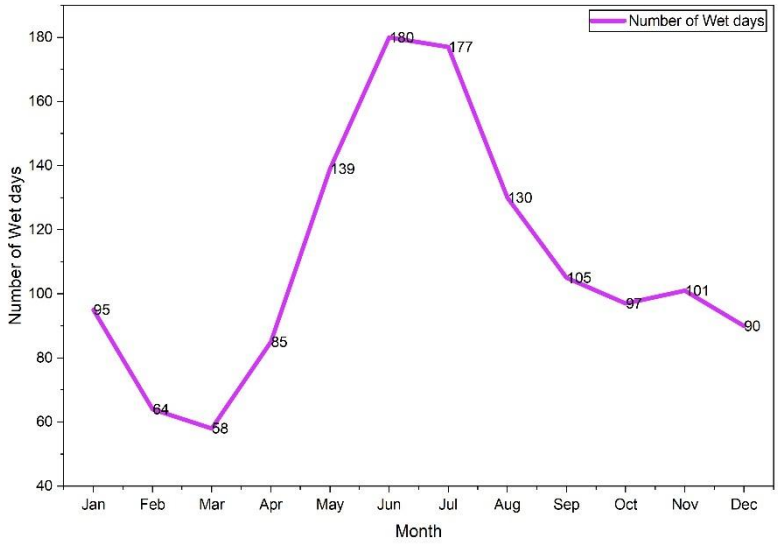


Figure 14. Line graph showing the simulated number of wet days from 1981-2022 for Ebini.

Table 11. Descriptive statistics for simulated precipitation over Black Bush Polder from 2023-2030.

Month	Mean	Median	STD deviation	Maximum	Percentiles	
					Q1	Q3
January	14.250	2.640	26.000	117.300	0.650	12.010
February	6.100	2.830	10.210	74.780	0.690	7.300
March	7.920	2.810	15.790	124.830	0.840	6.980
April	8.790	2.970	13.870	67.610	1.130	8.330
May	13.160	7.490	16.410	88.050	3.270	15.390
June	14.480	8.760	15.300	90.620	2.790	22.850
July	13.130	7.670	14.700	84.350	3.110	18.880
August	13.042	8.060	16.380	105.790	1.600	17.170
September	7.760	2.930	11.112	64.310	1.180	10.020
October	6.350	2.550	8.790	51.100	1.190	9.070
November	7.370	4.040	8.440	40.530	1.450	10.840
December	9.330	2.570	15.640	101.350	1.180	11.330

Table 12. Descriptive statistics for simulated precipitation over Ebini from 2023-2030.

Month	Mean	Median	Std. Deviation	Maximum	Percentiles	
					Q1	Q3
January	16.860	5.280	29.460	169.150	2.290	15.530
February	8.870	4.730	11.340	43.390	1.440	10.470
March	15.250	8.310	19.950	95.990	2.040	20.830
April	14.580	8.820	17.180	83.700	3.010	19.210
May	13.850	10.110	11.370	66.680	5.330	19.780
June	14.630	12.620	9.680	58.280	7.720	20.200
July	15.000	12.220	11.780	63.590	5.840	22.230
August	12.320	10.240	9.710	46.010	4.380	17.590
September	9.680	7.080	8.510	38.110	2.590	14.170
October	6.870	4.290	6.980	29.250	1.950	9.170
November	9.640	6.250	8.840	35.550	2.640	16.270
December	15.730	8.310	18.600	86.660	3.560	21.410

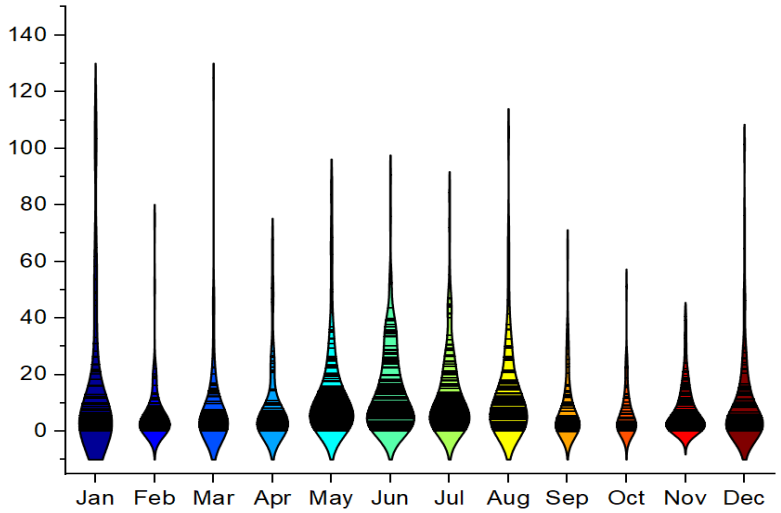


Figure 15. Violin plot of Simulated precipitation over Black Bush Polder from 2023-2030.

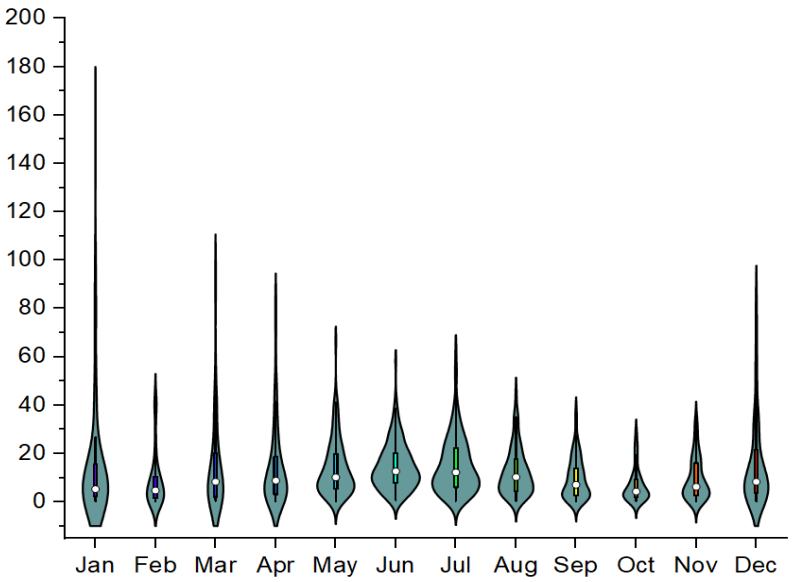


Figure 16. Violin with box plot of precipitation over Ebini from 2023-2030.

Table 11 and Table 12 provide insight into the simulated precipitation patterns over Black Bush Polder and Ebini from 2023 to 2030. In Black Bush Polder, the simulated precipitation from 2023 to 2030 shows notable variations across months. January and June had higher mean precipitation, indicating potential periods of increased rainfall.

Moreover, January showed substantial variability, with a higher standard deviation (25.997), suggesting significant fluctuations in precipitation levels. In contrast, October and November display lower variability, indicating more consistent precipitation during these months, which aligns with historical trends.

On the other hand, Ebini exhibited slightly different precipitation patterns over the 2023-2030 simulation period. During this timeframe, January, March, and December stood out for higher mean precipitation. While January and March exhibit greater variability, October and February show lower fluctuations in simulated precipitation.

Quartile analyses emphasize differences in data spreads between these regions. Ebini showed wider interquartile ranges in June and July, indicating more significant precipitation variability than the Black Bush Polder.

The distributional characteristics of the projections are further illustrated in Figures 15 and 16, which present violin plots of the projected monthly precipitation over Black Bush Polder and Ebini for the period 2023–2030. The plots visually highlight the seasonal contrasts identified in Tables 11 and 12, with broader distributions during the mid-year months and more concentrated rainfall during drier periods. In addition, the presence of elongated upper tails in several months reflects the occurrence of intense rainfall events within the projection period. Overall, the violin plots complement the tabulated statistics by emphasizing the variability and seasonal structure of projected precipitation at both locations.

4. DISCUSSION

The results of this study demonstrate that the first-order Markov chain model, combined with probability distribution fitting, is effective in capturing rainfall occurrence and intensity across the two selected regions in Guyana. The transition probabilities revealed clear seasonal patterns, with higher wet-to-wet transition probabilities during the rainy seasons, consistent with known climatic behavior in the country. The fitted probability distributions, particularly the Gamma and Exponential families, successfully reproduced the statistical properties of daily rainfall amounts, supporting the suitability of these models for rainfall simulation in tropical climates. Together, these findings indicate that stochastic rainfall models can serve as valuable tools for generating synthetic rainfall series to support agricultural planning, hydrological studies, and climate risk assessment in Guyana.

A key contribution of this research is addressing a gap in the existing literature: although rainfall variability and trends have been studied in Guyana [29], no prior work has applied stochastic simulation methods to generate synthetic rainfall sequences. By establishing a framework for rainfall simulation, this study not only advances local rainfall research but also provides a methodological foundation for future applications, such as runoff modeling, drought analysis, and crop yield forecasting.

Despite these contributions, several limitations must be acknowledged. First, the use of a first-order Markov chain, while effective, may underestimate the duration of prolonged dry or wet spells, as observed in other contexts [16]. Future research could explore higher-order models or hidden Markov models to better capture the persistence of rainfall patterns. Second, although multiple probability distributions were tested, extreme rainfall events remain challenging to model accurately, suggesting a potential role for heavy-tailed or mixed distributions in future studies.

Another limitation relates to data availability. Long-term, high-quality observed rainfall data were not available for all study locations. To address this, reanalysis data obtained through Climate Engine were used as a supplementary source. While reanalysis products are valuable for filling gaps, they rely on model assimilation. They may not always reflect localized rainfall extremes, especially in regions with sparse gauge networks such as Guyana. This reliance introduces some uncertainty into the simulations. Expanding the observational network in Guyana and validating reanalysis products against ground-based measurements should therefore be a priority for future work.

In summary, the study demonstrates the feasibility and utility of stochastic rainfall simulation for Guyana and identifies methodological and data-related limitations. By refining models and improving data availability, future research can further enhance the accuracy and applicability of rainfall simulations, providing more substantial support for water resource management and agricultural decision-making in the face of climate variability.

5. CONCLUSION

This study represents the first attempt to apply a stochastic rainfall simulation framework in Guyana, combining a first-order Markov chain for rainfall occurrence with probability distribution fitting for rainfall amounts. The results demonstrate that this approach effectively reproduces the seasonal variability and statistical characteristics of rainfall in two distinct regions: Black Bush Polder and Ebini. Transition probabilities captured the persistence of wet and dry spells, while fitted probability distributions, particularly the Gamma distribution, provided a good representation of daily rainfall intensity.

By filling a critical gap in the literature, this work establishes a methodological foundation for the use of stochastic rainfall simulations in Guyana. The ability to generate synthetic rainfall series has direct applications for hydrological modeling, agricultural planning, and climate risk assessment, offering decision-makers valuable tools for managing water resources under uncertain future conditions.

At the same time, limitations remain, including reliance on reanalysis data when long-term observed records are unavailable and challenges in accurately capturing extreme rainfall events. Future research should therefore focus on expanding observational networks, validating reanalysis datasets, and exploring higher-order or hybrid stochastic models that can better represent persistence and extremes.

Overall, this study highlights the feasibility and importance of stochastic rainfall modeling in data-sparse regions and provides a stepping stone for further research that can strengthen climate resilience and support sustainable agricultural and water resource management in Guyana.

Funding: This study received no specific financial support.
Institutional Review Board Statement: Not applicable.
Transparency: The authors state that the manuscript is honest, truthful, and transparent, that no key aspects of the investigation have been omitted, and that any differences from the study as planned have been clarified. This study followed all writing ethics.
Competing Interests: The authors declare that they have no competing interests.
Authors' Contributions: All authors contributed equally to the conception and design of the study. All authors have read and agreed to the published version of the manuscript.

REFERENCES

[1] J. Lee, D. Perera, T. Glickman, and L. Taing, "Water-related disasters and their health impacts: A global review," *Progress in Disaster Science*, vol. 8, p. 100123, 2020. https://doi.org/10.1016/j.pdisas.2020.100123

[2] WHO, "Floods," Retrieved: https://www.who.int/health-topics/floods?utm. [Accessed Dec. 02, 2025], 2025.

[3] B. Kang and J. A. Ramírez, "A coupled stochastic space-time intermittent random cascade model for rainfall downscaling," *Water Resources Research*, vol. 46, no. 10, p. 10534, 2010. https://doi.org/10.1029/2008WR007692

[4] F. Hussain, C. J. Langmead, Q. Mi, J. Dutta-Moscato, Y. Vodovotz, and S. K. Jha, "Automated parameter estimation for biological models using Bayesian statistical model checking," *BMC Bioinformatics*, vol. 16, no. Suppl 17, p. 1-14, 2015. https://doi.org/10.1186/1471-2105-16-S17-S8

[5] United Nations, ""Water – at the center of the climate crisis," Retrieved: https://www.un.org/en/climatechange/science/climate-issues/water?utm. [Accessed Dec. 02, 2025], 2025.

[6] M. Wu and N. Goodman, *Foundation posteriors for approximate probabilistic inference. In S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, & A. Oh (Eds.), Advances in Neural Information Processing Systems*. Red Hook, NY: Curran Associates, Inc, 2022.

[7] J. M. Gregory, T. Wigley, and P. Jones, "Application of Markov models to area-average daily precipitation series and interannual variability in seasonal totals," *Climate dynamics*, vol. 8, no. 6, pp. 299-310, 1993. https://doi.org/10.1007/BF00209669

[8] D. S. Wilks, "Interannual variability and extreme-value characteristics of several stochastic daily precipitation models," *Agricultural and Forest Meteorology*, vol. 93, no. 3, pp. 153-169, 1999. https://doi.org/10.1016/S0168-1923(98)00125-7

[9] E. H. Chin, "Modeling daily precipitation occurrence process with Markov chain," *Water Resources Research*, vol. 13, no. 6, pp. 949-956, 1977. https://doi.org/10.1029/WR013i006p00949

[10] A. Ibeje, J. Osuagwu, and O. Onosakponome, "A Markov model for prediction of annual rainfall," *International Journal of Scientific Engineering and Applied Science*, vol. 3, no. 11, pp. 1-5, 2018.

[11] S. K. Kannan and J. A. Farook, "Stochastic simulation of precipitation using Markov chain-mixed exponential model," *Applied Mathematical Sciences*, vol. 65, no. 9, pp. 3205-3212, 2015.

[12] R. Salakhutdinov, "Learning deep generative models," *Annual Review of Statistics and Its Application*, vol. 2, no. 1, pp. 361-385, 2015. https://doi.org/10.1146/annurev-statistics-010814-020120

[13] C. Taewichit, P. Soni, V. M. Salokhe, and H. P. Jayasuriya, "Optimal stochastic multi-states first-order Markov chain parameters for synthesizing daily rainfall data using multi-objective differential evolution in Thailand," *Meteorological Applications*, vol. 20, no. 1, pp. 20-31, 2013. https://doi.org/10.1002/met.292

[14] M. Tettey, F. T. Oduro, D. Adedia, and D. A. Abaye, "Markov chain analysis of the rainfall patterns of five geographical locations in the south eastern coast of Ghana," *Earth Perspectives*, vol. 4, no. 1, p. 6, 2017. https://doi.org/10.1186/s40322-017-0042-6

[15] P. Guttorp, *Stochastic modelling of scientific data*. New York: Chapman and Hall/CRC, 2018.

[16] P. Racsko, L. Szeidl, and M. Semenov, "A serial approach to local stochastic weather models," *Ecological Modelling*, vol. 57, no. 1-2, pp. 27-41, 1991. https://doi.org/10.1016/0304-3800(91)90053-4

[17] J. Schoof and S. Pryor, "On the proper order of Markov chain model for daily precipitation occurrence in the contiguous United States," *Journal of Applied Meteorology and Climatology*, vol. 47, no. 9, pp. 2477-2486, 2008. https://doi.org/10.1175/2008JAMC1840.1

[18] W. Kleiber, R. W. Katz, and B. Rajagopalan, "Daily spatiotemporal precipitation simulation using latent and transformed Gaussian processes," *Water Resources Research*, vol. 48, no. 1, pp. 1–17, 2012. https://doi.org/10.1029/2011WR011105

[19] D. Sonnadara and D. Jayewardene, "A Markov chain probability model to describe wet and dry patterns of weather at Colombo," *Theoretical and Applied Climatology*, vol. 119, no. 1, pp. 333-340, 2015. https://doi.org/10.1007/s00704-014-1117-z

[20] Z. Li and X. Shi, "Stochastic generation of daily precipitation considering diverse model complexity and climates," *Theoretical and Applied Climatology*, vol. 137, no. 1, pp. 839-853, 2019. https://doi.org/10.1007/s00704-018-2638-7

[21] R. Stern, "Analyse des täglichen Niederschlags in Samaru, Nigeria, mit Anwendung eines zweiteiligen Modells," *Archiv für Meteorologie, Geophysik und Bioklimatologie, Serie B*, vol. 28, pp. 123-135, 1980. https://doi.org/10.1007/BF02243840

[22] M. A. Sharma and J. B. Singh, "Use of probability distribution in rainfall analysis," *New York Science Journal*, vol. 3, no. 9, pp. 40-49, 2010.

[23] P. Jones, C. Harpham, A. Burton, and C. Goodess, "Downscaling regional climate model outputs for the Caribbean using a weather generator," *International Journal of Climatology*, vol. 36, no. 12, pp. 4141-4163, 2016. https://doi.org/10.1002/joc.4624

[24] K. Gutiérrez-García, A. Avilés, A. Nauditt, R. Arce, and C. Birkel, "Evaluating Markov chains and Bayesian networks as probabilistic meteorological drought forecasting tools in the seasonally dry tropics of Costa Rica: K. Gutiérrez-García et al," *Theoretical and Applied Climatology*, vol. 154, no. 3, pp. 1291-1307, 2023. https://doi.org/10.1007/s00704-023-04623-w

[25] M. E. Quesada and P. Waylen, "Variability of daily precipitation on the Caribbean Coast of Costa Rica," *Revista de Climatología*, vol. 20, pp. 61–74, 2020.

[26] M. Santos, P. S. Lucio, and A. C. Gomes, "Stochastic simulation for the precipitation frequency over some Brazilian cities through the Metropolis–Hastings algorithm," *Meteorological Applications*, vol. 23, no. 3, pp. 420-424, 2016. https://doi.org/10.1002/met.1566

[27] R. Kahana, K. Halladay, L. M. Alves, R. Chadwick, and A. J. Hartley, "Future precipitation projections for Brazil and tropical South America from a convection-permitting climate simulation," *Frontiers in Climate*, vol. 6, p. 1419704, 2024. https://doi.org/10.3389/fclim.2024.1419704

[28] E. I. Hsu and S. A. McIlraith, "VARSAT: Integrating novel probabilistic inference techniques with DPLL search," presented at the International Conference on Theory and Applications of Satisfiability Testing, Springer, 2009.

[29] S. O. Eastman and B. Khan, "Statistical analysis of rainfall data : A case study of georgetown, Guyana," *International Journal of Innovative Research in Multidisciplinary Project Studies*, vol. 10, no. 6, pp. 1–12, 2022. https://doi.org/10.37082/IJIRMPS.v10.i6.1612

[30] P. Dash, "A Markov Chain modelling of daily precipitation occurrences of Odisha," *International Journal of Advanced computer and Mathematical sciences*, vol. 3, no. 4, pp. 482-486, 2012.

[31] S. Geng, F. W. P. De Vries, and I. Supit, "A simple method for generating daily rainfall data," *Agricultural and Forest Meteorology*, vol. 36, no. 4, pp. 363-376, 1986. https://doi.org/10.1016/0168-1923(86)90014-6

[32] D. Zhang, Y. Xing, and P. Jiang, "Reliability demonstration test design method based on two-stage test with zero failure data," in *Proceedings of the 2023 Global Reliability and Prognostics and Health Management Conference (GRPHM)*, 2023.

[33] H. Akaike, "A new look at the statistical model identification," *IEEE Transactions on Automatic Control*, vol. 19, no. 6, pp. 716–723, 1974. https://doi.org/10.1109/TAC.1974.1100705