Asia Pac. J. Math. 2026 **13**:143

THE POWER GENERALIZED WEIBULL MIXTURE OF CHI-SQUARE DISTRIBUTION: DISTRIBUTIONAL THEORY AND SPATIAL RETURN-LEVEL MODELING OF EXTREME RAINFALL IN CARAGA, PHILIPPINES

JUNAR T. LINGO*, LORAINE M. GIRON, MILBURN O. MACALOS

Department of Mathematics, Caraga State University, Philippines

*Corresponding author: jtlingo@carsu.edu.ph

Received Jul. 17, 2026

Abstract. This paper develops and applies the *Power Generalized Weibull Mixture of Chi-Square Distribution* (*PGWMCD*) as a flexible probability model for nonnegative, right-skewed data with emphasis on extreme rainfall modeling. The proposed distribution is constructed through a continuous mixture framework in which a chi-square-type conditional distribution is mixed over a power generalized Weibull random parameter. Its probability density function is established, and the validity of the model is verified through nonnegativity and normalization. Integral and series representations are derived for the raw moments, mean, variance, skewness, kurtosis, and cumulative distribution function, thereby providing the theoretical basis for statistical inference and return-level estimation. For the empirical application, annual maximum daily precipitation series from 1981 to 2025 were constructed for the five provinces of the Caraga Region, Philippines. The PGWMCD was fitted together with the generalized extreme value, Weibull, Gamma, Lognormal, and Log-Pearson Type III distributions using numerical maximum likelihood estimation. Model performance was compared primarily through CDF-based root mean square error (RMSE), with Anderson–Darling bootstrap p-values and PDF–histogram overlays used as supplementary goodness-of-fit diagnostics. The PGWMCD produced the smallest RMSE in Agusan del Sur, Dinagat Islands, and Surigao del Norte, whereas the Lognormal and Gamma distributions were preferred in Agusan del Norte and Surigao del Sur, respectively. Rainfall return levels for 2-, 5-, 10-, 25-, 50-, and 100-year return periods were then computed from the RMSE-selected fitted distributions and mapped over the Caraga Region. The results show that the PGWMCD is a theoretically valid and empirically competitive model for annual maximum rainfall, particularly in provinces where additional flexibility in shape and tail behavior improves agreement with the empirical distribution.

2020 Mathematics Subject Classification: Primary 62E15; Secondary 62F10, 62P12.

Power Generalized Weibull Mixture of Chi-Square Distribution; mixture distribution; annual maximum rainfall; numerical maximum likelihood estimation; rainfall return period; CDF-based RMSE; return level mapping; spatial rainfall analysis.

DOI: 10.28924/APJM/13-143

1. Introduction

Extreme precipitation is a persistent source of hydrometeorological risk because the upper tail of the rainfall distribution controls the design magnitudes used in flood mitigation, drainage planning, reservoir safety, agricultural water management, and climate-adaptation decisions. In rainfall frequency analysis, the quantity of interest is often not the mean behavior of precipitation but the return level associated with a specified annual exceedance probability. If X denotes annual maximum daily rainfall and F is its fitted distribution function, the T-year return level is commonly expressed as

$$x_T = F^{-1}\left(1 - \frac{1}{T}\right),$$

so that $P(X > x_T) = 1/T$. This formulation links statistical distribution theory directly to an engineering interpretation: the estimated upper quantiles of the fitted model serve as design-relevant rainfall thresholds. Consequently, uncertainty in the selected probability model, especially in its right tail, propagates into uncertainty in return-period estimates and associated spatial risk assessments.

The statistical modeling of annual maximum rainfall has long been central to hydrology and climatology. [26] showed, using more than 15,000 global rainfall records, that the distributional behavior of annual maximum daily rainfall is sensitive to tail specification and sample length, and that naive adoption of a single extreme-value type can substantially distort inferred return periods. [30] similarly emphasized that rainfall frequency estimation remains methodologically diverse across countries, with no universally dominant procedure, particularly for long return periods estimated from relatively short observational records. These concerns are directly relevant to regional rainfall studies, where annual maxima are often limited to a few decades and where design quantiles must nevertheless be estimated for return periods exceeding the available record length.

Distribution selection is therefore not a peripheral step but a substantive component of extreme-rainfall inference. Studies of design rainfall have routinely compared families of distributions, such as the generalized extreme value (GEV), Gumbel, Weibull, Gamma, Lognormal, Pearson Type III, and Log-Pearson Type III, using goodness-of-fit tests and information criteria [2,8,20]. [2], for example, examined annual maximum rainfall in Qatar using multiple candidate distributions and goodness-of-fit tests, while [3] used graphical and numerical measures, including RMSE, AIC, and BIC, to identify probability models for annual maximum daily rainfall under current climatic conditions. [37] further demonstrated that goodness-of-fit tests differ in power and sensitivity across distributional settings, underscoring the need to view model comparison as multi-criteria rather than mechanical. Taken together, these studies justify comparative fitting frameworks in which likelihood-based indices, empirical-distribution diagnostics, and visual checks are interpreted jointly.

Recent work also indicates that the traditional three-parameter models used in rainfall frequency analysis may be insufficient in some climatic regimes. [24] showed that exploiting peaks-over-threshold information through a flexible Kappa distribution can improve rare quantile estimation relative to annual-maxima approaches in some settings. [29] likewise argued, using a global data set of annual maxima, that more flexible four-parameter models can better represent heavy-tail behavior where conventional GEV or Pearson-type models are constrained. These findings support the broader statistical motivation for developing and evaluating flexible probability distributions whose additional structure may improve tail representation without sacrificing interpretability.

The present article is positioned within that theoretical and applied context. Its theoretical contribution examines the Power Generalized Weibull Mixture of Chi-Square Distribution (PGWMCD), a model that combines the flexibility of the power generalized Weibull family with a chi-square mixture. Generalized Weibull-type models have been developed precisely to expand the range of attainable distributional and hazard-rate shapes [5,28]. In parallel, mixture representations involving chi-square distributions remain an active area of probability theory because they provide tractable mechanisms for inducing additional heterogeneity and tail flexibility [15]. A closely related development is the Gamma mixture of chi-square distribution of [17], where chi-square mixtures were used to model leptokurtic and right-skewed data and were applied to precipitation data in the Caraga Region. The present study extends this line of work by using a power generalized Weibull mixing mechanism and by linking the resulting distribution to rainfall return-level estimation and spatial mapping. For rainfall extremes, this flexibility is attractive because annual maximum precipitation may exhibit skewness, tail heaviness, and province-specific variation that are not always captured adequately by standard two- or three-parameter models.

The applied contribution of the article is to evaluate the practical usefulness of the PGWMCD in rainfall return-period analysis for the Caraga Region of the Philippines. The Philippines is a suitable setting because extreme rainfall is a recurrent and consequential hazard in the archipelago. [34] used GEV modeling of long-term station records to show that extreme rainfall in the Philippines has changed in relation to global mean temperature and ENSO variability. Subsequent studies have emphasized the role of large-scale and intraseasonal circulation features in shaping Philippine rainfall extremes, including the boreal summer intraseasonal oscillation, the Madden–Julian oscillation, tropical cyclones, cold surges, and related multiscale interactions [22,23,39]. At the watershed scale, [32] showed that spatiotemporal analysis of extreme rainfall and drought can inform local water-resource management in the Philippines. These studies establish both the hazard relevance and the spatial heterogeneity of Philippine extreme precipitation.

Spatial representation is increasingly important because decision makers require not only point estimates but also interpretable regional patterns of return-period rainfall. [9] provided an early statistical framework for mapping precipitation return levels and their uncertainty, while [14] reviewed the broader development of rainfall spatial estimation from interpolation to multi-source merging. [10]

specifically compared approaches for estimating extreme precipitation of any return period at ungauged locations and showed that the method used to spatialize extremes can materially affect mapped design values. More recently, [27] demonstrated that spatial GEV models supported by high-resolution climate information can improve the representation of 10-year and 100-year rainfall return levels. These works support the mapping component of the present article, where fitted return levels are translated into spatial rainfall-risk surfaces over the real provincial geometry of Caraga.

Accordingly, this study integrates distribution theory and applied spatial hydrology. First, it fits the PGWMCD alongside GEV, Weibull, Gamma, Lognormal, and Log-Pearson Type III models to annual maximum daily precipitation for the five provinces of Caraga. Second, it compares the fitted models using RMSE, Anderson–Darling goodness-of-fit diagnostics, and visual PDF–histogram overlays. Third, it computes rainfall return levels from the fitted CDFs using $x_T = F^{-1}(1 - 1/T)$, selecting the distribution with the smallest RMSE for the return-level application. Finally, it maps the resulting return-period rainfall estimates over the Caraga provinces to provide a spatially interpretable application of the theoretical model. By aligning the PGWMCD's distributional development with return-period estimation and real-map visualization, the article demonstrates how a newly derived probability model can be evaluated not only as a mathematical object but also as a tool for regional extreme-rainfall assessment.

2. Materials and Methods

This section describes the mathematical and empirical procedures used to evaluate the Power Generalized Weibull Mixture of Chi-Square Distribution (PGWMCD) as a model for extreme rainfall. The analysis proceeds in two connected stages. First, the distributional framework is established through mixture theory and the component distributions entering the PGWMCD. Second, the proposed model is applied to annual maximum daily precipitation in the Caraga Region and compared with standard probability models commonly used in rainfall frequency analysis.

2.1. **Mixture Distribution Framework.** Mixture distributions arise when the parameter of a baseline family varies according to another probability law. This construction is widely used for modeling heterogeneous populations and data-generating mechanisms that cannot be adequately represented by a single homogeneous distribution [18,33,38]. Let $f(x;\theta)$ be a density function indexed by a parameter $\theta \in \Theta$, and let $g(\theta)$ be the mixing density of Θ. Following the general mixture construction discussed by [36], the corresponding continuous mixture density is

$$h(x) = \int_{\Theta} f(x;\theta)g(\theta)\, d\theta, \tag{2.1}$$

provided that the integral exists and $h(x) \geq 0$ for all x in the support. If the mixing distribution is discrete with support points $\theta_1, \theta_2, \ldots, \theta_p$ and weights $w_1, w_2, \ldots, w_p$, the corresponding finite mixture is

$$h(x) = \sum_{k=1}^{p} w_k f(x;\theta_k), \qquad w_k \geq 0, \quad \sum_{k=1}^{p} w_k = 1. \tag{2.2}$$

This construction is useful in rainfall modeling because annual maximum precipitation may reflect heterogeneous generating mechanisms. A mixture model can therefore introduce additional flexibility in skewness, tail weight, and shape while retaining a probability-density interpretation [15,18].

2.2. **Component Distributions.** The power generalized Weibull distribution is used as the mixing distribution in the proposed model. This family extends the classical Weibull model [35] and provides additional shape flexibility through the generalized Weibull construction [4,28]. For parameters $\alpha > 0$, $\beta > 0$, and $\lambda > 0$, its probability density function is given by

$$g(\theta;\alpha,\beta,\lambda) = \lambda\alpha\beta\theta^{\alpha-1}\left(1+\lambda\theta^{\alpha}\right)^{\beta-1}\exp\left\{1-\left(1+\lambda\theta^{\alpha}\right)^{\beta}\right\}, \qquad \theta > 0, \tag{2.3}$$

and its cumulative distribution function is

$$G(\theta;\alpha,\beta,\lambda) = 1-\exp\left\{1-\left(1+\lambda\theta^{\alpha}\right)^{\beta}\right\}, \qquad \theta > 0. \tag{2.4}$$

Here, α and β are shape parameters and λ is a scale parameter. The additional shape parameter allows the model to accommodate distributional forms beyond the classical Weibull family [4,28].

The second component is the chi-square distribution, a nonnegative distribution that is fundamental in statistical inference and is commonly expressed through the Gamma function [7,12]. If X follows a chi-square distribution with $\nu > 0$ degrees of freedom, then its density is

$$f(x;\nu) = \frac{x^{\nu/2-1}\exp(-x/2)}{2^{\nu/2}\Gamma(\nu/2)}, \qquad x > 0. \tag{2.5}$$

The chi-square family is appropriate in the present construction because it is nonnegative, right-skewed, and analytically tractable through the Gamma function.

Combining Equation (2.3) and Equation (2.5) through the continuous mixture principle in Equation (2.1), the PGWMCD density may be expressed as

$$\begin{aligned} f_{\text{PGWMCD}}(x;\alpha,\beta,\lambda,\nu) = \int_0^{\infty} &\lambda\alpha\beta\theta^{\alpha-1}\left(1+\lambda\theta^{\alpha}\right)^{\beta-1}\exp\left\{1-\left(1+\lambda\theta^{\alpha}\right)^{\beta}\right\} \\ &\times \frac{x^{\nu/2+\theta-1}\exp(-x/2)}{2^{\nu/2+\theta}\Gamma(\nu/2+\theta)}\,d\theta, \quad x > 0. \end{aligned} \tag{2.6}$$

The parameter vector of the PGWMCD is therefore

$$\boldsymbol{\Theta} = (\alpha,\beta,\lambda,\nu)^T,$$

where α, β, and λ govern the power generalized Weibull mixing mechanism and ν governs the chi-square component.

2.3. **Auxiliary Mathematical Formulae.** Several derivations in the succeeding section require standard identities involving the Gamma function. For $n \in \mathbb{N}$ and $x > 0$, the Pochhammer symbol is [40]

$$(x)_n = x(x+1)(x+2)\cdots(x+n-1),$$

which can be written as

$$(x)_n = \frac{\Gamma(x+n)}{\Gamma(x)}. \tag{2.7}$$

In addition, for $\mu > 0$ and $a > 0$, the standard Gamma integral is [40]

$$\int_0^\infty x^{a-1} e^{-\mu x}\, dx = \frac{\Gamma(a)}{\mu^a}. \tag{2.8}$$

These formulae are used in deriving the raw moments and related distributional properties of the PGWMCD.

2.4. **Study Area and Rainfall Data.** The empirical component of the study uses daily precipitation data for the five provinces of the Caraga Region, namely Agusan del Norte, Agusan del Sur, Dinagat Islands, Surigao del Norte, and Surigao del Sur. Daily corrected precipitation, denoted by the NASA POWER variable $\mathrm{PRECTOTCORR}$ and measured in mm day^{-1}, was obtained from the NASA Prediction Of Worldwide Energy Resources (POWER) Daily API, version 2.9.4, with MERRA2 as the source product [21]. The extraction covered January 1, 1981 to December 31, 2025 and used local solar time. For each province, the requested point was placed at a representative provincial capital or administrative center: Cabadbaran City for Agusan del Norte, Prosperidad for Agusan del Sur, San Jose for Dinagat Islands, Surigao City for Surigao del Norte, and Tandag City for Surigao del Sur. The corresponding latitude and longitude values were retained in the project metadata and used in the spatial interpolation. This period provides 45 annual observations for each province after converting the daily series into annual maxima.

Let $R_{p,y,d}$ denote the observed daily precipitation for province p, year y, and day d. The annual maximum daily precipitation series is defined as

$$X_{p,y} = \max_{d \in y} R_{p,y,d}, \qquad y = 1981, 1982, \ldots, 2025. \tag{2.9}$$

The resulting series $\{X_{p,y}\}$ is the basis for rainfall frequency analysis because return-period estimation concerns the upper tail of annual extreme rainfall rather than ordinary daily rainfall behavior.

2.5. **Candidate Models and Parameter Estimation.** For each province, the PGWMCD was fitted together with five benchmark distributions: the generalized extreme value (GEV), Weibull, Gamma, Lognormal, and Log-Pearson Type III distributions. These models were selected because they are widely used in hydrologic frequency analysis and provide natural comparators for assessing the practical value of the PGWMCD [2,8,20,26,30].

The parameters of all candidate distributions were estimated by numerical maximum likelihood estimation. For a candidate model j with density $f_j(x;\boldsymbol{\theta}_j)$, the log-likelihood based on the annual maximum sample $x_1, x_2, \ldots, x_n$ is

$$\ell_j(\boldsymbol{\theta}_j) = \sum_{i=1}^{n} \log f_j(x_i;\boldsymbol{\theta}_j), \tag{2.10}$$

and the maximum likelihood estimator is

$$\widehat{\boldsymbol{\theta}}_j = \arg\max_{\boldsymbol{\theta}_j \in \Omega_j} \ell_j(\boldsymbol{\theta}_j), \tag{2.11}$$

where Ω_j denotes the admissible parameter space of model j [6]. Equivalently, the computation was carried out by minimizing the negative log-likelihood. For the PGWMCD, the parameter vector $(\alpha, \beta, \lambda, \nu)^T$ was estimated numerically subject to $\alpha > 0$, $\beta > 0$, $\lambda > 0$, and $\nu > 0$. For the Weibull, Gamma, and Lognormal distributions, two-parameter positive-support versions were fitted with the location fixed at zero. The GEV distribution was fitted as a three-parameter model, whereas the Log-Pearson Type III distribution was fitted by applying a Pearson Type III model to $Y = \log_{10}(X)$. Thus, the reported parameter values are likelihood-based numerical estimates, not method-of-moments estimates.

Let $x_{(1)} \leq x_{(2)} \leq \cdots \leq x_{(n)}$ be the ordered annual maximum observations for a given province, where $n = 45$. For a candidate model with fitted cumulative distribution function $\widehat{F}_j$, the empirical plotting position assigned to $x_{(i)}$ is

$$\widehat{P}_i = \frac{i - 0.5}{n}, \qquad i = 1, 2, \ldots, n. \tag{2.12}$$

The goodness of fit was primarily evaluated using the CDF-based root mean square error, a discrepancy measure between the fitted distribution function and the empirical plotting positions. Similar numerical fit measures are commonly used in rainfall frequency-model comparison [2,3,20]. Here, j indexes the candidate distribution being evaluated; in this study, j ranges over the PGWMCD, GEV, Weibull, Gamma, Lognormal, and Log-Pearson Type III models. Thus,

$$\mathrm{RMSE}_j = \left[\frac{1}{n} \sum_{i=1}^{n} \left\{ \widehat{F}_j(x_{(i)}) - \widehat{P}_i \right\}^2 \right]^{1/2}. \tag{2.13}$$

The preferred model for each province is the candidate distribution with the smallest value of RMSE_j. This criterion was used as the principal basis for model comparison because it directly measures the discrepancy between the fitted distribution function and the empirical distribution of the annual maxima. It is important to note that RMSE_j was used only after numerical maximum likelihood estimation; it was not used as the parameter-estimation objective.

The Anderson–Darling statistic was also computed as a goodness-of-fit diagnostic because it gives relatively greater weight to tail behavior than many central-distribution measures. This is important in rainfall frequency analysis, where the upper tail controls return-level estimates [26,37]. Since the parameters were estimated from the same data used for testing, bootstrap p-values were used to provide a more appropriate reference distribution for the Anderson–Darling statistic.

2.6. **Rainfall Return Levels and Spatial Mapping.** After fitting the candidate distributions, rainfall return levels were obtained from the fitted cumulative distribution functions. This quantile-based definition is standard in rainfall frequency analysis and extreme-rainfall mapping [9,10,30]. For a return period T, the return level x_T is defined as

$$x_T = \widehat{F}^{-1}\left(1 - \frac{1}{T}\right), \tag{2.14}$$

where $\widehat{F}$ is the fitted CDF of the selected distribution. Thus, x_T represents the annual maximum daily rainfall level exceeded with probability $1/T$ in any given year. In this study, $\widehat{F}$ was taken from the best-fitting model for each province according to the smallest CDF-based RMSE.

The estimated return levels were subsequently mapped over the real provincial boundaries of Caraga to provide a spatial interpretation of rainfall hazard. The provincial estimates were assigned to their representative geographic locations and interpolated to form continuous return-period rainfall surfaces, consistent with the broader use of spatial rainfall estimation and return-level mapping in hydrologic risk studies [9,10,14,27]. The resulting maps allow comparison of the magnitude and spatial pattern of design rainfall across the five provinces and provide an applied visualization of how the fitted probability models translate into regional rainfall-risk information.

3. Main Results

3.1. **Theoretical Derivations.** This subsection presents the theoretical development of the Power Generalized Weibull Mixture of Chi-Square Distribution (PGWMCD). The construction follows directly from the continuous mixture framework in Equation (2.1), where the mixing density is the power generalized Weibull density in Equation (2.3) and the conditional density is the chi-square-type density in Equation (2.5). The derivations establish the validity of the proposed distribution, obtain its raw moments, and express its principal descriptive measures in terms of the model parameters. These results provide the mathematical basis for the applied rainfall return-period analysis to be presented in the succeeding subsection.

Definition 1. A nonnegative random variable X is said to follow the Power Generalized Weibull Mixture of Chi-Square Distribution with parameter vector

$$\boldsymbol{\Theta} = (\alpha, \beta, \lambda, \nu)^T, \qquad \alpha > 0, \quad \beta > 0, \quad \lambda > 0, \quad \nu > 0,$$

if its probability density function is given by

$$f_{\mathrm{PGWMCD}}(x;\alpha,\beta,\lambda,\nu) = \int_0^\infty \lambda\alpha\beta\theta^{\alpha-1}\left(1+\lambda\theta^{\alpha}\right)^{\beta-1} e^{\left\{1-(1+\lambda\theta^{\alpha})^{\beta}\right\}} \times \frac{x^{\nu/2+\theta-1}e^{-x/2}}{2^{\nu/2+\theta}\Gamma(\nu/2+\theta)}\,d\theta, \qquad x>0. \tag{3.1}$$

and $f_{\mathrm{PGWMCD}}(x;\alpha,\beta,\lambda,\nu)=0$ for $x \le 0$.

The construction in Definition 1 is a direct specialization of the continuous mixture framework in Equation (2.1). In particular, conditionally on $\Theta=\theta$, the random variable X has a Gamma-form chi-square density with shape parameter $\nu/2+\theta$ and scale parameter 2. Thus, the model preserves the nonnegative support and right-skewed structure of the chi-square family while introducing additional shape flexibility through the power generalized Weibull mixing distribution [18,33,36].

Theorem 1. The function $f_{\mathrm{PGWMCD}}(x;\alpha,\beta,\lambda,\nu)$ in Equation (3.1) is a valid probability density function.

Proof. For $x>0$, the integrand in Equation (3.1) is nonnegative. This follows from $\alpha,\beta,\lambda,\nu>0$, $\theta>0$, and $\Gamma(\nu/2+\theta)>0$. Hence,

$$f_{\mathrm{PGWMCD}}(x;\alpha,\beta,\lambda,\nu) \ge 0.$$

To verify normalization, integrate Equation (3.1) over the support of X. By Tonelli's theorem for nonnegative integrands,

$$\int_0^\infty f_{\mathrm{PGWMCD}}(x;\alpha,\beta,\lambda,\nu)\,dx = \int_0^\infty g(\theta;\alpha,\beta,\lambda)\left[\int_0^\infty \frac{x^{\nu/2+\theta-1}e^{-x/2}}{2^{\nu/2+\theta}\Gamma(\nu/2+\theta)}\,dx\right]d\theta,$$

where $g(\theta;\alpha,\beta,\lambda)$ is the density in Equation (2.3). Using the Gamma integral in Equation (2.8), with $a=\nu/2+\theta$ and $\mu=1/2$, the inner integral equals one. Therefore,

$$\int_0^\infty f_{\mathrm{PGWMCD}}(x;\alpha,\beta,\lambda,\nu)\,dx = \int_0^\infty g(\theta;\alpha,\beta,\lambda)\,d\theta = 1.$$

Thus, f_{PGWMCD} satisfies both nonnegativity and normalization, and is therefore a valid probability density function. □

Theorem 2. If $X \sim \mathrm{PGWMCD}(\alpha,\beta,\lambda,\nu)$, then the r-th raw moment of X, for any positive integer r, is

$$\mu_r' = 2^r \int_0^\infty \left(\frac{\nu}{2}+\theta\right)_r g(\theta;\alpha,\beta,\lambda)\,d\theta, \tag{3.2}$$

where $(a)_r$ denotes the Pochhammer symbol and $g(\theta;\alpha,\beta,\lambda)$ is given in Equation (2.3).

Proof. Conditionally on $\Theta=\theta$, X has Gamma shape $\nu/2+\theta$ and scale 2. Hence,

$$E(X^r \mid \Theta=\theta) = 2^r\frac{\Gamma\left(\frac{\nu}{2}+\theta+r\right)}{\Gamma\left(\frac{\nu}{2}+\theta\right)}. \tag{3.3}$$

Using the identity in Equation (2.7), Equation (3.3) becomes

$$E(X^r \mid \Theta = \theta) = 2^r \left(\frac{\nu}{2} + \theta\right)_r.$$

Taking expectation with respect to the mixing density $g(\theta; \alpha, \beta, \lambda)$ yields

$$E(X^r) = E\{E(X^r \mid \Theta)\} = 2^r \int_0^\infty \left(\frac{\nu}{2} + \theta\right)_r g(\theta; \alpha, \beta, \lambda)\, d\theta,$$

which proves Equation (3.2). □

For compact notation, let

$$m_s = E(\Theta^s), \qquad s = 1, 2, 3, 4, \tag{3.4}$$

where Θ follows the power generalized Weibull distribution in Equation (2.3). By applying the transformation $y = (1 + \lambda\theta^\alpha)^\beta$, together with the generalized binomial expansion and the Gamma integral in Equation (2.8), the moment m_s may be written as

$$m_s = \frac{\mathrm{e}}{\lambda^{s/\alpha}} A_s, \qquad A_s = \sum_{k=0}^{\infty} (-1)^k \binom{s/\alpha}{k} \Gamma\left(\frac{s - \alpha k + \alpha\beta}{\alpha\beta}, 1\right), \tag{3.5}$$

for parameter values under which the series is convergent. Here, $\Gamma(a, z)$ denotes the upper incomplete Gamma function.

Using Equation (3.2), the first four raw moments are obtained as follows:

$$\mu_1' = \nu + 2m_1, \tag{3.6}$$

$$\mu_2' = \nu^2 + 2\nu + 4(\nu + 1)m_1 + 4m_2, \tag{3.7}$$

$$\mu_3' = \nu^3 + 6\nu^2 + 8\nu + 2(3\nu^2 + 12\nu + 8)m_1 + 12(\nu + 2)m_2 + 8m_3, \tag{3.8}$$

$$\begin{aligned}\mu_4' &= \nu^4 + 12\nu^3 + 44\nu^2 + 48\nu + 8(\nu^3 + 9\nu^2 + 22\nu + 12)m_1 \\ &\quad + 8(3\nu^2 + 18\nu + 22)m_2 + 32(\nu + 3)m_3 + 16m_4.\end{aligned} \tag{3.9}$$

Substituting Equation (3.5) into Equations (3.6)–(3.9) gives the raw moments in terms of α, β, λ, ν, and the series coefficients A_s. Define also

$$\Delta_2 = m_2 - m_1^2, \tag{3.10}$$

$$\Delta_3 = m_3 - 3m_1m_2 + 2m_1^3, \tag{3.11}$$

$$\Delta_4 = m_4 - 4m_1m_3 - 3m_2^2 + 12m_1^2m_2 - 6m_1^4. \tag{3.12}$$

Here, Δ_2, Δ_3, and Δ_4 are the second, third, and fourth cumulant-type combinations of the mixing variable expressed entirely through A_s.

From Equation (3.6), the mean is

$$\mu = \nu + 2m_1 = \nu + \frac{2\mathrm{e}}{\lambda^{1/\alpha}} A_1. \tag{3.13}$$

Using Equation (3.7), the variance is

$$\sigma^2 = 2\nu + 4m_1 + 4\Delta_2 \tag{3.14}$$

or, equivalently,

$$\sigma^2 = 2\nu + \frac{4\mathrm{e}}{\lambda^{1/\alpha}} A_1 + \frac{4}{\lambda^{2/\alpha}} \left(\mathrm{e}A_2 - \mathrm{e}^2 A_1^2\right). \tag{3.15}$$

The third central moment is

$$\mu_3 = 8\nu + 16m_1 + 24\Delta_2 + 8\Delta_3. \tag{3.16}$$

In terms of A_s, Equation (3.16) becomes

$$\begin{aligned} \mu_3 = 8\nu &+ \frac{16\mathrm{e}}{\lambda^{1/\alpha}} A_1 + \frac{24}{\lambda^{2/\alpha}} \left(\mathrm{e}A_2 - \mathrm{e}^2 A_1^2\right) \\ &+ \frac{8}{\lambda^{3/\alpha}} \left(\mathrm{e}A_3 - 3\mathrm{e}^2 A_1 A_2 + 2\mathrm{e}^3 A_1^3\right). \end{aligned} \tag{3.17}$$

The fourth central moment is

$$\mu_4 = 48\nu + 96m_1 + 176\Delta_2 + 96\Delta_3 + 16\Delta_4 + 3\sigma^4. \tag{3.18}$$

Equivalently, substituting $m_s = \mathrm{e}\lambda^{-s/\alpha} A_s$,

$$\begin{aligned} \mu_4 = 48\nu &+ \frac{96\mathrm{e}}{\lambda^{1/\alpha}} A_1 + \frac{176}{\lambda^{2/\alpha}} \left(\mathrm{e}A_2 - \mathrm{e}^2 A_1^2\right) \\ &+ \frac{96}{\lambda^{3/\alpha}} \left(\mathrm{e}A_3 - 3\mathrm{e}^2 A_1 A_2 + 2\mathrm{e}^3 A_1^3\right) \\ &+ \frac{16}{\lambda^{4/\alpha}} \left(\mathrm{e}A_4 - 4\mathrm{e}^2 A_1 A_3 - 3\mathrm{e}^2 A_2^2 + 12\mathrm{e}^3 A_1^2 A_2 - 6\mathrm{e}^4 A_1^4\right) + 3\sigma^4. \end{aligned} \tag{3.19}$$

Therefore, the coefficient of skewness and the kurtosis coefficient are, respectively,

$$\gamma_1 = \frac{8\nu + 16m_1 + 24\Delta_2 + 8\Delta_3}{(2\nu + 4m_1 + 4\Delta_2)^{3/2}}, \tag{3.20}$$

$$\beta_2 = \frac{48\nu + 96m_1 + 176\Delta_2 + 96\Delta_3 + 16\Delta_4 + 3\sigma^4}{\sigma^4}. \tag{3.21}$$

The excess kurtosis is $\gamma_2 = \beta_2 - 3$. Thus, Equations (3.20)–(3.21) display the dependence of skewness and kurtosis on the PGWMCD parameters through the A_s-series terms.

Theorem 3. If $X \sim \mathrm{PGWMCD}(\alpha, \beta, \lambda, \nu)$, then its cumulative distribution function is

$$F_X(x) = \int_0^\infty \lambda\alpha\beta\theta^{\alpha-1}(1+\lambda\theta^\alpha)^{\beta-1} e^{1-(1+\lambda\theta^\alpha)^\beta} \frac{\gamma\left(\frac{\nu}{2}+\theta, \frac{x}{2}\right)}{\Gamma\left(\frac{\nu}{2}+\theta\right)}\, d\theta, \qquad x \geq 0, \tag{3.22}$$

where $\gamma(a, z)$ denotes the lower incomplete Gamma function.

Proof. By definition,

$$F_X(x) = \int_0^x f_{\text{PGWMCD}}(u; \alpha, \beta, \lambda, \nu)\, du.$$

Substituting Equation (3.1) and interchanging the order of integration yields

$$F_X(x) = \int_0^\infty g(\theta; \alpha, \beta, \lambda) \left[\int_0^x \frac{u^{\nu/2+\theta-1} e^{-u/2}}{2^{\nu/2+\theta}\Gamma(\nu/2+\theta)}\, du \right] d\theta.$$

The inner integral is the Gamma cumulative distribution function with shape $\nu/2 + \theta$ and scale 2, which is

$$\frac{\gamma\left(\frac{\nu}{2} + \theta, \frac{x}{2}\right)}{\Gamma\left(\frac{\nu}{2} + \theta\right)}.$$

Substitution of this expression into the mixture integral gives Equation (3.22). □

3.2. **Rainfall Return-Period Analysis.** The empirical application translates the distributional development of the PGWMCD into rainfall-frequency estimates for the five provinces of the Caraga Region. The annual maximum daily rainfall series were constructed from the daily precipitation records according to Equation (2.9), yielding $n = 45$ observations for each province from 1981 to 2025. Table 1 shows that the annual maxima are positively skewed in all provinces, which supports the use of right-tailed probability models for return-period estimation. Mean annual maximum daily rainfall ranged from 74.38 mm day^{-1} in Agusan del Sur to 98.94 mm day^{-1} in Agusan del Norte, whereas the largest observed annual maximum ranged from 176.39 mm day^{-1} in Agusan del Sur to 224.33 mm day^{-1} in Dinagat Islands and Surigao del Norte.

Table 1. Descriptive statistics of annual maximum daily rainfall in the Caraga Region, 1981–2025.

Province	n	Minimum	Mean	SD	Skewness	Kurtosis	Maximum
Agusan del Norte	45	45.39	98.94	34.49	0.71	2.97	182.72
Agusan del Sur	45	33.14	74.38	26.37	1.57	6.98	176.39
Dinagat Islands	45	44.00	97.93	41.13	1.01	4.10	224.33
Surigao del Norte	45	44.00	97.93	41.13	1.01	4.10	224.33
Surigao del Sur	45	31.44	96.28	41.83	1.01	4.09	221.58

The candidate distributions were first fitted by numerical maximum likelihood estimation according to Equation (2.10) and Equation (2.11). The fitted models were then compared using the CDF-based RMSE in Equation (2.13), where the empirical probabilities were assigned through the plotting-position formula in Equation (2.12). Table 2 reports the numerical maximum likelihood estimates, Anderson–Darling bootstrap goodness-of-fit results, and CDF-based RMSE values for all fitted candidate distributions. The rows are arranged by province and then from the smallest to the largest RMSE within each province; therefore, the bold row in each province block identifies the distribution selected for return-level estimation. The Lognormal distribution produced the smallest RMSE for Agusan del Norte, the Gamma distribution produced the smallest RMSE for Surigao del Sur, and the PGWMCD produced the smallest RMSE for Agusan del Sur, Dinagat Islands, and Surigao del Norte. Thus, the proposed distribution was selected in three of the five provinces under the principal model-comparison criterion. For the three provinces in which the PGWMCD was preferred, the fitted CDF in Equation (3.22) was used directly in the return-level calculation in Equation (2.14). For the remaining provinces, the fitted CDF of the selected benchmark distribution was used.

Table 2. Numerical maximum likelihood estimates, Anderson–Darling bootstrap test results, and CDF-based RMSE values for all fitted candidate distributions.

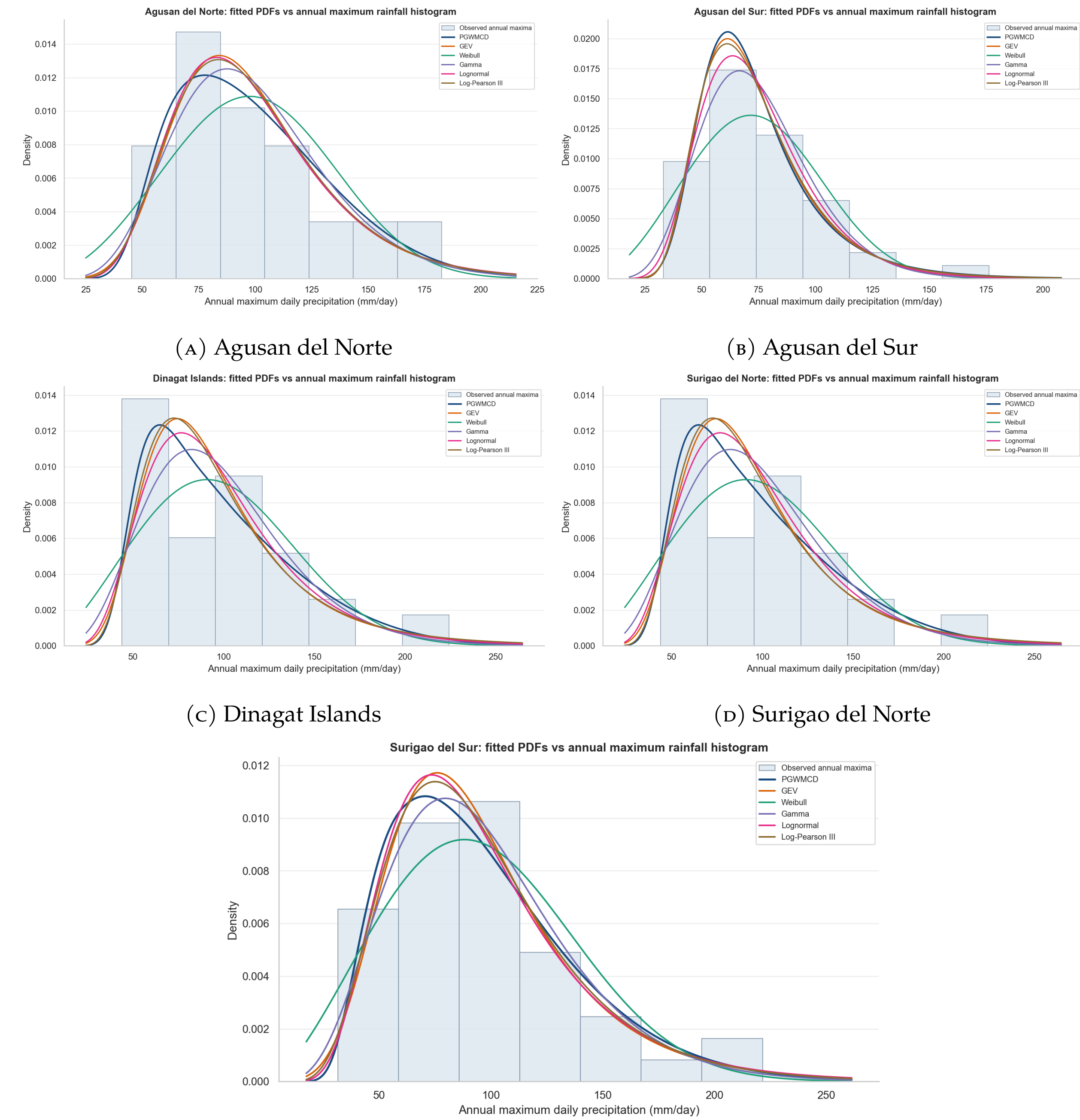
Province	Distribution	RMSE	AD statistic	AD bootstrap p	Parameter estimates
Agusan del Norte	**Lognormal**	**0.02495**	0.21335	0.980	$\widehat{\mu}_{\mathrm{LN}} = 4.53602,\ \widehat{\sigma}_{\mathrm{LN}} = 0.34351$
Agusan del Norte	Log-Pearson III	0.02518	0.21467	0.980	$\widehat{a}_{\mathrm{LP3}} = -0.00457,\ \widehat{b}_{\mathrm{LP3}} = 1066.45,\ \widehat{c}_{\mathrm{LP3}} = 6.84237$
Agusan del Norte	GEV	0.02537	0.22148	0.990	$\widehat{\mu}_{\mathrm{GEV}} = 83.4428,\ \widehat{\sigma}_{\mathrm{GEV}} = 27.6386,\ \widehat{\xi}_{\mathrm{GEV}} = -0.02427$
Agusan del Norte	PGWMCD	0.02765	0.23858	0.980	$\widehat{\alpha} = 1.13086,\ \widehat{\beta} = 5.17746,\ \widehat{\lambda} = 0.003517,\ \widehat{\nu} = 53.0787$
Agusan del Norte	Gamma	0.02917	0.26879	0.960	$\widehat{\alpha}_{\mathrm{G}} = 8.71848,\ \widehat{\beta}_{\mathrm{G}} = 0.08812$
Agusan del Norte	Weibull	0.04650	0.63388	0.647	$\widehat{\alpha}_{\mathrm{W}} = 110.740,\ \widehat{\beta}_{\mathrm{W}} = 3.08589$
Agusan del Sur	**PGWMCD**	**0.02260**	0.14058	0.995	$\widehat{\alpha} = 3.99372,\ \widehat{\beta} = 0.21890,\ \widehat{\lambda} = 0.000214,\ \widehat{\nu} = 38.5588$
Agusan del Sur	GEV	0.02340	0.14822	0.995	$\widehat{\mu}_{\mathrm{GEV}} = 62.4898,\ \widehat{\sigma}_{\mathrm{GEV}} = 18.4410,\ \widehat{\xi}_{\mathrm{GEV}} = 0.06404$
Agusan del Sur	Log-Pearson III	0.02423	0.16040	0.990	$\widehat{a}_{\mathrm{LP3}} = 0.01970,\ \widehat{b}_{\mathrm{LP3}} = 49.9088,\ \widehat{c}_{\mathrm{LP3}} = 0.86508$
Agusan del Sur	Lognormal	0.02947	0.22720	0.990	$\widehat{\mu}_{\mathrm{LN}} = 4.25565,\ \widehat{\sigma}_{\mathrm{LN}} = 0.32090$
Agusan del Sur	Gamma	0.04012	0.42032	0.856	$\widehat{\alpha}_{\mathrm{G}} = 9.50711,\ \widehat{\beta}_{\mathrm{G}} = 0.12782$
Agusan del Sur	Weibull	0.06577	1.24630	0.284	$\widehat{\alpha}_{\mathrm{W}} = 83.1973,\ \widehat{\beta}_{\mathrm{W}} = 2.87205$
Dinagat Islands	**PGWMCD**	**0.03486**	0.31993	0.945	$\widehat{\alpha} = 0.82432,\ \widehat{\beta} = 50.0000,\ \widehat{\lambda} = 0.000948,\ \widehat{\nu} = 51.2607$
Dinagat Islands	Gamma	0.03773	0.45680	0.781	$\widehat{\alpha}_{\mathrm{G}} = 6.28579,\ \widehat{\beta}_{\mathrm{G}} = 0.06418$
Dinagat Islands	Lognormal	0.04044	0.44944	0.791	$\widehat{\mu}_{\mathrm{LN}} = 4.50264,\ \widehat{\sigma}_{\mathrm{LN}} = 0.40307$
Dinagat Islands	Weibull	0.04153	0.64813	0.721	$\widehat{\alpha}_{\mathrm{W}} = 110.562,\ \widehat{\beta}_{\mathrm{W}} = 2.55038$
Dinagat Islands	Log-Pearson III	0.04507	0.50776	0.721	$\widehat{a}_{\mathrm{LP3}} = 0.02883,\ \widehat{b}_{\mathrm{LP3}} = 37.3037,\ \widehat{c}_{\mathrm{LP3}} = 0.87985$
Dinagat Islands	GEV	0.04571	0.53130	0.692	$\widehat{\mu}_{\mathrm{GEV}} = 77.6920,\ \widehat{\sigma}_{\mathrm{GEV}} = 29.1790,\ \widehat{\xi}_{\mathrm{GEV}} = 0.11193$
Surigao del Norte	**PGWMCD**	**0.03486**	0.31993	0.905	$\widehat{\alpha} = 0.82432,\ \widehat{\beta} = 50.0000,\ \widehat{\lambda} = 0.000948,\ \widehat{\nu} = 51.2607$
Surigao del Norte	Gamma	0.03773	0.45680	0.746	$\widehat{\alpha}_{\mathrm{G}} = 6.28579,\ \widehat{\beta}_{\mathrm{G}} = 0.06418$
Surigao del Norte	Lognormal	0.04044	0.44944	0.756	$\widehat{\mu}_{\mathrm{LN}} = 4.50264,\ \widehat{\sigma}_{\mathrm{LN}} = 0.40307$
Surigao del Norte	Weibull	0.04153	0.64813	0.662	$\widehat{\alpha}_{\mathrm{W}} = 110.562,\ \widehat{\beta}_{\mathrm{W}} = 2.55038$
Surigao del Norte	Log-Pearson III	0.04507	0.50776	0.786	$\widehat{a}_{\mathrm{LP3}} = 0.02883,\ \widehat{b}_{\mathrm{LP3}} = 37.3037,\ \widehat{c}_{\mathrm{LP3}} = 0.87985$
Surigao del Norte	GEV	0.04571	0.53130	0.711	$\widehat{\mu}_{\mathrm{GEV}} = 77.6920,\ \widehat{\sigma}_{\mathrm{GEV}} = 29.1790,\ \widehat{\xi}_{\mathrm{GEV}} = 0.11193$
Surigao del Sur	**Gamma**	**0.02962**	0.27472	0.955	$\widehat{\alpha}_{\mathrm{G}} = 5.76194,\ \widehat{\beta}_{\mathrm{G}} = 0.05985$
Surigao del Sur	PGWMCD	0.03077	0.24906	0.980	$\widehat{\alpha} = 1.45143,\ \widehat{\beta} = 1.14408,\ \widehat{\lambda} = 0.005360,\ \widehat{\nu} = 38.2879$
Surigao del Sur	Log-Pearson III	0.03128	0.25815	0.965	$\widehat{a}_{\mathrm{LP3}} = -0.01068,\ \widehat{b}_{\mathrm{LP3}} = 300.527,\ \widehat{c}_{\mathrm{LP3}} = 5.15441$
Surigao del Sur	GEV	0.03264	0.27985	0.955	$\widehat{\mu}_{\mathrm{GEV}} = 76.9516,\ \widehat{\sigma}_{\mathrm{GEV}} = 31.4076,\ \widehat{\xi}_{\mathrm{GEV}} = 0.03500$
Surigao del Sur	Lognormal	0.03317	0.27508	0.970	$\widehat{\mu}_{\mathrm{LN}} = 4.47797,\ \widehat{\sigma}_{\mathrm{LN}} = 0.42619$
Surigao del Sur	Weibull	0.04084	0.55601	0.662	$\widehat{\alpha}_{\mathrm{W}} = 108.788,\ \widehat{\beta}_{\mathrm{W}} = 2.46688$

Note. AD = Anderson–Darling. Rows are sorted by RMSE within each province. The bold row in each province block indicates the model selected under the CDF-based RMSE criterion in Equation (2.13). The AD statistic and bootstrap p-value are reported as supplementary goodness-of-fit diagnostics.

The Anderson–Darling bootstrap diagnostics in Table 2 provide supporting evidence that the RMSE-selected models are compatible with the annual maximum samples. All selected models yielded bootstrap p-values above 0.90, suggesting that the fitted distributions do not exhibit strong distributional conflict with the observed annual maxima under the bootstrap reference procedure. However, these diagnostics are interpreted as complementary checks only. The final model choice follows the RMSE criterion in Equation (2.13), because the objective of the application is to select the fitted CDF that most closely agrees with the empirical distribution of annual maxima.

Methodologically, these results are consistent with rainfall-frequency studies showing that the best-fitting probability law is often location-dependent and that no single distribution should be assumed to be uniformly optimal across rainfall stations or regions [2,20,26]. The selection of different models across the Caraga provinces therefore strengthens the case for a comparative fitting framework rather than an a priori choice of one standard model. At the same time, the strong performance of the PGWMCD in three provinces suggests that mixture-based flexibility can be useful when the annual maximum series exhibits skewness and tail behavior that may not be fully captured by simpler two- or three-parameter models. This interpretation agrees with broader rainfall-frequency practice, where model adequacy is commonly evaluated through a combination of fitted-distribution comparison, goodness-of-fit diagnostics, and hydrologic interpretability rather than through a single universal distributional rule [8,30,37].

Figure 1 provides the corresponding visual comparison of the fitted PDFs against the empirical histograms. The bars show the observed annual maximum rainfall values, whereas the curves show the fitted density functions implied by the numerical MLEs in Table 2. In Agusan del Norte, the Lognormal, Log-Pearson III, GEV, and PGWMCD curves are visually close over the central portion of the sample, which is consistent with their similar RMSE values. In Agusan del Sur, Dinagat Islands, and Surigao del Norte, the PGWMCD provides the smallest RMSE, whereas in Surigao del Sur the Gamma model is preferred and the PGWMCD remains a close competitor. The fitted-density plots are interpreted as graphical support only; the formal model ranking follows the CDF-based RMSE in Equation (2.13).



(A) Agusan del Norte (B) Agusan del Sur

(C) Dinagat Islands (D) Surigao del Norte

(E) Surigao del Sur

FIGURE 1. Histograms of annual maximum daily rainfall with fitted probability density functions for the six candidate distributions.

Using the selected fitted distributions, rainfall return levels were computed from Equation (2.14) for return periods of 2, 5, 10, 25, 50, and 100 years. The estimates in Table 3 increase monotonically with the return period, as expected from the quantile definition. Agusan del Sur has the lowest estimated return levels throughout the range of return periods, increasing from 69.08 mm day^{-1} for the 2-year event to 163.41 mm day^{-1} for the 100-year event. The largest 100-year estimates occur in Dinagat Islands and

Surigao del Norte, both equal to 218.57 mm day^{-1}, followed by Surigao del Sur at 213.12 mm day^{-1} and Agusan del Norte at 207.50 mm day^{-1}. In the processed NASA POWER extraction, Dinagat Islands and Surigao del Norte produced identical annual maximum series; consequently, their selected model, fitted parameters, and return-level estimates are also identical.

Table 3. Estimated rainfall return levels from the RMSE-selected model for each province.

Province	Selected model	2-year	5-year	10-year	25-year	50-year	100-year
Agusan del Norte	Lognormal	93.32	124.60	144.93	170.27	188.95	207.50
Agusan del Sur	PGWMCD	69.08	91.17	107.33	129.18	146.13	163.41
Dinagat Islands	PGWMCD	88.78	130.36	155.62	183.67	202.03	218.57
Surigao del Norte	PGWMCD	88.78	130.36	155.62	183.67	202.03	218.57
Surigao del Sur	Gamma	90.77	127.42	149.92	176.57	195.27	213.12

Note. Values are in mm day^{-1}. The selected model is the fitted distribution with the smallest CDF-based RMSE in Equation (2.13).

The spatial interpolation in Figure 2 provides a geographic synthesis of the province-level frequency estimates. The plotted surfaces represent estimated return levels in mm day^{-1} and were generated from the five representative NASA POWER points described in Subsection 2.4. For shorter return periods, the estimated design rainfall is relatively moderate over the region, with lower values centered around Agusan del Sur. As the return period increases, the return-level surface intensifies, especially over the northern and eastern portions of Caraga. The 50-year and 100-year panels show higher return-level rainfall over Dinagat Islands, Surigao del Norte, and Surigao del Sur than over the interior Agusan del Sur point. This spatial pattern is consistent with the fitted return-level estimates in Table 3 and suggests that the coastal and island provinces may require greater attention in design-rainfall and rainfall-risk interpretation.

The mapping component also has an important methodological implication. Previous spatial studies of extreme precipitation emphasize that return-level mapping is valuable because fitted quantiles are easier to interpret for planning and risk communication than distributional parameters alone [9,10,27]. In the present application, the map serves this same interpretive role by converting the RMSE-selected fitted CDFs into province-level design-rainfall surfaces. However, unlike fully spatial extreme-value models, the present map is an interpolated visualization from five representative provincial points. Hence, its contribution is primarily comparative and regional: it identifies relative spatial contrasts in estimated return levels across Caraga, while more densely distributed gauges or gridded validation data would be required for site-specific engineering design.

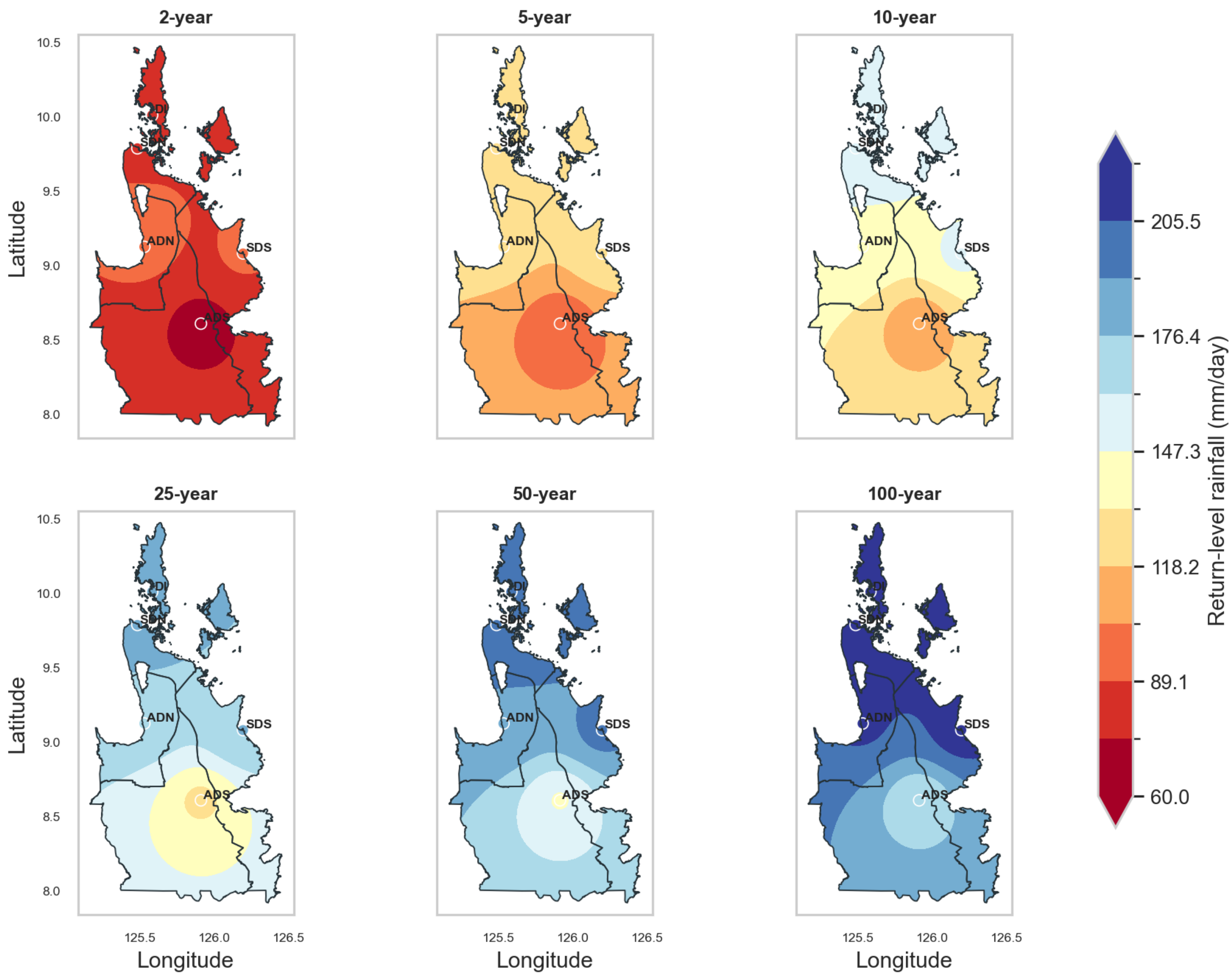


FIGURE 2. Spatial map of estimated rainfall return levels in the Caraga Region based on the RMSE-selected model for each province. Values are expressed in mm day^{-1}. Province labels indicate Agusan del Norte (ADN), Agusan del Sur (ADS), Dinagat Islands (DI), Surigao del Norte (SDN), and Surigao del Sur (SDS).

The return-period results should be interpreted as model-based estimates rather than direct observations. Because return levels are fitted quantiles, they depend on both the selected distribution and the available 45-year annual maximum record. The use of the RMSE criterion favors overall CDF agreement with empirical plotting positions, whereas very long return periods are governed by upper-tail behavior. Therefore, the 50-year and 100-year estimates should be read as extrapolations from the fitted models. Similarly, the mapped surfaces summarize spatial variation from five representative provincial points and should be treated as regional-scale visualizations rather than site-specific engineering design values. Within these limitations, the application demonstrates that the PGWMCD can be connected to an operational rainfall-frequency workflow: it can be fitted to annual maximum daily precipitation, compared against established hydrologic distributions using Equation (2.13), and converted through Equation (2.14) into interpretable return-period rainfall maps.

4. Conclusion

This article developed and applied the Power Generalized Weibull Mixture of Chi-Square Distribution (PGWMCD) as a flexible model for nonnegative and right-skewed data, with particular emphasis on annual maximum daily rainfall. The theoretical component defined the PGWMCD in Definition 1 through the continuous mixture construction in Equation (2.1), combining the power generalized Weibull mixing density in Equation (2.3) with the chi-square-type conditional density in Equation (2.5). The resulting density in Equation (3.1) was shown in Theorem 1 to be a valid probability density function, while the raw moment representation and cumulative distribution function were derived in Theorem 2 and Theorem 3, respectively. These results demonstrate that the proposed model has a coherent probabilistic foundation and that its distributional shape is governed jointly by the chi-square component and the power generalized Weibull mixing mechanism.

The empirical component connected these theoretical results to rainfall return-period analysis for the five provinces of the Caraga Region using annual maximum daily precipitation from 1981 to 2025. The PGWMCD was fitted together with the GEV, Weibull, Gamma, Lognormal, and Log-Pearson Type III distributions through numerical maximum likelihood estimation as described in Equation (2.10) and Equation (2.11). Model comparison was based principally on the CDF-based RMSE in Equation (2.13), with Anderson–Darling bootstrap diagnostics used as supplementary goodness-of-fit evidence. The results showed that the PGWMCD achieved the smallest RMSE in Agusan del Sur, Dinagat Islands, and Surigao del Norte, whereas the Lognormal model was preferred in Agusan del Norte and the Gamma model was preferred in Surigao del Sur. Hence, the proposed distribution was not uniformly dominant, but it performed competitively and was selected in three of the five provincial applications.

The return-period analysis further illustrated the practical value of the fitted models. Using the quantile relationship in Equation (2.14), the RMSE-selected distributions were converted into 2-, 5-, 10-, 25-, 50-, and 100-year rainfall return levels. The estimated return levels increased monotonically with return period, as expected, and revealed clear spatial contrasts across the region. Agusan del Sur consistently produced the lowest estimated return levels, whereas Dinagat Islands and Surigao del Norte produced the largest 100-year estimates. The spatial map of return-period rainfall translated these fitted quantiles into an interpretable regional surface, indicating greater design-rainfall magnitudes over the northern and eastern portions of Caraga than over the interior Agusan del Sur point.

These findings have two main implications. First, the derivations show that the PGWMCD extends the chi-square mixture framework in a mathematically tractable way, while allowing additional flexibility in skewness and tail behavior through the mixing distribution. Second, the rainfall application shows that a newly proposed distribution can be evaluated within a complete hydrologic workflow: construction of annual maxima, numerical likelihood-based estimation, RMSE-based model comparison, supplementary goodness-of-fit checking, return-level computation, and spatial visualization. This integration is important because distributional flexibility alone is insufficient for applied rainfall-frequency work; the fitted model must also provide empirical agreement with observed annual maxima and yield interpretable return-period estimates.

Several limitations should be acknowledged. The rainfall application used a 45-year annual maximum record for each province, so the 50-year and 100-year return levels necessarily involve upper-tail extrapolation beyond much of the observed record. The mapped return-period surfaces were also generated from five representative provincial estimates and should therefore be interpreted as regional-scale visualizations rather than site-specific engineering design values. In addition, the identical processed annual maximum series for Dinagat Islands and Surigao del Norte led to identical fitted parameters and return-level estimates for these two provinces. Future work may extend the analysis by incorporating additional ground-station observations, quantifying return-level uncertainty through bootstrap or Bayesian procedures, comparing the PGWMCD under alternative estimation criteria, and embedding the model within a fully spatial extreme-value framework. Within these qualifications, the results support the PGWMCD as a theoretically valid and empirically useful candidate distribution for rainfall return-period modeling and related applications involving nonnegative skewed extremes.

Conflicts of Interest. The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

[1] A.Z. Afify, D. Kumar, I. Elbatal, Marshall–Olkin Power Generalized Weibull Distribution with Applications in Engineering and Medicine, J. Stat. Theory Appl. 19 (2020), 223–237. https://doi.org/10.2991/jsta.d.200507.004.

[2] A. Al Mamoon, A. Rahman, Selection of the Best Fit Probability Distribution in Rainfall Frequency Analysis for Qatar, Nat. Hazards 86 (2017), 281–296. https://doi.org/10.1007/s11069-016-2687-0.

[3] V.M. Alofa, M.B. Hounsou, G.-D. Houeffa, Y. N'tcha M'po, D.H. Ahoton, E. Vissin, E. Agbossou, Selection of a Probability Model Adapted to the Current Climate for Annual Maximum Daily Rainfall in the Benin Mono-Couffo Basin (West Africa), Hydrology 12 (2025), 86. https://doi.org/10.3390/hydrology12040086.

[4] V. Bagdonavicius, M. Nikulin, Accelerated Life Models, Chapman and Hall/CRC, 2001. https://doi.org/10.1201/9781420035872.

[5] M. Bourguignon, R.B. Silva, G.M. Cordeiro, The Weibull-G Family of Probability Distributions, J. Data Sci. 12 (2014), 53–68. https://doi.org/10.6339/jds.201401_12(1).0004.

[6] G. Casella, R.L. Berger, Statistical Inference, Duxbury Press, 2002.

[7] R. Chattamvelli, R. Shanmugam, Central Chi-Squared Distribution, in: Continuous Distributions in Engineering and the Applied Sciences – Part II, Synthesis Lectures on Mathematics and Statistics, Springer, Cham, pp. 201–214, (2021). https://doi.org/10.1007/978-3-031-02435-1_5.

[8] M. Ciupak, B. Ozga-Zieliński, T. Tokarczyk, J. Adamowski, A Probabilistic Model for Maximum Rainfall Frequency Analysis, Water 13 (2021), 2688. https://doi.org/10.3390/w13192688.

[9] D. Cooley, D. Nychka, P. Naveau, Bayesian Spatial Modeling of Extreme Precipitation Return Levels, J. Am. Stat. Assoc. 102 (2007), 824–840. https://doi.org/10.1198/016214506000000780.

[10] S. Das, D. Zhu, Y. Yin, Comparison of Mapping Approaches for Estimating Extreme Precipitation of Any Return Period at Ungauged Locations, Stoch. Environ. Res. Risk Assess. 34 (2020), 1175–1196. https://doi.org/10.1007/s00477-020-01828-7.

[11] A.P. Dempster, N.M. Laird, D.B. Rubin, Maximum Likelihood from Incomplete Data via the EM Algorithm, J. R. Stat. Soc. Ser. B Stat. Methodol. 39 (1977), 1–22. https://doi.org/10.1111/j.2517-6161.1977.tb01600.x.

[12] C.C. Heyde, E. Seneta, I. J. Bienaymé: Statistical Theory Anticipated, Springer, New York, 1977. https://doi.org/10.1007/978-1-4684-9469-3.

[13] D. Hlubinka, Š. Hudecová, M. Maciak, M. Pešta, Asymptotic and Methodological Statistics: Festschrift in Honor of Marie Hušková, Springer, 2026. https://doi.org/10.1007/978-3-032-07178-1.

[14] Q. Hu, Z. Li, L. Wang, Y. Huang, Y. Wang, L. Li, Rainfall Spatial Estimations: A Review from Spatial Interpolation to Multi-Source Data Merging, Water 11 (2019), 579. https://doi.org/10.3390/w11030579.

[15] M.C. Jones, É. Marchand, A (Non-Central) Chi-Squared Mixture of Non-Central Chi-Squareds Is (Non-Central) Chi-Squared and Related Results, Corollaries and Applications, Stat 10 (2021), e398. https://doi.org/10.1002/sta4.398.

[16] H.S. Klakattawi, Survival Analysis of Cancer Patients Using a New Extended Weibull Distribution, PLoS One 17 (2022), e0264229. https://doi.org/10.1371/journal.pone.0264229.

[17] J.T. Lingo, J.V.B. Deluao, M.O. Macalos, Gamma Mixture of Chi-Square Distribution: Properties, Estimation, Simulation, and Application, Asia Pac. J. Math. 12 (2025), 111. https://doi.org/10.28924/APJM/12-111.

[18] G. McLachlan, D. Peel, Finite Mixture Models, Wiley, 2000. https://doi.org/10.1002/0471721182.

[19] W.Q. Meeker, L.A. Escobar, F.G. Pascual, Statistical Methods for Reliability Data, Wiley, 2021.

[20] R. Montes-Pajuelo, A.M. Rodríguez-Pérez, R. López, C.A. Rodríguez, Analysis of Probability Distributions for Modelling Extreme Rainfall Events and Detecting Climate Change: Insights from Mathematical and Statistical Methods, Mathematics 12 (2024), 1093. https://doi.org/10.3390/math12071093.

[21] NASA POWER Project, NASA Prediction of Worldwide Energy Resources (POWER) Daily API, NASA Langley Research Center, 2026.

[22] L.M.P. Olaguera, J.A. Manalo, J. Matsumoto, Influence of Boreal Summer Intraseasonal Oscillation on Rainfall Extremes in the Philippines, Int. J. Climatol. 42 (2022), 4656–4668. https://doi.org/10.1002/joc.7495.

[23] L.M.P. Olaguera, J.A. Manalo, A. Bathan, J. Matsumoto, Quantifying the Influence of the Madden–Julian Oscillation on Rainfall Extremes During the Northeast Monsoon Season of the Philippines, Atmos. Sci. Lett. 25 (2024), e1232. https://doi.org/10.1002/asl.1232.

[24] D. O'Shea, R. Nathan, A. Sharma, C. Wasko, Improved Extreme Rainfall Frequency Analysis Using a Two-Step Kappa Approach, Water Resour. Res. 59 (2023), e2021WR031854. https://doi.org/10.1029/2021WR031854.

[25] H.H. Panjer, Operational Risk, Wiley, 2006. https://doi.org/10.1002/0470051310.

[26] S.M. Papalexiou, D. Koutsoyiannis, Battle of Extreme Value Distributions: A Global Survey on Extreme Daily Rainfall, Water Resour. Res. 49 (2013), 187–201. https://doi.org/10.1029/2012WR012557.

[27] B. Poschlod, J. Koh, Convection-Permitting Climate Models Can Support Observations to Generate Rainfall Return Levels, Water Resour. Res. 60 (2024), e2023WR035159. https://doi.org/10.1029/2023WR035159.

[28] M.A. Selim, The Generalized Power Generalized Weibull Distribution: Properties and Applications, arXiv:1807.10763, 2018. https://doi.org/10.48550/arXiv.1807.10763.

[29] R. Strong, O. Borgstroem, R. Nathan, C. Wasko, D. O'Shea, Global Applicability of the Kappa Distribution for Rainfall Frequency Analysis, Water Resour. Res. 61 (2025), e2024WR039035. https://doi.org/10.1029/2024WR039035.

[30] C. Svensson, D.A. Jones, Review of Rainfall Frequency Estimation Methods, J. Flood Risk Manag. 3 (2010), 296–313. https://doi.org/10.1111/j.1753-318X.2010.01079.x.

[31] H. Teimourian, M. Abubakar, M. Yildiz, A. Teimourian, A Comparative Study on Wind Energy Assessment Distribution Models: A Case Study on Weibull Distribution, Energies 15 (2022), 5684. https://doi.org/10.3390/en15155684.

[32] A.T. Tejada Jr., P.A.J. Sanchez, F.J.F. Faderogao, C.B. Gigantone, R.A. Luyun Jr., Spatiotemporal Analysis of Extreme Rainfall and Meteorological Drought Events over the Angat Watershed, Philippines, Atmosphere 14 (2023), 1790. https://doi.org/10.3390/atmos14121790.

[33] D.M. Titterington, Some Recent Research in the Analysis of Mixture Distributions, Statistics 21 (1990), 619–641. https://doi.org/10.1080/02331889008802274.

[34] M.Q. Villafuerte, J. Matsumoto, H. Kubota, Changes in Extreme Rainfall in the Philippines (1911–2010) Linked to Global Mean Temperature and ENSO, Int. J. Climatol. 35 (2015), 2033–2044. https://doi.org/10.1002/joc.4105.

[35] W. Weibull, A Statistical Distribution Function of Wide Applicability, J. Appl. Mech. 18 (1951), 293–297. https://doi.org/10.1115/1.4010337.

[36] M.R. Zaman, D.N.R. Paul, N. Akhter, M.H. Howlader, M.S. Kabir, Chi-Square Mixture of Chi-Square Distributions, J. Appl. Sci. 6 (2006), 243–246. https://doi.org/10.3923/jas.2006.243.246.

[37] X. Zeng, D. Wang, J. Wu, Evaluating the Three Methods of Goodness of Fit Test for Frequency Analysis, J. Risk Anal. Crisis Response 5 (2015), 178–187. https://doi.org/10.2991/jrarc.2015.5.3.5.

[38] H. Zhang, Y. Huang, Finite Mixture Models and Their Applications: A Review, Austin Biom. Biostat. 2 (2015), 1013.

[39] W.-J. Zhang, R. Wu, P. Dai, Q. Gu, What Causes Extreme Precipitation in the Southern Philippines During El Niño Winters?, J. Clim. 38 (2025), 2047–2061. https://doi.org/10.1175/JCLI-D-24-0374.1.

[40] I.S. Gradshteyn, I.M. Ryzhik, Table of Integrals, Series, and Products, Academic Press, 2000.